LIANG CHENG
DAVID Y. ZHANG
JOHN N. EBLE
EDITORS

Molecular Genetic Pathology

Springer Reference

Molecular Genetic Pathology

Liang Cheng
David Y. Zhang
John N. Eble
Editors

Molecular Genetic Pathology

Second Edition

Springer Reference

Editors
Liang Cheng
Department of Pathology and Laboratory Medicine
Indiana University School of Medicine
Indianapolis, IN, USA

David Y. Zhang
Department of Pathology
Mount Sinai School of Medicine
New York, NY, USA

John N. Eble
Department of Pathology and Laboratory Medicine
Indiana University School of Medicine
Indianapolis, IN, USA

ISBN 978-1-4614-4799-3 ISBN 978-1-4614-4800-6 (eBook)
ISBN 978-1-4614-4801-3 (print and electronic bundle)
DOI 10.1007/978-1-4614-4800-6
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012952466

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Since the first edition of this textbook, the field of molecular genetic pathology has expanded exponentially. The advent of complete cancer genome sequencing, emerging tests for identification of infectious agents, new diagnostic molecular tests, and biomarkers for targeted therapy have demonstrated the importance of molecular tools both in patient care and in the practice of pathology. Furthermore, the development of many new technologies, including various microarray platforms, nanotechnology, and next-generation sequencing, has revolutionized the practice of molecular genetic pathology. The rapid growth of this field has led to an expansion of knowledge of molecular processes and many more clinical applications in molecular diagnostics. Thus, we recognize the need to update this textbook and hope that the second edition of the *Molecular Genetic Pathology* will continue to be the best available "quick reference" for pathologists, oncologists, geneticists, primary care physicians, and other medical professionals with an interest in this specialty. In addition, it is our continued belief that a concise overview of the field of molecular genetic pathology, emphasizing its clinical applications, will help not only practicing physicians, but also those in training, including residents and fellows in various disciplines.

The first edition of the *Molecular Genetic Pathology* has received much positive feedback from its readers. The two main features of the book commended by readers are the concise bullet-point format and the many useful charts and figures. These remain in the second edition. It should also be noted that over 100 leading experts have contributed to this new edition. The second edition of the *Molecular Genetic Pathology* is composed of seven sections, and 46 chapters. Sections 1 and 2 explore the basic principles and methodology of molecular genetic techniques. Sections 3 through 7 focus on disease-based information, including genetic disorders, molecular microbiology, hematologic malignancy assessment, diagnosis of hematological disorders, specialized applications of molecular tools, and management of the molecular laboratory.

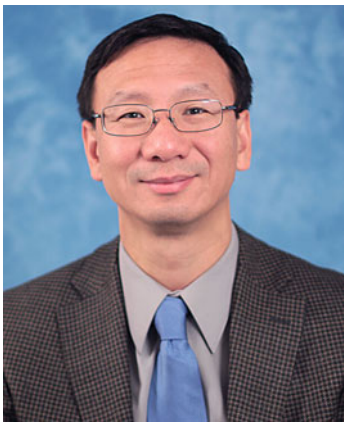
Assembling this diverse textbook has truly been a team effort, cutting across many traditional specialty boundaries. We are most grateful for all the contributors who have made this project possible and to the readers of the first edition who provided much valuable feedback. Our special thanks go to Ms. Tracey Bender, without whose outstanding work this textbook would have been an impossible achievement. We would like to thank Mr. Ryan P. Christy from the Multimedia Education Division of the

Department of Pathology at Indiana University, who has edited the illustrations for the handbook. We would also like to thank the staff at Springer, including Mr. Richard Hruska, Ms. Maureen Alexander, Ms. Neha Thapa, Ms. Saranya Jayakumar, and other colleagues for their assistance in the development and editing of this textbook.

January 2013

Liang Cheng, MD
David Y. Zhang, MD, PhD
John N. Eble, MD

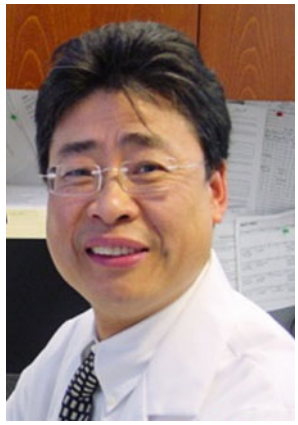
Biography



Liang Cheng, MD, Professor of Pathology and Urology, Director of Molecular Pathology Laboratory, Chief of Genitourinary Pathology Division, Director, Fellowship in Urologic Pathology, Indiana University School of Medicine, Indianapolis, IN, USA

Dr. Liang Cheng is professor of pathology and urology at the Indiana University School of Medicine, Indianapolis, Indiana, USA. Currently, he is the chief of the Genitourinary Pathology Service, director of the Urologic Pathology Fellowship Program, and director of Molecular Diagnostics and Molecular Pathology Laboratories. Dr. Cheng is board certified in molecular genetic pathology and anatomic and clinical pathology by the American Board of Pathology. He graduated from Beijing Medical University, China. Dr. Cheng completed his pathology residency at Case Western Reserve University, Cleveland, Ohio, and his fellowship at the Mayo Clinic, Rochester, Minnesota, in 1998. Dr. Cheng has received numerous prestigious awards including the Stowell-Orbison Award from the United States and Canadian Academy of Pathology (USCAP) and the Koss Medal Award from the International Society of Urological Pathology (ISUP). In 2006, he was the Arthur Purdy Stout Prize recipient from the Arthur Purdy Stout Society of Surgical Pathologists in recognition of outstanding contributions to the field of surgical pathology by a surgical pathologist less than 45 years old. Dr. Cheng has published more than 500 peer-reviewed articles in

high-impact scientific journals. He was also the author of over 60 book chapters and several books. Currently, he is an active member of over 30 editorial boards. His research focuses on molecular genetics and biological predictors of genitourinary cancers.



David Y. Zhang, MD, PhD, MPH, Director, Molecular Pathology Laboratory, Professor, Departments of Pathology and Oncological Sciences, Mount Sinai School of Medicine, New York, NY, USA

David Y. Zhang MD, Ph.D., MPH, is the director of the Molecular Pathology Division. He obtained his MD from Norman Bethune School of Medicine of Jilin University, Ph.D. from New York University School of Medicine, and MPH from Mount Sinai School of Medicine. He subsequently completed his residency training in anatomic and clinical pathology as well as fellowship training in cytopathology and clinical microbiology at Mount Sinai School of Medicine, New York. He also completed second residency in preventive medicine (occupational medicine) at the Selikoff Center, Mount Sinai School of Medicine. He is licensed to practice medicine in New York and New Jersey and certified in anatomic and clinical pathology, cytopathology, molecular genetic pathology by the American Boards of Pathology and Medical Genetics as well as in preventive medicine by the American Board of Preventive Medicine. He currently practices surgical pathology, cytopathology, and molecular pathology with a strong interest in solid tumor pathology. His major research interest is to develop genomic and proteomic technologies for biomarker discovery and molecular diagnosis. He currently holds more than 10 US and international patents and has published more than 50 scientific papers in major peer-reviewed journals and more than 10 book chapters.

John N. Eble MD, MBA, FRCPA, Nordschow Professor and Chairman, Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

Dr. John Eble obtained his medical degree from the Indiana University School of Medicine, where he also took his residency in anatomic and clinical pathology. Following residency, he joined the faculty of the Department of Pathology and Laboratory Medicine of Indiana University and became the head of the clinical laboratories at the Roudebush VA Medical Center at the Indiana University Medical Center. While at the VA, he established the first diagnostic molecular pathology laboratory in Indiana. A few years later, he attended the Indiana University Graduate School of Business and received an MBA degree in 1990. Dr. Eble continued as the VA chief of service until 2000. In 1998, Dr. Eble was appointed chairman of the Department of Pathology and Laboratory Medicine of the Indiana University School of Medicine and chief pathologist of Indiana University Health. In addition to his leadership activities, Dr. Eble currently signs out the urologic surgical pathology cases of Indiana University Health one week a month. While these events were developing, Dr. Eble used the rich clinical material of the VA Medical Center and IU Health as a foundation upon which to study urologic pathology. He has published more than 225 peer-reviewed papers, more than 30 book chapters, and 3 books devoted to diseases of the organs of the urinary tract and male genital system. With Jonathan Epstein, Guido Sauter, and Isabell Sesterhenn, he edited the 2004 World Health Organization "blue book" *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of the Urinary System and Male Genital Organs*.

Dr. Eble has shared his knowledge of urologic pathology with thousands of pathologists and urologists around the world. He served on the faculty of the pathology review course of the American Urologic Association from 1991 to 2000. He gave a course on renal neoplasia (with David Grignon) for the American Society of Clinical Pathologists from 2005 to 2008, directed their 73rd annual Slide Seminar on Anatomic Pathology in 2005, and is presently on the faculty of their annual course *Contemporary Issues in Urologic Pathology*. He served on the faculty of the annual Harvard course *Urologic*

Surgical Pathology for the Practicing Pathologist since its inception in 1997 to 2011. He became active in the International Society for Urologic Pathology at the time it was founded in 1990. He served as a councillor from 1991 to 1995, as its treasurer from 1995 to 2001, and its president-elect and president from 2001 to 2005. The Society awarded him the Leopold Koss Medal for Distinguished Service in 2003. With David Bostwick, Dr. Eble gave his first short course for the USCAP, *Incipient Neoplasia in Diagnostic Urologic Pathology*, from 1994 to 1998. This was followed by *Tumors of the Kidneys and Ureters* (with Stephan Störkel) from 1997 to 2000. Next was *Renal Neoplasia*: *Diagnostic Problems and Newly Recognized Entities* (with Holger Moch) from 2005 to 2008. Most recently, he has given *Handling and Reporting of Tumor-containing Specimens in Urologic Pathology* (with David Grignon). In 2008, he directed the Long Course, *Tumors of the Kidneys and Urinary Bladder* with David Grignon. When Dr. Mills announced he would be stepping down as editor of *Modern Pathology*, Dr. Eble had been editing the quarterly *Journal of Urologic Pathology* for 3 years and was looking for an opportunity to edit a major monthly pathology journal and so he applied for the position. The Academy selected him and he became the fourth editor of *Modern Pathology* in 2000. The Academy honored Dr. Eble with the F.K. Mostofi Award for distinguished service in 2010. The Royal College of Pathologists of Australasia made Dr. Eble an honorary fellow in 2001.

Contents

Volume 1

Section 1 Basic Principles in Molecular Genetic Pathology 1

1 Clinical Molecular Biology: Principles 3
Shaobo Zhang, Darrell D. Davidson, and Liang Cheng

2 Clinical Cytogenetics: Principles 43
Stuart Schwartz

3 Conceptual Evolution in Cancer Biology 77
Shaobo Zhang, Darrell D. Davidson, Rodolfo Montironi, Antonio Lopez-Beltran, Gregory T. MacLennan, and Liang Cheng

4 Genetic Inheritance and Population Genetics 111
Tatiana Foroud and Daniel L. Koller

5 Clonality Analysis and Tumor of Unknown Primary: Applications in Modern Oncology and Surgical Pathology 129
Liang Cheng, Shaobo Zhang, Federico A. Monzon, Timothy D. Jones, and John N. Eble

6 Modern Immunohistochemistry in Targeted Therapy 181
Allen M. Gown

7 Systems Pathology 197
José Costa and Michael J. Donovan

Section 2 Methodology in Molecular Genetic Pathology 209

8 Diagnostic Methodology and Technology in Molecular Genetic Pathology 211
Josephine Wu, Fei Ye, Miao Cui, Robert Shibata, Ruliang Xu, Liang Cheng, and David Y. Zhang

9 Clinical Flow Cytometry in Molecular Genetic Pathology 301
Magdalena Czader

10 Clinical Genomics in Oncology 335
Hugo M. Horlings, Thalia A. Farazi, and Marc J. van deVijver

11 Clinical Proteomics in Molecular Genetic Pathology 369
David H. Geho, Virgina Espina, Lance A. Liotta, Emanuel F. Petricoin, and Julia D. Wulfkuhle

12 Nanotechnology in Molecular Diagnostics 383
Andrew N. Young and Brad A. Kairdolf

13 Gene Therapy: Vector Technology and Clinical Applications 399
Sunyong Tang and Kenneth Cornetta

14 Instrumentation for Molecular Testing 413
Jordan Laser, Miao Cui, Bruce E. Petersen, Fei Ye, and David Y. Zhang

15 Tissue Microarrays and Biomarker Validation in Molecular Diagnostics 455
Martina Storz and Holger Moch

16 Laser Capture Microdissection in Molecular Diagnostics 465
Liang Cheng, Shaobo Zhang, Darrell D. Davidson, Matthew Kuhar, Mingsheng Wang, Sean R. Williamson, David Y. Zhang, and Gregory T. MacLennan

17 Human Tissue Biorepository 483
William E. Grizzle, Katherine C. Sexton, and Walter C. Bell

18 Pathology Informatics 499
Roy E. Lee, Long Phi Le, and John Gilbertson

Section 3 Genetic Disorders 515

19 Prenatal Cytogenetic and Cytogenomic Diagnosis 517
Ninette Cohen and Lisa Edelmann

20 Molecular Medical Genetics 529
Lisa Edelmann, Stuart Scott, Liu Liu, and Ruth Kornreich

Volume 2

21 Biochemical Genetics and Inborn Errors of Metabolism . . . 561
Chunli Yu, Melissa P. Wasserstein, and George A. Diaz

22 Molecular Genetics of Congenital Heart Disease 591
Abigail S. Carey and Bruce D. Gelb

23 Familial Cancer Syndrome 607
Beth T. Harrison, Rodolfo Montironi, Antonio Lopez-Beltran, Michelle P. Elieff, and Liang Cheng

24 Genetic Counseling 645
Kimberly A. Quaid and Lisa J. Cushman

Section 4 Molecular Microbiology 653

25 Molecular Virology 655
Fei Ye, Miao Cui, Rame H. Khasawneh, Robert Shibata, Josephine Wu, Mona Sharaan, and David Y. Zhang

26 Molecular Bacteriology and Mycobacteriology 733
Jeong Hwan Shin, Mona Sharaan, Josephine Wu, Miao Cui, David Y. Zhang, and Yi-Wei Tang

27 Molecular Parasitology and Mycology 781
Vincent J. LaBombardi and Wei Wang

28 Molecular Microbiology in Transplantation 795
Adnan A. Alatoom and Robin Patel

29 Molecular Microbiology Epidemiology 813
Miao Cui, Fei Ye, and David Y. Zhang

Section 5 Hematologic Malignancies 823

30 Immunoglobulin and T Cell Receptor Gene Rearrangement: Principle 825
Jeffrey E. Miller

31 Lymphoblastic Leukemia and Lymphoma: Molecular Diagnostics 857
Dan Jones and Zeqiu J. Han

32 Myeloid Neoplasms: Molecular Diagnostics 865
C. Cameron Yin and Dan Jones

33 Lymphoid Malignancies: Molecular Diagnostics 887
Vasiliki Leventaki and Francisco Vega

34 Myeloproliferative Neoplasms and Myelodysplastic Syndromes: Molecular Diagnostics 929
Gerlinde Wernig and Daniel A. Arber

35 Lymphoproliferative Disorders: Molecular Diagnostics ... 945
Timothy C. Greiner

Section 6 Hematological Disorders and Specialized Applications 953

36 Molecular Testing for Coagulation Abnormalities 955
Veshana Ramiah and Thomas L. Ortel

37 Molecular Hemoglobinopathies 975
Tina Y. Fodrie, Shaobo Zhang, Mina Patel, Jodi A. Parks, and Liang Cheng

38 HLA System and Transfusion Medicine: Molecular Approach 993
S. Yoon Choo

39 Clinical Pharmacogenomics 1009
Catalina López-Correa and Lawrence M. Gelbert

40 Molecular Forensic Pathology 1033
Stephen R. Dlouhy and P. Michael Conneally

41 Specimen Identification and Bone Marrow Engraftment Analysis 1049
Loren Joseph

42 Circulating Tumor Cells: Liquid Biopsy for Molecular Pathology 1065
Jeffrey S. Ross

Section 7 Laboratory Management in Molecular Genetic Pathology 1081

43 Molecular Testing: Ethical and Legal Issues 1083
Kimberly A. Quaid

44 Molecular Testing: Regulatory Issues 1091
Frank S. Ong, Wayne W. Grody, and Kingshuk Das

45 Molecular Pathology Reporting 1109
Stacey S. O'Neill and Margaret L. Gulley

46 Laboratory Management: Quality Assurance and Inspection 1117
Kristin M. Post and Liang Cheng

Index 1129

Contributors

Adnan A. Alatoom, MD, PhD, ABMM, FCAP Assistant Professor of Clinical Pathology and Microbiology, Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX, USA

Daniel A. Arber, MD Professor - Med Center Line, Pathology, Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA

William C. Bell, MD Associate Professor, Department Anatomic Pathology, University of Alabama School of Medicine, Birmingham, AL, USA

Abigail S. Carey, MD Department of Pediatrics, Mount Sinai School of Medicine, New York, NY, USA

Liang Cheng, MD Professor of Pathology and Urology, Director of Molecular Pathology Laboratory, Chief of Genitourinary Pathology Division, Director, Fellowship in Urologic Pathology, Indiana University School of Medicine, Indianapolis, IN, USA

S. Yoon Choo, MD Associate Professor of Pathology and Medicine, Director of Immunogenetics and Associate Medical Director of Transfusion Medicine, Mount Sinai School of Medicine, New York, NY, USA

Ninette Cohen, PhD Assistant Professor, Mount Sinai School of Medicine, Genetic and Genomic Sciences, New York, NY, USA

P. Michael Conneally, PhD Professor of Medical and Molecular Genetics, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

Kenneth Cornetta, MD Joe C. Christian Professor and Chair, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

José Costa, MD Professor of Pathology and Medicine, Vice Chairman, Department of Pathology, Yale University School of Medicine, New Haven, CT, USA

Miao Cui, MD Associate Researcher of Pathology, Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

Lisa J. Cushman, PhD, CGC Hayes, Inc., Landsale, PA, USA

Magdalena Czader, MD, PhD Director, Division of Hematopathology, Director, Clinical Flow Cytometry Laboratory, Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

Kingshuk Das, MD Assistant Clinical Professor, Department of Pathology and Laboratory Medicine, UCLA David Geffen School of Medicine, Los Angeles, CA, USA

Darrell D. Davidson, MD, PhD Clinical Assistant Professor, Indiana University School of Medicine, IU Health Pathology Laboratory, Indianapolis, IN, USA

George A. Diaz, MD, PhD Associate Professor, Department of Genetics & Genomic Sciences, Mount Sinai School of Medicine, New York, NY, USA

Stephen R. Dlouhy, PhD Associate Research Professor, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

Michael J. Donovan, PhD, MD Professor of Pathology, Director of Experimental Pathology, Mount Sinai School of Medicine, New York, NY, USA

John N. Eble, MD, MBA, FRCPA Nordschow Professor and Chairman, Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

Lisa Edelmann, PhD Associate Professor, Department of Genetic and Genomic Sciences, Mount Sinai School of Medicine, New York, NY, USA

Michelle Elieff, MD Forensic Pathologist, Sparrow Forensic Pathology Services, Lansing, MI, USA

Virgina Espina, MS Research Assistant Professor, George Mason University, Center for Applied Proteomics and Molecular Medicine, Manassas, VA, USA

Thalia A. Farazi, MD, PhD Instructor in Clinical investigation, Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, New York, NY, USA

Tina Y. Fodrie, MT, MP(ASCP) Laboratory Manager, Translational Research and Clinical Coordinator, Laboratory for Personalized Molecular Medicine, San Diego, California, USA

Tatiana Foroud, PhD Chancellor's Professor, P. Michael Conneally Professor of Medical and Molecular Genetics, Director, Division of Hereditary Genomics, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

David H. Geho, MD, PhD Associate Director, Imaging, Merck and Company, Inc., West Point, PA, USA

Bruce D. Gelb, MD Director, Child Health and Development Institute, New York, NY, USA; Professor of Pediatrics, Cardiology, Professor of Genetics & Genomic Sciences, Mount Sinai School of Medicine, New York, NY, USA

Lawrence Gelbert, PhD Managing Member at LMG PharmaTech, LLC, Indianapolis, IN, USA; Adjunct Professor of Pediatrics, Herman B Wells Center for Pediatric Research, Indiana University School of Medicine, Indianapolis, IN, USA

John Gilbertson, MD Associate Professor of Pathology, Harvard Medical School, Department of Pathology, The Massachusetts General Hospital, Boston, MA, USA

Allen M. Gown, MD Medical Director and Chief Pathologist, PhenoPath Laboratories, Seattle, WA, USA; Clinical Professor of Pathology, University of British Columbia, Vancouver, BC, Canada

Timothy C. Greiner, MD Professor, Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, USA

William E. Grizzle, MD, PhD Professor of Pathology, Head, Program for Translational Research in Neoplasia, Director, Tissue Procurement and Banking Facility, Senior Scientist, UAB Comprehensive Cancer Center; Center for Aging; Metabolic Bone Disease Research Center, Senior Editor, Clinical Cancer Research, Birmingham, AL, USA

Wayne W. Grody, MD, PhD Professor, Divisions of Medical Genetics and Molecular Pathology, Departments of Pathology and Laboratory Medicine, Pediatrics, and Human Genetics, UCLA School of Medicine, Los Angeles, CA, USA

Margaret L. Gulley, MD Professor of Pathology, University of North Carolina, Chapel Hill, NC, USA

Zeqiu J. Han, MD Department of Pathology, Quest Diagnostics Nichols Institute, Chantilly, VA, USA

Beth T. Harrison, MD Department of Pathology, NYU Langone Medical Center, New York, NY, USA

Hugo M. Horlings, MD, PhD Pathologist, Academic Medical Center, Amsterdam, The Netherlands

Dan Jones, MD, PhD Medical Director, Cancer Diagnostics Services, Quest Diagnostics Nichols Institute, Chantilly, VA, USA

Timothy D. Jones, MD Staff Pathologist, East Louisville Pathologists, PSC, Louisville, KY, USA

Loren Joseph, MD Director, Molecular Pathology Laboratory, Department of Pathology, University of Chicago Medical Center, Chicago, IL, USA

Brad A. Kairdolf, PhD Postdoctoral Fellow, Department of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA, USA

Rame H. Khasawneh, MD Molecular Genetic Pathology Fellow, Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

Daniel L. Koller, PhD Assistant Research Professor, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

Ruth Kornreich, PhD Research Assistant Professor of Human Genetics, Co-Director Genetic Testing Laboratory, Department of Human Genetics, Mount Sinai School of Medicine and the Mount Sinai Medical Center, New York, NY, USA

Matthew Kuhar, MD Associate Dermatopathologist, StrataDx & Strata Pathology Services, One Cranberry Hill, Lexington, MA, USA

Vincent J. LaBombardi, PhD Chief of Microbiology, New York Hospital-Queens, Flushing, NY, USA

Jordan Laser, MD Assistant Professor of Pathology, Director of Molecular Pathology, Department of Pathology and Laboratory Medicine, North Shore / Long Island Jewish Health System, Manhasset, NY, USA

Long Phi Le, MD, PhD Assistant Professor of Pathology, Harvard Medical School, Department of Pathology, The Massachusetts General Hospital, Boston, MA, USA

Roy E. Lee, MD Associate Medical Director for the Center of Pathology Informatics, Pathology and Laboratory Medicine Institute, The Cleveland Clinic, Cleveland, OH, USA

Vasiliki Leventaki, MD Hematopathology Fellow, Department of Hematopathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Lance A. Liotta, MD, PhD Professor, George Mason University, Center for Applied Proteomics and Molecular Medicine, Manassas, VA, USA

Antonio Lopez-Beltran, MD, PhD Professor of Anatomic Pathology, Unit of Anatomic Pathology, Department of Surgery, Cordoba University School of Medicine, Cordoba, Spain

Catalina López-Correa, MD, PhD Vice-présidente, Affaires scientifiques/ CSO, Génome Québec, Montréal (Québec)

Liu Liu, PhD, FACMG Associate Lab Director, Assistant Professor, Cancer Genetics Laboratory, Baylor College of Medicine, Houston, TX, USA

Gregory T. MacLennan, MD Professor of Pathology, Director, Anatomic Pathology, Case Western Reserve University, Cleveland, OH, USA

Jeffrey E. Miller, PhD Chief Scientific & Chief Executive Officer of Invivoscribe Technologies, Inc, and the Laboratory for Personalized Molecular Medicine, LabPMM, San Diego, CA, USA

Holger Moch, MD Professor and Chairman, Institute of Surgical Pathology, Department Pathology, University Hospital Zürich, Zürich, Switzerland

Rodolfo Montironi, MD, FRCPath Professor of Pathology, Institute of Pathological Anatomy and Histopathology, Polytechnic University of the Marche Region (Ancona), School of Medicine, United Hospitals, Ancona, Italy

Federico Monzon, MD Associate Professor, Director of Molecular Pathology, Cancer Genetics Laboratory at Baylor College of Medicine, Departments of Pathology & Immunology and Molecular & Human Genetics at Baylor College of Medicine, Houston, TX, USA

Stacey S. O'Neill, MD, PhD Molecular Genetic Pathology Fellow, University of North Carolina, Chapel Hill, NC, USA

Frank S. Ong, MD Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, USA

Thomas L. Ortel, MD, PhD Professor of Medicine and Pathology, Hemostasis and Thrombosis Center, Duke University Medical Center, Durham, NC, USA

Jodi A. Parks, MD Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

Mina Patel, MT, MP(ASCP) Supervisor, Molecular Pathology Laboratory, Roudebush VA Medical Center, Indianapolis, IN, USA

Robin Patel, MD(CM), FRCP(C), D(ABMM), FIDSA, FACP, F(AAM) Professor of Microbiology and Medicine, Chair, Division of Clinical Microbiology, Consultant, Divisions of Clinical Microbiology and Infectious Diseases, Mayo Clinic, Rochester, MN, USA

Bruce E. Petersen, MD Assistant Professor, Pathology and Medicine, Mount Sinai School of Medicine, New York, NY, USA

Emanuel F. Petricoin, PhD Professor, George Mason University, Center for Applied Proteomics and Molecular Medicine, Manassas, VA, USA

Kristin M. Post, MT Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, IU Health Pathology Laboratory, Indianapolis, IN, USA

Kimberly A. Quaid, PhD Professor of Medical and Molecular Genetics, Indiana University School of Medicine, Faculty Investigator, Indiana University Center for Bioethics, Indianapolis, IN, USA

Veshana Ramiah, MD Danville Regional Medical Center, Danville Hematology and Oncology, Danville, VA, USA

Jeffrey S. Ross, MD Cyrus Strong Merrill Professor and Chair, Department of Pathology and Laboratory Medicine, Albany Medical College, Albany, NY, USA; Medical Director, Foundation Medicine, Inc., Cambridge, MA, USA

Stuart Schwartz, PhD, FACMG Strategic Director, Cytogenetics, Cytogenetics Laboratory, Center for Molecular Biology and Pathology, Laboratory Corporation of America, Research Triangle Park, NC, USA

Stuart Scott, PhD Research Assistant Professor of Human Genetics, Department of Human Genetics, Mount Sinai School of Medicine and the Mount Sinai Medical Center, New York, NY, USA

Katherine C. Sexton, MBA Assistant Director, University of Alabama at Birmingham, Tissue Collection and Banking Facility, Southern Division, Cooperative Human Tissue Network, Birmingham, AL, USA

Mona Sharaan, MD Molecular Pathology Laboratory, Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

Robert Shibata, MD Molecular Genetic Pathology Fellow, Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

Jeong Hwan Shin, MD, PhD Associate Professor, Department of Laboratory Medicine, Inje University College of Medicine, Busan, South Korea

Martina Storz, MD Director of TMA Core Facility, Institute of Surgical Pathology, Department Pathology, University Hospital Zürich, Zürich, Switzerland

Sunyong Tang, PhD Indiana University Vector Production Facility, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

Yi-Wei Tang, MD, PhD, F(AAM), FIDSA Chief of Clinical Microbiology Service, Memorial Sloan-Kettering Cancer Center, Professor of Laboratory Medicine and Pathology, Cornell University Weil College of Medicine, New York, NY, USA

Marc J. van de Vijver, MD, PhD Head, Division of Diagnostic Oncology, NKI, Amsterdam; Head, Department of Pathology, NKI, Amsterdam; Pathologist, Department of Pathology, Leiden University Medical Center, The Netherlands; Professor of Pathology, Department of Pathology, The Netherlands; Cancer Institute/Antoni van Leeuwenhoek Hospital, Amsterdam, Netherlands

Francisco Vega, MD, PhD Associate Professor and Lymphoma Section Head, Department of Hematopathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Mingsheng Wang, MD Associate Scientist, Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

Wei Wang, PhD Research Associate, Mount Sinai Medical Center, New York, NY, USA

Melissa P. Wasserstein, MD Associate Professor, Mount Sinai School of Medicine, New York, NY, USA

Gerlinde Wernig, MD Department of Pathology, Stanford University School of Medicine, Stanford Hospital and Clinics, Stanford, CA, USA

Sean R. Williamson, MD Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

Josephine Wu, DDS, CLSP(MB), CLDIR Assistant Professor of Pathology, Mount Sinai School of Medicine, New York, NY, USA

Julia D. Wulfkuhle, PhD Professor, George Mason University, Center for Applied Proteomics and Molecular Medicine, Manassas, VA, USA

Ruliang Xu, MD, PhD Associate Professor of Pathology, Director of Molecular Pathology, Director of Gastrointestinal and Liver Pathology, Department of Pathology, New York University Langone Medical Center, New York, NY, USA

Fei Ye, PhD Assistant Professor of Pathology, Assistant Director of Molecular Pathology, Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

C. Cameron Yin, MD, PhD Associate Professor, Department of Hematopathology, UT MD Anderson Cancer Center, Houston, TX, USA

Andrew N. Young, MD, PhD Associate Professor, Department of Pathology & Laboratory Medicine, Emory University School of Medicine, Laboratory Director, Grady Health System, Atlanta, GA, USA

Chunli Yu, MD Associate Professor, Department of Genetics & Genomic Sciences and Pediatrics, Mount Sinai School of Medicine, New York, NY, USA

David Y. Zhang, MD, PhD, MPH Professor of Pathology, Director of Molecular Pathology Laboratory, Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

Shaobo Zhang, MD Associate Scientist, Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

Section 1

Basic Principles in Molecular Genetic Pathology

Clinical Molecular Biology: Principles

1

Shaobo Zhang, Darrell D. Davidson, and Liang Cheng

Contents

1.1 Deoxyribonucleic Acid (DNA) 4
1.1.1 Overview 4
1.1.2 Types of DNAs 4
1.1.3 DNA Replication 7
1.1.4 DNA Mutation 9
1.1.5 DNA Repair Mechanisms 13

1.2 Genes 15
1.2.1 Overview 15
1.2.2 Gene Structures 16
1.2.3 Functional Categories of Genes 17
1.2.4 Cancer-Related Genes 17
1.2.5 Regulation of Gene Expression 22
1.2.6 Signal Transduction 22

1.3 Chromosomes 23
1.3.1 Overview 23
1.3.2 Chromatin 23
1.3.3 Chromosomes 25

1.4 RNA and Protein 28
1.4.1 Overview 28
1.4.2 Types of RNA 29
1.4.3 Ribosome and Ribozyme 30
1.4.4 mRNA Processing 31
1.4.5 Protein Translation 33

1.5 Mitochondrial DNA 35
1.5.1 Overview 35
1.5.2 mtDNA Inheritance 35
1.5.3 Characteristics of mtDNA 36
1.5.4 Mitochondrial Genes and Gene Expression 36
1.5.5 mtDNA Replication 38
1.5.6 mtDNA Damage, Mutations, and Repair 38
1.5.7 Mitochondrial Disease 39

Further Reading 40

S. Zhang, MD • D.D. Davidson, MD, PhD •
L. Cheng, MD (✉)
Department of Pathology and Laboratory Medicine,
Indiana University School of Medicine, Indianapolis,
IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_2,

1.1 Deoxyribonucleic Acid (DNA)

1.1.1 Overview

- Definition
  - DNA (deoxyribonucleic acid), the hereditary material in humans and most other organisms, is a large nucleic acid polymer arranged in chromosomes for storage, expression, and transmission of genetic information
  - The genetic information is encoded by a sequence of nucleotide bases
- Components of DNA
  - Bases (nucleobase) are molecules containing carbon-nitrogen rings stacked like two piles of pancakes with their edges touching in DNA
    - Purines: adenine (A) and guanine (G) have two joined carbon-nitrogen rings
    - Pyrimidines: thymine (T) and cytosine (C) have one carbon-nitrogen ring
  - Nucleoside is a glycosylamine made up of a nucleobase and a ribose (RNA) or deoxyribose (DNA) sugar ring
  - Deoxyribose is the same sugar found in RNA but with oxygen removed from the 2′ carbon position
  - Nucleotide is made up of one to three phosphate groups, a pentose sugar (ribose or deoxyribose) and a nucleobase (Fig. 1.1)
    - The phosphodiester bond
      - The phosphate group forms strong covalent ester links between a phosphorous atom and two 5-carbon sugars of neighboring nucleosides in DNA or RNA
      - Phosphodiester bond links the 3′ carbon atom of one sugar molecule to the 5′ carbon atom of the adjacent nucleotide sugar, making up the DNA backbone (Fig. 1.2)
      - The phosphodiester bond determines DNA chain polarity (ends designated as either 5′ or 3′)
- DNA sequence refers to the order of the nucleotides in a DNA strand, which code for unique ordered sets of genetic codes for both proteins and regulatory segments
  - The deoxyribonucleotides in DNA differ only in the bases they carry, so the DNA sequence is denoted by a base sequence (e.g., -TTAGGGC-)
  - The base sequence is always presented from 5′ to 3′
  - DNA strands are pairs of complementary molecules, which entwine each other in an antiparallel direction (5′–3′ polarity paired to a 3′–5′ polarity)
  - Two strands of DNA wind around each other to form a spiral called the double helix (Fig. 1.3)
    - The deoxyribose-phosphate backbone is on the exterior of the DNA double helix
    - The interior of the DNA is formed by paired bases attached to each other by hydrogen bonds. G (Guanine) pairs with C (Cytosine) via three hydrogen bonds, and A (Adenine) pairs with T (Thymine) via two hydrogen bonds inside the double helix. Note that the three hydrogen bonds joining G to C (GC bond) are stronger than the two hydrogen bonds joining A to T (AT bond) (Fig. 1.4)
  - DNA has two DNA chains; one is oriented 5′ → 3, while the other strand is oriented 3′ → 5′ (antiparallel)
    - Sense is a DNA strand that has the sequence similar to its RNA transcript
    - Antisense is the complimentary strand of sense. Antisense works as template for the RNA transcription
    - Each DNA fragment appears to have a unique function: structural, regulatory, or coding

1.1.2 Types of DNAs

- The types of human DNA may be classified by function or cellular location. According to the location, the DNA could be chromosomal DNA or mitochondrial DNA. According to

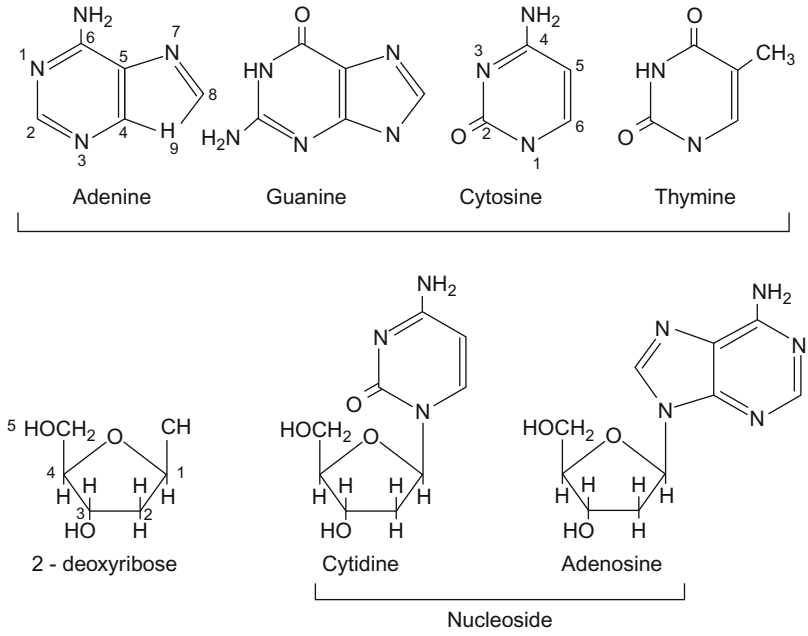


Fig. 1.1 DNA contains four bases: adenine (*A*), guanine (*G*), thymine (*T*), and cytosine (*C*). Adenine and guanine are purines with two aromatic rings and thymine and cytosine are pyrimidines with one heterocyclic ring. Deoxyribose is the sugar in DNA. The carbon atoms are numbered as indicated. Note there is no oxygen on site 2 of deoxyribose. A nucleoside molecule is composed of a base and deoxyribose. When a phosphate group is added to nucleoside linking the 3′ carbon of one sugar molecule and the 5′ carbon atom of another, the complex becomes a nucleotide base. Nucleotides are the basic building blocks of DNA

the function, DNA could be either coding or noncoding DNA

- Single copy DNA is a unique DNA sequence that is present only once in the genome
- Repetitive DNA is composed of DNA base cycles in which a specific DNA sequence is repeated multiple times and in multiple sites within the genome (Table 1.1)
- Moderately repetitive DNA refers to 10–10^5 copies of the base cycle per genome
  - Moderate repeated DNA is found primarily in noncoding sequences
- Highly repetitive DNA describes a DNA base cycle present in greater than 10^5 copies per genome
  - Highly repeated DNA is found primarily in centromere and telomere regions as tandem repeats
- Tandem repeat DNA contains a variable number of short DNA base cycles repeated many times in series. The number of repeats is unique to each individual and can be used for relationship testing
  - The tandem repeat pattern may vary from one base repeats (mononucleotide repeat) to several thousand base pair repeat sequences
  - These segments of DNA are called satellite DNA because they often form a minor satellite band near the major centrifugation fraction when DNA is separated in density gradient experiments
  - Clusters of such repeats are scattered on many chromosomes. Each variant is an allele that is inherited codominantly
  - Satellite DNAs consist of large arrays of tandem repeating, noncoding DNA that comprise about 15% of human DNA
    - The repeated sequence ranges from 5 to 170 bp, and a satellite DNA complex is about 100 kbp in length

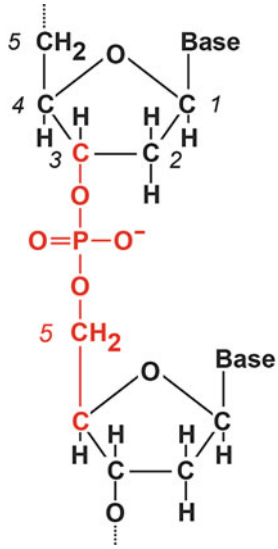


Fig. 1.2 3′ → 5′ Phosphodiester bonds join the units of the repetitive sugar-phosphate chain. Each nucleotide is linked by the 3′ carbon atom of the downstream ribose to the 5′ carbon of the upstream ribose in the 5′ → 3′ polarity. Phosphodiester bonds are central to all life forms, as they make up the backbone of the DNA and RNA strands in every organism

    - Satellite DNA is the main component of centromeres and forms the main structure of heterochromatin
  - Megasatellite DNAs are tandem repeat DNA segments with a length greater than 1,000 bp (1 kbp) repeated 50–400 times. The biological function of these repeats is to regulate heterochromatin formation
  - Minisatellite DNA (also called VNTR, variable number tandem repeats) is a repeating, GC rich complex spanning 0.1–20 kbp, containing a series of short repeat sequences 6–60 bp long and occurring at over 1,000 locations in the human genome
    - Hypervariable minisatellites are found at centromeres with repeat units 9–24 bp long
    - Telomeric minisatellite DNA is present in the telomere region with 6 bp units that repeat thousands of times
  - Microsatellite DNA loci (also called STR, short tandem repeats) contain sequences <15 bp in length that repeat 10–100 times without interruption for a total length up to 1,500 bp at each locus
    - There are approximately 200,000 microsatellite loci in the human genome (up to 10% of the genome)
    - The function of microsatellite DNA is not clearly known, but deletion of a microsatellite locus usually does not have a measurable effect on phenotype. Mutation, however, does cause harmful effects
- DNA conformations: B-DNA, A-DNA, and Z-DNA, as well as C-DNA and mitochondrial DNA (mtDNA)
  - B-DNA is the common form of chromosome DNA. B-DNA chains form a right-handed helix with 10 base pairs per turn
  - A-DNA is also a right-handed helix but has 11 base pairs per turn. A-DNA is biologically active in the cell, and it forms crystallized structures in lab experiments
  - Z-DNA is a left-handed helix and is believed to be biologically active. It often forms a zigzag pattern. Z-DNA has 12 base pairs per turn, and it is a transient structure
  - B-DNA, A-DNA, and Z-DNA differ in the tightness or pitch of the double helix turns. Z-DNA also differs from the other DNAs by the direction (chirality) of its helical turns
  - C-DNA is a form of DNA observed in conditions of low humidity or high Li + Mg++ ion concentration by x-ray crystallography
    - It initially presents as a paired DNA strand and its RNA template
    - It may be a physiologic form of DNA tightly coiled in a nucleosome
  - mtDNA is circular, double-stranded DNA found only within mitochondria (see Sect. 5)
- Loss of heterozygosity (LOH) in a cell represents the loss of one parent's gene allele contribution to a cell's genome, leaving only the other parent's allele
  - Microsatellite DNA loci are useful markers for the detection of LOH
  - LOH often suggests tumor suppressor gene loss near the microsatellite locus

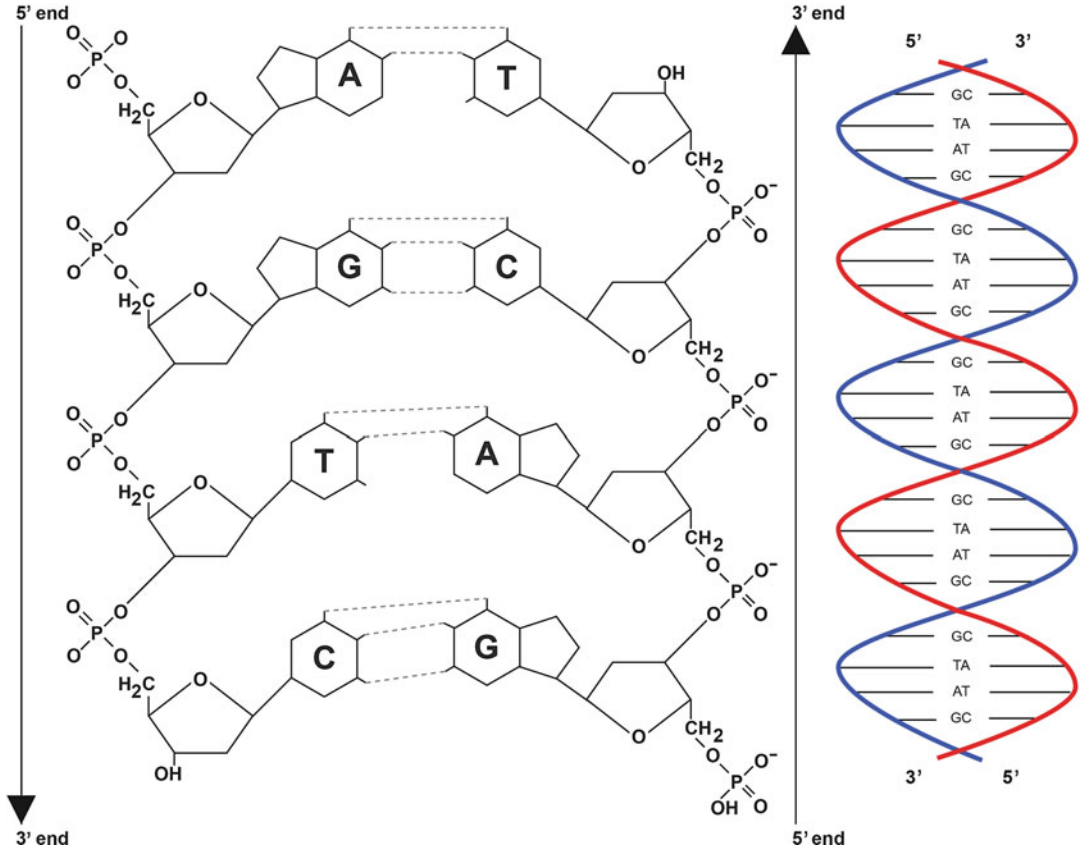


Fig. 1.3 Human genomic DNA contains two polynucleotide chains wound around each other to form a double-stranded helix. The two chains are "antiparallel," one running 5′ → 3′ and the other running in the 3′ → 5′ direction. The DNA strands are synthesized by DNA polymerase and transcribed by RNA polymerase in the 5′ → 3′ direction. The purine or pyrimidine attached to each deoxyribose projects into the center of the helix, and the aromatic rings are stacked like a leaning tower of pancakes. Base A pairs with T and G pairs with C through hydrogen bonds in the central axis

- LOH can arise through deletion, nonreciprocal DNA transfer, mitotic recombination, or chromosome loss
- LOH may be used to analyze the clonal origin of cancer-associated loci and to analyze the mechanisms of tumorigenesis
- When parents' contributions of certain microsatellite loci are of different sizes, forming two distinct bands by electrophoresis, these microsatellite loci are *informative*; otherwise the loci are un*informative* since these loci cannot be used to detect LOH
- Microsatellite instability at critical loci is a marker for malignancy or premalignant genetic change (see details in Chap. 3)

1.1.3 DNA Replication

- DNA replication occurs during S (synthesis) phase of the cell cycle
  - Cell cycle refers to a cycle of events in eukaryotic cells from one cell division to the next; it consists of G_0, G_1, S, G_2, and M phases (Fig. 1.5)
- Semiconservative replication means that each DNA molecule consists of one original and one newly synthesized chain (Fig. 1.6)
- Double-stranded DNA (dsDNA) is exactly duplicated prior to cell division so that each daughter cell is endowed with an exact replica of the parent cell DNA

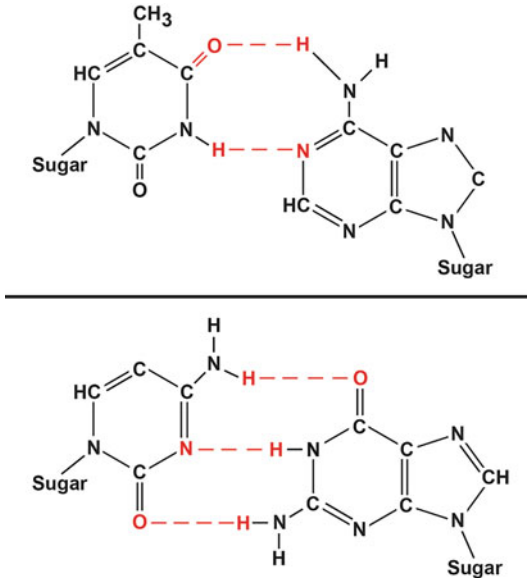


Fig. 1.4 The double helical structure of DNA is largely due to hydrogen bonding between the base pairs linking one complementary strand to the other. Hydrogen bonds are noncovalent, weak bonds between electron donors and recipients. There are two hydrogen bonds between A and T and three hydrogen bonds between G and C, thus the bonds between G and C are stronger than between A and T

- Protein factors involved in DNA replication
  - Topoisomerase unwinds the double helix spiral so the chains can separate
  - Helicase separates the DNA double helix into two single strands and splits the hydrogen bonds binding the two strands at the replication fork
  - Single-stranded binding protein keeps the single-stranded DNA from reannealing
  - Primase is an RNA polymerase that synthesizes the RNA primers to initiate replication
    - Primase catalyzes the synthesis of short RNA primers on single-stranded DNA templates used by DNA polymerase to initiate the synthesis of the leading strand and of Okazaki fragments on the lagging strand
    - Primase is an essential enzyme in all well-characterized systems of DNA replication because no DNA polymerase initiates DNA synthesis without a short nucleic acid primer
  - DNA polymerase synthesizes the new strand from nucleotides complementary to the sequence in the template strand
  - RNAse H removes the RNA primer that starts DNA synthesis of each DNA fragment
  - DNA ligase joins the short DNA fragments into a continuous DNA strand
- DNA polymerases
  - DNA polymerases are enzymes involved in DNA replication. Eukaryotic cells have five different DNA polymerases
    - DNA polymerase α and δ replicate nuclear DNA
    - DNA polymerase β and ε are involved in DNA repair
    - DNA polymerase γ replicates mitochondrial DNA (mtDNA)
- Semiconservative replication
  - The two original complementary strands in the DNA double helix separate
  - New strand synthesis uses each original strand as a template
  - Each newly synthesized double helix is a combination of one original and one new DNA strand
- Multiple replications mean that replication begins at multiple sites within both DNA original strands and proceed bidirectionally from multiple origins on each strand (Fig. 1.7)
  - In eukaryotic cells, DNA replication occurs bidirectionally from the two original strands

Table 1.1 Major characteristics of repetitive DNA

Form of DNA	Length (bp)	Number of repeats
Single copy	Vary	Single copy
Moderately repetitive	Vary	$10–10^5$
Highly repetitive	Vary	$>10^5$
Tandem repeat		
Megasatellite	>1,000	50–400
Satellite	5–170	500–2,000
Minisatellite	14–500	7–40
Microsatellite	<15	10–100

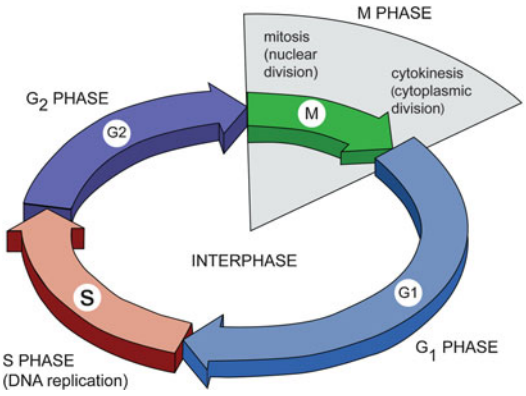


Fig. 1.5 The cell cycle, or cell-division cycle, is the series of events in a eukaryotic cell between one cell division and the next. The cell cycle consists of five phases, G_0, G_1, S, G_2, and M phase. G_0 is a period in which cells exist in a metabolically active but replicationally quiescent state. G_0, G_1, G_2, and S phase are collectively known as interphase with no light microscopic evidence of chromosome assembly. Cells in G_0 phase are resting cells unable to divide without a signal to reenter the cell cycle. DNA synthesis occurs during S phase, which is followed by a short G_2 phase for mitotic spindle protein synthesis. Mitosis and cytokinesis together are defined as the *M* (mitotic) phase, during which the mother cell divides into two daughter cells

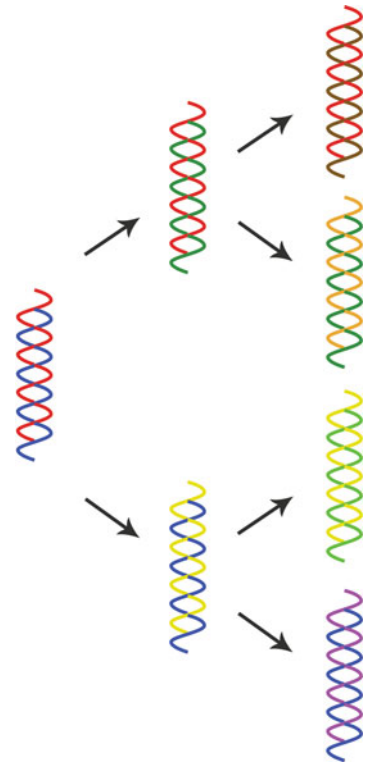


Fig. 1.6 Semiconservative DNA duplication occurs when DNA replicates, each parental chain is used as a template for synthesis of a complimentary daughter chain. Newly formed duplex strands contain one parental and one daughter chain as indicated by the color scheme

- The replication apparatus at each origin forms a bubble and extends toward both ends of the DNA molecule until it meets another bubble
- The leading strand is synthesized continuously in the 5′ → 3′ direction, since the replication fork moves in the 5′ → 3′ direction
  - Synthesis of the leading strand is catalyzed by DNA polymerase δ
- The 3′ → 5′ template duplicates in a complicated way to produce the lagging strand
  - The replication fork moves toward the 3′ end of the original strand
  - DNA synthesis proceeds in this reverse direction toward the 3′ end by a series of 5′ → 3′ replications of lagging strand segments
  - The lagging strand is synthesized as a series of short fragments, known as Okazaki fragments, polymerized in the 5′ → 3′ direction using RNA primers (Fig. 1.8)
  - The lagging strand polymerizes from 5′ to 3′ at the nucleotide level, but overall growth by ligation of Okazaki fragments is in the 3′ → 5′ direction by ligase joining of Okazaki fragments
  - The lagging strand is synthesized by DNA polymerase α

1.1.4 DNA Mutation

- DNA mutation is a permanent change in the genetic material sequence
- Single base pair substitution (point mutation) involves a single nucleotide, which is replaced by another nucleotide

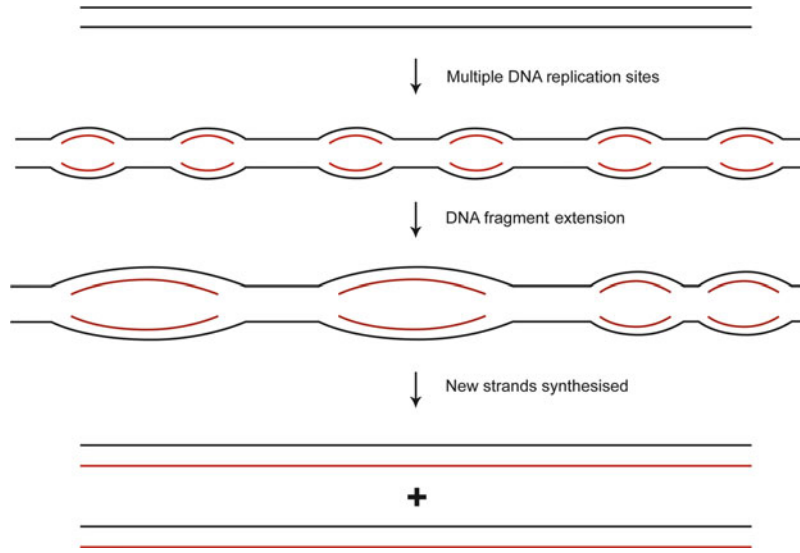


Fig. 1.7 Multiple replications are an efficient way to synthesize chromosomal DNA. DNA synthesis begins at many locations and proceeds bidirectionally from each location. Eventually, the replication bubbles merge and the DNA fragments are ligated to form two daughter DNA strands. Each of the newly synthesized double strands consists of one parental and one newly synthesized chain (semiconservative)

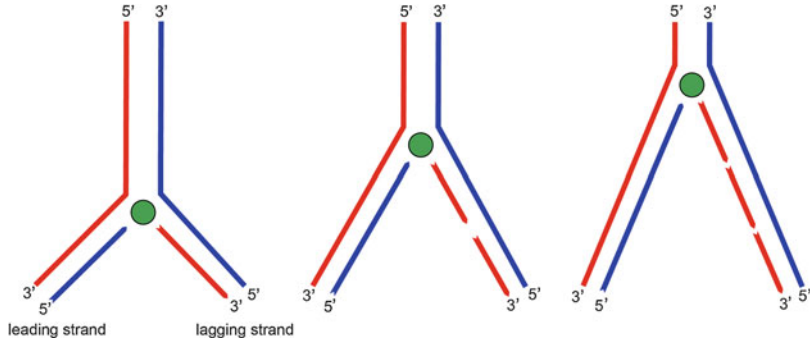


Fig. 1.8 The DNA double helix is unwound by the enzyme topoisomerase, and the strands separated by helicase before synthesis of new DNA chains begin usually in an AT rich region of the DNA. A DNA polymerase (shown in *green*) binds to the strand and moves along the strand assembling the leading strand (fragment inside the *left fork*). The lagging strand is synthesized in discontinuous polynucleotide segments called Okazaki fragments. A series of Okazaki fragments are linked by DNA ligase to form the lagging strand. In eukaryotic cells the leading and lagging strands are synthesized by DNA polymerase δ and α, respectively

- Point mutation is the most common form of mutation
- It happens most commonly in noncoding sequences, which comprise 98% of genomic DNA
- Point mutation is also the most frequent type of mutation associated with tumor suppressor gene inactivation
  - Transitions are the mutations that substitute a different purine for a purine or a pyrimidine for a pyrimidine
  - Transversions are mutations that substitute a purine for a pyrimidine or a pyrimidine for a purine
  - Synonymous (silent) mutation is a single base pair substitution yielding a different codon that still codes for the same amino acid
  - Silent mutations may change the rate of protein synthesis, alter the half-life of its mRNA, or modify posttranslational parameters such as protein folding

    - Missense mutation is a single base pair substitution that results in a different codon and a different amino acid
    - Nonsense mutation is a single base pair substitution that converts a codon specifying an amino acid into a stop codon
- Recombination involves motion of any length DNA from one physical location in the genome to another location
  - Recombination may be a harmless physiologic process that causes no mutation but allows for greater diversity of gene products or for repair of double-strand DNA breaks
  - Recombination may also be a pathologic process causing duplication of oncogenes or LOH inactivating tumor suppression
- Deletion is an irreversible mutation in which one or more nucleotides are removed from the DNA sequence
  - Deletions usually cause a shift of the reading frame
  - A one base deletion, for example, will shift all codons left, altering the amino acids for which they code in all downstream codons
- Insertion is a mutation that adds one or more nucleotides to the DNA sequence
  - An insertion in the coding region of a gene may also cause a shift in the reading frame
  - An insertion frequently alters splicing of messenger RNA (mRNA) (splice site mutation)
  - Amplification increases the dosage (number of redundant copies) of genes located within a locus by inserting multiple copies of the chromosomal region or by promulgating fragments of DNA containing the locus outside the chromosomes. Proteins produced from amplified genes are generally increased in number or concentration
- LOH is a DNA alteration in which one allele from one parent's contribution is lost, either by deletion or by a recombination event
  - LOH suggests gene loss in the region close to the microsatellite locus used to probe for LOH
  - The lost gene is often a tumor suppressor gene
  - The remaining copy of the tumor suppressor gene is also frequently mutated
  - One of the most frequently observed genes associated with LOH in sporadic cancer is *TP53*
- Recombination can be between homologous sites or nonhomologous sites involving motion of any length DNA from one physical location in the genome to another location
  - Recombination may be balanced and shows no loss or gain of DNA material
  - Homologous recombination may be involving the repair process of double-strand DNA breaks. Homologous recombination results in an exact exchange of genetic information
  - Recombination can also occur between nonhomologous DNA sequences
  - Recombination may also be a pathologic process causing duplication of oncogenes or loss of heterozygosity (LOH) inactivation of tumor suppressor genes

1.1.4.1 DNA Mutation and Disease

- DNA mutations cause errors in protein sequences, creating partially or completely nonfunctional proteins
- DNA mutation causes abnormal signal transduction, leading to aberrant activation or loss of function in related pathways
- DNA mutations in drug metabolizing enzymes or cytochromes alter the response to drugs
- DNA mutations in somatic cells give rise to mutations in germ cells and gametes that pass to at least 50% of the offspring
  - Human beings with a single allele mutation will transmit the mutation to half of the progeny
  - If both alleles are mutated, all progeny will inherit the mutation

1.1.4.2 Factors Related to DNA Aberrations

- Spontaneous chemical reactions
  - Spontaneous chemical reactions cause single base pair substitutions or single base

pair deletion through the following processes
- Tautomerism – base change by repositioning a hydrogen atom
- Depurination – loss of a purine base (A or G)
- Deamination – changes a normal base to an atypical base
  - Change from C → U
  - Spontaneous deamination of 5-methycytosine (irreparable) to thymine
  - Change from A → HX (hypoxanthine, a possible universal nucleobase)

- Induced mutations
  - Chemical mutagenesis can modify bases and cause either interstrand or intrastrand covalent cross-linking. Common chemical mutagens include:
    - Nitrosoguanidine (N-methyl-N′-nitrosoguanidine) is a powerful mutagen that causes random DNA mutations
    - Hydroxylamine (NH_2OH) causes G to C and A to T transitions
    - Base analogs (e.g., bromodeoxyuridine) only mutate DNA by being incorporated during S phase in replicating DNA in place of deoxythymidine (T)
    - Alkylating agents (e.g., cyclophosphamide) mutate both replicating and nonreplicating DNA. The alkylating agent transfers an alkyl group, often to the N7 position of guanine
    - Polycyclic hydrocarbons are converted within cells to highly active epoxy compounds that react with DNA (e.g., benzpyrenes found in internal combustion engine exhaust and cigarette smoke)
    - DNA intercalating agents insert themselves between the stacked bases at the center of the DNA strand (e.g., ethidium bromide)
    - DNA cross-linkers cause both interchain covalent bonds and stable bonds between the DNA strands and nuclear proteins (e.g., platinum)
    - Oxidative damage caused by oxygen radicals accelerates hydroxylation of guanine to 8-hydroxyguanine, causing mismatch pairing and G to T and C to A transversions after replication
  - Radiation
    - Ionizing radiation can cause individual base lesions, cross-linking, or strand breakage, sometimes mediated by oxygen radicals
    - Ultraviolet radiation causes covalent bonding between adjacent cytosine and thymine bases, creating pyrimidine dimers
  - Viral mutagenesis involves a DNA virus or cDNA reverse transcript of an RNA virus (retrovirus) integrating all or part of its sequence into the human genome
    - An episome is a DNA molecule separate from the chromosomal DNA and capable of autonomous replication. It is the usual status of viral particles during viral infection. Integration means that DNA fragments of viral origin have become inserted into chromosomal DNA
    - Epstein-Barr virus genes promote B lymphocyte division and have been associated with lymphoproliferative disorder in immunocompromised patients (such as posttransplantation), with Hodgkin disease, with B cell lymphomas, especially Burkitt lymphoma, and with nasopharyngeal carcinoma
    - Human herpesvirus 8 has been associated with Kaposi sarcoma, Castleman disease, body cavity lymphoma (primary effusion lymphoma), and multiple myeloma
    - Human papillomavirus has been associated with premalignant and malignant transformation of the uterine cervix, and with 15–25% of squamous cell carcinomas of the head and neck

- Inborn errors of metabolism
  - Inborn errors comprise a large class of genetic diseases involving metabolic disorders
  - The majority of inborn errors are due to single gene defects that code for enzymes to convert intermediary metabolites

1.1.5 DNA Repair Mechanisms

- DNA repair includes a collection of processes through which a cell identifies and corrects damage to its DNA. DNA repair is essential to cell survival
- If DNA damage is irreparable, then programmed cell death (apoptosis) should ensue
- Failure to correct molecular lesions in gamete-forming cells leads to progeny with congenital mutations
- Type of damage needing repair (see Sect. 1.1.4.2, "Induced mutations")
  - Oxidation, often of guanine to 8-oxoguanine and derivatives
  - Alkylation at guanine O-6 mispairing with thymine and other miscoding mutations
  - Hydrolysis of the glycosidic bond between sugar and base causing depurination an estimated 10,000 times a day
  - Adduct formation with strongly electrophilic proximate carcinogens, often produced in cells by polycyclic hydrocarbons
  - Mismatch of bases. Metals such as cadmium, chromium, and nickel reduce the fidelity of DNA replication
- Direct reversal
  - Cells can eliminate three types of damage by direct reverse chemical restoration
  - Only a few types of DNA damage are repaired in this way, particularly pyrimidine dimers
  - Direct reversal does not require a template
  - Direct reversal mechanisms do not involve breakage of the phosphodiester backbone
  - Mutations that can be fixed by this mechanism include pyrimidine dimers induced by UV light and DNA damage caused by alkylating agents
- Single-strand damage repair
  - The intact strand is used as a template in the single-strand damage repair
    - Base excision repair (BER) removes a damaged base and replaces it with a normal base (Fig. 1.9) selected by its complementary base in the undamaged strand
    - Nucleotide excision repair (NER) removes a short single-stranded DNA fragment containing the helix-distorting lesion and subsequently fills in the defect by DNA polymerase
    - Mismatch repair (MMR) removes a series of nucleotides, which have been mispaired with the corresponding base on the complementary parent chain, usually by DNA polymerase errors in the daughter strand. This repair process can remove up to 30 base insertions (Fig. 1.10)
  - Transcription-coupled repair (TCR), a type of NER, temporarily replaces the conventional RNA polymerase at a DNA lesion by one of a group of specialized polymerases that can replicate damaged DNA without displacing the RNA polymerase
  - Translesion synthesis (TLS) is a process during DNA replication that attempts to insert a "best guess" when a lesion such as a pyrimidine dimer or a basic site is found in the parent strand. The regular high fidelity DNA polymerase is temporarily switched out for a loose fitting translesion polymerase. The replication fork stalls only transiently for a solution to be found and then replication continues. It is better for the daughter cell to inherit a point mutation than to have part of a chromosome deleted, causing frameshift
- Double-strand break repair
  - Double-strand breaks in the DNA double helix are particularly harmful, for they may cause both gene deletion and genetic rearrangement

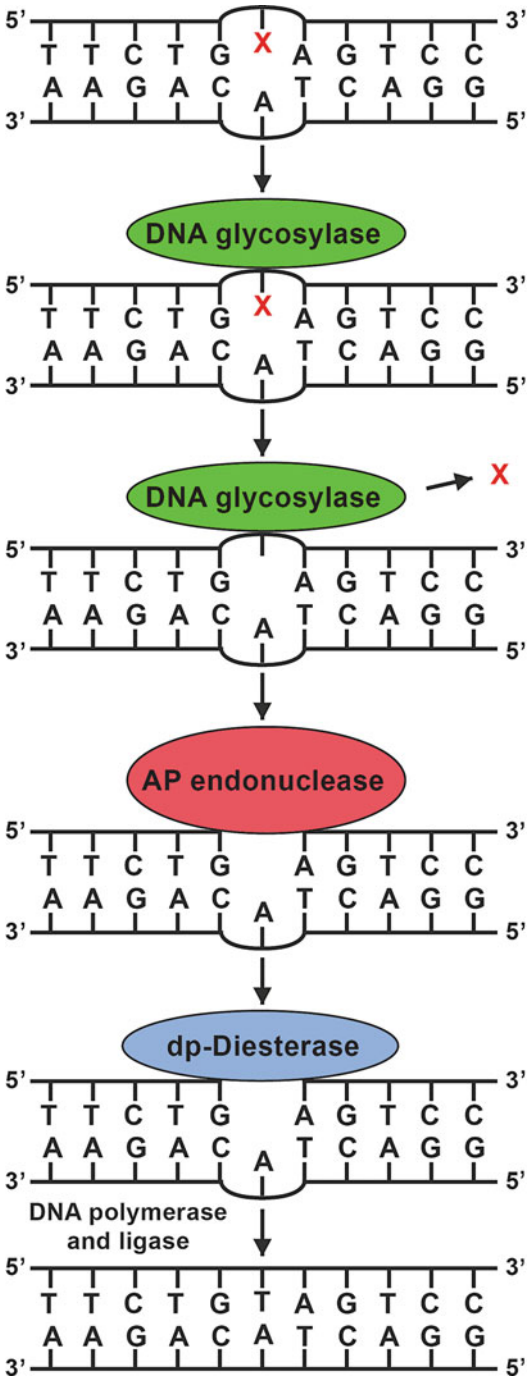


Fig. 1.9 A single base in DNA may be chemically mutated, for example, by deamination or alkylation, causing incorrect base pairing (*X*), and consequently, incorrect codons in the DNA. BER is initiated by DNA glycosylases linking particular types of chemically altered bases to the deoxyribose-phosphate backbone. This mutated base is excised as a free base, generating sites of base loss called apurinic or apyrimidinic (*AP*) sites. The AP sites are substrates for AP endonucleases that cleave the 3′ phosphodiester bond of the AP sugar. The ribose-phosphate backbone is then removed from the DNA by an exonuclease called deoxyribophosphodiesterase (dp-diesterase). Then the DNA polymerase and a ligase catalyze incorporation of a specific deoxyribonucleotide into the repaired site, enabling correct base pairing. A slightly different but analogous process, called long-patch repair, occurs when multiple tandem-mutated bases are replaced in series

- Homologous recombination (HR) uses an intact homolog sequence as a template for repair of a broken DNA strand (Fig. 1.11)
- In the process, the 3′ of broken strand invades the intact sister chromosome and uses it as a template for a DNA synthesis, by which it copies the sequence across the break
- HR is similar to chromosome crossover in meiosis and requires a homologous chromosome as a template

– Nonhomologous DNA end joining (NHEJ) repairs DNA strands directly without the need for a homologous template (Fig. 1.12)
  - NHEJ does not require a homologous sister chromosome, but the double-strand breaks are approximated by the end-binding protein and then sealed by a specialized DNA ligase
  - NHEJ can lead to loss of nucleotides or translocation
  - When the NHEJ is not possible, double-strand breaks can be repaired by a more error-prone pathway called microhomology-mediated end joining (MMEJ)
– Microhomology-mediated end joining (MMEJ) repairs the DNA by using base-pairing interactions between short terminal or embedded microhomologies
  - The process includes end resection, microhomologies annealing, and ligation of the broken strands
  - MMEJ is more error prone than NHEJ because ends are matched by approximate pairing of the 5–25-base overhang tails followed by excision of nonpaired flaps

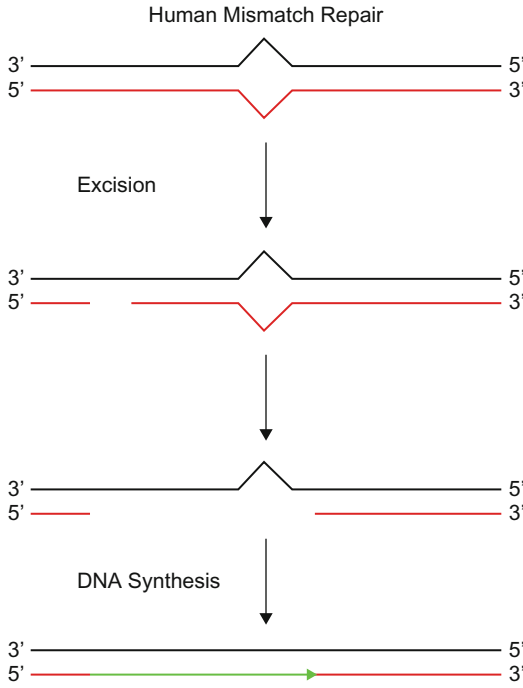


Fig. 1.10 Mismatch repair (*MMR*) is a cellular procedure for recognizing and repairing insertion, deletion, or misincorporation resulting from errors during DNA replication or recombination. The DNA damage is repaired by excising the misincorporated base or segment and synthesizing a new stretch of DNA to replace the excised segment. This process involves more than just the mismatched nucleotide itself and can lead to the removal and synthesis of a significant piece of DNA. Also, there are multiple excision-repair systems in a normal human cell

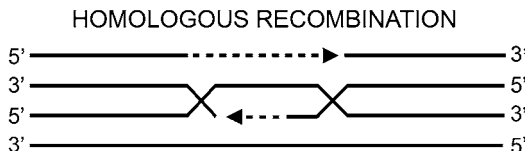


Fig. 1.11 Homologous recombination (*HR*) repairs a double-stranded break. It allows the precise replacement of a sequence from one allele with a sequence from the homologous allele. The breaks in dsDNA use a homologous dsDNA molecule as a template. Homologous recombination requires a homologous sequence to guide the repair. This type of repair leads to loss of heterozygosity, since one parental allele is the template for replacing the other damaged allele

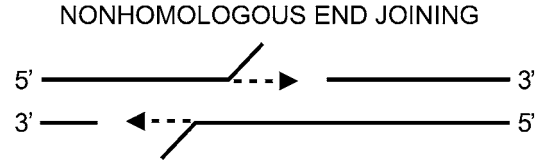


Fig. 1.12 Nonhomologous end joining (*NHEJ*) can also repair double-strand breaks in DNA. NHEJ directly ligates the breaks without a homologous template. NHEJ typically utilizes short homologous DNA sequences, termed microhomologous, to guide the repair but still results in some DNA sequence information being lost or "spliced out." Microhomologies often prevent frameshift but cannot eliminate deletion mutations resulting from the repair

and replacement of noncomplementary bases
- MMEJ results in deletion of DNA sequence
- The cell uses MMEJ only when the NHEJ is not possible

1.2 Genes

1.2.1 Overview

- Genes are DNA sequences that encode heritable biologic characteristics
  - The human genome is divided into two categories, nuclear and mitochondrial genome
- Human nuclear DNA contains about 23,000 protein-coding genes, fewer than previous estimates of around 100,000 genes before completion of the genome map
  - Protein-coding genes encode protein factors including cell structure proteins, enzymes, receptors, transporters, and immunoglobulin
  - RNA genes encode RNAs not translated into proteins, including tRNA, ribosomal RNA, microRNA, long nuclear RNAs that regulate gene expression, and ribozyme RNA that catalyzes chemical reactions
- DNA in the human mitochondrial genome encodes 37 genes

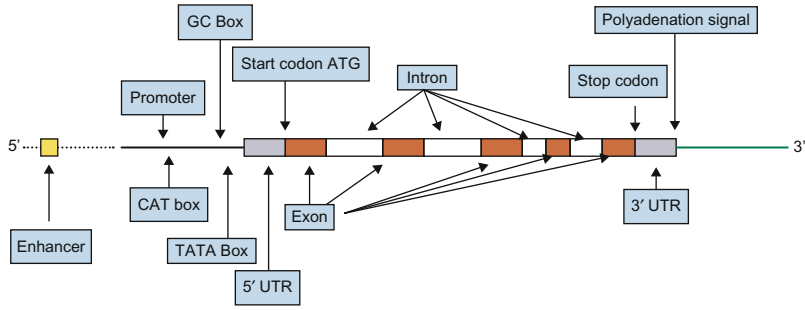


Fig. 1.13 A gene consists of both coding and noncoding sequences. The coding sequence (open reading frame, *ORF*) extends from a start codon to a stop codon. Introns are noncoding sequences that will be spliced out after transcription. 5′ untranslated region (*UTR*) is a part of mRNA located between the cap site and the start codon. 3′ untranslated region is also a part of mRNA following the coding sequence after the stop codon. A promoter and different regulatory motifs are located upstream of a gene (toward the 5′ end in the 5′ *UTR*). Enhancer or silencer may be located either upstream or downstream of the gene it regulates

- The transcription initiation site (TSS for transcription start site) is the first base of the first three-base codon for the first amino acid in the protein for which the gene codes. It is called base1

1.2.2 Gene Structures

- Promoter is a DNA segment to which RNA polymerase binds for initiation of transcription (Fig. 1.13)
  - Core promoter directs the basal transcription complex to initiate transcription of the gene
  - Promoter is identified by negative (upstream) base values at –30, –75, and –90
- TATA box is a short sequence located within the promoter of most genes
  - TATA box has a core 5′-TATAAA-3′ sequence
  - The TATA box is usually found as the binding site of RNA polymerase II and several other factors of the preinitiation complex
- CCAAT Box (CAT box) is located at −75 and serves as a modulator for the basal transcription, enhancing the rate of mRNA production
  - CCAAT box has a core 5′-CCAATC-3′ sequence
  - CCAAT box is the binding site of nuclear factor 1 (NF-1) and CCAAT box binding factor (CBF)
- GC box is also called Sp1 box. It has consensus sequences GGGCGG and is found within 100 bp of the transcription initiation site
  - GC boxes serve as a modulator to the basal transcription of the core promoter enhancing gene transcription during early development of an organism
- CpG sites are regions of DNA with a high frequency of cytosine-guanine sequence pairs found in about 40% of human gene promoter regions
  - The "p" in CpG refers to the phosphodiester bond between the cytosine and the guanine
  - These islands of high CG content may be 300–3,000 base pairs in length
  - The genes with CpG islands are expressed if the CpG islands are not methylated to contain 5-methylcytosines
- Enhancers are DNA sequences that when bound by certain factors increase the transcription activity of genes
  - Unlike promoters, enhancers do not have to be within or near the genes they activate, or even located on the same chromosome
- Silencer is a DNA sequence that can bind regulators of transcription called repressors

- The binding of repressor prevents RNA polymerase from initiating transcription; RNA synthesis is decreased or fully suppressed Exons are regions in a gene that encode amino acids of a protein. Exons of many eukaryotic genes interleave with segments of noncoding DNA (introns). Mature mRNA contains only sequentially linked exons
- Introns are sections of noncoding DNA located between exons, which are transcribed into RNA but spliced out to form mRNA
- Open reading frame (ORF) is the sequence of DNA or mRNA molecule from the start codon (ATG) to a stop codon (TAA, TAG, or TGA). An open reading frame codes for amino acid codons that can be translated into a protein and may contain both exons and introns
- Boundary elements create domain boundaries and also mark the border of a gene
  - Boundary elements insulate the gene expression from regulatory elements occurring in different domains
  - They also work as a barrier against heterochromatin condensation proteins spreading onto active euchromatin
- Gene expression is the process by which a gene's DNA sequence is converted into the structures and functions of a cell
  - Protein-coding genes are translated into proteins
  - Nonprotein-coding genes frequently code for RNAs, including rRNA, tRNA, microRNA, and catalytic RNA (ribozyme)

1.2.3 Functional Categories of Genes

- Housekeeping genes are genes required for the maintenance of basic cellular function. They are transcribed at a relatively constant level and remain unaffected by environmental conditions
  - Housekeeping gene products are necessary for cell maintenance
    - Since their expression is typically unaffected by external conditions, they may be used for normalization of inducible gene expression levels in the cell
    - Housekeeping genes often lack the CCAAT and TATA boxes
  - Beta actin (ACTB), and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) are examples of housekeeping genes commonly used as controls for mRNA quantitation
- Facultative genes are transcribed only when needed
- Inducible genes express the products either in response to environmental signals or in dependence on the stage of the cell cycle
- Pseudogenes are multiple copy genes characterized by homology to a known gene and nonfunctional, mostly truncated, product
  - Pseudogenes arise from gene duplication or retrotransposition
- RNA genes transcribe mRNA or noncoding RNAs as their end products without protein translation

1.2.4 Cancer-Related Genes

- Commonly found cancer-related genes include oncogenes and tumor suppressor genes (Table 1.2)
- Tumor suppressor genes prevent cell overgrowth (neoplasia) and spontaneous mutation
  - They are involved in cell cycle control, cell differentiation, and apoptosis
  - Tumor suppressor gene products generally promote genomic stability
  - Tumor suppressor gene inactivation mechanisms include point mutation, deletion, and epigenetic inactivation of the gene
- Proto-oncogenes are normal genes that code for proteins to regulate cell growth and differentiation and cause a malignant phenotype either by mutation or by increased expression
  - A proto-oncogene becomes an oncogene when mutated, inappropriately expressed, or overexpressed, transforming the cell by unregulated growth and inappropriate differentiation
- Oncogenes are frequently mutated proto-oncogenes, characterized by the ability to transform cells when activated

Table 1.2 Selected tumors with commonly found chromosomal anomalies

Tumor	Common chromosome anomalies	Genes involved
Epithelial tumors		
Basal cell carcinoma of the skin	9q22.3	*PTCH*
Clear cell renal carcinoma	3p25–26	*VHL*
Translocation renal cell carcinoma	t(X;1)(p11.2;q21)	*PRCC-TFE3*
	t(X;17)(p11.2-q25)	*ASPL-TFE3*
	t(X;1)(p11.2;p34)	*PSF-TEF3*
Papillary renal cell carcinoma	Gain 7, 17, loss of Y	–
Hereditary papillary renal cell cancer	7q31	*c-MET*
Breast cancer	17q21	*BRCA1, Her-2/neu*
	13q12	*BRCA2*
	1q	–
	del(16q)	–
Colorectal cancer	12p12	*KRAS*
	del(17p)	*TP53*
	3p14	*FHIT*
	5q21–22	*APC*
	18q21	*DCC, SMAD4*
Lung cancer	7p11.2	*EGFR*
	inv(2)(p22-p21p23)	*EML4-ALK*
	del(3p)	*FHIT*
	13q	*RB*
	9p21	*P16*
	17p	*TP53*
Prostate cancer	del(21)(q22.3q22.3)	*TMPRSS2-ERG* fusion
	del(8p12–21)	*NKX3.1*
	Iq24	*HPC1*
	Xq27–28	*HPCX*
	Xq11	*AR*
	del(10q24)	*PTEN*
	Trisomy 7	–
	Loss of Y	–
Bladder urothelial carcinoma	gain 3, 7, 17, del(9p21) (UroVysion panel)	*P53, P16, FGFR3*
Medullary thyroid carcinoma	10q11.2	*RET*
Papillary thyroid carcinoma	10q11–q13	*RET*
	inv(1)	*NTRK1-TPM3 (TRK)*
	inv(10)(q11.2q11.2)	*RET-NCOA4*
	inv(10)(q11.2q21)	*RET-CCDC6*
Follicular thyroid carcinoma	t(2;3)(q13;p25)	*PAX8-PPARG*
Mesothelioma	del(3p21)	*CTNNB1*
Ovarian papillary cystadenocarcinoma	t(6;14)	–
Granulosa cell tumor and Brenner tumor	trisomy 12	–
	i(12p)	–
Testicular germ cell tumors	i(12p)	–
	12p overrepresentation	–
Wilms tumor	del(11p13)	*WT1*

(*continued*)

Table 1.2 (continued)

Tumor	Common chromosome anomalies	Genes involved
Soft tissue tumors		
Alveolar soft-part sarcoma	t(X;17)(p11;q25)	*TFE3-ASPL*
Alveolar rhabdomyosarcoma	t(2;13)(q35;q14)	*PAX3-FKHR*
	t(1;13)(p36;q14)	*PAX7-FKHR*
Clear cell sarcoma (melanoma of soft part)	t(12;22)(q13;q12)	*EWS-ATFI*
Dermatofibrosarcoma protuberans and giant cell fibroblastoma	t(17;22)(q22;q13)	*COL1A1-PDGFB*
Myxoid chondrosarcoma	t(9;22)(q22;q12)	*EWS-CHN*
Lipoma	t(3;12)(q27;q13)	*HMGIC-LPP*
Lipoblastoma	8q rearrangement	–
Myxoid liposarcoma	t(12;16)(q13;p11)	*CHOP-FUS*
	t(12;22)(q13;q12)	*EWS-CHOP (DDIT3)*
Well-differentiated liposarcoma	Ring chromosome 12	–
Ewing sarcoma/primitive neuroectodermal tumor	t(11;22)(q24.1-q24.3;q12.2)	*EWS-FL11, EWS-ERG*
	t(21;22)(q22.3;q12.2)	*ERG-EWSR1*
	t(21;22)(q22;q12)	
Desmoplastic small round cell tumor	t(11;22)(p13;q12)	*EWS-WT1*
Synovial sarcoma	t(X;18)(p11;q11)	*SYT-SSX1 (biphasic), SYT-SSX2 (monophasic)*
	t(X;20)	–
Infantile fibrosarcoma and congenital mesoblastic nephroma	t(12;15)(p13;q25)	*ETV6-NTRK3*
Inflammatory myofibroblastic tumor	t(1;2)(q22;p23)	*TPM3-ALK*
	t(2;19)(p23;p13)	*TPM4-ALK*
Gastrointestinal stromal tumor	4q11–21	*c-kit* exon 11
Hemangiopericytoma	t(12;19)	–
Uterine leiomyoma	t(12;14)(q13–15;q24.1)	*HMGIC*
Endometrial stromal sarcoma	t(7;17)(p15–p21;q12–q21)	*JAZF1-JJAZ1*
Leiomyosarcoma	del(1p)	–
Pleomorphic adenoma	t(3;8)(p12;q12)	*FGFR1-FIM*
Aneurysmal bone cyst	17p rearrangement	–
Desmoplastic fibroblastoma and fibroma of tendon sheath	t(2;11)	–
Melanoma	7q34	*BRAF*
	del(9p21)	*CDKN2*
	12q14	*CDK4*
	del(22q)	–
Neural/neuroendocrine tumors		
Acoustic neuroma/Schwannoma	22q12.2	*NF2*
	del(22q13)	*NF2*
Meningioma	del(22q11-q13)	*NF2*
	Monosomy 22	–
Glioblastoma multiforme	7p11	*EGFR* gene amplification
	Trisomy 7	
	Monosomy 10, 22	
	t(10;19)(q24;q13)	

(*continued*)

Table 1.2 (continued)

Tumor	Common chromosome anomalies	Genes involved
Medulloblastoma	del(10q)	*PTEN* and *DMBT1* (10q)
	i(17q)	–
	del(17p13.2)	–
	Trisomy 8	–
Neuroblastoma	2p	*N-Myc* amplification (homogenous staining region and double minutes)
	del(1p31–32)	
Oligodendroglioma	del(1p36)	
	del(19q13)	
Retinoblastoma	del(13q14)	*RB*
Pheochromocytoma	del(22q13)	*SLC1*
	del(1p11–36)	*RIZ1*
Lymphomas/leukemia		
Burkitt lymphoma	t(8;14)(q24;q32)	*c-myc-IgH*
	t(8;22)(q24;q11)	*c-myc-IgL*
	t(2;8)(p12q24)	*c-myc-IgK*
Follicular lymphoma	t(14;18)(q32;q21)	*IgH-BCL2*
Mantle cell lymphoma	t(11;14)(q13;q32)	*IgH-BCL1* (cyclin D1)
	11q	*ATM* mutation or deletion
	del(13q14)	–
	Trisomy 12	–
	del(17p)	–
Extranodal marginal zone of lymphoma of MALT	t(11;18)(q21;q21) Trisomy 3	*AP12-MALT1*
Diffuse large cell lymphoma	t(14;18)(q32;q21)	*IgH-BCL2*
	t(3;14)(q27;q32)	*BCL6-IgH*
Anaplastic large cell lymphoma	t(2;5)(p23;q35)	*NPM-ALK*
Lymphoplasmacytic lymphoma	t(9;14)(p13;q32)	*PAX5*
Multiple myeloma	Monosomy or partial	*IgH-BCL1* (cyclin D1)
	t(4;14)(p16.3;q32.33)	*WHSC1-IGHG1*
	deletion of 13 (13q)	
	t(11;14)(q13;q32)	–
Myelodysplastic syndrome	del(5q)	–
	Trisomy 8	–
	Monosomy 7	–
	del(7q)	–
	del(17p)	–
Myelodysplasia with ring sideroblasts	2q33	*SF3B1*
Myeloid neoplasm associated with eosinophilia	del(20q)	–
	del(4)(q12q12)	*FIP1L1-PDGFRA*
	t(5;12)(q31-q32;p13)	*PDGFRB-ETV6*
Chronic myelogenous leukemia	t(9;22)(q34;q11)	*BCR-ABL*
AML with recurrent genetic abnormalities	t(8;21)(q22;q22)	*RUNX1/RUNX1T1*
	inv(16)(p13.1q22) or t(16;16)(p13.1;q22)	*CBFB/MYH11*
	t(15;17)(q22;q21)	*PML/RARA*
	t(9;11)(p22;q23)	*MLLT3/MLL*

(*continued*)

Table 1.2 (continued)

Tumor	Common chromosome anomalies	Genes involved
	t(6;9)(p23;q34)	*DEK/NUP214*
	inv(3)(q21q26.2) or t(3;3)(q21;q26.2)	*RPN1/EVI1*
	t(1;22)(p13;q13)	*RBM15/MKL1*
	mutated NPM1	
	mutated CEBPA	
Acute myelogenous leukemia (AML) (formerly M2)	t(8;21)(q22;q22)	*RUNX1-RUNX1T1 (formerly AML-ETO)*
Acute promyelocytic leukemia (formerly M3)	t(15;17)(q22;q21)	*PML-RARα*
AML (formerly M4eo)	inv(16)(p13;p22)	*MYH11-CBFB*
AML (formerly M4, M5)	t(11;19)	*MLL-ENL*
	11q23	*MLL*
AML (formerly M5)	t(9;11)(p22;q23)	*MLL-AF9*
Acute megakaryoblastic leukemia (formerly M7)	t(1;22)(p13;q13)	*RBM15-MKL1*
Pediatric AML	translocation involving 11q23	*MLL-AF9*
Acute lymphoblastic leukemia	episome(9q34.1)	*NUP214-ABL1*
	t(12;21)(p13;q22.3)	*ETV6-RUNX1*
Pre-B ALL	t(1;19)	*E2A-PBX1*
B-ALL	t(12;21)	*TEL-AML1*
Infantile ALL	t(4;11)	*AF4-MLL*
ALL	t(12;21) (most common translocation in ALL)	*TEL-AML1*
Chronic lymphocytic leukemia/small lymphocytic lymphoma(CLL)	del (13q)	–
	del (11q)	–
	del (17p)	–
	trisomy 12q	–
	t(11;14)(q13;q32)	*IgH-BCL1* (cyclin D1)

- Oncogene products include:
  - Growth factor receptors on the cell membrane
  - Receptor tyrosine kinases, which phosphorylate target proteins that regulate the activity of cellular pathways
    - Receptor tyrosine kinases are membrane-bound enzymes to transfer a phosphate group from ATP to a tyrosine residue in a regulatory protein
    - The tyrosine kinase-binding hormones and growth factors are generally growth-promoting and mitogenic agents, such as epidermal growth factor (EGF) whose receptor is called EGFR
    - This process usually results in a functional change of the target protein, resulting in changed enzyme activity, altered cellular location, or modified association with other proteins
    - Cytoplasmic tyrosine kinases are nonreceptor tyrosine kinases (TK) that also regulate many cellular processes
  - Regulatory GTPases activate the following growth pathways:
    - Signal transduction at the intracellular domain of transmembrane receptors

  - Protein biosynthesis at the ribosome
  - Control of differentiation during cell division
  - Translocation of proteins through membranes
  - Transport of vesicles within the cell
  - *RAS* oncogene produces a small regulatory GTPase, important as a molecular switch for a variety of signal pathways, including cytoskeletal integrity, cell proliferation, adhesion, apoptosis, and migration
- Transcription factors which promote or repress gene transcription by sequence-specific DNA binding
  - Transcription factors mediate the binding of RNA polymerase to DNA and initiation of transcription
- Cytoplasmic serine/threonine kinase phosphorylates the hydroxyl group of serine or threonine. The *RAF* kinase and cyclin-dependent kinases belong to the serine/threonine kinase family
- Adaptor proteins are small accessory proteins, which lack intrinsic enzymatic activity but bind signal transduction pathway components, driving the formation of active protein complexes

1.2.5 Regulation of Gene Expression

- Regulation of gene expression controls the amount and appearance agenda of a gene's functional product
  - All steps of gene expression can be modulated, including preinitiation, initiation, promoter clearance, elongation, and termination
  - Regulation of gene expression is the basis for cell differentiation, diversity, and adaptation
- *Cis*-action factors are short regulatory sequences located within the promoter or in the vicinity of a gene's structural portion. *Cis*-sequences facilitate the transcription of adjacent polypeptide-encoding sequences
- *Trans*-action factors bind to the *cis*-acting sequences to control gene expression
- Enhancer is a short region of DNA that upregulates gene transcription levels. Enhancer sequences are active when bound to *trans*-action factors
- Response element is a short sequence of DNA within the promoter region that binds to a specific hormone receptor complex and regulates transcription of genes subject to that hormone

1.2.6 Signal Transduction

- Signal transduction involves signal molecules, their receptors, second messengers, and effector processes
- A signal transduction pathway is a sequence of enzymes and second messengers by which a receptor communicates with the cell nucleus
- Signal transduction is initiated by the signal molecule binding to its receptor
- The receptor in turn alters intracellular molecules that pass the signal to the downstream effectors
  - A signal molecule activates a specific receptor
  - The second messenger passes the signal to downstream factors, which induce a variety of responses
  - G protein-coupled receptors comprise a large family of transmembrane proteins that bind to ligand molecules outside the cell and activate signal transduction pathways inside the cell
  - Receptor tyrosine kinase binds its ligand with the extracellular domain inducing dimer formation, causing autophosphorylation of tyrosines in the intracellular domain and activating the signal cascade of downstream factors
  - Integrins lack kinase activity; therefore, integrin signal transduction depends upon other protein kinases and adaptor molecules. The integrins regulate the cell survival, proliferation, motility, and differentiation

- Transcription factors
  - A transcription factor is a molecule that initiates transcription of specific genes in DNA of the eukaryotic nucleus
  - Transcription factors interact with promoter or enhancer sequences either by binding directly to DNA or by interacting with other DNA-bound proteins
  - *MYC* is an example of a transcription factor that activates expression of many genes by binding to consensus sequences of conserved bases in promoter regions for up to 15% of all *mRNA* genes in the genome
- Programmed cell death (apoptosis) and apoptosis regulators
  - The process of programmed cell death is triggered by either extracellular or intracellular molecules
  - Activation of apoptosis regulator molecules will either positively or negatively influence the cell death pathway
  - The intrinsic pathway is initiated by the signals resulting from DNA damage, a defective cell cycle, detachment from the extracellular matrix, hypoxia, loss of cell survival factors, or severe cell stress
    - The release of proapoptotic proteins of the intrinsic pathway activates caspase enzymes from the mitochondria
  - The extrinsic pathway is activated by proapoptotic ligands. The extrinsic pathway triggers apoptosis independently from p53
  - Alternatively, apoptotic mechanisms may be initiated by increased calcium concentration within a cell

1.3 Chromosomes

1.3.1 Overview

- A chromosome is an enormous macromolecule into which somatic DNA is packaged in eukaryotic cells. Three billion nucleotide base pairs (a complete set of DNA) are divided among 46 chromosomes, each containing many genes, regulatory elements, and intervening nucleotide sequences (Fig. 1.14)

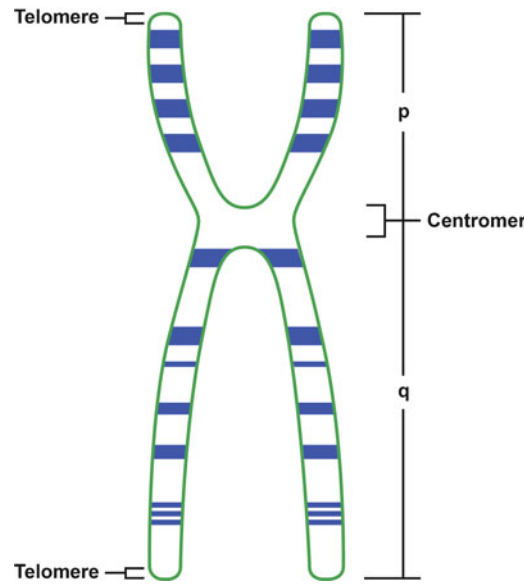


Fig. 1.14 Structure of a typical human chromosome

- Chromosomes are found only in the eukaryotic nucleus and can be seen in condensed form only during nuclear division (mitosis)
- During interphase, the genetic material occupies areas of nuclei in the form of chromatin, and individual chromosomes cannot be distinguished except by fluorescent probes complementary to specific chromosome sequences
- In eukaryotes, the basic function of chromatin is to package and compress the DNA, exposing specific genes for transcription during certain phases of the cell life span and condensing the rest in a highly efficient, volume conserving compact form
- Typical eukaryotic chromosomes are large and linear, but mitochondria in most eukaryote cells contain small circular chromosomes containing only 37 genes

1.3.2 Chromatin

- Chromatin is the orderly combination of DNA and proteins, the main content of the cell nucleus
- Chromatin is a packaged state of DNA in a small volume to strengthen the DNA, to

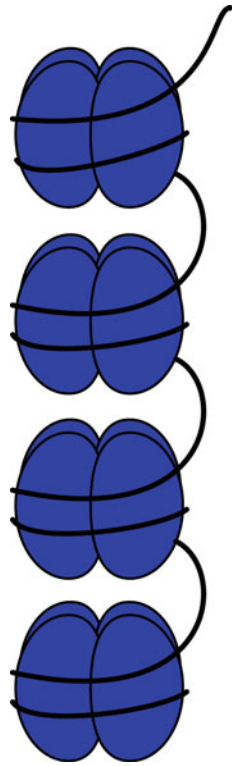


Fig. 1.15 Nucleosomes are the fundamental repeating subunits of all eukaryotic chromatin. They package DNA into chromosome supercoils inside the cell nucleus and control gene expression. The DNA winding around the nucleosome core particle consists of about 146 bp of dsDNA wrapped in 1.65-left-handed superhelical turns around protein disks composed of four pairs of histone proteins known as the histone octamer. The DNA hanging between two nucleosome cores is typically 55 bp long and is known as linker DNA

allow mitosis and meiosis, and to serve as a mechanism for expression control

- Chromatin is seen with the light microscope after staining with nuclear stains
- Chromatin functions as a gene regulator
  - The structure of chromatin depends on the stage of the cell cycle
    - In interphase, much of the chromatin is structurally loose to allow transcription of DNA (euchromatin)
    - Epigenetic molecules regulate transcription activity by modifying the structural proteins in chromatin through methylation and acetylation
    - In mitosis or meiosis, the chromatin is packaged more tightly to form individual chromosomes, visible by optical microscopy during metaphase
    - Bookmarking is an epigenetic mechanism to transmit the "memory" of pre-mitotic gene activity to daughter cells; thus, hepatocytes divide into more hepatocytes
  - Euchromatin is a loosely packed form of chromatin that is involved in active transcription or regulation and is lightly stained by nuclear stains
    - Euchromatin refers to actively transcribed regions of the cell's DNA
  - Heterochromatin is a darkly staining and tightly packed form of DNA
    - Heterochromatin refers to regions that are inactive for transcription either to mRNA, regulatory RNA, or functional RNA
  - The layout of the genome within the nucleus is not random even in interphase but is not yet well characterized except that one X chromosome in mammalian females aggregates to the nuclear membrane in a Barr body
  - A chromatid is one-half of the two identical copies of DNA joined at the centromere to make up a chromosome during mitotic phase
    - The centromere attaches the metaphase chromosome to tubulin fibers of the mitotic spindle and brings about the orderly separation of identical chromatids during mitosis
    - During mitosis, the two halves of the x-shaped chromosome separate, and each of the strands is called a daughter chromosome
- Chromatin composition
  - Histones are the major chromatin binding proteins. Histone octamers act as spools around which the DNA double helix winds
    - Histones play a role in gene regulation
    - Histones H2A, H2B, H3, and H4 form octamers (two of each) with a cylindrical shape (Fig. 1.15)
  - Double-stranded DNA wraps around histone proteins to form nucleosomes
    - When DNA winds 1.65 times around a histone octamer, a nucleosome results.
      - Each nucleosome contains 146 bp
  - Multiple nucleosome chained by DNA form strands with a "beads on a string" structure
    - An array of nucleosomes is stacked and further coiled into a 30-nm fiber, which

makes up the chromosome residing in the interphase cell nucleus
    - Linker DNA is the DNA hanging between two nucleosomes, typically 55 bp long
  - The overall structure and extent of condensation depends upon the stage of cell cycle
  - Chromatin structure alters DNA packing and regulates transcription activity
    - Histone-DNA interaction regulates gene expression. Acetylation of histone modulates gene expression and leads to transcription activation. The extent of interaction between histone and DNA is affected by the degree of histone acetylation
    - In histone acetylation, charged lysine side chains are acetylated, leading to reduced affinity between histone and DNA. After acetylation, RNA polymerase and transcription factors have better access to the promoter

1.3.3 Chromosomes

- Chromosome structure
  - Each chromosome has two short arms (p), two long arms (q), one centromere, and four telomeres at the ends of the four arms
    - The centromere is the constricted region of a chromosome, which has a special sequence and structure for attachment to the spindle filament during mitotic phase and for separation of chromosomes during mitosis (Fig. 1.14)
    - Centromeres are chromosome specific and therefore used for chromosome enumeration by fluorescent probes
    - The centromere divides the chromosome into four arms
      - The two equal short arms are designated "p" (*petite*)
      - The two equal long arms are designated "q" (follows p in the Latin alphabet)
      - Telomere (see Chap. 3, Telomere section)
- Chromosome grouping
  - Chromosomes are numbered and grouped according to the ratio of p-arm to q-arm length and total length from largest to smallest
  - The position of the centromere determines chromosome grouping
    - Group A has nearly equal p and q arms, whereas group E has the centromere almost at the telomere
    - Acrocentric chromosomes have such short p arms that the centromere appears as a "peak" of the two q arms (groups D and G, chromosomes 13–15, 21, 22)
  - Chromosome identification is confirmed by the banding pattern unique to each chromosome (Fig. 1.16)
- Chromosome number
  - Human cells contain 46 chromosomes (23 from each parent) including 22 pairs of autosomes and one pair of sex chromosomes, XX in female and XY in male
  - The number of chromosomes doubles during cell division
    - Meiosis is a process in eukaryotes allowing one diploid cell to divide in a special way into four haploid gamete cells
    - Mitosis is the process by which a cell separates its duplicated chromosomes into two daughter cells with identical chromosome sets
    - Ploidy is the number of homologous sets of 23 human chromosomes in a eukaryotic cell
      - Haploidy (monoploidy) is the number of chromosomes in the gamete of an individual (23 in a human). Haploid chromosomes have only one short and one long arm
      - Diploidy is the normal state of chromosomes in a somatic cell, with two sets of 23 chromosomes, one from each parent. The two chromosomes in a pair are said to be homologous
      - Tetraploidy is the normal state of chromosomes in a cell after replication prior to mitosis. Cells in metaphase have two short and two long

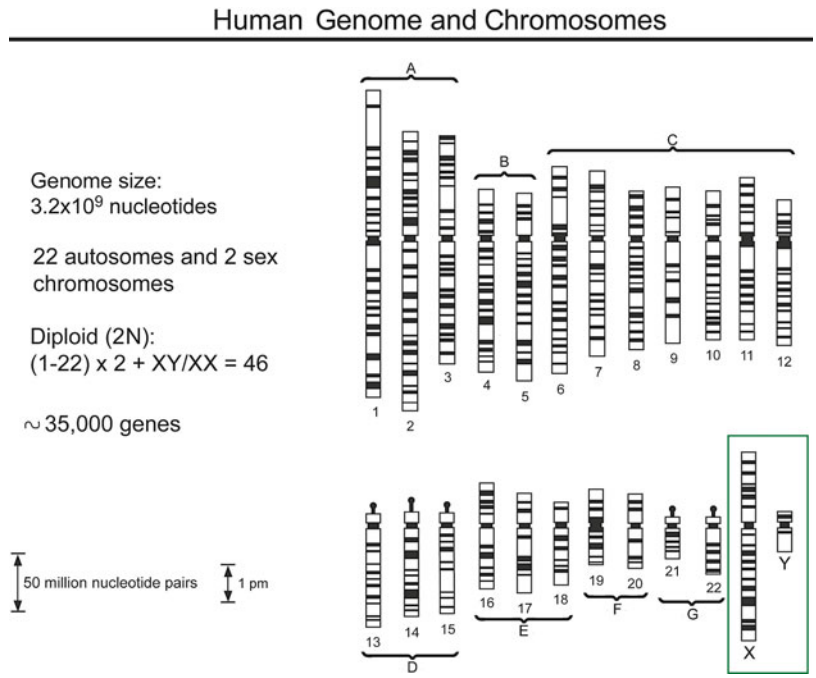


Fig. 1.16 Chromosomes are grouped according to their relative size, the position of their centromere, and banding patterns

arms on each of the 46 chromosome dimers
  - Polyploidy is the state of cells with integer multiples of 23 chromosomes beyond the basic set of paired chromosomes
  - Aneuploidy is a condition in which the number of chromosomes is abnormal owing either to extra or missing chromosomes. The number of chromosomes in an aneuploid cell cannot be a multiple of the haploid set (Fig. 1.17)
  - Monosomy is a type of aneuploidy with at least one missing parental chromosome (Fig. 1.18)
  - Trisomy is a type of aneuploidy with one extra chromosome added to a pair of homologous chromosomes
- Structural alterations of chromosomes (see details in Chap. 2)
- The four major types of chromosomal structural alterations are deletion, duplication, inversion, and translocation (Fig. 1.19)
  - Deletion is the loss of a chromosome segment that could involve one or many nucleotide pairs. Deletion can be terminal or interstitial, and the symbol **del** is used to denote deletions, both terminal and interstitial
  - Duplication is a mutation by formation of extra copies of a chromosomal region, sometimes a whole gene or even an entire chromosome. The symbol **dup** is used to denote DNA segment duplication. The symbol may be preceded by **dir** or **inv** to further define whether the duplication is direct or inverted
  - Inversion occurs when a chromosome segment is flipped end to end. Inversion is designated by the symbol **inv**
  - Translocation is a chromosome abnormality caused by rearrangement of DNA fragments between nonhomologous chromosomes. A fusion gene may be created when the translocation joins two otherwise separated genes. Translocations are denoted by the symbol **t** followed by parentheses showing the exchanged chromosome breakpoints separated by a semicolon
  - The designation t(A;B)(p1;q2) is used to denote a reciprocal translocation

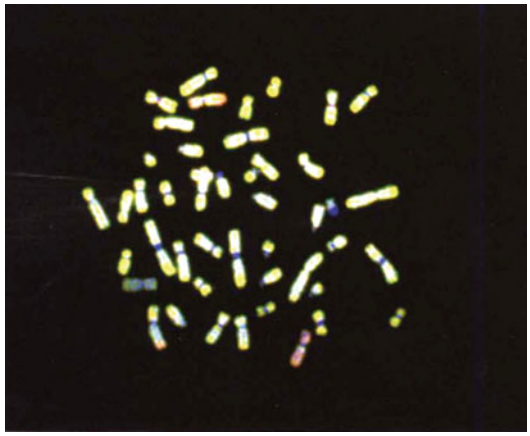


Fig. 1.17 Aneuploidy is a condition in which the number of chromosomes is not a multiple of the haploid number due to gaining or losing chromosomes. The figure shows fluorescent probe chromosome painting of tumor cell chromosomes featuring an assortment of chromosome gains and loses

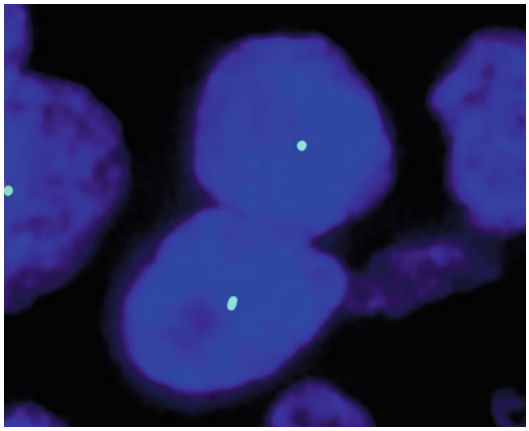


Fig. 1.18 Monosomy is a type of aneuploidy with loss of one chromosome from a pair in the cell's diploid chromosome set. Fluorescence in situ hybridization using probes for the centromeres shows the loss of one chromosome in these cells indicated by having only one signal per nucleus

between chromosome A and B, in which the p-arm of chromosome A is broken off at band 1 and moved to the q-arm of chromosome B at band 2. The B chromosome segment beginning at band q2 is reciprocally moved to the p1 position of chromosome B. The numbers refer to regions by banding dye in a metaphase karyotype

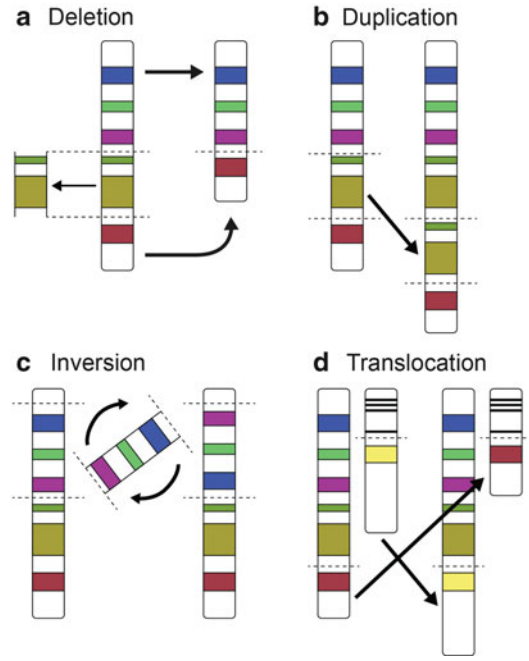


Fig. 1.19 Four major types of chromosomal structural alterations: deletion (**a**), duplication (**b**), inversion (**c**), and translocation (**d**)

- A reciprocal translocation is an exchange of material between nonhomologous chromosomes. The translocation does not involve net gain or loss of DNA material
- A Robertsonian translocation is a fusion between the centromeres of two acrocentric chromosomes (13, 14, 15, 21, and 22) with loss of the short arms, forming a chromosome with two long arms

– Isochromosome is a chromosome which has lost one set of its arms, either p or q, and replaced them with an exact copy of the other arms. Isochromosomes thus have four identical arms, either p or q. Isochromosome is denoted by the symbol **i**
– Ring chromosome formation occurs through breaks in the chromosome arms with fusion of the proximal broken ends, leading to loss of distal material or by telomere dysfunction causing fusion of the chromosome ends without significant loss

of genetic material. A ring chromosome is denoted by the symbol **r**
  - Centromeric fission results from horizontal instead of vertical breakage of the centromere, leading to two derivative symmetric chromosomes from the two short arms in one daughter cell or two long arms in the other daughter cell. The denotation **fis** is the symbol for the alteration
  - Insertion is denoted with the symbol **ins** indicating that chromosome material is inserted into a chromosome location. For example, ins(2)(p13q21q31) indicating that genetic material from chromosome band 2p13 is inserted into the q-arm of chromosome 2 between bands q21 and q31
- Fragile sites are heritable chromosome regions that are poorly connected to the rest of the chromosome and tend to break when the cell is exposed to replication stress
  - Fragile sites are often rich in CGG or CGC repeats, inherited like a gene with a tendency to break away frequently
  - Double chain breaks in fragile sites lead to the loss of genetic material
  - Fragile sites are classified as "common" or "rare" according to the frequency in the population
  - Over 120 fragile sites have been identified in the human genome
  - Common fragile sites are normal chromosome structures present in all individuals. Under normal conditions, most common fragile sites are not especially prone to spontaneous breaks
  - Fragile sites, however, are prone to breakage when cells are cultured under conditions that inhibit DNA replication or repair

1.4 RNA and Protein

1.4.1 Overview

- What is RNA?
  - RNA (ribonucleic acid) is a single-stranded nucleic acid polymer consisting of nucleotide monomers joined by phosphodiester links as in DNA
  - The five-carbon sugar in the RNA chain is ribose (containing a 3′ hydroxyl group) instead of deoxyribose
  - Bases in RNA are A (adenine), G (guanine), C (cytosine), and U (uracil), which takes the place of T (thymine) in DNA
  - RNA folds back on itself to form hairpin or loop structures and double-stranded helical segments via intramolecular hydrogen bonds (Table 1.3)

Table 1.3 The differences between DNA and RNA

	DNA	RNA
Bases	A, G, C, T	A, G, C, U
Strand	Double stranded	Single stranded
Structure	Antiparallel helix	Hairpin and loops
Sugar	Deoxyribose	Ribose
Location	Nuclear or mitochondrial	Nuclear or cytoplasmic
Lifetime	Long	Short
Process	Transcription	Translation
Types	Nuclear DNA, mtDNA	mRNA, tRNA, rRNA, miRNA, siRNA, ribozyme

- What is protein?
  - Proteins are chains of amino acids joined by peptide bonds
  - The peptide bond forms when the carboxyl group of one amino acid residue is joined to the amino group of the next amino acid residue, designated "CONH" in chemical structures
- Synthesis of RNA in cells
  - Transcription is the process of complementary RNA synthesis according to a DNA template sequence
  - The DNA sequence is enzymatically copied by RNA polymerase in a process analogous to DNA chain duplication
  - In RNA transcription, A from the DNA template determines U in RNA instead of T (Fig. 1.20)
  - RNA polymerases are a group of nucleotidyltransferases that polymerize ribonucleotides

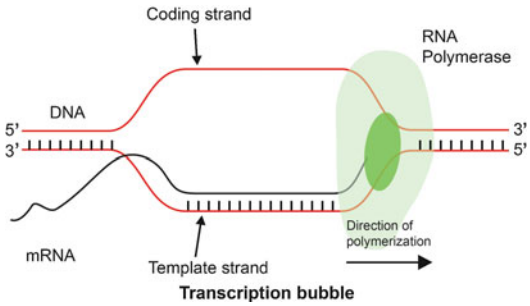


Fig. 1.20 Transcription is the process through which a DNA sequence is copied by RNA polymerase II to produce a complementary RNA chain. The RNA polymerase II proceeds along one strand of DNA moving in the 3′ → 5′ direction and assembles ribonucleotides into a strand of RNA. Synthesis of the mRNA product proceeds in the 5′ → 3′ direction until it reaches the stop codon

of RNA in accordance with the information present in a DNA template
  - RNA polymerase I transcribes genes encoding ribosomal RNA (rRNA)
  - RNA polymerase II transcribes genes encoding proteins (mRNA) and certain small nuclear RNAs (snRNA)
  - RNA polymerase III transcribes genes encoding transfer RNA (tRNA) and other small RNAs (5SRNA in ribosomes)
- RNA primers are de novo synthesized short RNA strands to duplicate the parent DNA strand through primase "reads" of the DNA and synthesis of short RNA primers
  - Since DNA polymerases must only add new nucleotides to an existing 3′ end of a nucleotide strand, primer is required for all DNA synthesis
  - Okazaki fragments form through extension of many of these primers, and leading strand requires fewer primer "starts"

1.4.2 Types of RNA

- mRNA is transcribed from a DNA template to carry sequence information from the nucleus to the cytoplasm, usually for protein synthesis
  - Transcription is the process of copying sequence information from DNA to RNA
  - Translation is the process of converting mRNA sequence information into a protein chain
- mRNA constitutes about 5% of the total cellular RNA
  - mRNA components
    - Exons are sequences in mRNA that encode amino acid sequences
    - Introns are noncoding sequences between exons, which are spliced out of mRNA by exonucleases prior to translation (Figs. 1.13 and 1.24)
    - Codons are sets of three base "words" conveying genetic information for each amino acid to be added to a growing peptide sequence
    - Polyadenylation is the covalent linkage of 50–250 adenosine ribonucleotide units to the 3′ end of mRNA at the polyadenylation signal (AAUAA). This structure is known as the polyadenosine [poly (A)] tail
      - Adding poly (A) is one of the steps needed to produce mature mRNA required for translation
      - Poly (A) stabilizes mRNA and protects the mRNA molecule from exonucleases
      - Poly (A) is necessary for transcription termination and for export of the mRNA from the nucleus to the cytoplasm
    - 5′ cap is a modified guanine nucleotide added to the 5′ end of a eukaryotic mRNA
      - The 5′ cap is critical for mRNA recognition by the ribosome and for protection from ribonucleases (RNases)
      - 5′ UTR (5′ untranslated region) is a noncoding section of mRNA located between the 5′ cap and the start codon at the 5′ end
      - It is not translated
      - It affects the mRNA stability
      - It regulates gene expression in response to iron
      - It facilitates the initiation of translation

    - 3′ UTR (3′ untranslated region) is a consensus section of mRNA located between the stop codon and the poly-A tail at the 3′ end
        - It contains a polyadenylation signal sequence, usually AAUAAA, or a slight variant
        - It is a binding site for proteins that affect mRNA stability and local transcript concentration in the cell
        - It contains binding sites for microRNAs (miRNAs). The miRNAs bind to 3′ UTR-specific sequence targets and interfere with translation
- tRNA is an amino acid-specific adaptor molecule required for protein translation
    - Each tRNA molecule contains three hairpin loops and an amino acid–accepting site at the 3′ end
    - tRNA contains a three-nucleotide anticodon in the second loop, which constitutes the sequence complementary to the three-base codon on the mRNA transcript
    - Each tRNA is specific for only one amino acid and anticodon sequence
    - tRNA constitutes about 10% of total cellular RNA
    - tRNA transfers a specific amino acid to a growing polypeptide chain at the aminoacyl-tRNA synthetase ribosomal site of protein synthesis during translation at a rate of 6–9 amino acid residues per second (Fig. 1.21)
    - tRNA molecules are encoded by RNA genes in both nuclear and mitochondrial DNA

1.4.3 Ribosome and Ribozyme

- rRNA (ribosome RNA) is the RNA component of a ribosome
    - rRNAs in eukaryotes include 5S, 5.8S, and 18S components included in the large subunit, and 18S rRNA in the small subunit. Each rRNA subcomponent is essential for ribosome structure and function
    - rRNA is the most abundant and stable RNA species in the cell and constitutes about 50% of the total cellular RNA
    - Ribosomes are the protein manufacturing machinery of all living cells
        - Ribosomes are composed of rRNA and ribosomal proteins
        - Ribosomes translate mRNA into polypeptide chains
- Ribozyme (catalytic RNA) is an RNA molecule that catalyzes a chemical reaction analogous to a protein enzyme and has an active site consisting entirely of RNA
    - Natural ribozymes cleave their own phosphodiester bonds or cleave bonds in other RNAs
    - Some ribozymes also appear to catalyze the aminoacyl-tRNA synthetase reaction of the ribosome
    - A ribozyme consists of a conserved catalytic core motif which is required for *trans*-cleavage of a phosphodiester bond within an RNA target
    - Ribozyme-based cancer therapy uses specially designed ribozymes to knock down oncogene mRNA (Fig. 1.22)
    - Ribozyme also cuts and liquates other RNA molecules, as when introns are spliced out of mRNA
- Noncoding RNA
    - Many RNA genes encode RNA that is not translated into protein
        - The human nuclear genome contains about 3,000 unique RNA genes (<10% of total gene number)
    - miRNA (see Chap. 3, MicroRNA section)
- Double-stranded RNA (dsRNA)
    - dsRNA is RNA with two complementary strands
    - dsRNA forms the genetic material of some viruses
    - In eukaryotes, it acts as a trigger to initiate the process of RNA interference
    - dsRNA is an intermediate step in the formation of siRNAs (small interfering RNAs)
- Catalytic RNA (see Sect. 1.4.3)

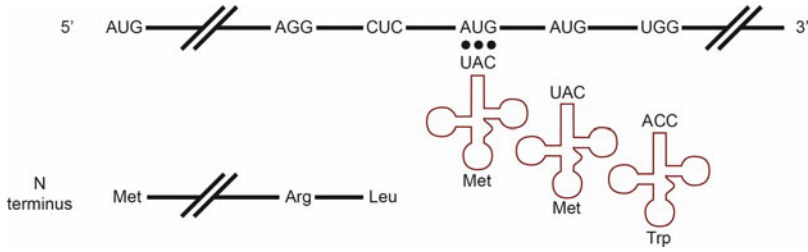


Fig. 1.21 tRNA transfers a specific amino acid to a growing polypeptide chain at the ribosomal site of protein synthesis during translation. It has a site for amino acid attachment and a 3-base region called the anticodon that recognizes the corresponding 3-base codon region on mRNA via complementary base pairing

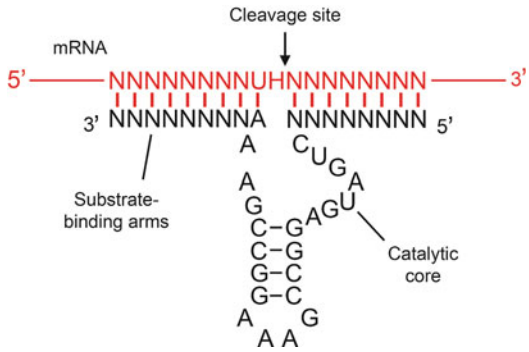


Fig. 1.22 The figure illustrates an mRNA (in *red*) and a ribozyme (in *black*) with secondary structures of substrate-binding arms, catalytic core, and cleavage site. A ribozyme is an RNA molecule catalyzing a chemical reaction. Ribozymes function by binding to the target RNA and cleaving the phosphodiester backbone at a specific cutting site. Five classes of ribozymes have been described based on unique characteristics in their sequences as well as by their 3-dimensional structures. The binding arms of a ribozyme include sequences complementary to the target RNA. Ribozymes recognize a target via interactions between the binding arms of the ribozyme and the mRNA. Ribozymes can cleave any RNA substrate that matches the binding sequence and contains an NUH triple; *N* is any nucleotide and H could be A, C, or U

1.4.4 mRNA Processing

- Capping is the process of adding a guanosine nucleotide to the 5′ end of mRNA via a 5′–5′ triphosphate linkage followed by guanosine methylation (Fig. 1.23)
  - It is critical in regulation of mRNA export from the nucleus
  - It prevents the mRNA from degradation by exonucleases
  - It promotes protein translation by prolonging tRNA half-life
  - It promotes excision of the proximal 5′ intron
- Polyadenylation (see "mRNA" in Sect. 4.2)
- mRNA splicing is the process of removing introns from the primary transcript and joining the exons via an intermediate usually described as a "lariat" structure (Fig. 1.24) formed by a ribozyme complex called a spliceosome
- The intron generally starts with GU and ends with AG
  - A "lariat" structure is formed by 5′ G of the intron joining in a 2′, 5′-phosphodiester bond to an adenosine near the 3′ end of the intron
  - Two transesterifications are needed for splicing to occur between RNA nucleotides, forming an excised intron loop (the "lariat" structure) and an exon to exon joined mRNA segment
    - First, the 2′ OH of a specific branch-point nucleotide within the intron is defined during spliceosome assembly and performs a nucleophilic attack on the 5′ phosphodiester bonds of the first nucleotide in the upstream intron at the 5′ splice site, forming the lariat intermediate
    - Second, the 3′ OH of the released 5′ exon performs a nucleophilic attack on

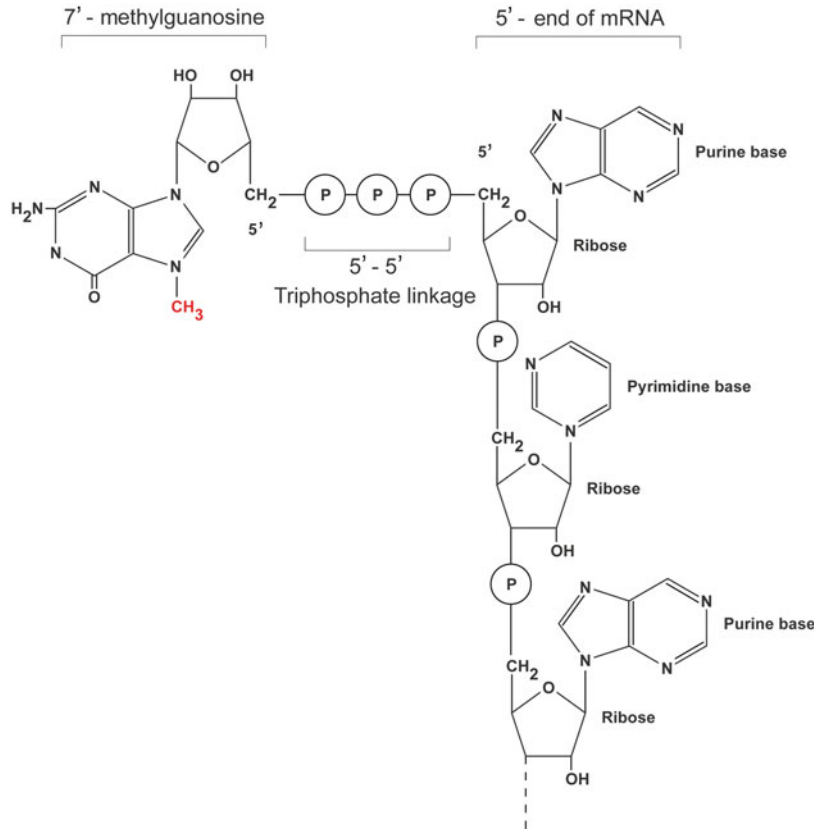


Fig. 1.23 Capping adds a guanosine nucleotide to the 5′ end of mRNA via a 5′–5′ triphosphate linkage followed by guanosine methylation. The process of 5′ capping is critical to create mature mRNA, which is then able to undergo translation. Capping stabilizes the mRNA during protein synthesis so that truncated proteins are not produced by ribonuclease digestion of mRNA. Capping is a highly regulated process occurring in the nucleus

the 5′ phosphodiester bond of the last intron nucleotide at the 3′ splice site, thus joining the exons and releasing the intron lariat
    - The function of the lariat structure is to form a stable hairpin, excise the intron, and join the exons
- Mutations that affect mRNA splicing include the 5′ and 3′ splicing sites
    - The GU sequence marks the 5′ splice site and the AG sequence marks the 3′ splice site
    - Mutation of a splice site results in a premature stop codon, loss of an exon, or inclusion of an intron
    - The intron generally starts with GU and ends with AG. If the initial G changes to A, the site is cut at the upstream intron, but the reaction stops and ligation fails. A truncated mRNA molecule is the result of this error
    - Mutation of a splice site results in variation in the splice location, causing insertion or deletion of amino acids, or shift of reading frame
    - Splicing errors are under surveillance of a mechanism termed "nonsense-mediated mRNA decay"
    - If the 5′–2′ second transesterification is incorrect, then 3′ cutting does occur and the intron is not excised
    - Other mutations block the spliceosome reaction altogether
- Alternative splicing means that a primary mRNA transcript from one gene may be spliced at different locations within the sequence to produce different mature mRNA molecules and therefore to produce different proteins
    - Alternative splicing of pre-mRNAs is a powerful and versatile regulatory mechanism for quantitative control of gene

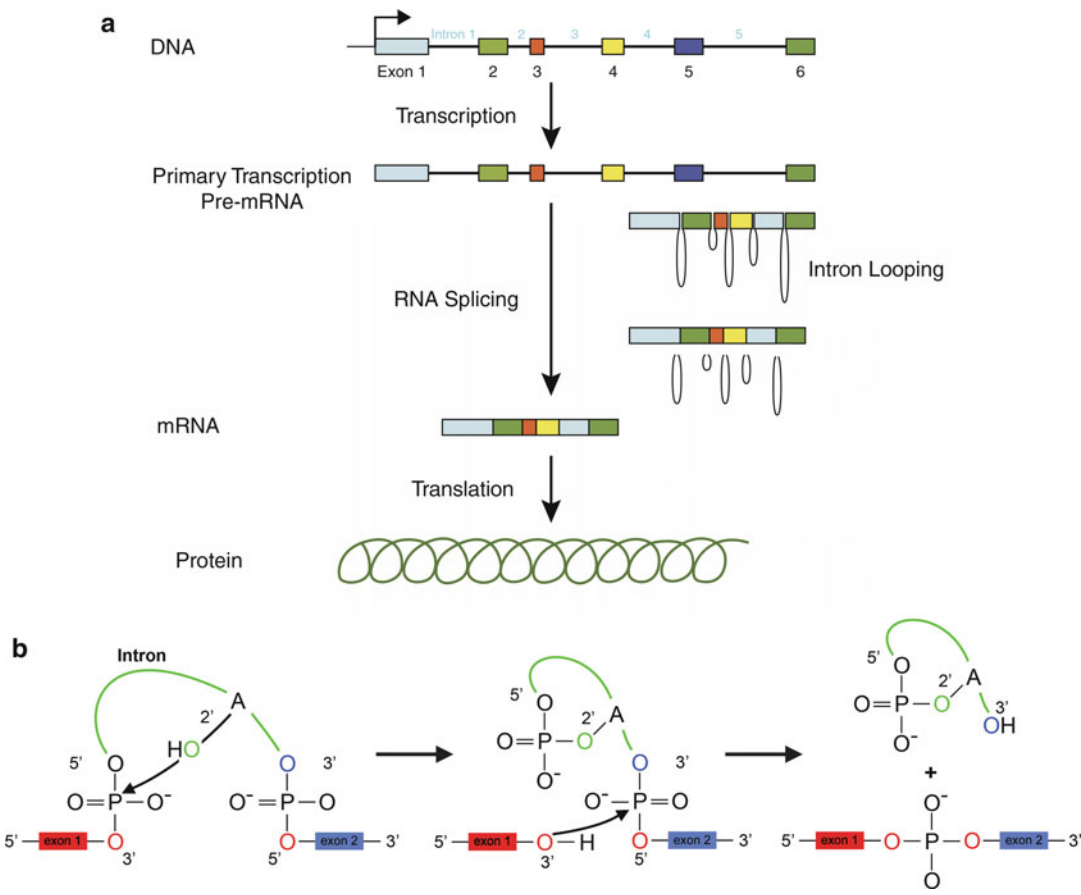


Fig. 1.24 RNA splicing is a modification of the molecule after transcription, in which introns of precursor messenger RNA (pre-mRNA) are removed and exons are joined. Splicing only occurs in eukaryotes. In *panel A* the exons are shown in colored boxes and introns are shown in lines. After transcription the introns form loops following the GU-AG rule; the loops are removed and the exons are joined to form mature mRNA. *Panel B* shows the two-step biochemical process of RNA splicing. Both steps involve transesterification reactions occurring between approximated RNA nucleotides

expression and functional diversification of protein production
  - It contributes to major developmental decisions as well as to fine tuning of gene function
- RNA editing is the molecular process by which the information content in mRNA is altered through a chemical change in the base makeup
  - U editing is achieved by deamination of cytosine to uracil. Through this editing process, the codon CAA (Gly) is changed to UAA (stop codon)
  - A-I editing modifies adenosine (A) to inosine (I), which serves in a codon as if it were guanine (G) during protein translation, although it indiscriminately pairs with adenine (A), thymine (T), or cytosine (C) during in vitro polymerase chain reaction (PCR)

1.4.5 Protein Translation

- The synthesis of proteins is known as translation. Protein translation occurs in the cytoplasm where ribosomes reside
- The ribosome travels down the mRNA one codon at a time, and the amino acids are added one by one (Fig. 1.25)

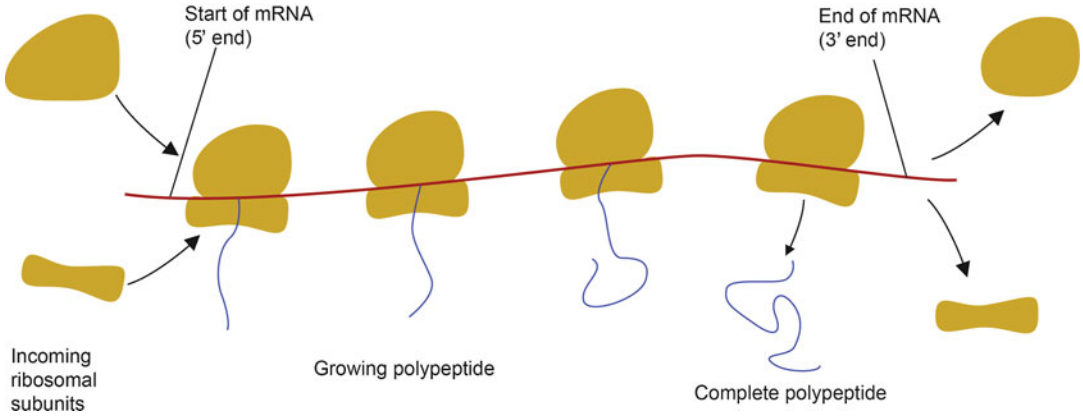


Fig. 1.25 Translation occurs in the cytoplasm where ribosomes are located either free or attached to rough endoplasmic reticulum membranes. Translation is the process that converts an mRNA sequence into a chain of amino acids to form a protein. Translation proceeds in four phases: activation, initiation, elongation, and termination. In activation the amino acid to be added is bonded by its carboxyl group to the 3′ OH of the proper tRNA by an ester bond. Initiation involves the ribosome small subunit binding to 5′ end of mRNA with the help of initiation factors and other proteins. Elongation occurs when the next aminoacyl-tRNA in line binds to its codon in the mRNA lying in the ribosome groove along with GTP and an elongation factor. Synthesis is terminated when the polypeptide meets a stop codon

- During protein translation, mRNA is read in the 5′ → 3′ direction until the ribosome reaches a stop codon
  - Protein translation begins at a start codon (*AUG*)
  - Protein translation ends at a stop codon (*UGA, UAG*, and *UAA*)
  - Protein is made from N-terminal (amino terminus) to C-terminal (carboxyl terminus)
- Genetic code is the set of rules by which information encoded in genetic material (DNA or RNA sequence) is synthesized into protein (amino acid sequence) in living cells
- Protein translation relies on aminoacyl-tRNA to carry a specific amino acid and recognize the corresponding codon in mRNA by anticodon base pairing
  - Ribosomes move along the mRNA molecule in the 5′ → 3′ direction, adding amino acid residues one aminoacyl-tRNA at a time
- The peptide bond is formed between two amino acid residues. The group from one amino acid reacts with the amino group of another amino acid hydroxyl group of a terminal carboxylic acid, releasing a molecule of water in a condensation reaction. This reaction is facilitated by a ribozyme-containing enzyme (Fig. 1.26)
- Posttranslational modification is the chemical modification of a protein after its translation. It is one of the later steps in protein biosynthesis for many proteins
  - One common posttranslation modification is the enzyme-catalyzed covalent addition of a chemical group, usually an electrophile, to amino acid side chains in a protein
    - Phosphorylation adds phosphate groups to a protein, usually to serine, threonine, tyrosine, and histidine residues
    - Protein phosphorylation is the most important regulatory event for enzymes, receptors, and signal transduction molecules
    - Many enzymes and receptors are switched "on" or "off" by phosphorylation and dephosphorylation
    - Phosphorylation is catalyzed by specific protein kinases, and dephosphorylation is catalyzed by protein phosphatases
    - Glycosylation modifies the protein by covalent attachment of oligosaccharides.

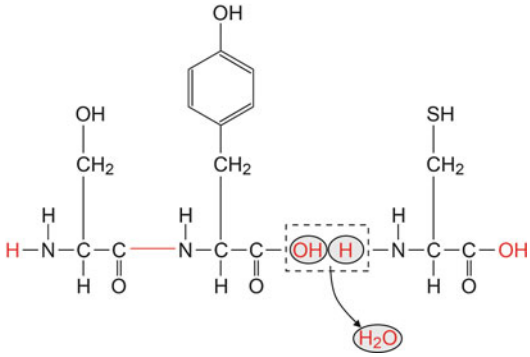


Fig. 1.26 A peptide bond is formed between two amino acids by the carboxyl group of one amino acid reacting with the amino group of the other amino acid, releasing a molecule of water (H_2O). Polypeptides and proteins are chains of amino acid residues held together by peptide bonds

- Glycoproteins are crucial for protein sorting, immune recognition, receptor binding, inflammation, and structural integrity

– A second common posttranslation modification is cleavage of a peptide fragment by protease or, less commonly, by autocatalytic cleavage
  - Protease removes amino acids from the amino end of the protein
  - Methionyl amino peptidase removes the initiation methionine always added by the start codon AUG, which is also the only codon for methionine
  - Many proteins are synthesized as inactive precursors that must be activated by removal of polypeptide blocking fragments to become active
  - Insulin is an example of a blocked protein since it is secreted from pancreatic islet beta cells as a hormonally inactive signal peptide. Proinsulin is a circular peptide with two disulfide bridges joining the overlapping amino- and carboxy-terminal ends. Active insulin forms with the cleavage of the 24-amino acid signal peptide needed for export through the endoplasmic reticulum. The active insulin is composed of two peptides with linkage through disulfide bonds
  - Some enzymes are synthesized as inactive precursors called zymogens. Subsequently, they are activated by proteolytic cleavage of the inactivating segment, having used this segment to prevent autodigestion or to promote proper folding of the enzyme

– Posttranscriptional modification may also involve addition of functional groups
  - The chemical nature of a protein is sometimes changed by adding groups, such as acetate, lipid, or carbohydrate
  - Protein structure is also changed by forming disulfide bridges between two cysteine residues, generating a cystine dimeric amino acid linking two peptide chains

1.5 Mitochondrial DNA

1.5.1 Overview

- mtDNA is a double-stranded circular DNA located in the mitochondrial matrix
- Mitochondrial DNA is proposed to have a separate evolutionary origin from somatic DNA
- Each mitochondrion is estimated to contain 2–10 mtDNA copies or 100–10,000 copies per cell
- The majority of the proteins present in mitochondria are encoded by nuclear DNA

1.5.2 mtDNA Inheritance

- All mtDNA is inherited from the maternal zygote
- Sperm carries the father's mtDNA in its tail, which is lost during fertilization
- mtDNA inheritance is non-Mendelian because Mendelian inheritance presumes that half the genetic material of an embryo derives from each parent
- mtDNA passes unchanged from mother to offspring by this mechanism and is conserved for hundreds, or even thousands, of generations

Table 1.4 The differences between nuclear DNA and mtDNA

	Nuclear DNA	mtDNA
Location	Nucleus	Mitochondrial matrix
Size	3,200 Mb	16.6 kb
Structure	Antiparallel double helix, linear	Double strand, circular
Introns	Present	Absent
Noncoding sequence	Many noncoding sequences (98%)	Few noncoding sequences (7%)
Transcription	Monogenic	Multigenic
Copy number	One set per cell	Hundreds to thousands per cell
Number per cell	46	Several copies per mitochondrion
Associate proteins	Histone and nonhistone	Largely protein free
Encoding genes	30,000	37
Transcription	Individual gene transcription	Bulk transcription for whole strand
Codon	Universal codon	Mitochondrial codon (see Table 1.5)
Inheritance	Mendelian, from both parents	Maternal

1.5.3 Characteristics of mtDNA

- The two strands of mtDNA have significantly different compositions from nuclear DNA and from one another (Table 1.4)
  - The heavy (H) strand is rich in purines (adenine and guanine)
  - The light (L) strand is rich in pyrimidines (thymine and cytosine)
- mtDNA is highly conserved, so it is useful for phylogenetic study
- mtDNA is devoid of introns
- mtDNA contains only a few noncoding intergenic regions
- Over 92% of the mtDNA genome has encoding function, and therefore most mtDNA mutations lead to functional anomalies
- The entire molecule is regulated by only one regulatory region which contains the origins of replication of both heavy and light strands
- The genetic code of mitochondrial genes is not completely similar as nuclear genes, since mtDNA codes for 22 unique tRNA molecules that are different from the cytoplasmic tRNA molecules encoded by nuclear RNA genes
- Mutations in nuclear DNA may also have a wide array of effects on mtDNA replication owing to dependence of mitochondria on DNA repair enzymes encoded by nuclear genes
- The size and number of mtDNA
  - Human mtDNA is 16,569 bp in length
  - Chromosomes, by contrast, range in size from 47 to 247 million base pairs
  - mtDNA is subjected to oxidative stress because of reactive oxygen species generated by the respiratory chain in mitochondrial cristae

1.5.4 Mitochondrial Genes and Gene Expression

- mtDNA genes
  - mtDNA encodes 37 genes
    - There are 13 peptide coding genes, including the following peptides
      - Transcription factor A
      - The mtRNA processing ribonuclease P
      - The transcription termination factor
      - The mitochondrial electron transport proteins are synthesized on mitochondrial ribosomes
    - The heavy strand (H strand) encodes 12 polypeptides, 2 rRNAs, and 14 tRNAs
    - The light strand (L strand) encodes one polypeptide and eight tRNAs
    - Because of its abundance, ease of isolating mitochondria from cell lysate, and

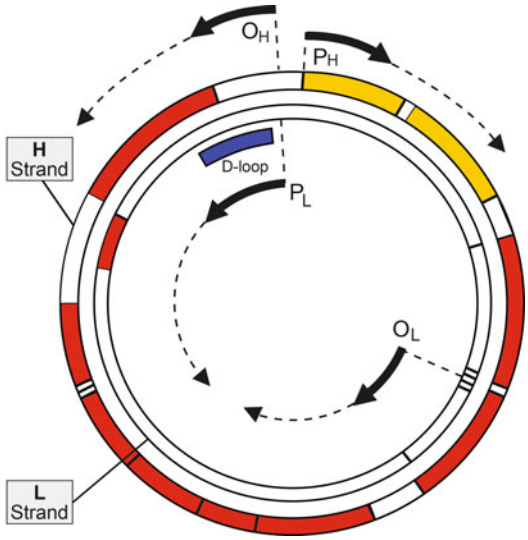


Fig. 1.27 mtDNA is DNA located in the mitochondrion. In human beings, 100% of the mtDNA contribution to an embryo is inherited from the mother. Each mtDNA consists of a heavy (*H*) strand, which is rich in adenines and guanines, and a light (*L*) strand, which is rich in thymine and cytosine. Human mtDNA is present at 100–10,000 copies per cell, with each circular molecule consisting of 16,569 bp coding for 37 genes: 13 peptides (*red*), 22 tRNAs (*black*), and two rRNAs (*yellow*). H strand encodes two ribosomal RNAs, 12 peptides, and 14 tRNAs, whereas L stand encodes one peptide and eight tRNAs. There are no introns in mtDNA. Transcription of H strand originates from two closely located promoters shown as P_H and P_L in the D-loop region, a triple-stranded structure. The origin and direction of transcription of H and L strands is shown by arrows at P_H and P_L. The mtRNA transcripts come from both heavy and light chains. The mitochondrial genes lack introns and are transcribed in full length then processed to the natural products. Origins O_H and O_L are the beginning points for replication of mtDNA

small size, mtDNA is usually the first DNA to be fully sequenced from an organism

- mtDNA expression
  - Transcription initiation sites of mtDNA
    - The promoters of H and L strands (termed P_H and P_L) are both located in the D-loop region and 150 bp apart (see Sect. 5.5 at the end of this section) (Fig. 1.27)
      - H strand transcription starts at nucleotide 561
      - L strand transcription starts at nucleotide 407
    - Transcription of mtDNA starts from the promoters in the D-loop region and continues in opposing directions for the two strands around the circle to generate large multigenic transcripts
    - Transcription initiation in mitochondria involves three types of proteins
      - The mtRNA polymerase (POLRMT)
      - Mitochondrial transcription factor A (TFAM)
      - Mitochondrial transcription factors B1 and B2 (TFB1M, TFB2M)
      - POLRMT, TFAM, and TFB1M or TFB2M assemble at the mitochondrial promoters and begin transcription
  - Promoters of mitochondrial gene expression
    - Heavy-strand 1 (H1) promoters initiate transcription of the entire heavy strand
    - Heavy-strand 2 (H2) promoters initiate transcription of the two mitochondrial rRNAs
    - Light-strand (L) promoter initiates transcripts of the entire light strand
- Mitochondrial mRNA
  - Mitochondrial mRNAs are small molecules
  - Full-length transcripts are cut into functional tRNA, rRNA, and mRNA molecules
  - Mitochondrial mRNA lacks a 5′ cap structure
  - Mitochondrial mRNAs lack both a 5′ and a 3′ UTR
  - The first codon specifies N-formylmethionine and is located at or very near the 5′ end
- Regulation of mtDNA expression
  - mtDNA expression depends on a large number of proteins encoded by nuclear DNA
  - The regulatory proteins are synthesized in the cytosol and enter the mitochondria via specialized pores
  - mtDNA replication is regulated by only one regulatory region, controlling both the heavy and the light strands

- Posttranscriptional modification of mitochondrial mRNA
  - Mitochondrial mRNAs are processed by mitochondrial ribonuclease (mtRNase) cleavage of the transcript
  - The light strand may produce either short transcripts, which serve as primers for mtDNA replication, or a long transcript for peptide and tRNA production
  - The production of primer occurs by processing of light-strand transcripts with the mtRNase P
  - The H- and L-strand mRNA molecules are polyadenylated by a mitochondrial poly(A) polymerase, imported from the cytosol
- Mitochondrial mRNA translation
  - The protein components necessary for mitochondrial translation, including ribosomal proteins, tRNA synthetases, ribonucleases, initiation, and elongation factors, are all encoded by nuclear genes
  - Mitochondrial protein synthesis and DNA replication are thus under nuclear regulatory control
  - Mitochondrial translation is bacteria-like both in its sensitivity to antibiotics that act on the ribosome and in its use of N-formylmethionyl-tRNA for initiation
  - Mitochondrial ribosomes are smaller than those found in the cytosol and have a sedimentation coefficient of 55S instead of the denser 80S sedimentation coefficient for cytosolic ribosomes or 70S coefficient for bacterial ribosomes

1.5.5 mtDNA Replication

- mtDNA replication is an asynchronous process, which begins at the origin of the H strand
  - mtDNA replication is controlled by genes in the nucleus based on how many mitochondria the particular cell needs at that time
  - When the replication apparatus meets the origin of the L strand, it is forced into a single-strand configuration by the extending daughter H strand, and L-strand replication begins at this point
  - RNA derived from the L-strand promoter serves as a primer for H-strand DNA replication
- The D-loop (displacement loop) is an 1,123-base stretch of DNA, often triple-stranded, which contains sites for DNA-binding proteins that control mtDNA replication and transcription
  - The D-loop contains the promoters for both the H-strand and L-strand transcripts
  - mtDNA replication causes the D-loop to move along the heavy strand as mtDNA polymerase-γ produces a complimentary replica strand
    - Heavy-strand DNA replication begins at the D-loop and proceeds in a $5' \rightarrow 3'$ direction until returning to the origin of replication
    - DNA polymerase γ begins in the reverse direction to produce a complimentary replica of the light strand when replication of the heavy strand reaches the light-strand replication origin (O_L)
    - Two identical double-strand mtDNA molecules are the result of this process
    - When mitochondria have enough copies of mtDNA, sufficient mitochondrial proteins, and adequate surface area, a nuclear protein may permit the mitochondrion to divide by fission into two daughter mitochondria

1.5.6 mtDNA Damage, Mutations, and Repair

- mtDNA damage
  - mtDNA is susceptible to insult by all the same processes that damage nuclear DNA
  - mtDNA is especially susceptible to insult by reactive oxygen species, which are prevalent in mitochondria
    - Because mtDNA is not bound to histones, it is exposed to damage caused by free oxygen radicals produced by

Table 1.5 Genetic code of mtDNA

UUU Phe	UCU Ser	UAU Tyr	UGU Cys
UUC Phe	UCC Ser	UAC Tyr	UGC Cys
UUA Leu	UCA Ser	UAA STOP	**UGA Trp (STOP)**
UUG Leu	UCG Ser	UAG STOP	UGG Trp
CUU Leu	CCU Pro	CAU His	CGU Arg
CUC Leu	CCC Pro	CAC His	CGC Arg
CUA Leu	CCA Pro	GAA Gln	CGA Arg
CUG Leu	CCG Pro	CAG Gln	CGG Arg
AUU Ile	ACU Thr	AAU Asn	AGU Ser
AUC Ile	ACC Thr	AAC Asn	AGC Ser
AUA Met (Ile)	ACA Thr	AAA Lys	**AGA STOP (Arg)**
AUG Met	ACG Thr	AAG Lys	**AGG STOP (Arg)**
GUU Val	GCU Ala	GAU Asp	GGU Gly
GUC Val	GCC Ala	GAC Asp	GGG Gly
GUA Val	GCA Ala	GAA Glu	GGA Gly
GUG Val	GCG Ala	GAG Glu	GGG Giy

() indicated the universal code. Bold marked the different mitochondrial codons

electron transfer during oxidative phosphorylation of the respiratory chain
    - mtDNA also undergoes the same types of mutation as nuclear DNA including spontaneous modifications and replication errors
- mtDNA mutations
  - The rate of mutation in mtDNA is calculated to be about 10 times greater than that of nuclear DNA
  - The mtDNA mutations may be either acquired or inherited
  - Several different mutations of mtDNA may present clinically as the same disease
  - Large deletions and duplications of mtDNA increase with age
    - This may account for some aging processes in oxygen-dependent organs, such as brain, kidney, muscle, and heart
    - Mutant electron transfer proteins may release more oxygen-free radicals into the mitochondrial matrix, accelerating the aging process in some cases of Alzheimer and coronary artery disease
  - There are hypervariable segments (HV1 and HV2) located at base 57–372 and base 16,024–16,383, respectively. The rate of mutation in these regions is significantly higher than in the rest of mtDNA
- mtDNA repair
  - mtDNA does not code for any DNA repair proteins
  - Proteins from the cytosol under nuclear control enter the mitochondrion through specialized membrane pores
  - Recent evidence has suggested that mitochondria have enzymes to proofread mtDNA and fix mutations owing to free radicals
  - Evidence for nucleotide excision repair, direct damage reversal, mismatch repair, and recombinational repair mechanisms have also been found in mitochondria (see "DNA Repair" section)
  - As with nuclear DNA repair, the ability of mitochondria to repair DNA damage declines with age

1.5.7 Mitochondrial Disease

- Mitochondrial diseases result from failures of processes in the specialized compartments for oxidative phosphorylation, which are present in every cell of the body except red blood cells

  - About 1 in 4,000 children in the United States will develop a mitochondrial disease by the age of 10 years
  - 1,000–4,000 children per year in the United Sates are born with some type of congenital mitochondrial disease
  - Mitochondrial diseases may either be observable at birth or asymptomatic until late adulthood
  - Heteroplasmy refers to a phenomenon in which the number of mutant versus wild-type mitochondria varies from cell to cell and from tissue to tissue
  - When a tissue reaches a certain ratio of mutant to wild-type mitochondria, symptoms of a disease become manifest
  - Mitochondrial disease may be caused either by mtDNA mutations (inherited or acquired) or by mutations in nuclear DNA coding for mitochondrial components
  - Types of mutations
    - Homoplasmic: similar distribution of mtDNA mutation in all tissues
    - Heteroplasmic: variable distribution of mtDNA mutation in different cells or tissues
- Typical symptoms of mitochondrial disease include
  - Loss of muscle coordination, muscle weakness
  - Neurologic problems, seizures
  - Visual and/or hearing problems
  - Developmental delays, learning disabilities
  - Heart, liver, or kidney disease
  - Gastrointestinal disorders and severe constipation
  - Diabetes
  - Increased risk of infection
  - Thyroid and/or adrenal dysfunction
  - Autonomic dysfunction
  - Neuropsychologic changes characterized by confusion, disorientation, and memory loss
- The diagnosis of mitochondrial disease is problematic
  - There is no reliable and consistent means of diagnosis
  - Evaluating the patient's family history is essential
  - Diagnosis may require consultation with one of the few physicians who specialize in mitochondrial disease
  - Diagnosis can be made by blood DNA testing and/or muscle biopsy, but neither of these tests is completely reliable
- Mitochondrial code is similar to the universal code with four exceptions highlighted in Table 1.5

Further Reading

Anderson S, Bankier AT, Barrell BG. Sequence and organization of the human mitochondrial genome. Nature. 1981;290:457–65.

Baylin SB, Jones PA. A decade of exploring the cancer epigenome – biological and translational implications. Nat Rev Cancer. 2011;11:726–34.

Blackburn EH. Structure and function of telomeres. Nature. 1991;350:569–73.

Brenner S, Jacob F, Meselson M. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. Nature. 1961;190:576–81.

Buratowski S. Mechanisms of gene activation. Science. 1995;270:1773–4.

Chakravarti A, Kapoor A. Genetics. Mendelian puzzles. Science. 2012;335:930–1.

Darnell Jr JE. Variety in the level of gene control in eukaryotic cells. Nature. 1982;297:365–71.

Eschenmoser A. Chemical etiology of nucleic acid structure. Science. 1999;284:2118–24.

Felsenfeld G, Groudine M. Controlling the double helix. Nature. 2003;421:448–52.

Friedberg EC. DNA damage and repair. Nature. 2003;421:436–40.

Grunstein M. Histone acetylation in chromatin structure and transcription. Nature. 1997;389:349–52.

Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74.

Horn PJ, Peterson CL. Molecular biology. Chromatin higher order folding – wrapping up transcription. Science. 2002;297:1824–7.

Ionov Y, Peinado MA, Malkhosyan S, et al. Ubiquitous somatic mutations in simple repeated sequences reveals a new mechanism for colonic carcinogenesis. Nature. 1993;363:558–61.

Kasinski AL, Slack FJ. Epigenetics and genetics. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy. Nat Rev Cancer. 2011;11:849–64.

Krishnan KJ, Reeve AK. What causes mitochondrial DNA deletions in human cells? Nat Genet. 2008;40:275–9.

Lewis JD, Tollervey D. Like attracts like: getting RNA processing together in the nucleus. Science. 2000;288:1385–8.

Lindahl T, Wood RD. Quality control by DNA repair. Science. 1999;286:1897–905.

Moore MJ. From birth to death: the complex lives of eukaryotic mRNAs. Science. 2005;309:1514–8.

Noller HF. RNA structure: reading the ribosome. Science. 2005;309:1310–514.

Papaemmanuil E, Cazzola M, Boultwood J, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med.. 2011;365:1384–95.

Qiu J. Epigenetics: unfinished symphony. Nature. 2006;441:143–5.

Rajagopalan H, Lengauer C. Aneuploidy and cancer. Nature. 2004;432:338–41.

Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature. 2006;444:444–54.

Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science. 2001;293: 1089–92.

Rouse J, Jackson SP. Interfaces between the detection, signaling, and repair of DNA damage. Science. 2002;297:547–51.

Sharp PA. Splicing of messenger RNA precursors. Science. 1987;235:766–71.

Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458:719–24.

Tubio JM, Estivill X. Cancer: When catastrophe strikes a cell. Nature. 2011;470:476–7.

Watson JD, Crick FH. Molecular structure of nucleic acid; a structure for deoxyribose nucleic acid. Nature. 1953;171:737–8.

Wolffe AP, Matzke MA. Epigenetics: regulation through repression. Science. 1999;286:481–6.

Zou L. DNA repair: a protein giant in its entirety. Nature. 2010;467:667–8.

Clinical Cytogenetics: Principles

2

Stuart Schwartz

Contents

2.1 Historical Overview 43
2.1.1 Historical Aspects 43
2.1.2 Chromosome Structure 44
2.1.3 Methodology 44
2.1.4 Microarray Technology 50
2.1.5 Indications for Cytogenetic Studies 51

2.2 Autosomal Abnormalities 51
2.2.1 Types of Abnormalities 51
2.2.2 Polyploidy 51
2.2.3 Aneuploidy 52
2.2.4 Aneuploidy: Causes 54

2.3 Structural Rearrangements: Intrachromosomal 56
2.3.1 Intrachromosomal Rearrangements 57

2.4 Structural Rearrangements: Interchromosomal 62
2.4.1 Interchromosomal Rearrangements 62

2.5 Contiguous Gene Syndromes 66
2.5.1 Definition 66
2.5.2 Microdeletions 66
2.5.3 Rare Deletions 68
2.5.4 Duplication Syndromes 68
2.5.5 Microarray Syndromes 69
2.5.6 Microarray Detected: Deletion and Duplications 69

2.6 Sex Chromosome Aberrations 70
2.6.1 Sex Determination 70
2.6.2 Sex Chromosome Abnormalities 71
2.6.3 Lyon Hypothesis 75

Further Reading 76

S. Schwartz, PhD, FACMG
Cytogenetics Laboratory, Center for Molecular Biology and Pathology, Laboratory Corporation of America, Research Triangle Park, NC, USA

2.1 Historical Overview

2.1.1 Historical Aspects

- Ages of cytogenetics
  - I – Dark ages <1952
  - II – Hypotonic 1952–1958
  - III – Trisomy 1959–1969
  - IV – Banding era 1970–1976
  - V – High resolution 1976–1988
  - VI – Molecular cytogenetics 1988 till present
- 1921 – Painter (48 chromosomes)
  - Chromosome number stayed at 48 chromosomes from 1921 to 1956.
- 1959–1960 – Trisomies found for a number of chromosomal syndromes
  - 1959 – Lejeune et al. – down syndrome
  - 1959 – Ford et al. – turner syndrome
  - 1959 – Jacobs and Strong – 47,XXY
  - 1959 – Jacobs et al. – 47,XXX
  - 1960 – Patau et al. – trisomy 13
  - 1960 – Edwards et al. – trisomy 18
  - 1960 – Nowell and Hungerford – Ph[1]
- Lymphocyte cultures developed in 1960
  - Short-term culture technique developed lymphocyte cultures
  - Effective due to phytohemagglutinin
    - Stimulates cell in G_0 to undergo cell division
  - Previously cytogenetic analysis done on either fibroblasts or bone marrow

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_3,

- Chromosome banding
  - Prior to 1970 – Everything solid stained
  - 1970 – Caspersson developed quinacrine banding
  - 1976 – Yunis initiated high-resolution studies
  - 1988–Pinkel/Ward–independently introduced fluorescence in situ hybridization (FISH)
- Multiple tissues can be used for chromosomal analysis
  - Lymphocyte cultures
  - Bone marrow
  - Solid tumors
  - Fibroblasts
  - Amniotic fluid
  - Chorionic villi

2.1.2 Chromosome Structure

- A chromosome consists of a primary constriction (centromere) connected to both a short arm (p arm) and a long arm (q arm) (Fig. 2.1)
- Based on the placement of the centromere, the chromosome could be one of the following:
  - Metacentric
  - Submetacentric
  - Acrocentric
    - An acrocentric chromosome has both satellites and stalks on its short arm.
- Chromosomes can be classified into seven groups
  - Group A (1–3) – metacentric
  - Group B (4,5) – submetacentric
  - Group C (6,7,8,11,X) – metacentric (9,10,12) – submetacentric
  - Group D (13–15) – acrocentric
  - Group E (16) – metacentric (17,18) – submetacentric
  - Group F (19,20) – metacentric
  - Group G (21,22,Y) – acrocentric
- A karyotype is the particular chromosome complement of an individual as defined by the number and morphology of their chromosomes

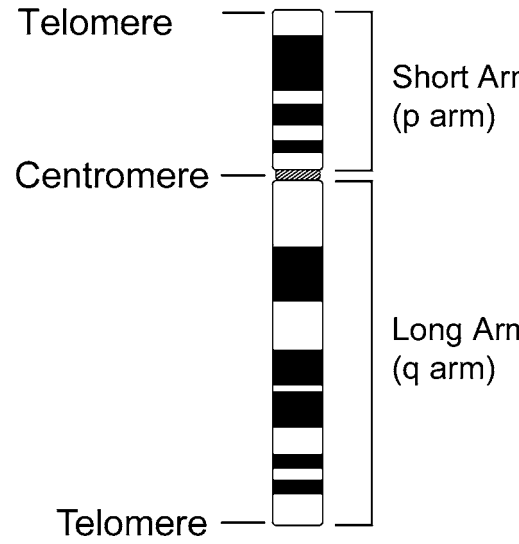


Fig. 2.1 An image showing both chromosome structure and morphology

2.1.3 Methodology

2.1.3.1 Cell Culture

- Chromosome cultures
  - An understanding of the cultures for chromosome analysis is based on a good knowledge of both the cell cycle and mitosis
  - Cell cycle (Fig. 2.2)
    - This is the period between successive mitosis that lasts between 16 and 24 h
    - During interphase the chromosomes are thin and extended
    - DNA replication occurs and chromatid is replicated to become two chromatids
    - The chromosomes begin to condense in preparation for the next mitotic division
  - Mitosis (Fig. 2.3)
    - This is the process of somatic cell division during which the nucleus divides.
    - Each chromosome divides into two daughter cells
    - One chromosome segregates into each of two daughter cells
    - The number of chromosomes in the nucleus remains unchanged
    - Mitosis consists of five major steps and lasts for about 1–2 h

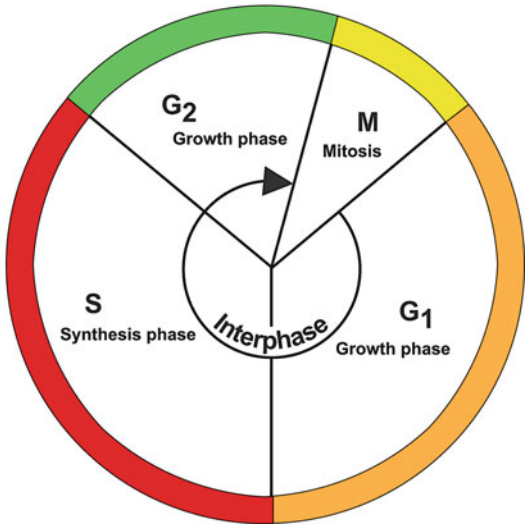


Fig. 2.2 A diagram of the cell cycle illustrating the G_1, S, and G_2 phases along with the stages of mitosis

- Interphase
  - The chromosomes begin to condense
  - Mitotic spindle begins to form
  - Two centrioles form
  - Microtubules radiate and move toward opposite poles
- Prometaphase
  - Nuclear membrane disintegrates
  - Chromosomes spread around cell
  - Each becomes attached at its centromere to a microtubule of the mitotic spindle
- Metaphase
  - The chromosome becomes oriented along the equatorial plate.
  - Each chromosome becomes attached to a centriole by microtubules.
  - The chromosomes are easily visible, are maximally contracted, and resemble an X in its configuration
- Anaphase
  - The centromere of each chromosome divides longitudinally
  - Two daughters separate to opposite poles
- Telophase
  - Chromatids – independent chromosomes
  - Two groups of daughter chromosomes enveloped in a new nuclear membrane
  - Cell cytoplasm also separates
  - Two new daughter cells

– Suspension cultures
  - In most cases for routine analysis, peripheral blood culture is used as this is the most readily available
    - This can be obtained by venipuncture, finger stick, or heal stick
  - Blood must be collected in an anticoagulant (e.g., heparin)
  - Blood can be kept at room temperature or at 4 °C and set up in culture several days later if transport of the sample is slow
  - Blood cannot be frozen
  - The blood can either be set up by macromethods where the white blood is separated out or it can be set up as whole blood by micromethods
  - The white blood cells are stimulated to go from G_0 to G_1 by the use of phytohemagglutinin, which is a T-cell antigen
  - Cells are harvested to obtain metaphase spreads by utilizing colcemid or colchicine
    - Colcemid is a mitotic inhibitor
    - It prevents formation of mitotic spindles and prevents cells from entering anaphase
  - During the harvesting protocol, the cells are treated with a hypotonic solution (usually NaCl)
    - With the spindle apparatus gone, the chromosomes are held together by cytoplasmic membranes
    - The hypotonic solution causes the cells to swell and disperse the chromosomes due to a concentration gradient between the cytoplasm and hypotonic solution
  - The final step of the harvesting procedure is to treat the cells with a fixative, usually a methanol to acetic acid ratio of 3:1
    - This treatment removes water from the cells
    - It enhances the morphology of the chromosomes and its ability to pick up stain

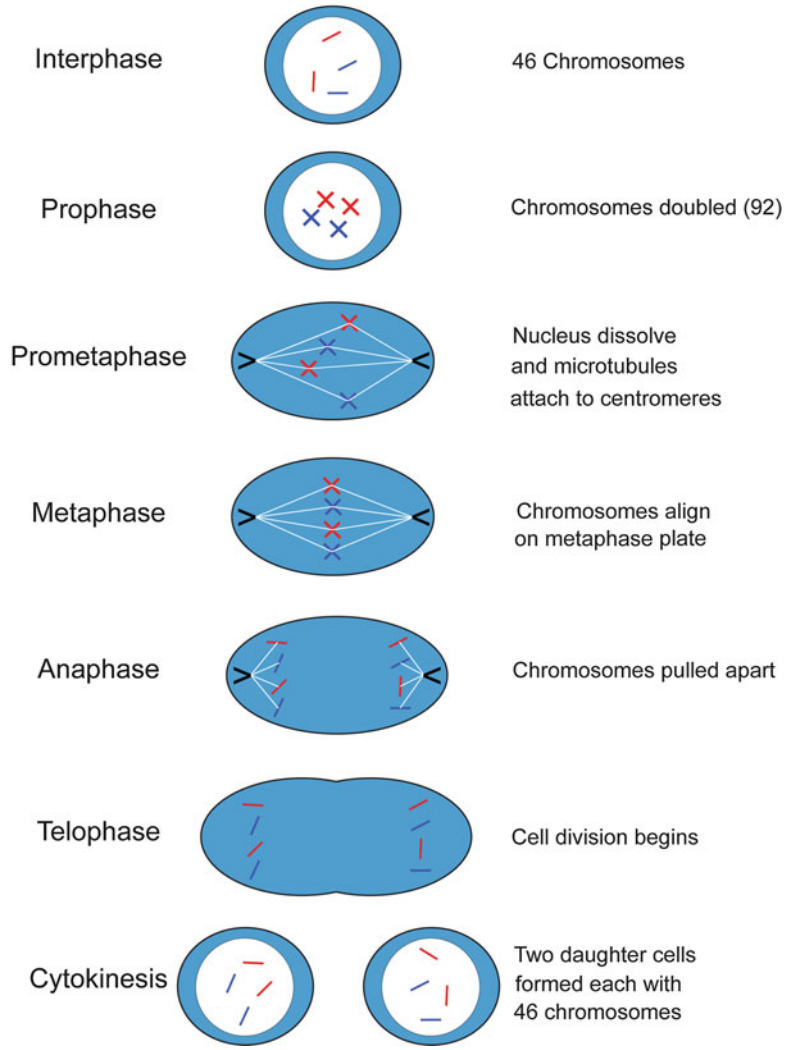


Fig. 2.3 An illustration of the stages of mitosis showing the movement of chromosomes

2.1.3.2 G-Banding

- Chromosome banding
  - Allowed each chromosome to be individually identified
    - Ideograms (diagrammatic representation) of each chromosome allow identification of each region and subregion of the chromosome
  - G-banding
    - This is the most common type of permanent stain, making discrimination of bands easy
    - It can be accomplished by either pretreatment by a proteolytic enzyme (such as trypsin) or chemical (e.g., acetic saline solution)
  - Q-banding
    - Initial banding, described by T. Caspersson
      - Researcher at Karolinska Institute (Sweden) and consultant to Cancer Research Foundation in Boston
      - Authority on fluorescence and interferometry
      - Used alkylating agent to fluorescent molecule, might cross-link guanine
    - Fluorescent bands visible after staining with quinacrine mustard, quinacrine dihydrochloride, or similar compound

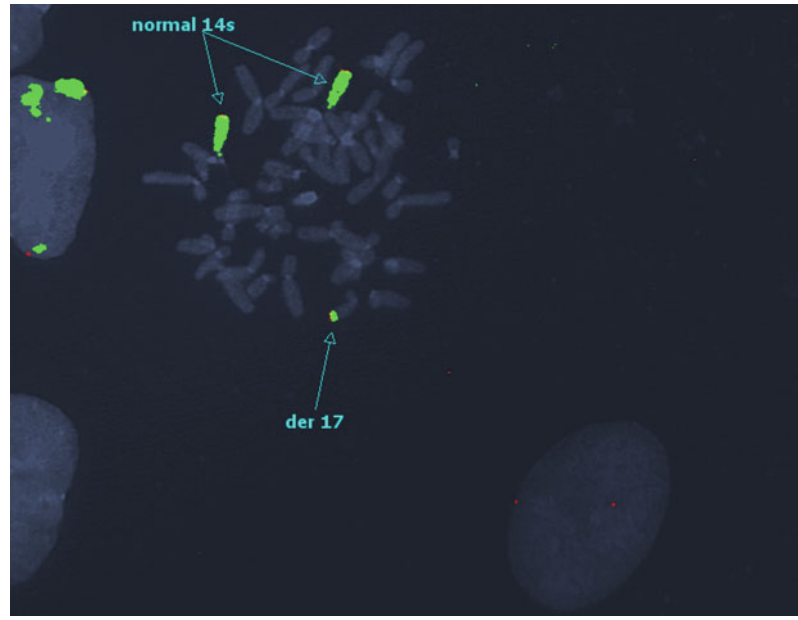


Fig. 2.4 Picture of FISH with a chromosome 14 paint demonstrating extra material on a derivative 17

    - The occurrence of bands related to AT regions, which show brightness due to fluorochrome, which intercalates in DNA
      - Increased fluorescence with runs of AT bases
      - Proteins can modify fluorescence
      - GC-rich regions quench fluorescence
      - G-dark and Q-bright bands are similar
  - R-banding (reverse)
    - This banding is the reverse of G- and Q-banding
    - Staining is usually accomplished by denaturation by heat
  - C-banding
    - This banding stains constitutive heterochromatin; therefore, it stains the centromeric and pericentromeric regions of chromosomes

2.1.3.3 Fluorescence In Situ Hybridization

- General information
  - FISH
    - Allows for the identification of sequences (unique or repetitive) on metaphase or interphase chromosomes
  - Utilization of FISH
    - Identification of deletions
    - Identification of marker chromosomes
    - Identification of duplications
    - Identification of subtle translocations
    - Characterization of chromosome structure
    - Phenotype/karyotype correlations
    - Interphase cytogenetics
    - Characterization of rearrangements
    - Quantification of mosaicism
    - Gene mapping
    - Replication analysis
    - Evolutionary studies
- Materials for FISH
  - DNA FISH probes
    - Chromosome-specific sequences (library-paint) (Fig. 2.4)
      - Identification of translocations, markers, and duplications
    - DNA-repetitive sequences (α-, β-, classical satellite, or telomeres)
      - Centromeric probes – identification of aneuploidy and marker chromosomes
      - Telomeric probes – identification of deletions and cryptic translocations
    - Unique sequences – cosmids, YACs (yeast artificial chromosomes), BACs

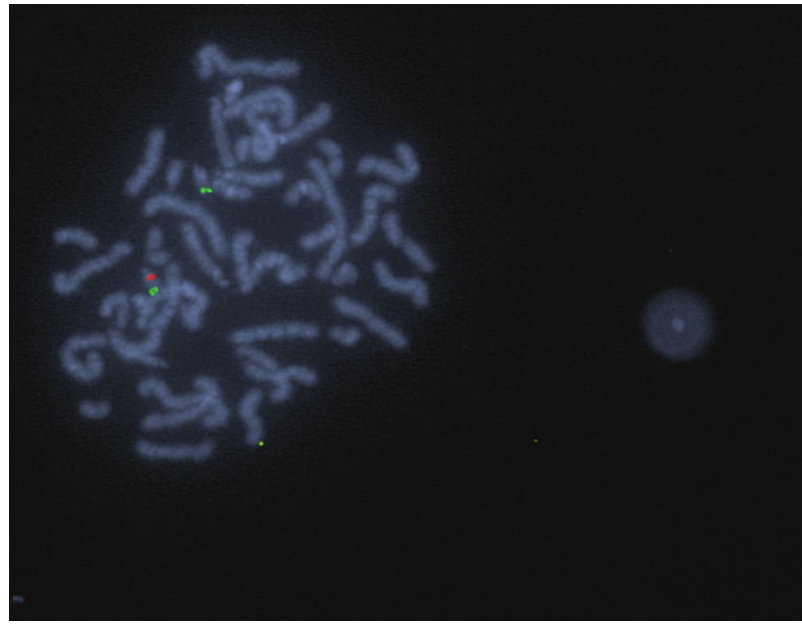


Fig. 2.5 Picture of FISH with a TUPLE1 probe elucidating a microdeletion of 22q11.21

(bacterial artificial chromosomes), fosmid clones
  - Microdeletion probes – identification of specific syndromes (Fig. 2.5)
  - Single-copy probes – detection of present or absence of specific DNA
  - Fusion probes – detection of specific translocations in leukemia
- Labeled by nick translation or random priming or polymerase chain reaction

- Types of cells that can be used for FISH
  - Lymphocytes
  - Bone marrow
  - Amniocytes (cultured or uncultured)
  - Chorionic villus material
  - Fibroblasts
  - Blood smear
  - Buccal smear
  - Paraffin sections
  - FISH analysis can be done on either metaphase chromosomes or interphase cells
- Advantages of metaphase analysis
  - Can visualize all of the chromosomes
  - Can characterize position of signal on chromosomes
  - "Gold standard" of analysis
- Drawbacks of metaphase analysis
  - Often limited number of spreads available
  - Morphology may be compromised and affect hybridization
  - Requires metaphases to be present
- Advantages of interphase analysis (Fig. 2.6)
  - Increased number of cells to examine
  - Allows for investigation of nuclear organization
  - Can save what might have been a failure due to lack of metaphases
  - Replication studies
  - Allows study of nondividing cells

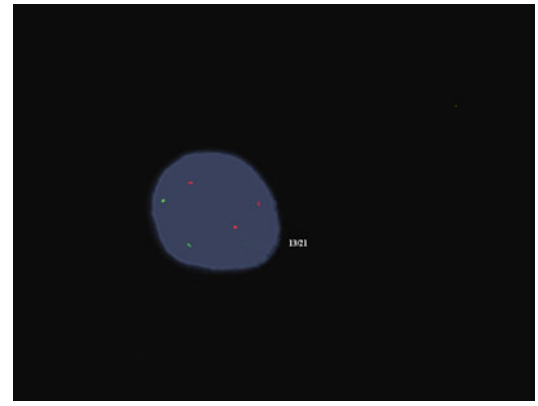


Fig. 2.6 Picture of interphase FISH with locus-specific probes on chromosomes 13 and 21 demonstrating three signals indicating the presence of trisomy 21

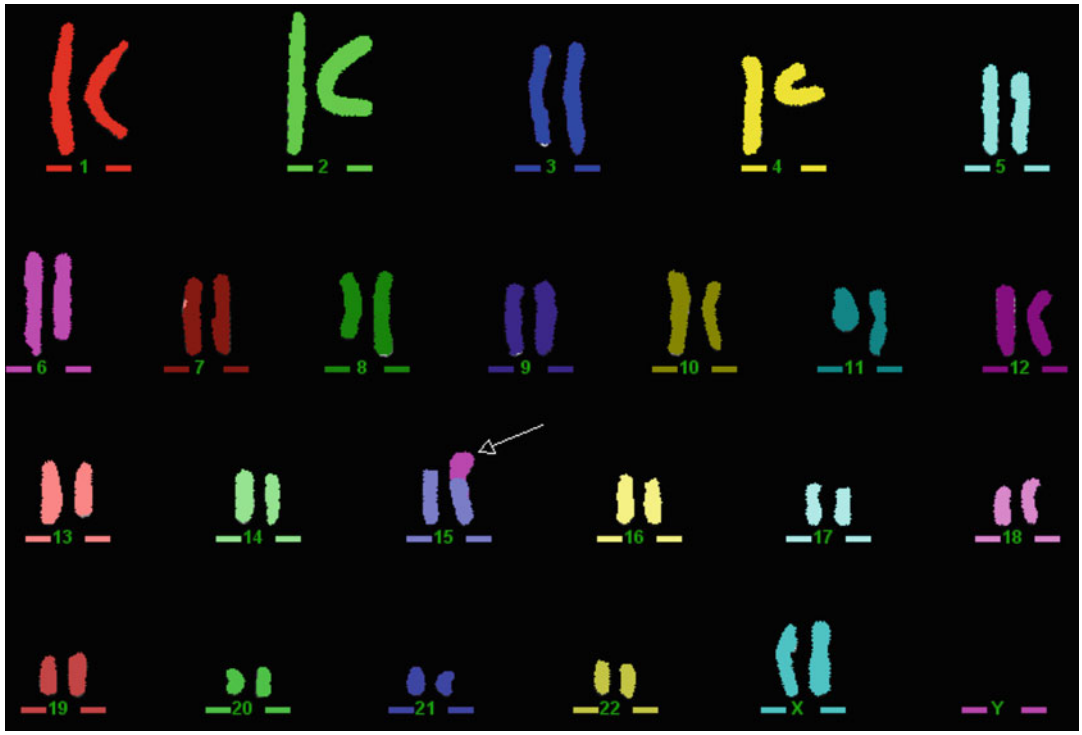


Fig. 2.7 Picture of m-FISH demonstrating that the extramaterial on 15p originated from a Y chromosome

- Disadvantages of interphase analysis
  - Cannot appreciate chromosomes
  - Uncertainty in the form of inefficient hybridization
  - Limited in the number of probes/experiment
- FISH protocol
  - Slide with unbanded cells fixed
  - DNA probe labeled
  - Most are now directly labeled with fluorophore
  - Cot-1 DNA suppression of extraneous DNA
  - Slide dehydrated in ethanol
  - DNA target dehydrated with heat and formamide
  - DNA probe denatured
  - Cold ethanol used to "fix" single strands
  - Probe applied to slide
  - Cover slip and seal slide
  - Slide and probe hybridized overnight (4–20 h) at 37 °C
  - Excess probe washed off
  - Slide counterstained with DAPI (4′,6-diamidino-2-phenylindole) or propidium iodide
  - Visualized with fluorescent microscope
  - Image captured
- Other FISH technologies
  - Fiber FISH
  - Primed in situ labeling (PRINS)
  - Reverse painting
  - Spectral karyotyping or M-FISH (Fig. 2.7)
    - Combinatorially or ratio-labeled probes
    - Are used to create a distinct "color" for each chromosome
    - Chromosomes are studied simultaneously
    - Computer software detects the probes
    - Pseudocolors the chromosomes for analysis
    - Especially useful for complex rearrangements
  - Comparative genomic hybridization
    - Control DNA is labeled red
    - Test DNA is labeled green

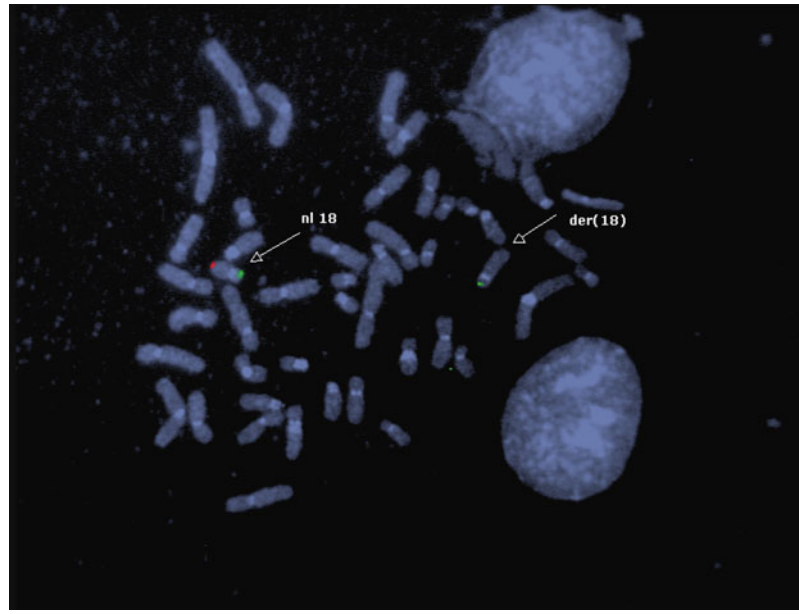


Fig. 2.8 Picture of FISH with telomere locus probes for chromosome 18 (p arm telomere – *green*; q arm telomere – *red*) demonstrating the deletion in 18q

- They are equally mixed and hybridized to normal chromosomes
- The red/green ratio is analyzed by computer software and detects gains and/or losses of material from the test DNA

– Subtelomeric probes – detection of abnormal phenotype (Fig. 2.8)
- Increase in causes of mental retardation detected with subtelomeric probes
- Associated with loss of subtelomeric/telomeric segments
- Cryptic telomeric rearrangements detected
- Individuals with unexplained MR (mental retardation) (3%) explained by use of telomeric probes

2.1.4 Microarray Technology

- Outgrowth of comparative genomic hybridization (CGH) FISH technology
- Utilizes a combination of previously learned CGH FISH technology combined with information learned from the Human Genome Project
  – All array technology employs high-throughput technology
  – All types of arrays can detect subtle copy number changes
    - Changes can be as low as ~25–50 kb
- Several types of arrays
  – Comparative genomic hybridization arrays (CGH)
    - Uses two-color fluorescent system to detect copy number changes
      – Control DNA – red
      – Patient DNA – green
    - Hybridized to DNA probes affixed to a glass slide
      – Exact location of probes known through the Human Genome Project
    - After hybridization
      – Normal amount of DNA – yellow
      – Patient with deleted material – red
      – Patient with duplicated material – green
    - BAC–CGH
      – Original CGH array methodology
      – Utilizes mapped BACs from the Human Genome Project
      – Can detect subtle changes but less sensitive than other methodologies
    - Oligo-CGH
      – Similar to BAC arrays

  - Uses smaller probes (60mer) than BACs
  - More sensitive and can detect smaller changes than BACs
- SNP array technology
  - Most utilize both SNP probes along with copy number (structural) probes
    - SNP – single nucleotide polymorphisms
      - One base-pair substitution of one nucleotide
      - Substitution not considered a mutation
      - Substitution must be in population at least at a frequency of greater than 1.0%
    - Approximately 25 base pairs
  - Utilizes single fluorescent system
  - Probes affixed to silicon wafer
    - Exact location of probes known through the Human Genome Project
  - Detects copy number changes by measuring quantitative changes
    - Does not employ comparative genomic hybridization
  - Can detect copy number changes, similar to CGH methodologies
  - Also detects copy-neutral changes
    - Uniparental disomy
    - Consanguinity

2.1.5 Indications for Cytogenetic Studies

- Why are chromosomes studied?
  - There are a number of indications with respect to why chromosomes are studied:
    - Prenatal diagnosis
    - Confirmation or exclusion of chromosomal syndrome (e.g., trisomy 21)
    - Unexplained psychomotor retardation with/without dysmorphic features
    - Monogenic disorders associated with mental retardation and/or dysmorphic features
    - Abnormalities of sexual differentiation and development
    - Infertility
    - Recurrent miscarriages or stillbirths
      - Parents
      - Fetus
    - Neoplastic conditions
      - Leukemia
      - Lymphoma
      - Solid tumors

2.2 Autosomal Abnormalities

2.2.1 Types of Abnormalities

- Numerical
  - Polyploidy
  - Aneuploidy
- Interchromosomal – abnormalities involving more than one chromosome
  - Reciprocal translocations
  - Robertsonian translocations
- Intrachromosomal
  - Deletion
  - Inversion
  - Ring
  - Duplication
- Numerical abnormalities
  - Polyploidy
    - Multiple of the haploid number (n)
  - Aneuploidy
    - Not a multiple of the haploid number
    - Loss or gain of particular chromosome
  - Mixoploidy
    - Two or more cell lines which differ in chromosome number
  - Mosaic
    - Two or more different cell lines derived from a single zygote
  - Chimera
    - Two or more cell lines that originate from different zygotes

2.2.2 Polyploidy

- Numerical changes – polyploidy
  - Haploid (n)
  - Diploid ($2n$)

2.2.2.1 Triploidy

- Triploid (3*n*)
- 2/3 of human triploids arise by fertilization of a single egg by two sperm
- Fertilization is between a normal haploid gamete and a diploid gamete
- Involvement of a diploid sperm most common occurrence
- 20% of spontaneously aborted fetuses are triploid
- Phenotypic features
  - Voluminous placenta
    - Hydatidiform changes
  - Craniofacial dysmorphology
    - Deformed skull
    - Ocular anomalies
    - Palatal abnormalities
    - Severe micrognathia
    - Low-set poorly folded ears
  - Neck, thorax, and abdomen
    - Diastasis recti
    - Omphalocele (occasional)
  - Limbs are usually deformed
    - Syndactyly – toes and fingers
  - Often severe malformations are seen
    - Cerebral malformations, including holoprosencephaly
  - Cardiac, digestive, kidney, and internal genitalia malformations are often seen

2.2.2.2 Tetraploidy

- Tetraploidy (4*n*)
- Tetraploidy is most often due to the failure of the first cell division of a zygote.
- 2–3% of all fertilized eggs are polyploid
  - Majority of these are spontaneously lost
- 6% of spontaneously aborted fetuses are tetraploid
- Some polyploidy pregnancies do result in live births; these die very early during the first few hours or days

2.2.3 Aneuploidy

- Numerical changes – aneuploidy
  - Aneuploid changes
    - Nullosomic ($2n - 2$)
    - Monosomic ($2n - 1$)
    - Trisomic ($2n + 1$)
    - Tetrasomic ($2n + 2$)
    - Double trisomy ($2n + 1 + 1$)

2.2.3.1 Trisomy 21

- Trisomy 21 – down syndrome (Fig. 2.9)
- Frequency, at birth, is approximately 1/650–1/700
- The sex ratio is approximately three males/two females
- The most frequent cause leading to trisomy 21 is advanced maternal age
- Phenotypic features
  - Newborn period
    - Hypotonia
    - Sleepy
    - Excess nuchal skin
  - Craniofacial abnormalities
    - Brachycephaly
    - Epicanthal folds
    - Protruding tongue
    - Upward-slanting palpebral fissures
    - Small ears
  - Limbs abnormalities
    - Single palmar crease
    - Small middle phalanx of the fifth finger
    - Wide gap between first and second toe
  - Cardiac abnormalities
    - Atrial and ventricular septal defect
    - Common atrioventricular canal
    - Patent ductus arteriosus
  - Other abnormalities
    - Anal atresia
    - Duodenal atresia
    - Short stature
    - Strabismus
- Natural history of trisomy 21
  - Broad range of intellectual ability
    - IQ between 25 and 75
  - Social skills
    - Well-advanced skills
    - Happy and very affectionate
  - Adult height, approximately 150 cm (4 ft 11 in.)
  - Life expectancy – good (~60 years)
    - Except for severe cardiac anomaly
      - Early death in 10–20% of the cases

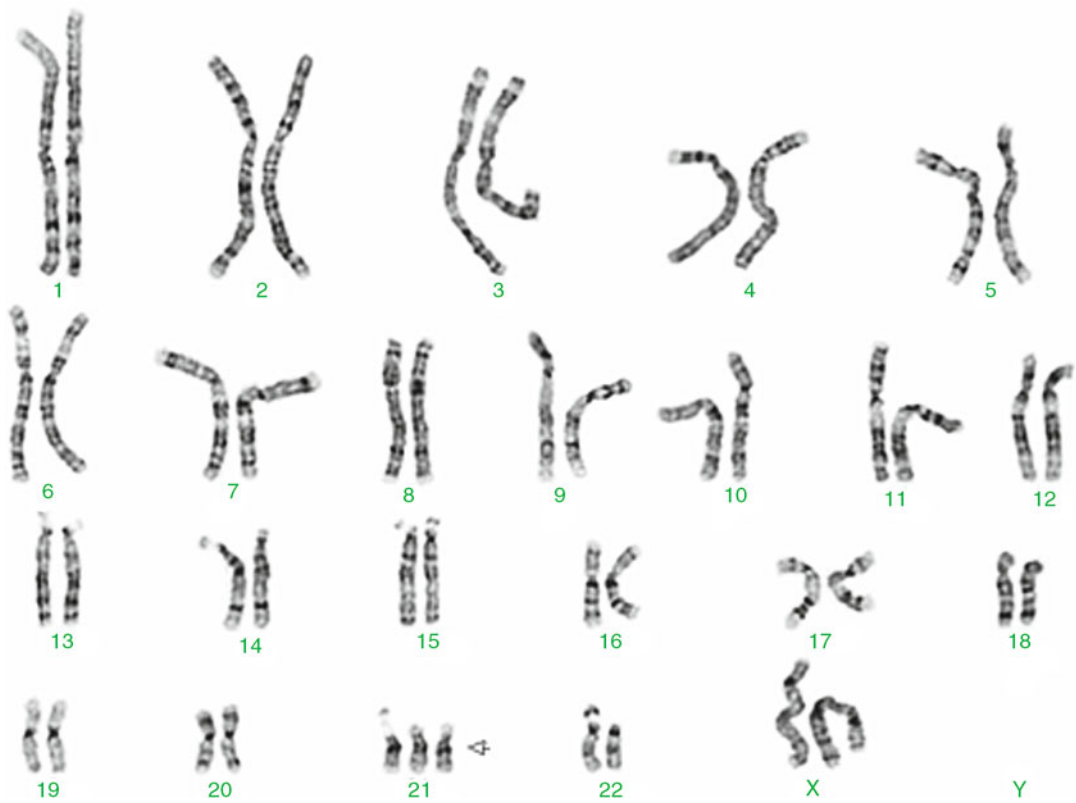


Fig. 2.9 G-banding karyotype demonstrating trisomy 21

    - Life expectancy approximately 9 years in 1930
  - 40% of the individuals have significant heart defect
    - Atrioventricular canal defects
  - Five percent of the individuals have serious gastrointestinal anomalies
    - Duodenal stenosis is the most common
  - Children with down syndrome have a 15–20-fold increase risk of leukemia
    - Overall, this is a frequency of approximately 1%
  - Most affected adults develop Alzheimer disease
    - All >35 years old develop it
- Overlap of phenotypic features with the general population
  - Most features in trisomy 21 can be found in individuals in the general population
  - The phenotypic features are especially important when all of the features are grouped together
    - For example, single palmar crease
      - In 50% of the cases of down syndrome
      - Seen in 2–3% general population
- Trisomy 21 – chromosome findings
  - 95% are due to an extra free trisomy 21 (resulting from a meiotic error)
  - 3–4% result from a Robertsonian translocation
  - 1–2% result from mosaicism (resulting from a mitotic error)
- Trisomy 21 – recurrence risk
  - This risk is related to maternal age
    - Approximately 1/100–1/200

2.2.3.2 Trisomy 18

- Trisomy 18 – Edwards syndrome
  - The frequency of this syndrome is about 1/5,000–1/8,000

- The sex ratio at birth is approximately four female/one male
- Overall survival of affected individuals
  - 30% die within first month
  - 50% die within 2 months
  - 90% die within 1 year
- Phenotypic information
  - General features
    - Many pregnancies demonstrate post-maturity with delivery at 42 weeks
    - Low birthweight with severe growth retardation
    - Hypoplasia of skeletal muscle
    - Newborns can be hypotonic or hypertonic
    - Affected individuals have severe mental retardation
  - Craniofacial dysmorphology
    - Dolichocephaly
    - Protuberant occiput
    - Protuberant nose
    - Short palpebral fissures
    - Low-set malformed ears
    - Micrognathia
  - Neck, thorax, and abdomen abnormalities
    - Short neck, excess skin
    - Short sternum
    - Narrow pelvis
  - Limbs abnormalities
    - Clenched hands
    - Overlapping fingers
    - Absent distal flexion creases
    - Rocker bottom feet
  - Genitalia abnormalities
    - Cryptorchidism
    - Clitoral hypotrophy
    - Hypoplasia of the labia majora
  - Other malformations are seen in >95% of patients including:
    - Cardiac malformations
      - This is often responsible for death
    - Gastrointestinal malformations
    - Renal malformations

2.2.3.3 Trisomy 13

- Trisomy 13 – Patau syndrome
  - The frequency of this syndrome is about 1/10,000–1/15,000
  - Overall survival of affected individuals
    - 30% die within first month
    - 50% die within 2 months
    - 90% die within 1 year
  - Phenotypic information
    - General features
      - Newborns have failure to thrive
      - Seizures
      - Newborns are usually hypotonic
      - Affected individuals have severe mental retardation
    - Craniofacial dysmorphology
      - Holoprosencephaly
      - Microcephaly
      - Microphthalmia
      - Iris colobomata
      - Cleft lip and/or palate
      - Hemangiomas
    - Neck, thorax, and abdomen abnormalities
      - The last rib is either hypoplastic or absent
    - Limbs abnormalities
      - Hexadactyly
      - Rocker bottom feet
    - Genitalia abnormalities
      - Cryptorchidism
      - Scrotal abnormalities
      - Clitoral hypertrophy
      - Bicornuate uterus and double vagina
    - Other malformations seen
      - Cardiac malformations
      - Digestive malformations
      - Ocular malformations
        - Microphthalmia
        - Anophthalmia
      - Visceral malformations
      - Cerebral malformations
        - Holoprosencephaly
      - Urinary malformations
        - Polycystic kidneys

2.2.4 Aneuploidy: Causes

- Causes of aneuploidy
  - Nondisjunction
    - Can be either meiotic or mitotic

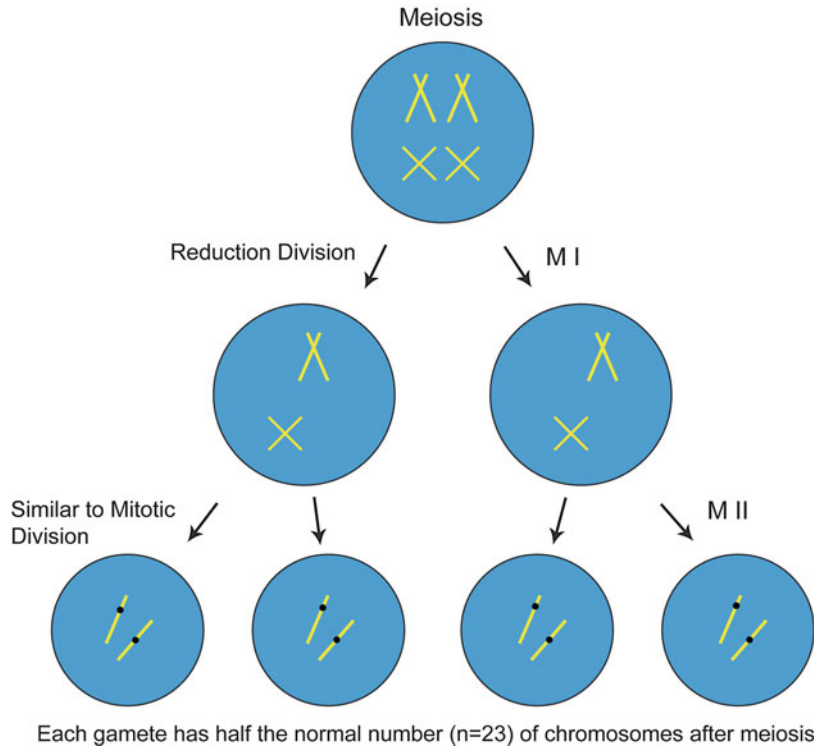


Fig. 2.10 A diagram of normal meiosis demonstrating normal segregation of chromosomes

- Failure of paired chromosomes to separate (disjoin) at meiosis I
- Failure of paired sister chromatids to disjoin at meiosis II or mitosis
- Conjoined chromosomes/chromatids migrate to one pole
  - The other pole has no chromosome

– Anaphase lag
- Failure of incorporation of a chromosome into one of the daughter nuclei following cell division
- Occurs due to delayed movement of the chromosome during anaphase and chromosome is subsequently lost

– Meiosis (Fig. 2.10)
- This is the process where the diploid count is halved
  - From 46 to 23 chromosomes (and becomes haploid)
- Occurs only at the final division of gamete maturation
- This is a two-step process and involves two cell divisions

– Meiosis I
- Referred to as the stage of reduction division becaus the chromosome number is halved
- Prophase I
  - Homologous chromosomes pair
  - Crossing over (recombination) occurs between nonsister chromatids
  - Prophase I is relatively lengthy consisting of five stages
    - Leptotene
    - Zygotene
      - Synaptonemal complexes are formed
    - Pachytene
      - Pairs of homologous chromosomes
      - Bivalents formed
      - Crossing over occurs
    - Diplotene
      - Chromosomes separate
      - The chromosomes are attached by chiasma

        - Diakinesis
            - Separation of the chromosomes proceeds
    - Metaphase I
        - Chromosomes are attached to spindle
    - Anaphase I
        - Chromosomes separate and go to opposite poles
    - Telophase I
        - Two new daughter cells are formed
- Meiosis II
    - This is essentially similar to mitotic division
    - Each chromosome (pair of chromatids)
        - Becomes aligned along equatorial plate
        - Forms two new daughter gametes
- What are the consequences of meiosis?
    - Two major objectives are achieved
        - The diploid number of chromosomes is halved
            - Haploid
        - Meiosis provides extraordinary potential for generating genetic diversity
- Comparison between mitosis and meiosis
    - Location
        - Mitosis: all tissues
        - Meiosis: only in testis and ovary
    - Products
        - Mitosis: diploid somatic cells
        - Meiosis: haploid sperm and egg cells
    - DNA replication and cell division
        - Mitosis: normally one round of replication per cell division
        - Meiosis: only one round of replication (in meiosis I), but two cell divisions
    - Length in prophase
        - Mitosis: short (~30 min in human cells)
        - Meiosis: long and complex in meiosis I; can take years to complete
    - Pairing of homologs
        - Mitosis: none
        - Meiosis: yes (in meiosis I)
    - Recombination
        - Mitosis: rare and abnormal
        - Meiosis: normally at least once for each pair of homologs
    - Relationship between daughter cells
        - Mitosis: genetically identical
        - Meiosis: different (recombination and independent assortment of homologs)
- What are the causes of aneuploid?
    - Aneuploidy results from nondisjunction
        - Failure of chromosomes to separate normally during cell division
            - In either meiosis or mitosis
    - Parental origin of nondisjunction
        - Origin of nondisjunction determined by parental polymorphisms
            - Chromosomal heteromorphisms have been used
            - DNA markers used now
                - For example, microsatellite markers
        - Origin of nondisjunction
            - Trisomy 21
                - Maternal – 88% of time
                - Paternal – 8%
                - Mitotic – 3%
            - Trisomy 13
                - Maternal – 95%
                - Paternal – 5%
            - Trisomy 18
                - Maternal – 89%
                - Paternal – 0%
                - Mitotic – 11%

2.2.4.1 Nondisjunction

- Two major causes of nondisjunction (Fig. 2.11)
    - Advanced maternal age
        - Primary oocyte can remain in a state of suspended inactivity for up to 50 years
            - Stays in dictyotene stage
        - Well-documented association between advanced maternal age and nondisjunction (Table 2.1; Fig. 2.12)
        - No association with advanced paternal age
    - Altered recombination

2.3 Structural Rearrangements: Intrachromosomal

- Interchromosomal
    - Reciprocal translocations
    - Robertsonian translocations

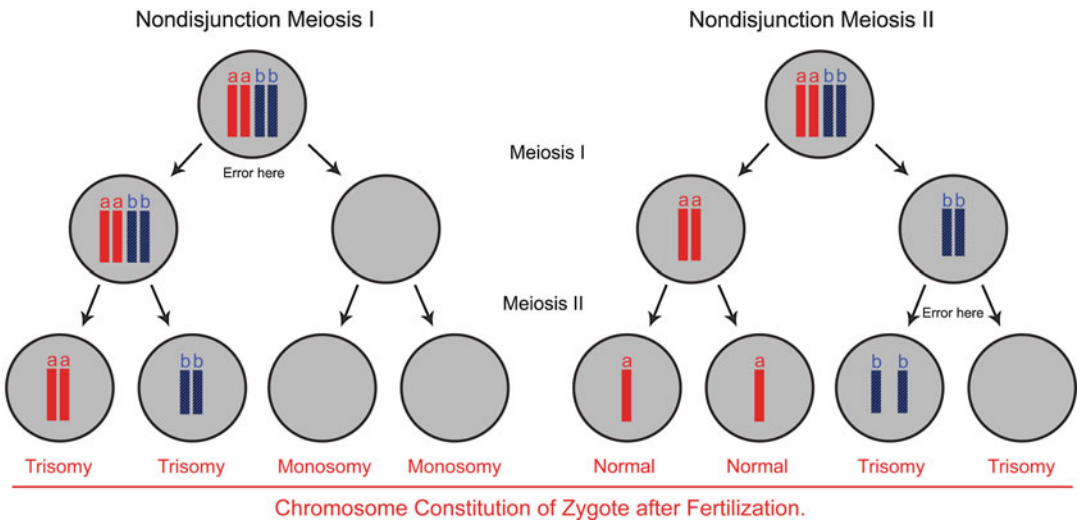


Fig. 2.11 An example of nondisjunction resulting in monosomic, euploid, and trisomic gametes

Table 2.1 Association of advanced maternal age and increased risk of nondisjunction

Maternal age	Risk of nondisjunction
20 years old	1/1,500
25 years old	1/1,350
30 years old	1/900
35 years old	1/400
40 years old	1/100
45 years old	1/30

- Intrachromosomal (Fig. 2.13)
  - Deletion
  - Ring
  - Isochromosome
  - Duplication
  - Insertion
  - Inversion
  - Dicentric, acentric

2.3.1 Intrachromosomal Rearrangements

2.3.1.1 Deletions

- Involve loss of part of a chromosome (Fig. 2.14)
  - Results in monosomy of that segment of the chromosome
  - Large deletions will be incompatible with survival to term
  - Any deletion of a loss of >2% of the haploid genome will usually be lethal
- Types of deletions (I)
  - Deletions visualized under the microscope
    - Wolf-Hirschhorn syndrome
    - Cri du chat syndrome
  - Microdeletions, contiguous gene deletions
    - Prader-Willi syndrome
    - Velocardiofacial syndrome
- Types of deletions (II)
  - Terminal deletion
    - Single break
    - Acentric terminal fragment lost
    - Telomeres
    - May be reattached or reformed
  - Interstitial deletion
    - Two breaks
    - Telomere retained
    - Interstitial acentric fragment lost
- Deletions – phenotypes
  - Deletions seen in most chromosome arms
  - Certain deletions are seen more frequently
  - Del(4p), del(5p), del(9p), del(11p), del (11q), del(13q), del(18p), and del(18q)
  - Wide variability in phenotypes

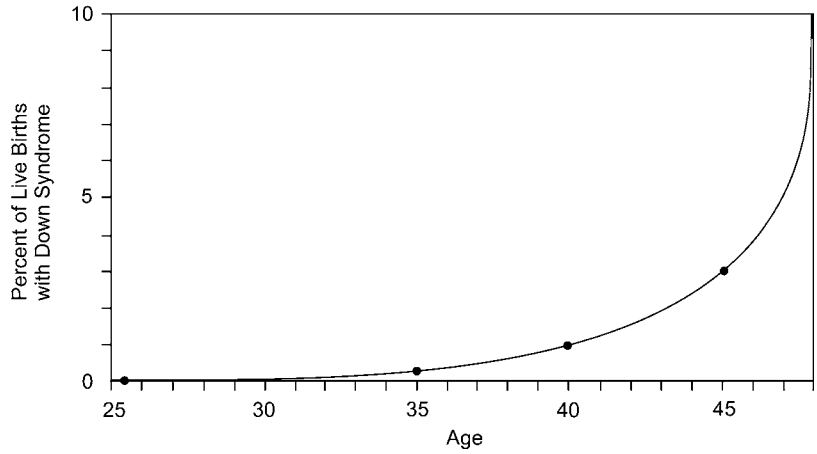


Fig. 2.12 A maternal age curve showing the increase in the percent of live births with Down syndrome with increasing maternal age

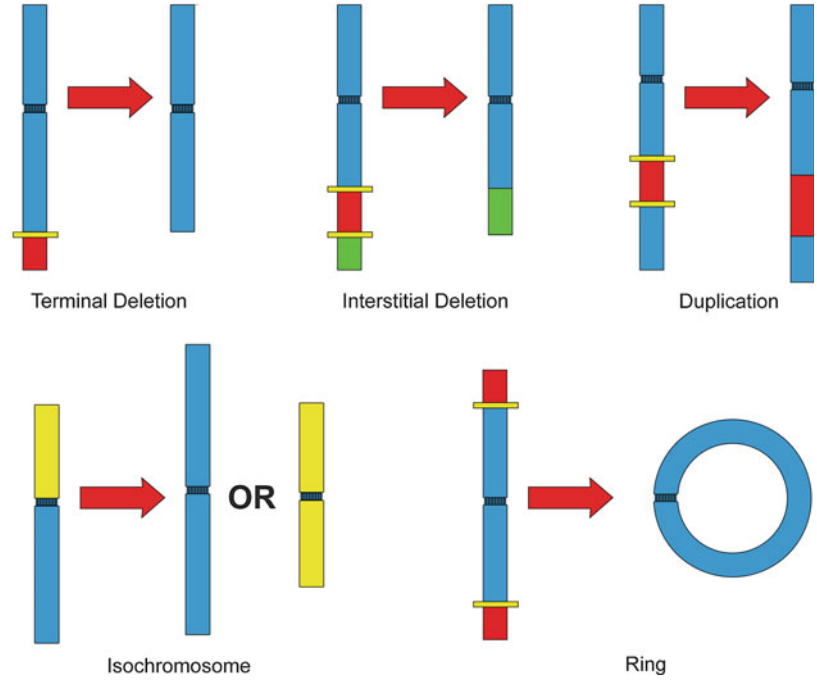


Fig. 2.13 Diagrammatic representation of several different structural abnormalities including a terminal deletion, interstitial deletion, duplication, isochromosome, and ring

5p Deletion

- Deletion (5p) – general
  - Cri du chat syndrome
  - Deletion – short arm of chromosome 5 (p14p15)
  - Frequency – 1/45,000–1/50,000
  - First reported – Lejeune et al. in 1963
- Deletion (5p) – phenotype
  - General
    - Cry – mewing of a kitten
    - Craniofacial dysmorphism
    - Microcephaly
    - Moon-like face
    - Hypertelorism
    - Micrognathia
    - Malformations rare
  - Larynx
    - Laryngomalacia, laryngeal stridor
    - Distinctive cry
  - Mental retardation
    - Usually severe
    - IQ often <20

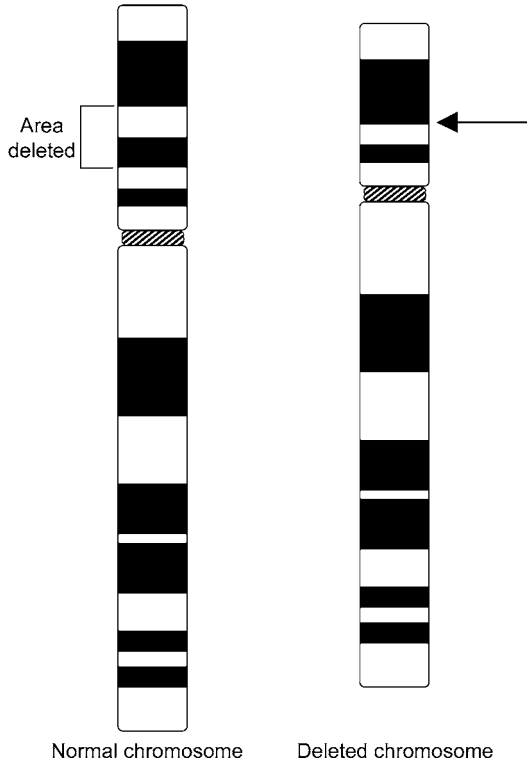


Fig. 2.14 Diagrammatic representation of an interstitial deletion

- Frequency of disorder higher among individuals with mental retardation (1.5/1,000)

4p Deletion

- Deletion (4p) – general
  - Wolf-Hirschhorn syndrome
  - Deletion – short arm of chromosome 4 (p16)
  - Frequency – 1/45,000–1/50,000
  - First reported – Wolf et al. in 1965
- Deletion (4p) – phenotype
  - General
    - Severe growth retardation
    - Severe mental retardation
  - Craniofacial dysmorphism
    - Microcephaly
    - "Greek warrior helmet"
  - Genitalia abnormalities (male and female)
  - Malformations
    - Defects in closure of scalp
    - Cleft lip and/or palate
    - Coloboma
    - Cardiac defects (50%)
  - Mental retardation
    - Very pronounced
    - IQ <20
  - Cytogenetics findings
    - De novo deletion – 90% of cases
    - Parental mosaicism or translocation
      - 10% of cases
- Deletions – nomenclature
  - Interstitial deletion
    - 46,XX,del(5)(p14p15)
  - Terminal deletion
    - 46,XX,del(5)(p14)

2.3.1.2 Ring Chromosomes

- Breaks occur in each arm of the chromosome
- Two "sticky" ends
- Reunites as a ring
- Two distal chromosome fragments lost
- Have been found for all human chromosomes
- Rings – effects
  - If autosomal
    - Effects can be serious
    - Depends on amount of material deleted
    - Difficult to make phenotype/karyotype correlations
  - Sex chromosome
    - r(X), r(Y)
- Rings – cytogenetics
  - Often unable to complete mitotic divisions
  - Leads to mosaicism
  - Loss of ring
  - Double rings
  - Extra rings
- Rings – nomenclature
  - 46,XX,r(4)(p14q34)

2.3.1.3 Isochromosomes

- Loss of one arm with duplication of the other arm
- Isochromosome – formation
  - Originally thought to be due to centromere misdivision
    - Centromere divides transversely rather than longitudinally

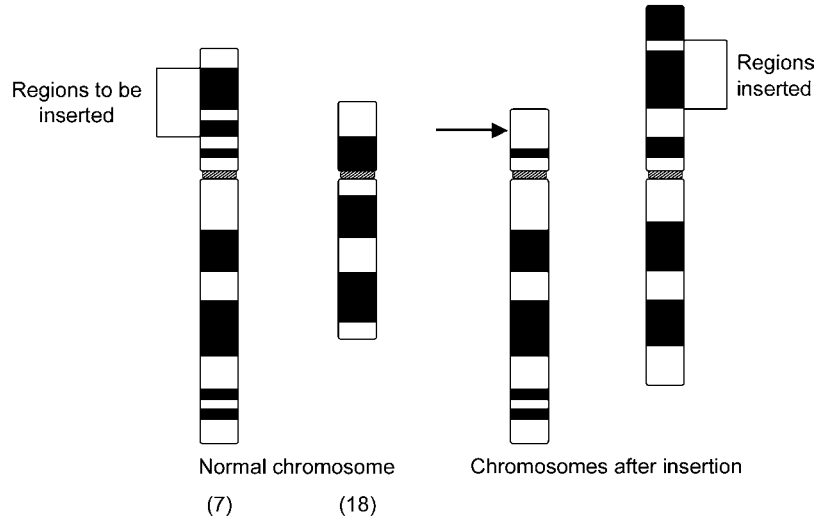


Fig. 2.15 Diagrammatic representation of a balanced insertion

  - More recent studies
    - Most breakpoints not in the centromere
      - FISH
        - Dicentric – two centromeres present
- Isochromosome – phenotype
  - For most (but not all) autosomal chromosomes
    - Not viable
    - i(18q) – viable
      - Monosomy and trisomy for either chromosome arm
  - i(9p), i(18p), i(18q), and i(12p)
    - These isochromosomes are viable when present as accessory chromosomes
  - Sex chromosomes
    - i(Xq) – see Sect. 2.6.2.1
- Isochromosome – nomenclature
  - 46,XX,i(18)(p11)

2.3.1.4 Insertions

- Insertions involve three breaks (Fig. 2.15)
  - Two breaks to delete segments
  - One to reinsert segment
  - Insertion involves transfer to another chromosome
    - Will discuss with translocations
  - When in the same chromosome
    - Shift
  - Either inserted straight or inverted

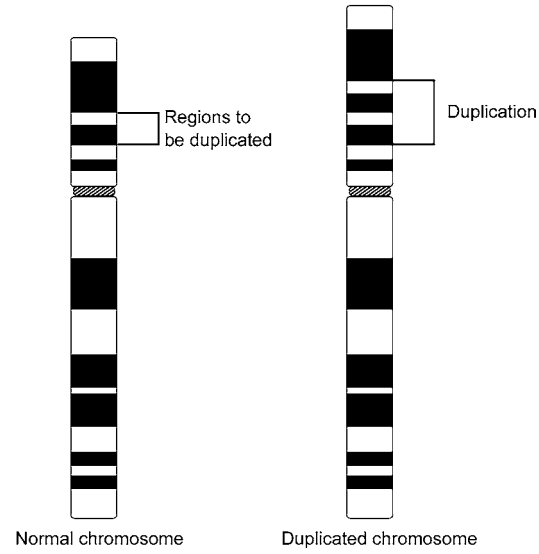


Fig. 2.16 Diagrammatic representation of a duplication

2.3.1.5 Duplications

- Can be de novo or familial (Fig. 2.16)
  - Familial
    - Due to abnormal segregation of familial balanced rearrangement
    - Most duplications are due to abnormal segregation of parental translocation, insertion, or inversion
  - De novo – many types
    - Tandem
    - Inverted tandem

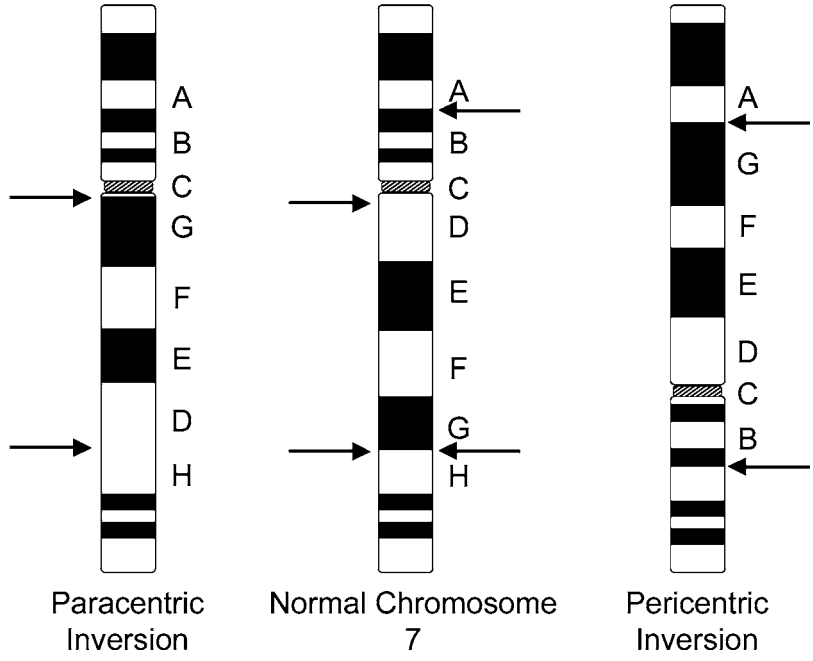


Fig. 2.17 Diagrammatic representation of both a paracentric and a pericentric inversion

- Isochromosome
- Accessory chromosome

2.3.1.6 Inversions

- Two-break rearrangement involving a single chromosome in which a segment is reversed in position (i.e., inverted) (Fig. 2.17)
- Two types
  - Paracentric
  - Pericentric
- Frequency: 1/100–1/1,000
- Inversions – types
  - Pericentric
    - Involves both chromosome arms and includes the centromere
  - Paracentric
    - Involves only one of the chromosome arms and does not include the centromere
  - Inversions – phenotype
    - Balanced rearrangement which rarely causes problems in carriers
      - Unless breakpoint is in an important functional gene
      - Can lead to significant chromosome imbalance in offspring with important clinical consequences

Segregation

- Inversion – segregation
  - During meiosis, inverted chromosome will pair by forming an inversion loop
    - If crossing over occurs in inversion loop
    - Unbalanced gametes
      - Lead to duplications and deletions
  - Pericentic inversion – segregation
    - Small inversion
      - Less likely for crossing over to occur
      - Greater degree of imbalance if crossing over
      - More likely to spontaneously abort
    - Large inversion
      - More likely for crossing over to occur
      - If crossing over occurs – smaller imbalance
      - More likely for live birth
  - Pericentric inversion – risk
    - Balanced pericentric inversion
      - 5–10% risk for having a live-born child with a viable imbalance
        - If ascertained due to a previous abnormal child
      - 1% risk
        - If ascertained due to a history of multiple spontaneous abortions

2.4 Structural Rearrangements: Interchromosomal

2.4.1 Interchromosomal Rearrangements

2.4.1.1 Reciprocal Translocations

- Definition
  - Frequency – 1/1,000 live births
  - Involve two chromosomes
    - An interchromosomal rearrangement
  - Involve breaking and exchanging segments
  - New morphologically different, but recognizable, rearrangement of genetic material
  - Different types of translocations

Translocation Formation

- Formation
  - Mutual breakage and exchange of material
  - Exchange may not be mutual
  - Terminal translocation
    - Most common type of translocation involving terminal regions
  - Insertional translocation
    - Less common type of rearrangement involving exchange of interstitial regions inserted into chromosomes
  - Cryptic translocation
    - Translocation not seen with typical banding analysis
      - Usually need FISH to visualize
- Inheritance
  - Familial
  - De novo

Translocation Segregation

- Meiotic segregation
  - Translocations do not disjoin as normal homologous pairs of chromosomes
  - Forms a quadrivalent in pachytene stage of meiosis
    - Chiasma keeps the configuration together during meiosis
- Types of segregation
  - Alternate segregation
  - Adjacent-1
    - Segregation involving nonhomologous centromeres
  - Adjacent-2
    - Segregation involving homologous centromeres
  - 3:1 disjunction
- Outcomes of segregations
  - Alternate or adjacent-1 segregation
    - Most frequent types of segregation
    - Can have normal/balanced or unbalanced segregants
    - Dependent on chiasma placement
  - Adjacent-2 or 3:1 segregation
    - Will always lead to unbalanced segregants
- Unbalanced segregants
  - Both monosomic and trisomic segments present
    - Makes phenotype/karyotype correlations difficult
    - Counseling difficult but imperative
- Unbalanced segregants – risks (I)
  - Overall an ~11–12% risk of unbalanced segregant
  - Empiric data from prenatal diagnosis studies
    - Ascertainment dependent
      - Previous child with unbalanced translocation
        - ~20% risk
      - Previous multiple spontaneous abortions (SABs)
        - ~3–4% risk
      - Other ascertainment
        - ~7% risk
- Unbalanced segregants – risks (II)
  - If ascertainment is because of multiple SABs – lower risk
    - Larger translocation exchange segments involved
      - Unbalanced segregants usually not viable
  - If ascertainment is because of a previous unbalanced child – higher risk
    - Smaller translocation exchange segments involved
      - Increased viability of unbalanced segments

- Translocation – types
  - Most translocations not similar
  - Leads to difficulties in counseling
  - Usually use overall group risks based on ascertainment
  - One exception
    - t(11;22)(q23;q11.2)
      - Model for 3:1 segregation
- 3:1 disjunction
  - Usually one acrocentric chromosome involved
  - Disparity of length of chromosomes involved
  - Best example – t(11;22)(q23;q11.2)
  - Short interstitial region between the centromere and point of breakage
  - Most frequently maternal transmission
  - – t(11;22)(q23;q11.2) – risks
    - Empiric information based on >100 individuals
    - Risk of unbalanced segregants (5–6%)
    - Most common – 47, +der(22)t(11;22)(q23;q11.2)
    - Increased risk of pregnancy loss – 35%
    - Overall risk for recognized pregnancy >40%
- Parental translocation – segregation risks
  - Alternate and adjacent-1 segregation most commonly observed
    - Frequency depends on distance between the centromere and breakpoint and relative length of the chromosome arms
  - Adjacent-2 segregation – large imbalances present
    - Results in nonviable gametes or early embryonic deaths
  - 3:1 segregation
    - Commonly observed if one translocation product is a small acrocentric
  - Risks
    - Empiric risks often used
    - Best to obtain individual family histories
    - Depends on type of translocation and viability of unbalanced products
    - Familial pattern for inherited translocation with unbalanced segregants may be distinctive
  - Once a rearrangement is found in proband
    - Study parents
    - Study other family members including siblings

De Novo Translocations

- Serious clinical implications
- Unbalanced translocation
  - Obvious clinical implications
    - Need to delineate abnormality
      - FISH
- Balanced translocation
  - Incidence of mental retardation
    - Increased in de novo rearrangements
      - 20% of translocations in newborns de novo
      - 55% of translocations in MR de novo
    - 3/1,000 individuals with MR with de novo translocations
  - Incidence of congenital anomalies
    - Increased in de novo rearrangements
- Balanced de novo translocations – why clinical findings?
  - Possible position effect
    - Genes turned on/off in rearranged chromosome
  - Breakage within a gene
    - For example, Duchenne muscular dystrophy
  - Subtle deletion/duplication
    - FISH, molecular studies, and array studies
  - Uniparental disomy
- Prenatal diagnosis translocation
  - 6–9/10,000 amniocentesis
    - Determine if parent has a rearrangement
    - If de novo determine paternity
  - If de novo
    - 8% risk of phenotypic abnormality
      - Based on empiric data
    - Perform high-level ultrasound

2.4.1.2 Robertsonian Translocations

- Definition (Fig. 2.18)
  - Centric fusion – Robertson described in 1916
    - Most common structural abnormality

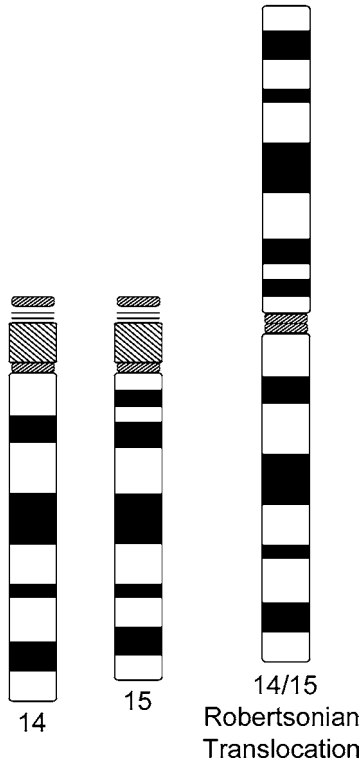


Fig. 2.18 Diagrammatic representation of the formation of a Robertsonian translocation

- Frequency – 1/1,000 live births
  - Involves two acrocentric chromosomes
- Different types of rearrangements
  - Homologous chromosomes
  - Nonhomologous chromosomes

Translocation Formation

- Formation
  - Not a centric fusion of acrocentric chromosomes
  - Arise from pericentromeric exchanges
    - Within short arms, stalks, and satellite regions
  - Location of breakage can vary
    - Most occur within proximal satellite III DNA
  - Dicentric vs monocentric Robertsonian translocations
- Translocations involving homologous chromosomes
  - 10% of all Robertsonian translocations
  - Involves homologous chromosomes
    - For example, rob(13;13) (q10;q10)
  - Are they isochromosomes?
    - Studies revealing homozygosity of molecular markers suggest that they are isochromosomes
  - Most are monocentric
  - Segregation
    - Leads to trisomies or monosomies
- Nonhomologous translocations
  - 90% of all Robertsonian translocations
  - Involves nonhomologous chromosomes
    - For example, rob(13;14) (q10;q10)
  - Preferential involvement of specific types
    - For example, rob(13;14) (q10;q10); rob(14;21) (q10;q10) seen most frequently
  - Most are dicentric
- Ascertainment – overview
  - Couples with multiple spontaneous abortions
  - Unbalanced probands
  - Translocation trisomy 13 or 21
  - Prenatal/newborn studies
  - Neoplasia
- Ascertainment – multiple spontaneous abortions
  - There is an increase in rob(13;14) over other types of translocations
  - There is an increase in homologous translocations
- Ascertainment – unbalanced proband
  - Predominance of rob(13;14) and rob(14;21) over other types of translocations
    - Ascertained through trisomy 13 and 21 offspring
  - Translocations usually involve chromosomes 13 and 21
- Ascertainment – prenatal/newborn studies
  - Ascertainment by chance
  - Serendipitous finding or through survey
    - Unbiased findings
    - High frequency of rob(13;14) and rob(14;21)
- Inheritance
  - Familial
  - De novo

Translocation Segregation

- Meiotic segregation (Fig. 2.19)
  - Translocations do not disjoin as normal homologous pairs of chromosomes
  - Forms a chain of three (trivalent) in pachytene stage of meiosis
  - Types of segregation
    - Alternate segregation
    - Adjacent-1
    - No adjacent-2
    - No 3:1 disjunction
  - Outcomes of segregations
    - Alternate segregation
      - Normal/balanced segregants
    - Adjacent-1 segregation
      - Unbalanced segregants
        - Either monosomic or trisomic segments
        - Different than reciprocal translocation
        - Phenotype/karyotype correlations straightforward
        - Counseling imperative but more straightforward
- Unbalanced segregants – risks
  - Dependent on:
    - Type of translocation
      - Homologous versus nonhomologous
      - Chromosomes involved
      - Sex of carrier
  - Unbalanced segregants risks nonhomologous
    - Risk based on empiric data
      - Theoretical risk of abnormal live birth – 33%
      - rob(14;21) (q10;q10) – empiric risk
        - Female carrier – 10–15%
        - Male carrier – 2%
      - rob(13;14) (q10;q10) – empiric risk
        - Female carrier – 1%
        - Male carrier – probably <1%
  - Unbalanced segregants risks – homologous
    - Theoretical and empiric risks similar
    - For example, rob(21;21) (q10;q10)
      - 100% of offspring abnormal
  - Once a rearrangement is found
    - Examine family history
    - Study parents

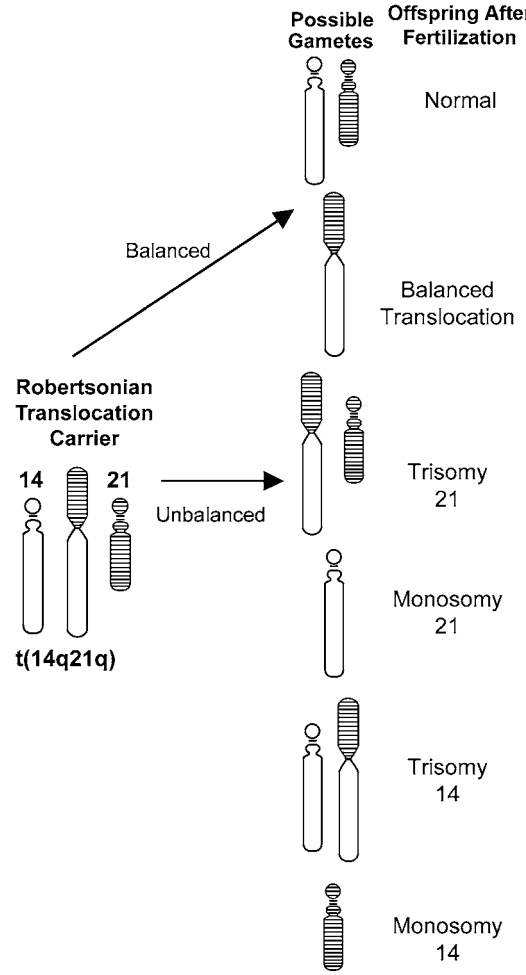


Fig. 2.19 Diagrammatic representation of the segregation of a 14/21 Robertsonian translocation

    - Study other family members including siblings
- De novo translocations
  - Unbalanced translocation
    - Obvious clinical implications
    - "trisomy 21" – 3% of down syndrome cases
    - "trisomy 13"
    - Spontaneous abortions
      - +13, +14, +15, +21, and +22
  - Balanced translocation de novo
    - Incidence of mental retardation
      - Increased in de novo rearrangements
        - Reciprocal translocations – yes
        - Robertsonian translocations – no

    - Need to be concerned about uniparental disomy
      - Chromosome 15
        - Prader-Willi syndrome
        - Angelman syndrome
- Prenatal diagnosis of a Robertsonian translocation
  - 1/10,000 amniocentesis – de novo
  - Determine if parent has a rearrangement
  - Check paternity
  - Do uniparental studies
    - If translocation involves chromosome 14 or 15

2.5 Contiguous Gene Syndromes

2.5.1 Definition

- First defined by Schmickel in 1986
- Include microdeletion syndromes
- Sometimes defined as segmental aneusomy
- Deletion of contiguous stretch of DNA including multiple genes on a chromosome
- Include syndromes that are caused by involvement of genes that are located physically close to one another on a specific chromosome
- "Idiopathic" dysmorphic syndromes that were originally delineated clinically and subsequently found to have microscopic or submicroscopic chromosome abnormalities

2.5.2 Microdeletions

- These include syndromes that are clinically recognized
- Syndromes are usually sporadic
  - Cytogenetic abnormalities are sometimes detected
    - High-resolution chromosomes needed
    - FISH often needed
    - Some patients – demonstrate submicroscopic molecular deletions
  - Often specific features of syndrome demonstrate Mendelian inheritance
  - Involve multiple unrelated loci that are contiguous
- Types of classical contiguous deletions
  - Xp21 deletion
    - Duchenne muscular dystrophy
    - Chronic granulomatous disease
    - McLeod phenotype
    - Retinitis pigmentosa
    - Mental retardation
    - Glycerol kinase deficiency
    - Adrenal hypoplasia
    - Aland eye disease
  - Xp22.3 deletion
    - X-linked ichthyosis
    - Kallmann syndrome
    - Chondrodysplasia punctata
    - Mental retardation
    - Short stature
    - Ocular albinism
  - Xq21 deletion
    - Choroideremia
    - Mental retardation
    - Deafness
    - Cleft lip and palate
- Microdeletion syndromes
  - Prader-Willi syndrome
  - Angelman syndrome
  - Velocardiofacial syndrome
  - Williams syndrome
  - Miller-Dieker syndrome
  - Smith-Magenis syndrome
  - Langer-Giedion syndrome
  - Aniridia-Wilms tumor association

2.5.2.1 Prader-Willi Syndrome

- Neonatal hypotonia
- Feeding difficulties
- Genital hypoplasia
- Hyperphagia/obesity (1–2 years)
- Short stature, small hands and feet
- Hypopigmentation
- Mental retardation
- Involves a paternal deletion in 15q11–q13
  - 60–70% of patients
  - Maternal disomy
  - 30% of patients

2.5.2.2 Angelman Syndrome

- Microcephaly
- Macrosomia with prominent tongue

- Hypotonia
- Ataxic gait
- Excessive laughter
- Seizures
- Hypopigmentation
- Severe mental retardation
- Laboratory findings
  - Involves a maternal deletion 15q11–q13
    - 60–70% of patients
  - Paternal disomy
    - 3–5% of patients
  - UBE3A mutation
    - 10% of patients

2.5.2.3 Miller-Dieker Syndrome

- Type I lissencephaly
- Dysmorphic facies
- Visible deletions
  - 50% of patients
- FISH or molecular testing needed to detect all cases

2.5.2.4 Velocardiofacial Syndrome

- Chromosome 22q11.21 deletions
  - DiGeorge syndrome
    - Abnormalities in the development of the third and fourth branchial arches
    - Thymic hypoplasia
    - Parathyroid hypoplasia
    - Conotruncal cardiac defects
    - Facial dysmorphism
  - Velocardiofacial syndrome
    - Palatal defects
    - Hypoplastic alae nasi, long nose
    - Learning disorders or mental retardation
    - Congenital heart defects
    - Conotruncal defects
  - Both syndromes part of the same spectrum involving defects including:
    - Cardiac defects, abnormal facies, thymic hypoplasia, palatal abnormalities, hypocalcemia, and 22q11.21 deletion
    - Variable phenotypes associated with similar deletions
    - Ascertainment of patients
      - Congenital heart defects
      - Hypocalcemia
  - Deletion can be detected cytogenetically
    - FISH must always be used to be sure (Fig. 2.5)

2.5.2.5 Familial Cases

- About 8% of cases due to familial deletions

2.5.2.6 Langer-Giedion Syndrome

- Langer-Giedion syndrome
  - Trichorhinophalangeal syndrome
    - Sparse scalp hair
    - Bulbous/pear-shaped nose
    - Cone-shaped phalangeal epiphysis
  - Multiple cartilaginous exostoses
  - Mental retardation
  - Deletion – 8q24.1

2.5.2.7 Aniridia-Wilms Tumor Association

- Aniridia
- Wilms tumor
- Genitourinary dysplasia
- Mental retardation
- Deletion – 11p13

2.5.2.8 Smith-Magenis Syndrome

- Dysmorphic facial features
  - Brachycephaly
  - Flat midface
  - Prognathism
- Hoarse, deep voice
- Short broad hands
- Delayed speech
- Behavioral abnormalities
  - Self-destructive behavior
    - Seen in 75% of Smith-Magenis syndrome patients
    - More apparent in older children and adults
    - Onychotillomania
      - Pulling out of fingers and toenails
    - Polyembolokoilamania
      - Insertion of foreign objects
    - Head banging
    - Wrist biting
- Peripheral neuropathy
- Sleep disorders
- Mental retardation

- Deletion – 17p11.2
  - Most of these deletions can be detected with cytogenetics; FISH can be used to confirm.

2.5.2.9 William Syndrome

- Developmental disorder – central nervous system and vascular connective tissue
- Dysmorphic facial features
- Infantile hypercalcemia
- Congenital heart diseases
- Gregarious personality
- Premature aging of skin
- Mental retardation
- Deletion – 7q11.23
  - Involves the elastin gene
  - Not visible with high-resolution analysis
  - Not visible without FISH

2.5.2.10 Deletion 1p Syndrome

- Malformed ears with hearing loss
- Broad root of nose
- Thin lips
- Fifth finger clinodactyly
- Overlapping toes
- Dorsal hirsutism
- Chronic seizures
- Central nervous system abnormalities
- Congenital heart defect
- Seizures
- Hypotonia
- Abnormalities of the skull and brain
- Deeply set eyes
- Malformed, malpositioned digits
- Growth and psychomotor retardation
- Deletion involves loss of 1p36.3

2.5.3 Rare Deletions

- Rubinstein-Taybi syndrome
  - Dysmorphic facial features
  - Beaked nose
  - Prominent columella
  - Hypoplastic maxilla
  - Down-slanted palpebral fissures
  - Broad thumbs and first toes
  - Mental retardation
  - Some involve submicroscopic deletions of 16p13
    - Most are due to a mutation
- α-Thalassemia and MR
  - α-Thalassemia
    - (Hemoglobin H)
  - Facial dysmorphism
  - Mental retardation
  - Deletion – 16p13.3
- Alagille syndrome
  - Dysmorphic facial features
  - Chronic cholestasis
  - Vertebral arch defects
  - Peripheral pulmonic stenosis/hypoplasia
  - Autosomal dominant
  - Some involve deletions – 20p11.23 – 20p12.2
    - Most are due to mutations
- Greig cephalopolysyndactyly syndrome
  - Craniosynostosis
  - Polysyndactyly
  - Mental retardation occasionally
  - Deletion in 7p13

2.5.4 Duplication Syndromes

- Microduplication syndromes
  - Beckwith-Wiedemann syndrome
    - Macrosomia, macroglossia
    - Omphalocele
    - Hypoglycemia
    - Transverse earlobe creases
    - Hemihypertrophy
    - Advanced bone age
    - Increased risk of malignancy
    - A small percent of cases due to a paternal duplication of 11p15
    - About 10% of cases due to paternal disomy
  - Duplication 17p11.2p12
    - Hypotonia
    - Decreased reflexes
    - Club foot
    - CMT1A (PMP22)
    - Absence of REM sleep

- Cat eye syndrome
  - Coloboma of the iris
  - Anal atresia
  - Ear abnormalities
  - Cardiac defects
  - Mental retardation in some cases
  - Duplication of 22q

2.5.5 Microarray Syndromes

- Many new syndromes have been delineated by microarray technology
  - Most syndromes have been delineated by genotype first analysis
  - Most will not be clinically recognized and need to be delineated by genotype
- Many of the new syndromes (although not all) appear to be susceptibility syndromes
  - Probands have an alteration along with phenotype
  - Parents often have alteration but no phenotype
  - Alterations not seen in general population
  - Alterations appear to be necessary but not sufficient for phenotype
- May involve just one gene or only a small group of genes

2.5.6 Microarray Detected: Deletion and Duplications

- Numerous deletions and duplications have been detected (short list given below)
  - 1q21.1 microdeletion and microduplications
  - 1q21.1 TAR microdeletion
  - 15q13.3 microdeletions
  - 16p11.2 obesity microdeletion
  - 16p11.2 autism microdeletion and microduplication
  - 16p13.11 autism microdeletion and microduplication
  - 17q12 microdeletion and microduplication
  - 17q21.31 microdeletion and microduplication

2.5.6.1 1q21.1 Microdeletion and Microduplications

- Both syndromes very variable
- Both often inherited with parents with milder features
- 1q21.1 microdeletion
  - Mild to moderate mental retardation
  - Neuropsychiatric disorders
  - Microcephaly
  - Dysmorphic features
  - Cardiac abnormalities
  - Cataracts
- 1q21.1 microduplication
  - Mental retardation/autism spectrum disorder
  - Neuropsychiatric disorders
  - Macrocephaly
  - Dysmorphic features

2.5.6.2 1q21.1 TAR Microdeletion

- Variable syndrome
- Often inherited from parents with no phenotype
- Mode of inheritance is unknown
- 1q21.1 TAR microdeletion – thrombocytopenia-absent radius syndrome
  - Bilateral absence of the radii
  - Thrombocytopenia
  - Other anomalies of skeleton, heart, and genitourinary system can occur

2.5.6.3 15q13.3 Microdeletion

- Syndromes very variable with patients at increased risk for wide range of clinical manifestations
- Mostly associated with 1.5 Mb deletion
  - Syndrome with just deletion of CHRNA7 often seen but not as well clinically defined
  - 15q13.3 duplication – not well defined
- 15q13.3 microdeletion
  - Mild mental retardation in about 50% of patients
    - May be more severe
  - Neuropsychiatric disorders
    - Including autism and schizophrenia
  - Behavior problems
  - Seizures

2.5.6.4 16p13.11 Microdeletion and Microduplications

- Both syndromes very variable
- Both often inherited with parents with milder or no features
- Deletion much more established than duplication
  - Phenotypic abnormalities associated with duplication still controversial
  - Phenotypic features not always penetrant
- 16q13.11 microdeletion and microduplications
  - Neuropsychiatric disorders
    - Intellectual disabilities, autism, schizophrenia, epilepsy, and attention-deficit hyperactivity disorder
  - Dysmorphic features
  - Congenital heart defect
  - Multiple congenital anomalies
    - Skeletal manifestations
    - Craniosynostosis
    - Microcephaly

2.5.6.5 16p11.2 Microdeletion and Microduplications

- Both syndromes characterized by autism
- Both often inherited with parents with no autism
- Other syndromes involving this region
  - Microdeletion of 16p11.2-p12.2
    - Autism
    - Minor facial anomalies
    - Significant delay in speech development
    - Recurrent ear infections
  - Microdeletion of 16p11.2 involving SH2B1 gene
    - Developmental delay
    - Autism
    - Early-onset severe obesity

2.5.6.6 17q12 Microdeletion and Microduplications

- Both syndromes are very variable
- Both often inherited with parents with no or milder features
- 17q12 microdeletion
  - Variable developmental delay and cognitive impairment
  - Seizures
  - Renal abnormalities
  - Mature-onset diabetes
- 17q12 microduplications
  - Cognitive impairment
  - Behavior abnormalities
  - Epilepsy
  - Renal abnormalities

2.5.6.7 17q21.31 Microdeletion and Microduplications

- Both syndromes very variable
- Not susceptibility syndrome – parents do not have 17q21.31 alteration
- MAPT gene appears to be critical gene
- 17q21.31 microdeletion and microduplication
  - Mental retardation
  - Microcephaly
  - Facial dysmorphic features
  - Microduplication appears to have milder manifestation than microdeletion

2.6 Sex Chromosome Aberrations

- X chromosome
  - ~6% of total genomic DNA
    - 995 X-linked genes
- Y chromosome
  - ~1% of total genomic DNA
    - ~56 Y-linked genes
    - Sex-determining region of the Y chromosome (SRY) plays a critical role in determining gonadal sex

2.6.1 Sex Determination

- Presence of Y chromosome
  - Leads to maleness
    - Regardless of number of X chromosomes present
- Absence of Y chromosome
  - Results in female development
- Sex differentiation
  - Gonads undergo sexual differentiation at 6–7 weeks
    - If Y chromosome is present
      - Gonads – testis

  - Testis produces testosterone
  - Wolffian duct proliferation
  - Testis produces mullerian inhibitory substance
    - Regression of mullerian duct structures
- Originally thought sex was determined in humans as in *Drosophila*
  - By the ratio of X chromosomes to autosomal chromosomes
  - 1959 – determined that the Y chromosome determined sex
    - SRY
  - 1990 – testis-determining factor
    - Localized on the short arm of the Y chromosome
    - Close, but centromeric, to the pseudoautosomal region
    - SRY and testis-determining factor shown to be the same
    - SRY sequences found in XX males

2.6.1.1 Pseudoautosomal Region

- 46,XX males (Fig. 2.20)
  - In normal meiosis cro ssovers between the tip of the short arm of the X and Y chromosomes
    - Pseudoautosomal region (2.5 Mb)
    - SRY centromeric to pseudoautosomal region
    - If crossover distal (below SRY) leads to
      - XX male
      - XY female
- Androgen insensitivity syndrome
  - Formerly known as the testicular feminization syndrome
  - Frequency: 1/20,000
  - 46,XY karyotype
  - Phenotypic features
    - Normal female external genitalia
    - No uterus and blind vagina
    - Axillary and pubic hair – sparse
    - Testes present
      - In abdomen or inguinal canal
  - End-organ unresponsiveness
    - Testes secrete androgens normally
    - Absence of androgen receptors in appropriate target
    - Defect in X-linked androgen receptor locus
  - This is an X-linked disorder
    - It has associated counseling dilemmas
      - Individuals can develop gonadoblastoma
      - Phenotypic females but XY female

2.6.2 Sex Chromosome Abnormalities

2.6.2.1 Turner Syndrome

- Clinical description in 1938
- Chromosomal basis defined in 1959
- Common cause of short stature
- Most common cause of primary amenorrhea
- Incidence
  - 1 in 1,500–5,000 live-born females
  - 1 in 30–50 (3%) of all female conceptuses
  - 1 in 100 45,X conceptuses survive to delivery
- 2/3 of patients retain maternal X
- Recurrence risk is very low
- See Table 2.2 for karyotypes seen
- Height considerations
  - Tendency toward small size at birth
  - By 2–3 years most <5th percentile for height
  - Final height based on genetic (familial) potential as altered by genetic abnormality
  - Incidence of short stature, 100%
  - Mean height – 143 cm (pregrowth hormone [GH]) – 153 cm (post-GH and anabolic steroids)
- Major phenotypic findings
  - Short stature, 100%
  - Normal intelligence, 95%
  - Edema, hands and feet, 80%
  - Broad chest, hypoplastic nipples, >80%
  - Narrow maxilla, >80%
  - Unusual ears, >80%
  - Low posterior hairline, >80%
  - Micrognathia, >70%
  - Cubitus valgus, >70%
  - Nail dysplasia, >60%
  - Webbed neck, 50%

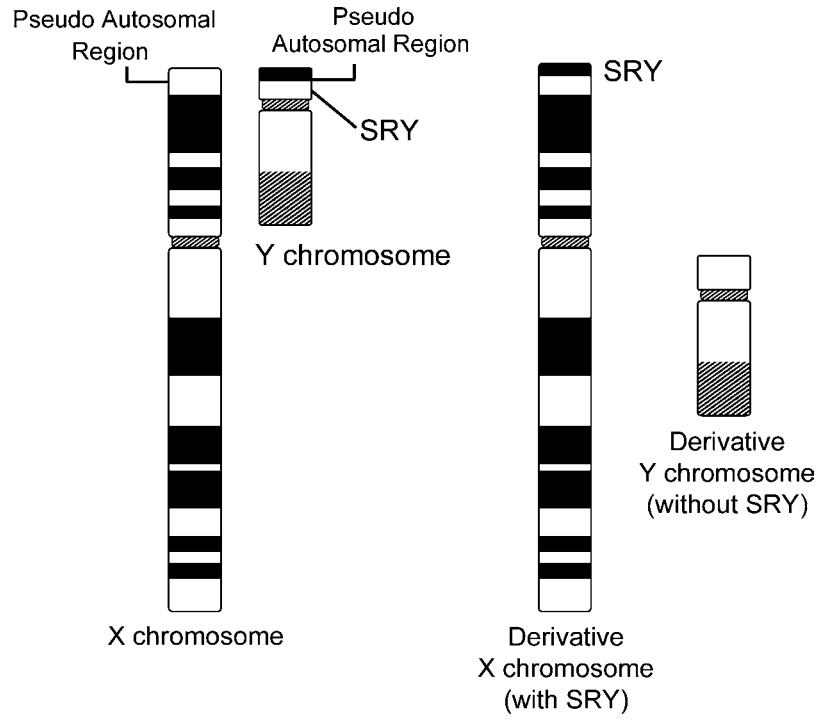


Fig. 2.20 A representation of the normal pairing of the X and Y chromosome along with a diagram showing abnormal crossing over leading to XX males and XY females

Table 2.2 Turner syndrome karyotypes

45,X	50%
Isochromosome X	12–20%
Mosaic	30–40%
45,X/46,XX	10–15%
45,X/46,XY	2–5%
Other	18–20%

- Internal anomalies
  - Gonadal dysgenesis, >90%
  - Tibial exostoses, >60%
  - Cardiac anomalies, 40–60%
  - Renal (mostly horseshoe kidney), >60%
  - Sensorineural hearing loss, >50%
- Less common include:
  - Scoliosis
  - Hip dysplasia
  - Vertebral abnormalities
  - Tendency to obesity
  - Hemangiomata
- Cardiac anomalies
  - The single cause of increased early mortality associated with Turner syndrome
    - Bicuspid aortic valve, 50%
    - Coarctation of aorta (includes hypoplastic arch), 20%
    - Aortic root dilation, 9%
    - Mitral valve prolapse, common
    - Hypoplastic left heart, rare
- Gonadal dysgenesis
  - >90% have abnormal gonads
  - Follicles develop normally by 2–3 months gestation, many still functional by birth.
  - Follicles nonfunctional by age 2 years
    - Menopause before menarche
  - 45,X/46,XY mosaics have an approximately 30% risk of developing gonadoblastoma
- Intelligence and behavior
  - Normal intelligence
    - Except ring X
      - When XIST (X inactive-specific transcript) is not present on the ring
  - Specific cognitive defects
  - Spatial perception (difficulty driving, sorting, and so on.)
  - Visuomotor coordination
  - Mathematics problems
  - Auditory verbal learning

  - Increased incidence of attention deficit disorder
- Management
  - Newborn period
    - Echocardiogram
    - Renal ultrasound
    - Rule out chromosomal mosaicism
    - Counsel family
      - Recurrence risk very low
  - Growth
    - Frequently cannot show GH deficiency
    - Start GH when height drops <3% ile for age (usually age 2–5 years)
    - Usually require 125% of dose used for GH deficiency
  - Does not appear to affect aortic root dilation
  - Low-dose anabolic steroid around start of puberty
  - Start estrogen at end of bone growth to improve secondary sex characteristics
    - Avoid early estrogen as it suppresses growth
  - Childhood
    - Monitor blood pressure regularly
    - Echocardiogram at 3–5 years looking for aortic root dilation
    - Annual thyroid function test
    - Consider cosmetic surgery for webbing, epicanthal folds
    - Ensure adequate calcium intake
    - Monitor hearing
- Fertility
  - As many as 8% of 45,X females may have spontaneous menses
    - Possibly higher in mosaics
    - At least 21 reported pregnancies in 13 45,X women
      - 8/21 (38%) fetal loss
      - 2/13 (15%) of liveborns malformed.
      - 10/13 (85%) "normal"
      - However, when mosaics are considered, numbers are less reassuring
    - At least 62 patients with 45,X or mosaicism – 135 pregnancies
      - SAB or stillborn
        - 44/135 (32%)
      - Anomalies (liveborns or late SABs)
        - 25/87 (29%)
      - Anomalies (liveborns only)
        - 23/80 (29%)
      - Anomalies included sex chromosome aneuploidy, trisomy 21, and single malformations
- Adult issues
  - Hypertension
  - Osteoporosis
  - Insulin resistance/diabetes mellitus
  - Obesity
  - Lymphedema
  - Inflammatory bowel disease
  - Hormone replacement

2.6.2.2 XXX Syndrome

- Incidence 1 in 1,000 live-born females
- No recognizable pattern of malformations
- Usually tall for their family (mean 172 cm = 5 ft 8 in.)
- Head size and IQ typically low for their family
- Delayed motor milestones, poor coordination
- Speech and language abnormalities common
- Behavior problems (depression, conduct disorders, and poor socialization)
- Normal sexual development with no apparent increased risk of abnormal offspring

2.6.2.3 XXXX Syndrome

- Variable phenotype – usually moderate mental retardation
- Other findings variable
  - Tall stature
  - Midface hypoplasia
  - Upslanting palpebral fissures
  - Hypertelorism
  - Epicanthal folds
  - Fifth digit clinodactyly
  - Narrow shoulder girdle
  - Amenorrhea

2.6.2.4 XXXXX (Penta X) Syndrome

- IUGR (intrauterine growth retardation)/poor growth/short stature
- Severe mental retardation
- Facial abnormalities
  - Hypertelorism
  - Upslanted palpebral fissures

  - Epicanthal folds
  - Malocclusion
- Congenital heart disease
- Multiple joint dislocations
- Small hands/single palmar crease/fifth digit clinodactyly

2.6.2.5 Klinefelter Syndrome

- Clinical description by Klinefelter in 1942
- Shown to have Barr body in 1956
- 47,XXY karyotype demonstrated in 1959
- Incidence 1 in 500–600 male newborns
- No evidence of increased recurrence risk
- 1/2 paternal meiosis I errors (no age affect noted)
  - Remainder maternal M I or M II errors
    - Age affect seen with M I errors only
- XXYY similar, but more mental retardation
- XXY – Neurocognitive profile
  - Full scale IQ tends to be low normal, occasional mild mental retardation
  - Performance IQ/verbal IQ difference greater than expected
  - Specific defects include impairments of
    - Verbal memory
    - Fluency
    - Speed of verbal processing
    - Overall language skills
  - Reading disabilities are common
  - Math skills usually normal
- Behavior profile
  - Early studies suggested that XXY boys were prone to criminal and antisocial behavior
  - Later, prospective studies (Denver, Toronto, Denmark, Winnipeg) demonstrated that criminality is not a part of the phenotype
- Behavior profile
  - XXY males tend to
    - Be withdrawn
    - Have difficulty with socialization skills
    - Have poor judgment
    - Have difficulty with adaptation to adult life
    - Have difficulty separating from family
    - Have poor self image
    - These boys are felt to be at increased risk for poor outcomes in dysfunctional families
- Growth
  - Normal birthweight and length
  - Tendency toward small stature (range 25–99%ile with mean at 75th %ile)
  - Normal weight and OFC (occipital frontal circumference)
  - Upper/lower segment ratio decreased
  - Occasionally marfanoid habitus
  - Arm span may be >3 cm greater than height
- Physical findings
  - Small testes – postpuberty
  - Normal facies, sparse facial hair
  - Congenital malformations in 18% without specific pattern
  - Gynecomastia
  - Scoliosis
  - Low muscle tone
  - Tend to enter puberty normally, but testicular insufficiency develops soon after
  - Phallus usually normal
  - Untreated hypoandrogenization leads to "eunuchoid" habitus
  - Varicose veins and hypostatic leg ulceration
- Management
  - At risk for developmental delay, learning disability
  - Slightly increased risk of mediastinal germ cell tumors
  - Risk of breast cancer 66 times normal males (approaches rate for females)
  - Testosterone from early puberty (adult levels usually 1/2 normal)
    - Improved body image
    - Improved strength and endurance
    - Improved mood and concentration
- Fertility
  - Extremely rare to be fertile
  - Fibrosis of seminiferous tubules
  - Exogenous testosterone probably does not improve fertility
  - In the few reported cases of successful reproduction, the offspring have been normal

2.6.2.6 XXXXY Syndrome

- Low birthweight, poor growth, microcephaly
- Hypotonia
- Severe hypogenitalism
- Severe mental retardation
- Some phenotypic overlap with trisomy 21, especially in newborn period

2.6.2.7 XXXY Syndrome

- Very rare Klinefelter variant
- Mental retardation is the rule
- Other findings
  - Hypotonia
  - Facies – low nasal bridge, epicanthal folds, upslanted palpebral fissures, abnormal ears
- Limited abduction at elbows
- Small phallus and testes
- Fifth digit clinodactyly
- Flat feet

2.6.2.8 XYY Syndrome

- Incidence 1 in 1,000 live-born males
- No physical phenotype
- No increased incidence of congenital malformations
- IQ tends to be in low normal range
- Tend to be tall with normal proportions.
- Behavior is impulsive, with temper tantrums and poor emotional control
- Usually normal fertility
- Majority of offspring are normal
- Few reports of XYY offspring

2.6.3 Lyon Hypothesis

- X inactivation
  - Males
    - Hemizygous for the X chromosome
    - Only a single copy of the X chromosome
  - Female
    - Two X chromosomes
    - Mean amounts of gene products of X-linked genes; same in females as males
      - Why?
  - Mechanism of dosage compensation – Lyon hypothesis
    - In somatic cells, X inactivation occurs early in embryonic life
    - Inactivation is random
      - Either the paternal or maternal X chromosome
    - Inactivation is complete
    - Inactivation is permanent and clonally propagated
  - Lyon hypothesis – exceptions
    - Inactivation not always random
      - With structurally abnormal X nonrandom inactivation seen
    - Inactivation not complete
      - A number of genes known to escape activity
      - May be up to 15% of all X-linked genes escape inactivation in some way
    - Inactivation reversible in development of germ cells
  - Lyon hypothesis – evidence
    - Genetic studies have helped to confirm the Lyon hypothesis
      - Tortoiseshell mouse
      - Calico cats (Fig. 2.21)
    - Cytologic evidence
    - Barr body
      - Darkly staining chromatin body
      - Number of barr bodies = number of X chromosomes − 1

2.6.3.1 Significance

- Lyon hypothesis – significance
  - Explains manifestations of X-linked disorders
  - Explains variability of clinical manifestations in females
  - Explains difficulty in biochemical carrier detection in female carriers
- Mechanism of X inactivation
  - Involves altered chromatin structure
  - Differential methylation of DNA
  - CpG islands in silenced genes – methylated
  - X inactivation center At Xq13.2
    - Cis for X inactivation
  - XIST
    - Uniquely expressed from the inactive, but not the active X

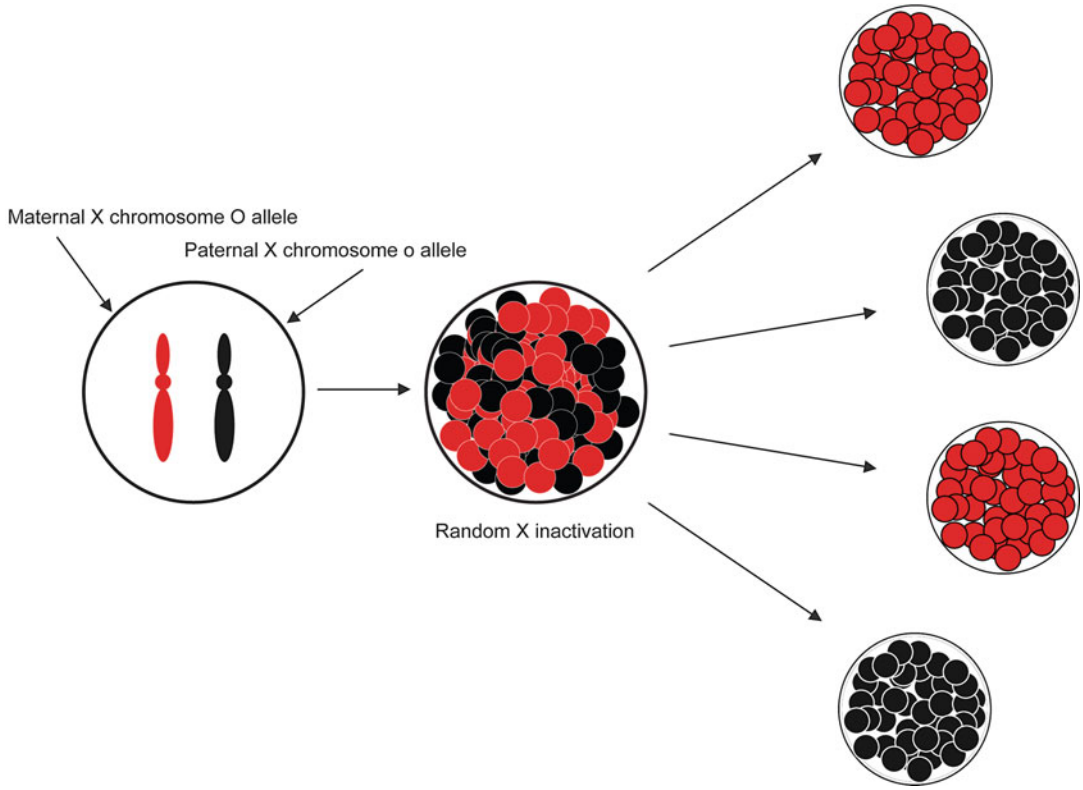


Fig. 2.21 An illustration showing a representation of random X inactivation leading to *red* and *black patterns* as seen in a calico cat

Further Reading

Ferguson-Smith MA. Cytogenetics and the evolution of medical genetics. Genet Med. 2008;10:553–559.

Gardner RJM, Sutherland GR, Shaffer LG. Chromosome abnormalities and genetic counseling. 4th ed. New York: Oxford University Press; 2011.

Hassold T, Hunt P. Maternal age and chromosomally abnormal pregnancies: what we know and what we wish we knew. Curr Opin Pediatr. 2009;21:703–708.

Jones KL. Smith's recognizable patterns of human malformation. 6th ed. Philadelphia: Elsevier Saunders; 2006. p. 3–81.

Keagle MK, Gersen SL. Principles of clinical cytogenetics. 3rd ed. Totowa: Humana press; 2011.

Lupski JR, Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet. 2005;1:E49.

Manning M, Hudgins L, Professional Practice and Guidelines Committee. Array-based technology and recommendations for utilization in medical genetics practice for detection of chromosomal abnormalities. Genet Med. 2010;12:742–745.

Morey C, Avner P. The demoiselle of X-inactivation: 50 years old and as trendy and mesmerising as ever. PLoS Genet. 2011;7:e1002212.

National Center for Biotechnology Information (US). Genes and Disease [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 1998. http://www.ncbi.nlm.nih.gov/books/NBK22183/

Nussbaum RL, McInnes RR, Willard HF. Thompson and Thompson genetics in medicine. 7th ed. Philadelphia: W.B. Saunders; 2007.

Savage MS, Mourad MJ, Wapner RJ. Evolving applications of microarray analysis in prenatal diagnosis. Curr Opin Obstet Gynecol. 2011;23:103–108.

Shaffer LG, Slovak ML, Campbell LJ, eds. ISCN 2009: an international system for human cytogenetic nomenclature. Switzerland: S. Karger AG; 2009.

Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. Annu Rev Med. 2010;61:437–455.

Conceptual Evolution in Cancer Biology

3

Shaobo Zhang, Darrell D. Davidson, Rodolfo Montironi, Antonio Lopez-Beltran, Gregory T. MacLennan, and Liang Cheng

Contents

3.1 Stem Cells 78
3.1.1 Overview 78
3.1.2 Cancer Models 79
3.1.3 Cancer Stem Cells 79
3.1.4 Embryonic Stem Cells 83
3.1.5 Fetal Stem Cells 83
3.1.6 Adult Stem Cells (Somatic Stem Cells) 84
3.1.7 Comparison Between Normal Stem Cells and Cancer Stem Cells 84

3.2 Epigenetics 84
3.2.1 Overview 84
3.2.2 Molecular Basis of Epigenetics 85
3.2.3 Gene Silencing 85
3.2.4 Other Epigenetic Processes 85

3.3 DNA Methylation 87
3.3.1 Overview 87
3.3.2 DNA Methylation Enzymes 87
3.3.3 DNA Methylation and Genetic Regulation 88
3.3.4 DNA Methylation and Cancer 88
3.3.5 Methylation and Potential Clinical Interventions 88
3.3.6 Methods for DNA Methylation Analysis 89

3.4 Microsatellite Instability 89
3.4.1 Overview 89
3.4.2 Definition and Mechanisms of Microsatellite Instability 89
3.4.3 Bethesda Panel for Microsatellite Instability 91
3.4.4 Microsatellite Instability in Hereditary Nonpolyposis Colorectal Cancer 91
3.4.5 Methods for Microsatellite Instability Analysis 92

3.5 Chromosomal Instability 92
3.5.1 Overview 92
3.5.2 Mechanisms of Chromosomal Instability 92
3.5.3 Clinical Implications of Chromosomal Instability 93
3.5.4 Chromosomal Instability Syndromes 94
3.5.5 Methods for Chromosomal Instability Analysis 94

3.6 Gene Imprinting 95
3.6.1 Overview 95
3.6.2 Regulation of Gene Imprinting 96
3.6.3 *H19* Gene Imprinting 96
3.6.4 Imprinting and Disease 97

3.7 MicroRNA 98
3.7.1 Overview 98
3.7.2 Clinical Implications 100
3.7.3 Methods for MicroRNA Discovery 100

3.8 RNA Interference 100
3.8.1 Overview 100
3.8.2 Biologic and Clinical Implications of Small Interfering RNA 101

3.9 Telomere 102
3.9.1 Overview 102
3.9.2 Telomere Structure and Maintenance 103

S. Zhang, MD (✉) • D.D. Davidson, MD, PhD • L. Cheng, MD
Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

R. Montironi, MD
Institute of Pathological Anatomy and Histopathology, Polytechnic University of the Marche Region (Ancona), School of Medicine, United Hospitals, Ancona, Italy

A. Lopez-Beltran, MD, PhD
Unit of Anatomic Pathology, Department of Surgery, Cordoba University School of Medicine, Cordoba, Spain

G.T. MacLennan, MD
Department of Pathology, Case Western Reserve University, Cleveland, OH, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_4,

3.9.3 Cell Senescence and Telomere Shortening 104
3.9.4 Telomere Shortening and Cancer 105
3.9.5 Telomerase Summary 105
3.9.6 Telomerase and Cancer 107
3.9.7 Diagnostic and Therapeutic Implications of Telomerase 107
3.9.8 Methods for Telomere Analysis 108

Further Reading 108

3.1 Stem Cells

3.1.1 Overview

- Stem cells are undifferentiated cells, which are the foundation of every organ, tissue, and cell in the body
- Stem cells come from both embryos formed during the blastocyst phase of embryological development (embryonic stem cells) and from adult tissue (adult stem cells)
- Self-renewal is the property that distinguishes stem cells from most other terminally differentiated nonstem cells
- Stem cells also have the potential to differentiate into other types of cells
- Potency refers to the ability of stem cells to differentiate into different cell types. Stem cells could be categorized according their potency
  - Totipotent stem cell: a cell that can differentiate into all possible cell types and can form an entire organism
  - Pluripotent stem cell: more differentiated embryonic stem cell that has the ability to differentiate into all cell types from each of the three embryonic layers
  - Multipotent stem cell: cells with the potential to generate to cells from multiple but a limited number of lineages. Hematopoietic stem cells are multipotent cells that can produce all cell types of the blood
  - Oligopotent: cells with ability to differentiate into a few cell types, such as vascular stem cells differentiating into endothelial or smooth muscle cells
  - Unipotent: cells with ability to produce only one cell type, but maintaining the property of self-renewal. A synonym for unipotent cell is precursor cell
- Stem cells have unique properties
  - They renew themselves through division to maintain a population of cells with the same properties as the original cells
  - Stem cells propagate by asymmetric differentiation allowing a stem cell to generate one stem cell and one differentiated progenitor cell (Table 3.1, Fig. 3.1)

Table 3.1 Characteristics of stem cells and progenitor cells

	Stem cell	Progenitor cell
Self-renewal	Unlimited	Can be limited
Plasticity	Pluripotent	Unipotent or multipotent

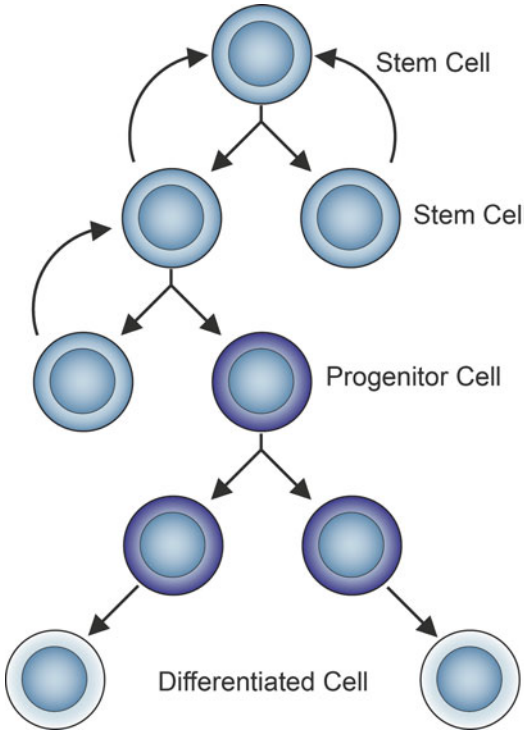


Fig. 3.1 Stem cells may renew themselves through division to maintain a population of cells with the same properties as the original cells. They also may propagate with asymmetric differentiation, depending on environmental factors and on whether the organism needs to expand or maintain a constant stem cell population. A cycle with asymmetric differentiation yields one stem cell and one differentiated progenitor cell. The progenitor cells can further differentiate into various committed cell types

- Stem cells may be classified into different groups according to their origin
- Stem cells can be used to generate healthy and functional specialized cells to replace diseased or dysfunctional cells
- Cancer stem cells (CSC) are the origin of cancer

3.1.2 Cancer Models

3.1.2.1 Traditional Cancer Model

- *All* tumor cells can form new tumors and are therefore equally tumorigenic
- Unregulated growth is due to serial acquisition of genetic events leading to the expression of genes that promote cell proliferation with concomitant silencing of growth inhibitory genes and blunting of cell death genes
- Cancer is a proliferative disorder in general cells
- Any cell type can be targeted in carcinogenesis

3.1.2.2 Modern Cancer Model

- Only a minority of tumor cells can form new tumors
- Unregulated cell growth in tumors results from disruption of stem cell self-renewal regulatory mechanisms
- Cancer is a stem cell regulatory disorder and not a simple augmentation of proliferation signals
- Only stem cells or progenitor cells are carcinogenic targets since they possess the self-renewal pathway

3.1.3 Cancer Stem Cells

3.1.3.1 Definition and Properties of Cancer Stem Cells

- Cancer stem cells (CSC) are cancer cells and the cellular source of cancer, characterized by critical stem cell properties of self-renewal and asymmetric differentiation
- CSCs maintain themselves through self-renewal and propagate by asymmetric differentiation
- The origin of a CSC is not clear
  - The CSC may be a normal stem cell transformed by genetic alterations
  - Some CSCs may arise from mutation of committed progenitor cells
  - CSCs could even derive from a nonstem somatic cell, which acquires stem-like attributes by genetic mutations or epigenetic dysregulation

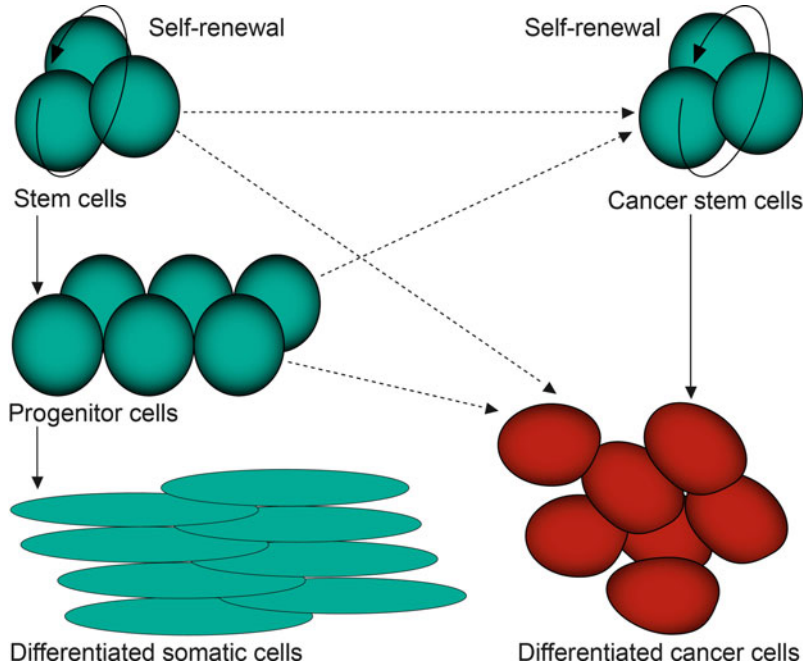


Fig. 3.2 Cancer stem cells (CSCs) are the cellular source of cancer and have the characteristics of stem cells. A CSC may result from transformation of either a normal stem cell or a progenitor cell through genomic mutation. CSCs proliferate and expand the population of malignant cells related to the original malignant clone. Dashed lines indicated the possible pathways of carcinogenesis

- The CSC is responsible for a cancer's growth, spread, and relapse
- A CSC is highly drug resistant and radiation tolerant
- CSCs possess the character both of stem cells and of cancer cells
  - CSCs can self-renew and differentiate into the heterogeneous clones that make up a malignant neoplasm
  - CSCs can be maintained in culture in an undifferentiated state
  - CSCs proliferate and expand the population of malignant cells related to the original malignant clone
  - CSCs initiate tumor growth after xenotransplantation into athymic mice
- Cancer cells derived from a single CSC have the clonal characteristics of that stem cell
  - The isolated CSC differentiates into new cancers that are phenotypically indistinguishable from the original tumor
  - Derivation of a CSC from a committed progenitor cell results from reactivation of self-renewal mechanisms
  - CSC transformation is the result of genetic and epigenetic changes in oncogenes and tumor suppressor genes (Fig. 3.2)

3.1.3.2 Cancer Stem Cell Pathways

- Normal or somatic stem cells may be transformed into CSCs through dysregulation of self-renewal and differentiation pathways, usually resulting in wanton proliferation and autonomous differentiation
  - BMI1 (polycomb group transcriptional repressor BMI1 for "B lymphoma Mo-MLV insertion region 1 homolog") is a transcriptional repressor activated in lymphoma, in neural stem cells, and in CSC of pediatric brain tumors
    - BMI1 is an oncogene regulating p16 and p19, important cell cycle inhibitor proteins in neural crest and hematopoietic stem cells
    - BMI1 expression leads to repression of cell cycle regulators or activation of stem cell maintenance genes, which play a role in stem cell fate decisions including self-renewal, senescence, and differentiation
    - It has been shown that BMI1 is necessary for efficient self-renewing of adult hematopoietic stem cells and neural stem cells
    - Loss of BMI1 leads to impaired DNA repair and radiation sensitivity, and its

absence in mouse embryos causes severe defects in hematopoiesis, skeletal development, neurological function, and cerebellar maturation

- NOTCH pathway regulates proliferation in hematopoietic, neural, and mammary stem cells
  - NOTCH is a transmembrane heterooligomer with a large extracellular portion and a small intracellular region
  - It signals by direct cell–cell contact (juxtacrine) with progenitor cells important in neural development and angiogenesis
  - NOTCH signaling promotes proliferation during neurogenesis
  - Mammals possess four NOTCH receptors, NOTCH1, NOTCH2, NOTCH3, and NOTCH4
  - Components of the NOTCH pathway may act as oncogenes in mammary and hematopoietic neoplasia
  - NOTCH ligand binding upregulates genes related to cell proliferation
  - Its dysregulation is implicated not only in CSCs but also in tetralogy of Fallot, multiple sclerosis, and Alagille syndrome
- Sonic hedgehog homolog (SHH) is the ligand of the hedgehog signaling pathway, which regulates vertebrate organogenesis and embryonic morphogenesis
- Sonic hedgehog and WNT developmental pathways are so intimately interrelated that defects in one almost always associate with changes in the other pathway
- WNT signaling pathway is a complex network of proteins involved in embryogenesis, cancer, and normal adult physiologic processes. The name WNT was a combination of *WG* and *INT*, two genes associated with fruit fly development and murine carcinogenesis
  - Both sonic hedgehog and WNT pathways are required to sustain tumor growth and thus are commonly hyperactivated in tumors
  - Cyclopamine is a sonic hedgehog blocker, which inhibits the hedgehog signaling pathway, causing cyclopia in lambs born in 1957 on a farm in Idaho where wild corn contained this alkaloid
- Both genetic and epigenetic changes transform normal stem cells to CSC

3.1.3.3 Cancer Stem Cell Markers

- Each CSC retains markers of the stem cell from which it derives
- CSC markers are not related to carcinogenesis per se
- CSCs are more accurately defined by functional dysregulation than by specific phenotypic markers (Table 3.2)
  - Hoechst 33342 dye efflux defines a small and homogeneous population of cells consistent with stem cells, since one of the most common properties of stem cells is drug transporter

3.1.3.4 Cancer Stem Cell Functional Profiling

- CSCs produce all the malignant cells in both primary and metastatic tumors, including self-renewing CSCs and terminally differentiated tumor cells
- CSCs may acquire features associated with tumor progression
  - CSCs may become genetically unstable
  - CSCs resist several chemotherapeutic agents through efflux of the drugs by drug transporter proteins
  - CSCs are relatively radiation resistant
  - CSCs compose a small reservoir of drug/radiation-resistant cells, which are responsible for tumor relapse after chemotherapy or radiation therapy
  - CSCs are founder cells that give rise to distant metastases
  - CSCs express angiogenesis factors, such as vascular endothelial growth factor and CXCR4 (Chemokine [c-x-c motif] receptor 4), which enhance angiogenesis
- Niche is the microenvironment in which stem cell populations become established

Table 3.2 Selected CSC markers

Tumor	Tissue-specific marker
Breast stem cell	$CD44^+CD24^{-/low}$, $CD29^{hi}CD24^+$ ESA^+, ABCG2
Neuroectodermal stem cell	Nestin, SOX2
Brain stem cell	CD133
Prostate CSC	CD133, CD34, $CD24^+CD49f^-$, SCA-1
Hematopoietic stem cell	$CD34^+CD38^-$, CD133, ABCG2
Leukemia stem cell	$CD44^+$, OCT4
Testicular germinal stem cell	OCT4, STELLAR, NANOG, GDF3 SSEA

- Niche is necessary for CSC to form a cancer (Fig. 3.3)
- CSC niche provides the CSC with supports needed for self-renewal and differentiation
- Niche constitutes a basic unit of tissue physiology, integrating signals that mediate CSC maintenance, and response to environmental stress such as ischemia and therapeutic interventions
- The niche must have both anatomic and functional dimensions rather than a simple location where the CSCs were found
  - Niche aids the CSC by juxtacrine, paracrine, and probably autocrine mechanisms

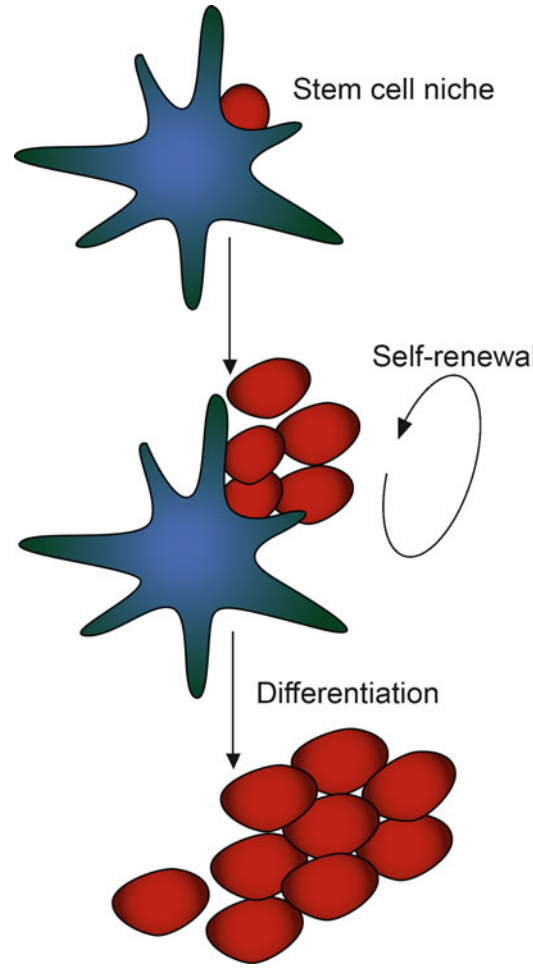


Fig. 3.3 Niche is a specialized microenvironment that provides cancer stem cells (CSCs) with the support needed for self-renewal. Niche is necessary for the CSC population to expand and form a cancer. In this specialized microenvironment, CSCs possess self-renewal and differentiation. Signals from the stem cell niche may determine whether stem cells differentiate symmetrically or expand asymmetrically

3.1.3.5 Clonal Expansion of Cancer Stem Cells

- CSCs represent 2–3% of the diverse tumor cell populations in a malignancy
- CSCs can differentiate into heterogeneous cancer cell lines
- Asymmetric differentiation is a characteristic of CSC replication, through which a CSC divides into a new CSC and a progenitor cell
  - The progenitor cell loses self-renewal capacity but gains the ability to proliferate and further differentiate into the various cell types of the malignancy
  - The daughter CSC retains the ability to self-renew by asymmetrical differentiation
- Mutations accumulated in the self-renewing population allow progression toward increasingly malignant behavior
- Progenitor cells divide into differentiated cancer cells but can also acquire new malignant features and polyclonality by mutation
- Progenitor cells sometimes regain self-renewal ability due to new mutations in self-renewal and differentiation pathways (Fig. 3.2)

3.1.3.6 Clinical Implications of Cancer Stem Cells

- The growth of tumors depends upon the presence of neoplastic stem cells
- CSC-targeting therapy aims to deplete the CSC pool to cure the cancer
  - Surface molecule-targeting therapy uses a cytotoxic drug conjugated to antibodies against CSC surface molecules; however, similar surface molecules exist on both CSC and normal stem cells
  - Another type of therapy targets tumor-specific oncoproteins expressed by CSCs, such as *HER2*, *RAS*, *KIT*, *BCR/ABL*, and *PML/RARA*
  - Other therapies would target CSC pathway downstream signal molecules, such as BMI1, NOTCH, WNT, and sonic hedgehog agents
- Challenges for CSC-targeted therapy
  - CSC-targeted therapy should selectively eradicate CSCs without harming normal stem cells
  - Such therapies call for functional assays for CSCs to monitor the effectiveness of the targeted treatment
  - There is a need for more knowledge of mechanisms by which CSCs evade or become resistant to targeted therapy

3.1.4 Embryonic Stem Cells

- Embryonic stem cells (ESCs) are the stem cells derived from the inner cell mass of a blastocyst 4–5 days after fertilization
- ESCs are pluripotent and could potentially differentiate into cell types from all three primary germ layers, ectoderm, endoderm, and mesoderm
  - ESCs are potential sources for tissue replacement after injury or disease because of pluripotency
  - ESCs may also remain in an undifferentiated state and can divide indefinitely
  - ESCs are defined by the presence of several transcription factors and cell surface molecules, such as *NANOG, OCT4*, and *SOX2*, which ensure the maintenance of pluripotency and suppression of genes that lead to differentiation
- ESC regenerative therapies have been difficult to develop because of neoplastic transformation risk in cells with highly activated proliferation and differentiation pathways
- ESCs are similar to adult stem cells with added capacity for pluripotency and high regenerative capacity

3.1.5 Fetal Stem Cells

- Fetal stem cell (FSC) populations are prominent in the fetus but dwindle with completion of organogenesis
- FSCs are easily isolated and induced to differentiate into mature cell types
  - FSCs have the ability to self-renew
  - FSCs readily differentiate in vitro and retain their cell lineage properties when transplanted in vivo
  - FSCs express the adhesion molecule, AA4, in three major cell types
    - Endothelial cells
    - Aorta-associated hematopoietic clusters
    - Liver-associated hematopoietic cells
- FSCs can be isolated from different origins
  - From fetal blood
  - From hematopoietic organs in early pregnancy
  - From a variety of somatic organs in the fetus
  - From amniotic fluid and placenta throughout gestation
- FSCs proliferate more rapidly than adult stem cells
- FSCs may thus represent an intermediate cell type in the current debate focusing on dichotomized adult vs ESCs and may prove advantageous as a source for regenerative and replacement therapies
- FSCs have properties similar to adult stem cells, with higher proliferative rates and greater plasticity (see Sects. 3.1, 3.1.1)

3.1.6 Adult Stem Cells (Somatic Stem Cells)

- Adult stem cells are found in human tissues throughout the body after fetal development
- Adult stem cells are a tiny population of undifferentiated reserve cells in organs with regenerative capacity, dividing to replenish dying cells and to replace damaged tissues
- Adult stem cells are derived from adult tissue samples without destroying embryos and are thus less controversial than ESCs
- Adult stem cells generate all the functional cell types of the organ from which they originate through self-renewal and differentiation
- Adult stem cells from different organs possess different tissue-specific markers
- Adult stem cells are instrumental for organogenesis, tissue homeostasis, and carcinogenesis
- Adult stem cells share the common properties of all stem cells
  - Self-renewal allows adult stem cells to experience numerous cycles of asymmetric cell division, reproducing themselves and propagating progenitor cells with differentiation plasticity
  - Adult stem cells have multidifferentiation potential (multipotency) through which they may differentiate into progeny of several distinct cell types
  - Tissue-specific adult stem cells can generate a spectrum of cell types of other tissues. This is the basis of metaplasia and transdifferentiation, in which the metaplastic cells cannot revert to the original cell type
  - Adult stem cells have the property of multidrug resistance. They are capable of actively pumping a diversity of organic molecules out of the cell
  - Adult stem cells express ATP-binding cassette transporter (ABC-transporter) proteins that move a diversity of organic molecules out of the cell

3.1.7 Comparison Between Normal Stem Cells and Cancer Stem Cells

- Similarities between normal and CSC
  - Normal and cancer cells share the ability to self-renew and to differentiate
  - CSCs share remarkable phenotypic and functional similarities with normal stem cells
  - Proposed markers of normal stem cells are also expressed in CSCs
  - Normal and CSCs are relatively few in number, compared to terminally differentiated cells
- Differences between normal and CSCs
  - Self-renewal is tightly regulated in normal stem cells but is dysregulated in CSCs
  - CSC differentiate to produce malignant cells, but normal stem cells differentiate to produce mature tissues
  - One major difference between normal stem cells and CSCs is the degree of dependence on the stem cell niche

3.2 Epigenetics

3.2.1 Overview

- Epigenetics is a complex phenomenon controlling the switching "on" and "off" of certain genes
- Epigenetics refers to stable changes in gene expression that are not due to changes in DNA structure
- Epigenetic changes are preserved when cells divide
- Specific epigenetic processes of interest include gene silencing, paramutation, bookmarking, imprinting, X chromosome inactivation, position effect, reprogramming, transvection, maternal effect, and histone modification
- Epigenetic processes are associate with events in an individual organism's development
- Some epigenetic changes are inherited

- Some genetic mutations may influence offspring through changes in epigenetic programming

3.2.2 Molecular Basis of Epigenetics

- Epigenetic processes often involve chemical modification of DNA, most importantly through methylation at CpG sites, converting cytosine to 5-methylcytosine
- Epigenetic processes also modify chromatin proteins associated with DNA
- DNA methylation and chromatin remodeling are the major processes of epigenetics
- Epigenetic regulation of gene activity may also involve messenger RNA (mRNA) methylation or small noncoding RNA molecules
- Epigenetic controls can also be mediated by the production of different spliced forms of mRNA
- Epigenetic processes can be modulated by formation of double-stranded RNA (miRNA)

3.2.3 Gene Silencing

- Gene silencing describes the phenomenon of "switching off" genes by any mechanism other than genetic mutation (Fig. 3.4)
  - Transcription level gene silencing
    - DNA methylation (see Sect. 3.3)
    - Histone modifications include
      - Acetylation, substitution of an acetyl group for an active hydrogen atom of a hydroxyl group. As a consequence, the condensed chromatin is transformed into a transiently relaxed structure, which allows genes to be transcribed. Deacetylation leads to gene silencing
      - Methylation is the transfer of methyl groups to histone arginine or lysine amino acids. Histones which are methylated on certain residues act epigenetically to repress or activate gene expression
      - Histone phosphorylation is crucial for chromosome condensation and cell cycle progression; the combination of phosphorylation at serine-10 and acetylation of lysine-14 enables transcription of genes as a consequence of various cell signaling events
    - These modifications signal whether a region of the chromosome is condensed into silent heterochromatin or remains active for transcription as euchromatin
  - Posttranscription level gene silencing involves RNA interference (RNAi) machinery to stop mRNA translation and to induce mRNA degradation (see Sect. 3.7)
- Gene silencing protects the genome from transposons and viruses. Thus, it represents an intracellular system protecting from infectious DNA particles

3.2.4 Other Epigenetic Processes

- Paramutation is an interaction between two alleles at a single locus, resulting in a heritable change of one of the alleles. This phenomenon, which violates Mendel's law of independent transmission of traits, allows for varying penetrance or continuous variation of a monogenetic trait, especially in plants
- Bookmarking is an epigenetic mechanism for transmitting cellular memory of gene expression pattern to daughter cells. This biologic phenomenon causes daughter cells to maintain the phenotype of a cell lineage
- Imprinting causes a subset of all the genes to be expressed according to their parental origin, whether maternal or paternal
- X chromosome inactivation is a DNA methylation process causing one of the two X chromosomes in a female mammal to be randomly inactivated. This compensates for the dose doubling of X chromosome genes in females

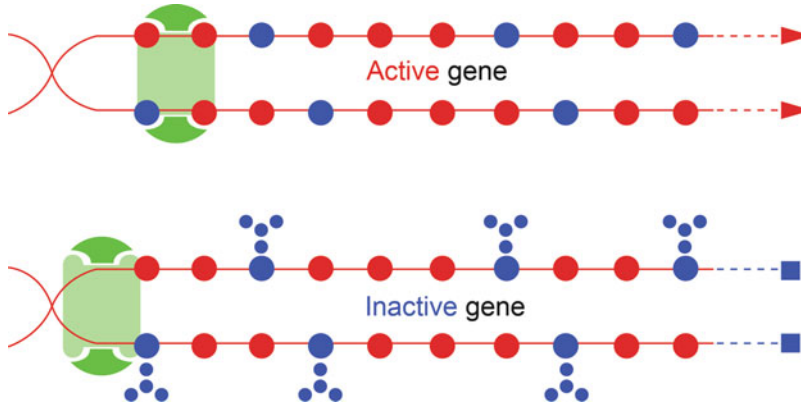


Fig. 3.4 Epigenetic effects are stable changes in gene expression without genomic mutation. One important epigenetic process is gene silencing, the phenomenon of gene "switching off" by any mechanism other than genetic mutation. DNA methylation is the most frequent mechanism for gene silencing. This process impedes transcription by adding a methyl group (purple dots) to cytosine (larger blue dots) in the promoter region of a gene to eliminate its expression

- Position effect is the expression effect of gene location in a chromosome. Expression is often changed by translocation without deletion, amplification, or mutation
- Reprogramming refers to demethylation and reestablishment of DNA methylation during mammalian fetal development
- Transvection is an epigenetic "paramutation-like" interaction between an allele on one chromosome and the corresponding allele on the homologous chromosome. This interaction may lead to either gene activation or repression in animals that is not always passed on through gametes
- Maternal effect is the phenomenon in which the genotype of the mother is expressed unaltered in the phenotype of the offspring. This generally occurs when maternal mRNA or mitochondrial DNA influences early stages of embryonic development
- Pathologic epigenetic changes are nonsequence-based alterations, which interrupt or stimulate gene function and cause disease. Examples include
  - Hypermethylation of promoter regions for tumor suppressor genes
  - Hypomethylation of promoter switches on oncogenes
  - Histone modification causing heterochromatin in regions of tumor suppressor gene loci
  - Loss of gene imprinting is associated with many pediatric tumors. Loss of *IGF2* imprinting accounts for half of all Wilms tumors
- Epigenetic carcinogens are not mutagenic but result in increased incidence of tumors. Many of these have previously been classified as chemical carcinogen promoters. Examples include diethylstilbestrol, arsenite, hexachlorobenzene, and nickel compounds
- Teratogens influence fetal development by epigenetic mechanisms
  - Reversibility of epigenetic changes exposes pathologic epigenetic changes to possible cure by targeted therapy
- Examples of "epigenetic therapy" include
  - DNA methyltransferase (DNMT) inhibitors
    - Nucleoside analogs with a reactive modified cytosine ring that is incorporated into DNA or RNA where the modified base binds to DNMT and leads to cell death (Goffin et al., Ann Oncol 2002)

- Nonnucleoside analogs that inhibit DNA methylation by binding to the catalytic region of the methyltransferase enzyme

– Histone deacetylase inhibitors
  - Short-chain fatty acids that inhibit cell growth and induce apoptosis
  - Hydroxamic acids (R-CONHOH) are potent inhibitors of histone deacetylase. These induce cell differentiation and inhibit proliferation
  - Cyclic tetrapeptides are irreversible histone deacetylase inhibitors
  - Benzamides are antitumor agents that bind to the catalytic region of histone deacetylase
– Epigenetic therapies may be applicable whenever aberrant heterochromatin regions in tumor chromosomes involve tumor suppressor genes or other critical regulatory genes

3.3 DNA Methylation

3.3.1 Overview

- DNA methylation is an epigenetic chemical modification of DNA which impedes transcription of a gene
  – DNA methylation adds a methyl group to the 5 position of the cytosine, which reduces gene expression of the gene involved
  – DNA methylation can be inherited through cell division
  – DNA methylation plays a crucial role in carcinogenesis
- Methyl groups are added to the number 5 carbons of cytosine pyrimidines in a CpG island in the promoter region of a gene (Fig. 3.5)
- CpG islands are DNA regions with frequent cytosine–guanosine dinucleotide pairs in the 5′–3′ direction. CpG islands mark the start of about 50% of human exons
- Developmental gene activation often correlates with a loss of DNA methylation at tissue-specific genes in the relevant somatic lineages
  – Only 1% of DNA bases are subject to DNA methylation
  – In adult somatic tissues, DNA methylation typically occurs only in CpG dinucleotide pairs
  – Non-CpG methylation is prevalent in ESCs
  – When a CpG site is methylated, it is methylated on both strands
  – In many disease processes, including familial cancers, gene repression due to tumor suppressor gene promoter hypermethylation is a heritable trait

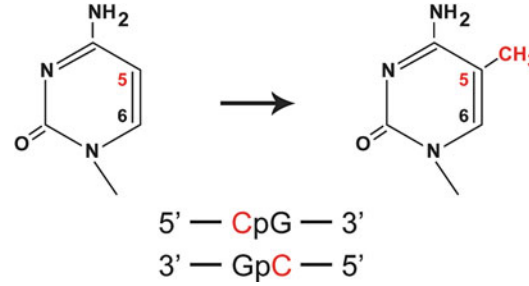


Fig. 3.5 DNA methylation is a chemical modification of cytosine pyrimidines in a CpG island in which a methyl group is added to the number 5 carbon of the base. Methylated DNA generally is "turned off" and does not permit transcription of the exon

3.3.2 DNA Methylation Enzymes

- In humans, the process of DNA methylation may be carried out by one of three isoenzymes
  – DNA methyltransferase 1 (DNMT1) methylates hemimethylated DNA
  – DNA methyltransferase 3a (DNMT3a) is more highly expressed in ESC and appears to methylate genes that are critical in late development
  – DNA methyltransferase 3b (DNMT3b) is also more active in ESC, and mutations of DNMT3b are associated with the rare ICF syndrome (immunodeficiency, centromeric instability, facial anomalies)

3.3.3 DNA Methylation and Genetic Regulation

- DNA methylation is an important contributor to gene silencing
  - Methylation of DNA may itself physically impede the binding of transcriptional proteins to the gene, blocking initiation of transcription
  - The methylated DNA binds proteins known as methyl-CpG-binding domain proteins (MBD1), which recruit additional proteins to the promoter and silence transcription
  - DNA methylation also affects histone structure and forms heterochromatin, which is transcriptionally inactive

3.3.4 DNA Methylation and Cancer

- Many tumor suppressor genes in cancer cells are silenced by hypermethylation of CpG islands
- The high incidence of C → T transitions found in the *TP53* tumor suppressor gene is attributed to the spontaneous deamination of 5-methylcytosine residues to form thymine bases (Fig. 3.6)
- C → T transition in pluripotent tissues is also mediated by activated cytidine deaminase
- Observations linking DNA methylation to cancer also suggest a model in which there is a high rate of mutation at CpG dinucleotides due in part to methyltransferase-facilitated deamination
- However, an overall methyl deficit is observed in tumor cells consistent with global hypomethylation of tumor cell DNA and increased expression of many nonsuppressor genes
  - This occurs despite high levels of DNA methyltransferase expression
  - Hypomethylation leads to increased gene transcription, frequently causing increased expression of oncogenes
  - Hypomethylation induces chromosomal instability (CIN)

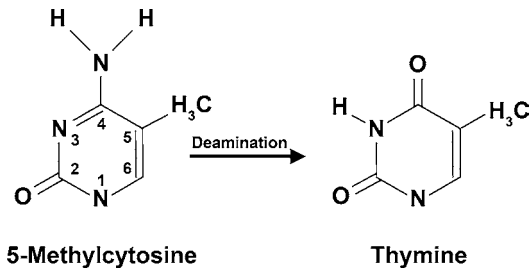


Fig. 3.6 Spontaneous deamination of 5-methylcytosine is responsible for the high incidence of C → T found in the *TP53* tumor suppressor gene in malignancy. When an amino group is removed from 5-methylcytosine, the base changes from cytosine to thymine. In DNA, this reaction cannot be corrected by the DNA repair mechanisms

- Hypermethylation and hypomethylation are frequently found in promoter regions of primary cancer cell genes
- The high rate of 5-methylcytosine single nucleotide mutation compared to other nucleotides has been widely documented in the literature
- Changes in global and regional methylation patterns are among the most frequent and earliest events to occur in cancer
- Both mutational and epigenetic changes alter DNA methylation and have a direct impact on neoplastic transformation

3.3.5 Methylation and Potential Clinical Interventions

- One of the most characteristic features of cancer is the inactivation of tumor suppressor genes by hypermethylation of the CpG islands located in their promoter regions
- Compounds with DNA-demethylating capacity have potential for treatment of cancer
- Discovery of DNA-demethylating capacity was a decisive event in the clinical trials that have merited the approval of 5-azacytidine by the US Food and Drug Administration for the treatment of myelodysplastic syndrome

3.3.6 Methods for DNA Methylation Analysis

3.3.6.1 Polymerase Chain Reaction-Based Methods

- Methylation-specific polymerase chain reaction (PCR) is a frequently used method for studying DNA methylation (Fig. 3.7)
  - Genomic DNA is modified using sodium bisulfite to deaminate nonmethylated cytosine to uracil prior to PCR. The methylated cytosine in genomic DNA cannot be converted by sodium bisulfite
  - Methylation-specific primers are designed according to nonconvertible sequences
  - Only methylated sequences are amplified by these methylation-specific primers
- PCR amplifies bisulfite-modified genomic DNA using strand-specific primers, and the PCR product is sequenced to determine the uracil content that represents nonmethylated cytosine
- Bisulfite PCR followed by restriction analysis
  - The amplified bisulfite-modified regions contain sites that are cleaved by restriction enzymes specific for uracil–guanosine-containing sequences
- Real-time PCR allows quantitative analysis of DNA methylation using methylation-specific primers
- Methylation-sensitive single-strand conformation analysis
  - Methylated and nonmethylated sequences form different secondary structures and have different electrophoretic mobility

3.3.6.2 Non-PCR-Based Methods

- Single nucleotide primer examination (SNuPE) anneals primers designed to end at the GC sequence on the complimentary bisulfite-modified DNA strand in the presence of labeled deoxynucleotide triphosphate (dNTP)
  - Two primer extension reactions, one with dCTP and another with dTTP, determine that the missing base was methylated or nonmethylated, respectively
- DNA high-resolution melting (HRM) analysis is based on melting properties of DNA in solution. Bisulfite modification changes the DNA sequence and lowers the DNA melting temperatures when a significant number of methylcytosines have been converted to uracil
- Enzymatic regional methylation assay is a quantitative method for determining the methylation density of a DNA region
  - *E. coli* cytosine methyltransferase (Sss1) and *E. coli* adenine methyltransferase (dam) add differently labeled methyl groups to total and methylated DNA strands
  - The ratio of the two signals is proportional to methylation density
- Methylation-specific oligonucleotide array uses bisulfite-modified DNA hybridized to a methylation-specific microarray to analyze multiple methylation sites

3.4 Microsatellite Instability

3.4.1 Overview

- Microsatellites are short repeated nucleotide sequences virtually unique for each individual
- The repeating sequences are usually only four or five nucleotides in length
- Not only is the sequence of microsatellites at a locus highly variable among people, but the length of microsatellites is also individual specific and constant within an individual's normal somatic cells
- DNA repair defects cause the number of repeats to vary in abnormal or transformed cells

3.4.2 Definition and Mechanisms of Microsatellite Instability

- Microsatellite instability (MSI) is the condition of having longer or shorter microsatellite regions than normal cells of the individual

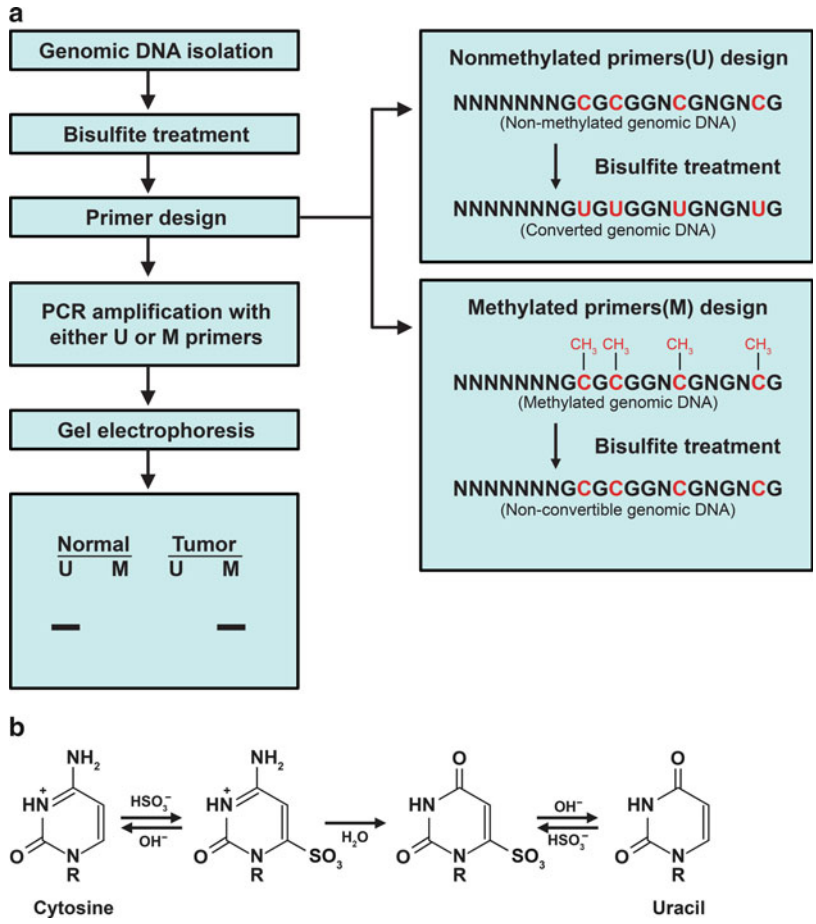


Fig. 3.7 Methylation-specific PCR is a commonly used method for DNA methylation analysis. Genomic DNA is extracted from tissue. Chemical conversion of any nonmethylated cytosine to uracil (U) by hydroquinone and sodium bisulfite permits all nonmethylated cytosine bases can be identified (panel B). Methylation-specific primers are designed according to the uracil-containing sequences after bisulfite conversion. Methylcytosine resists deamination by sodium bisulfite, so any methylcytosine bases are unchanged in the reaction mixture. All the nonmethylated cytosine would be converted into uracil. Therefore, specific primers for methylation of nonmethylated sequences are designed (*right* panel A). Specific primers for nonmethylated sequence cannot amplify methylated (nonconvertible) DNA, and the primer for methylated sequence cannot amplify converted (nonmethylated) sequences. The gel electrophoresis shows the different methylation states of tissue. In the example, normal tissue exhibits a nonmethylated (U) PCR product, whereas the tumor tissue shows a methylated (M) PCR product (panel A)

- MSI refers solely to novel length alleles and is distinct from the loss of heterozygosity (LOH) or allelic imbalance, in which one of the preexisting alleles has been lost
  - Alteration in the length of microsatellites due to deletion or insertion of single nucleotides or repeating units is found in tumor DNA when compared with normal/germline DNA from the same locus
  - The DNA instability is due to failure of the mismatch repair (MMR) system to correct errors in the transcription of microsatellite short sequence repeats, mistakes normally fixed by proofreading enzymes

  - Defective *MMR* genes produce enzymes that cannot correct nucleotide mismatches due to insertion, deletion, or misincorporation of bases during DNA replication

3.4.2.1 DNA Mismatch Repair Enzymes and Regulation Factors

- Inactivation of the *MMR* genes, including *hMLH1*, *hMSH2*, *hMSH6*, *hPMS*, and *hPMS2*, results in MSI
- Epigenetic inactivation of *hMLH1* by promoter methylation is also one of the major causes of MSI
- Tumors with MSI respond differently to treatments than those without MSI
  - Colon cancers with high MSI respond to first-line chemotherapy better than MSI-stable tumors

3.4.2.2 Implications of Microsatellite Instability

- The type of microsatellite change found in MSI can be detected only if many cells are affected
- Detectable MSI is thus an indicator of clonal expansion typical of neoplasia

3.4.3 Bethesda Panel for Microsatellite Instability

- A 1997 National Cancer Institute consensus workshop recommended a 5-microsatellite marker panel for the detection of MSI
- The panel recommended loci include (BAT)-25, (BAT)-26, D2S123, D5S346, and D17S250
- The tumors were divided into three subgroups according to the MSI status
  - MSI-high (MSI-H): samples with variability in two or more of five markers. The majority of MSI-H tumors are near diploid and have few karyotypic abnormalities
  - MSI-low (MSI-L): cancers show variability in only one of the five microsatellite markers
  - MSI-stable (MSI-S): cancers show no MSI in any of the five markers. Tumors with abnormal cytogenetic analysis are frequently MSI-S
- Interpretation of the Bethesda panel
  - Many colorectal cancers show MSI-H
    - MSI-H is associated with hereditary nonpolyposis colorectal cancer (HNPCC)
    - MSI-H is also found in 15–20% of sporadic colorectal cancers
    - The presence of MSI-H is associated with a more favorable prognosis
  - MSI-L cancers may or may not represent a biologically distinct category
    - MSI-L cancers have a worse prognosis in some, but not all, studies
  - Cancers with intact MMR (MSI-S) have a significantly better prognosis than those with MSI (MSI-H)

3.4.4 Microsatellite Instability in Hereditary Nonpolyposis Colorectal Cancer

- HNPCC is a hereditary cancer syndrome associated with increased risk for various cancers, predominantly colorectal cancers
- Defects in the DNA *MMR* system result in accelerated accumulation of mutations affecting critical genes that lead to malignant progression
- The syndrome is caused by germline mutation in *hMLH1* and *hMSH2* in 90% of cases
  - Mutation in one DNA *MMR* gene allele is inherited in the germline
  - MSI follows only if there is inactivation of the other *MMR* allele
  - MSI also occurs in the absence of germline MMR mutation as the result of epigenetic inactivation of *MMR* genes by promoter methylation
- Tumors in *MMR* mutation carriers typically exhibit MSI. These are called MSI-H tumors if MSI is detected in multiple microsatellite markers
- 85–90% of HNPCC-associated colorectal cancers are characterized by MSI

- MSI status is useful for prognosis and therapeutic decision making
  - Many studies have demonstrated improved prognosis for MSI-H cancer relative to the MSI-stable cancer
  - MSI-positive colorectal cancers are less responsive to fluorouracil (5FU)-based adjuvant chemotherapy
- See Chap. 23 for detailed description

3.4.5 Methods for Microsatellite Instability Analysis

- PCR-gel electrophoresis uses isotope-labeled nucleotides and primers designed to amplify microsatellite markers from genomic DNA of normal and tumor cells (Fig. 3.8)
  - False-negative results are caused mainly by contamination from noncancer cells. For a reliable MSI analysis, at least 70% of the cells examined should be tumor cells
  - It is recognized that microdissection should be used to enrich the tumor cell populations, which enhances the likelihood of detecting shifted microsatellite markers
- Fluorescence-based methods use labeled primers for coamplification of multiple markers subsequently separated by capillary electrophoresis
- Denaturing high-performance liquid chromatography uses high-performance liquid chromatography instead of gel electrophoresis to analyze PCR-amplified microsatellite DNA from tumor and normal cells
- Real-time PCR followed by melting point analysis reveals alterations in the length of repetitive sequences within 2 h

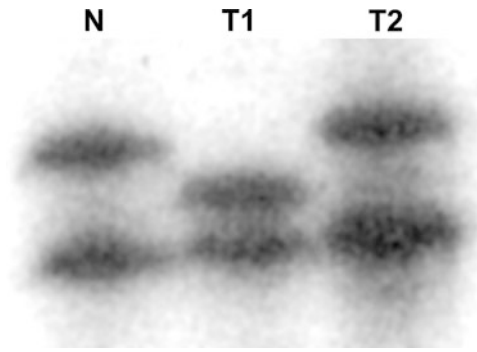


Fig. 3.8 MSI is the condition of having longer or shorter microsatellite regions in the tumor cells than in the normal cells of the individual. The figure illustrates a typical MSI pattern. N, normal control; T1, T2 are different tumors from same patient. The figure shows a lengthened upper allele of T1 and a shortened upper allele of T2. MSI does happen in cases with noninformative normal control, since double bands in tumor samples indicate microsatellite alteration of an allele

3.5 Chromosomal Instability

3.5.1 Overview

- Chromosomal instability (CIN) is a state of continuous new chromosome structure formation at a rate higher than in normal cells
- CIN is not synonymous with aneuploidy
- Neither cytogenetic complexity nor cytogenetic heterogeneity per se is conclusive evidence for CIN
- CIN describes mutation at the gross chromosomal level, including structural and numerical instability

3.5.2 Mechanisms of Chromosomal Instability

- CIN involves changes in the number and structure of chromosomes, including chromosomal losses, gains, and rearrangements
- CIN can occur secondary to defects affecting different steps of chromosome segregation, including:
  - The spindle checkpoint
  - Kinetochore–microtubule attachments
  - Sister chromatid cohesion
  - Centrosome duplication
  - Bipolar spindle assembly
- Structural rearrangement is caused by DNA damage and recombination
  - DNA double-strand breaks can result in a number of different structural

rearrangements, including translocations, inversions, ring chromosome formation, insertions, and deletions
  - Downregulation or inactivation of a cell cycle checkpoint system, such as p53 or p21, allows a cell with damaged DNA to escape from apoptosis
  - Since telomere caps normally prevent crossing over of chromosome ends with chromosome midportions, telomere dysfunction results in segmental gains or losses of chromosomes by telomeric fusion between chromosomes
- Breakage–fusion–bridge cycle structural abnormalities
  - Dicentric chromosomes have two centromeres, so they are pulled apart by the spindle during mitosis, producing two chromosome fragments with uncapped ends
  - The two broken chromosome ends often fuse into novel dicentrics and rings, which break again at the next cell division
  - Concurrent breaks in two different chromosomes may give rise to either translocations or dicentrics
  - Centromere malfunctions lead to numerical instability (aneuploidy) with gains or losses of entire chromosomes (Fig. 3.9)
- Asymmetrical segregation of chromosomes at the metaphase–anaphase transition
  - Tumors exhibit asymmetrical segregation of chromosomes at the metaphase–anaphase transition, resulting in unequal or incomplete distribution of the genetic material to the daughter cells
  - Abnormal number, structure, or function of the centrosome, consisting of the spindle apparatus and centrioles, leads to asymmetrical segregation of the chromosomes
  - An abnormal number of centrosomes or mitotic spindles causes uneven distribution of the chromosomes during mitosis
  - Failure of the centromere to bind the mitotic spindle leads to permanent loss of the chromosome in the next cell division
- Inactivation of genes that control the timing of mitotic chromosome segregation results in some chromosomes being "left behind"

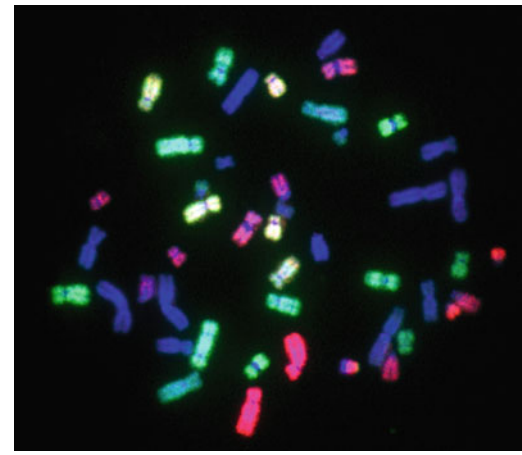


Fig. 3.9 CIN is a state of continuous chromosome addition, deletion, or rearrangement at a rate higher than normal. CIN is mutation at the gross chromosomal level, including both structural and numerical alterations. The figure shows a typical picture of multiple gains and losses by chromosomal painting. Chromosomes with different color bands or arms have had insertions or translocations from other chromosomes. Color coding allows the cytogeneticist to quickly identify pairs of chromosomes. If paired chromosomes have different lengths, there has been a deletion from the shorter chromosome resulting from chromosome instability

- Failure of genome surveillance machinery, such as by the *BRCA1*-associated genome surveillance complex, allows cells with gain or loss of large genetic segments to escape apoptosis

3.5.3 Clinical Implications of Chromosomal Instability

- The pathogenetic role of CIN, the underlying cellular defects that give rise to CIN, and the contribution of CIN to the progression of cancer remain to be established
- Most malignant tumors exhibit both structural and numerical chromosome abnormalities
- With progression of tumors toward higher malignancy, chromosome instability becomes more frequent
- CIN causes an imbalance in chromosome number, which is an important mechanism of tumor suppressor gene inactivation

  - Almost all cancer cells have gains or losses of chromosomes and frequent rearrangements. However, scientists have argued for nearly a century about whether this abnormality is the cause of cancer or merely collateral damage
  - In most cancers, tumor cells within the neoplasm share similar cytogenetic abnormalities, indicating a stepwise accumulation of heritable chromosomal changes has occurred during tumor growth
  - CIN appears to be both an epiphenomenon and a cause of cancer
  - CIN produces other traits favoring genomic variation and inviting rapid selection for deathless phenotype
- Stable aneuploidy, such as Down syndrome (trisomy chromosome 21), can occur without CIN, but aneuploidy observed in cancer is often caused by CIN
- Tumors with CIN are usually microsatellite stable
- CIN can be exploited as quarry for targeted therapy

3.5.4 Chromosomal Instability Syndromes

- CIN syndromes are a group of inherited conditions associated with CIN and breakage. Each is associated with a tendency to develop certain types of malignancy
  - Ataxia-telangiectasia is a primary immunodeficiency disorder characterized by progressive cerebellar ataxia, oculocutaneous telangiectasia, progressive cerebellar dysfunction, and recurrent pulmonary infections. About 20% develop cancer, usually acute lymphocytic leukemia or lymphoma
  - Bloom syndrome is a rare inherited disorder characterized by CIN because of a mutation in the *BLM* gene which codes for a DNA helicase protein essential for maintaining genomic stability during DNA unwinding for replication. Affected individuals have a 150–300 times increased risk of malignancy, usually acute leukemia, lymphoma, or gastrointestinal cancer
  - Nijmegen breakage syndrome is a recessive syndrome characterized by CIN due to mutations in the complex that manages double-strand DNA breaks. It is characterized by microcephaly, short stature, immunodeficiency, radiation sensitivity, and a strong predisposition to lymphoid malignancy
  - Fanconi anemia is an autosomal recessive condition associated with bone marrow failure and high sensitivity to DNA cross-linking chemicals. About 10% of patients develop leukemia, 6% myelodysplastic syndrome, and 10% solid malignancies of liver, esophagus, and vulva
  - Xeroderma pigmentosum is a defect in the nucleotide excision repair gene that renders these individuals exquisitely sensitive to ultraviolet radiation. They have a 1,000-fold increase in nonmelanoma skin cancer. After coauthoring the original article describing this syndrome in 1874, Kaposi went on to name the condition in 1882 for the dry, pigmented skin changes usually seen from infancy in persons with this genetic syndrome

3.5.5 Methods for Chromosomal Instability Analysis

- Metaphase karyotype analysis. Dividing cells with condensed chromosomes are swollen in hypotonic solution and gently burst open to deposit the chromosomes together on a slide
- Banding karyogram analysis. The chromosomes may be studied in greater detail by enzymatic digestion and special stains to reveal condensed and loose bands of the chromosomes
- Fluorescence in situ hybridization (FISH) chromosome painting. A cocktail of fluorescent-labeled probes incubated with the chromosomes allows even more specific identification of specific chromosome

segments, even if misplaced on the wrong chromosome
- Comparative genomic hybridization. Equal amounts of normal and tumor DNA tagged with different fluorescent dye-binding molecules are allowed to hybridize. Special computer software detects zones of mismatch in the chromosomes after painting (Fig. 3.10)

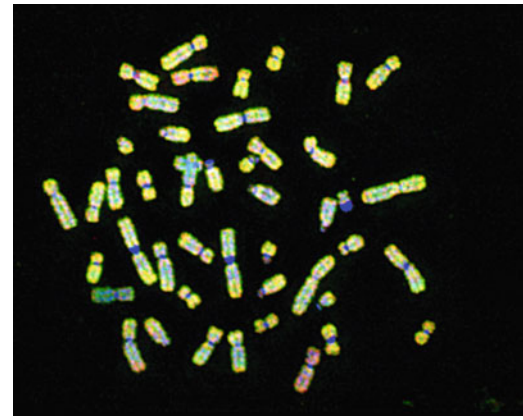


Fig. 3.10 The figure is a typical comparative genomic hybridization of tumor DNA. Chromosome spreads from a normal individual are painted with one color fluorescent probe (usually *red*). Tumor DNA painted with a different color probe mixture (usually *green*) is layered over the normal chromosomes and allowed to hybridize. DNA with neither red nor green probe painting binds a blue fluorescent dye. Loss of tumor DNA yields more *red* color, whereas amplification of tumor DNA results in more *green* color. When *red* and *green* are about equal, the net color appears *yellow*. Computer software is used to quantitatively analyze the gains and losses for each chromosome region. This technique is complementary to chromosome painting or banding, since it fails to show inversions, reciprocal translocations, or changes without gain or loss of DNA

3.6 Gene Imprinting

3.6.1 Overview

- An individual inherits two copies of their genes, one from each parent
- Genomic imprinting is the phenomenon by which the active copy of a gene depends on the parent of origin, resulting in a functional difference between homologous chromosomal regions
  - Gene imprinting is an epigenetic phenomenon
- Establishment of imprinting at a locus requires that the two alleles be differentially marked in oogenesis and spermatogenesis
- Parental allele-specific expression involves a small subset (<100 genes) of all the genes (about 30,000 genes) in the genome expressed according to their parent of origin
  - The imprinted genes tend to cluster within specific regions of chromosomes. Two major clusters of imprinted genes are at chromosome 11p15 and in the region from 15q11 to 15q13
- The pattern of parental allele-specific expression is stably transmitted during cell division. Some imprinted genes are expressed from a maternally inherited chromosome and silenced on the paternal chromosome; while other imprinted genes show the opposite expression pattern
- Hypermethylation on one of the two parental alleles is the major mechanism for most imprinting. Differential methylation of CpG island promoters causes one parent's allele to have higher expression than the other's, based on arrival through ovum or spermatozoa
- Selective gene silencing by hypermethylation of CpG island promoters or by RNAi directed toward one parental allele now explains some patterns of inheritance which were formerly difficult to classify
- Currently <100 imprinted genes (about 75) have been identified, most of which are protein-coding genes
- Reprogramming refers to erasure and reestablishment of DNA methylation during gamete development. Reprogramming occurs in the parent germ cell when egg or sperm is maturing (Fig. 3.11)
  - During gametogenesis, the primordial germ cells must have their original biparental DNA methylation patterns erased and reestablished based on the sex of the transmitting parent
  - The maternal and paternal genomes are differentially marked and must be properly

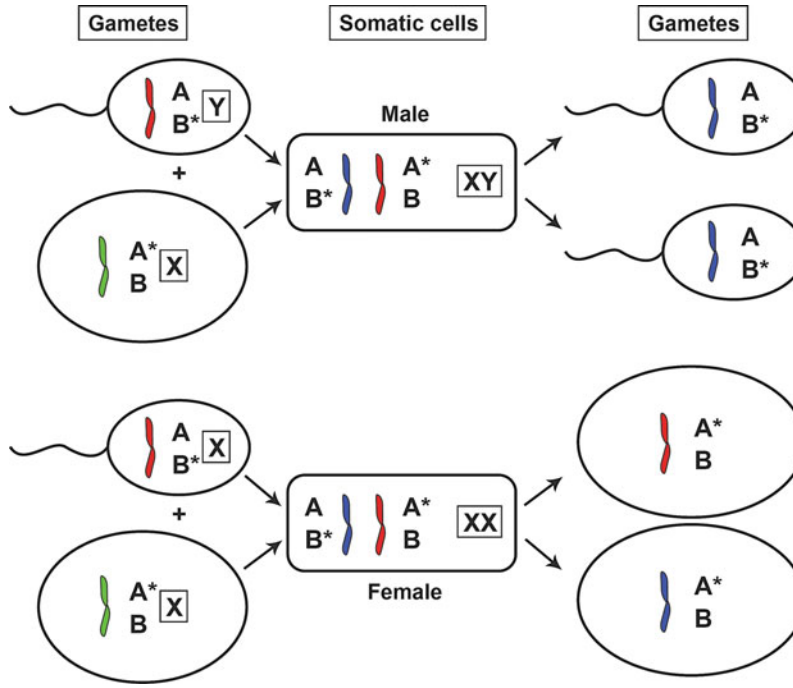


Fig. 3.11 Gene imprinting is an epigenetic phenomenon that causes a functional difference between homologous mammalian chromosomal regions based on their parental origin. Maternal and paternal genomes are differentially marked and must be properly reprogrammed every time they pass through the germline to restore full maternal or paternal marking in the gametes. During gametogenesis, the primordial germ cells must have their original biparental DNA methylation patterns erased and reestablished based on the sex of the transmitting parent. This process is referred to as reprogramming

reprogrammed every time they pass through the germline
- Reprogramming is achieved through reestablishment of DNA methylation in regulatory sequences for imprinted genes
- Reprogramming results in gene silencing based on the sex of the parent
- Phosphorylation or other chemical modification of histone proteins also contributes to gene imprinting and gamete-specific differential gene silencing
  - DNA rich in methylated CpG islands is associated with hypoacetylated histone cores and increased histone H1
  - DNA containing nonmethylated CpG islands is associated with hyperacetylated histone cores and less histone H1
  - Histone core acetylation modulates the expression of numerous genes

3.6.2 Regulation of Gene Imprinting

- The imprinting process is controlled by other genes, named imprinting control centers, typically located on the same chromosome near the imprinted genes
- The imprinting control center is a differentially methylated region (DMR) producing noncoding RNA that regulates multiple genes on the same chromosome
- The imprinted region is differentially methylated on one allele
  - Binding of a nonmethylated imprinting control region by zinc finger proteins forms a chromatin barrier

3.6.3 *H19* Gene Imprinting

- *H19* is a noncoding gene with no protein production and is frequently overexpressed in cancers
- *H19* is a maternally expressed noncoding RNA located at chromosome 11p15.5
- The characteristics of *H19* gene
  - The final product of the *H19* gene is an RNA
  - It demonstrates maternal monoallelic expression in fetal tissue
  - Postnatally, H19 is expressed at high levels in cells of many types of cancer

  - It is the first designated oncofetal RNA
  - It functions as a riboregulator causing both posttranscriptional-increased oncogene expression and decreased tumor suppressor mRNA effectiveness as a translation template
  - *H19* enables tumor cell survival under stress conditions by promoting angiogenesis and cancer progression
- Functional grouping of genes modulated by *H19* RNA indicates that cellular migration, angiogenesis, and metastasis are favored by *H19* imprinting
- Parentally methylated regions in the germline are present upstream of the *H19* promoter in normal cells
- Mutation in the methylation region for maternal *H19* gene locus causes silencing of *H19* gene
- Recently, the *H19* transcript was shown to be the pre-miRNA of miR-675 whose target mRNA is the tumor suppressor RB1 (Refer to Sect. 3.7 below)

3.6.4 Imprinting and Disease

- Disturbance of imprinting is an important reason for disease even when the patient has an intact DNA gene related to the imprinting disease
- The imprinting diseases may also be caused by gene mutations or microdeletions in some cases
  - The deregulation of imprinted genes has been associated with human diseases
    - Beckwith–Wiedemann syndrome (BWS) is a pattern of overgrowth of body parts, midline abdominal wall defects, and increased risk of developing Wilms tumor or hepatoblastoma that has been linked to the imprinted region at 11p15
    - Angelman syndrome is a complex of developmental delay and seizure-like motor disorders that is associated with loss of maternal genes in regions of chromosome 15 where the paternal genes are silenced by imprinting
  - The characteristic of these diseases is that they do not follow normal Mendelian patterns of inheritance as well as showing parental origin effects
- Alterations of imprinting patterns are found in a number of cancers
- Since they are complex epigenetic diseases, imprinting syndromes are more economically diagnosed by symptom criteria than by genetic testing

3.6.4.1 Imprinting Syndromes

- Imprinted genes are highly variable in individual symptom traits since a single or epigenetic alteration can critically change the function of many genes
- The incidence of these disorders is increased in humans conceived by assisted reproductive technology
- Beckwith–Wiedemann syndrome
  - The incidence of BWS is 1 in 13,700
  - The disorder is found equally in males and females
  - The disease has been identified in different ethnic groups
  - BWS maps to 11p15, a differentially methylated region clustering imprinted genes
  - BWS is characterized by general overgrowth with symptoms including hemihypertrophy, macroglossia, and visceromegaly
  - 10–20% of BWS individuals are predisposed to embryonal tumors, the most frequent of which are Wilms tumor and adrenocortical carcinoma
  - The most common molecular event occurring in BWS patients that do not have cytogenetic abnormalities is the biallelic expression of *IGF2* due to loss of imprinting
  - Loss of imprinting at the *IGF2* locus may be accompanied by the methylation and/or silencing of the active maternal allele of *H19*
- Prader–Willi (PWS) and Angelman syndrome (AS)
  - PWS and AS are two clinically distinct genetic diseases associated with genomic imprinting on chromosome 15q11-q13

- The disease is now known to occur with a frequency of 1 in 15,000–20,000 persons worldwide
- Although the symptoms of these two disorders are quite different, they are both caused by deletions on chromosome 15
  - PWS results from the loss of a group of paternally inherited genes [del(15)(q11q13)pat]
  - AS results from loss of a maternally inherited gene in the same chromosomal region
- Both paternal and maternal deletions alter *SNRPN* (small nuclear ribonucleoprotein polypeptide N) promoter methylation and prevent expression of its paternal allele

3.6.4.2 Gene Imprinting and Human Cancer

- Numerous tumors are associated with the preferential loss of a particular parental chromosome, indicating the involvement of imprinted genes
  - Neuroblastoma is associated with imprinting disorders of maternal chromosome 1p36 and paternal chromosome 2
  - Acute myeloblastic leukemia is associated with an imprinting disorder of paternal chromosome 7
  - Wilms tumor is associated with maternal chromosome 11p15.5
    - 70% of Wilms tumors have biallelic *IGF2* expression
- Loss of imprinting (LOI) is an alteration in cancer that involves loss of parental origin-specific gene expression
  - LOI may involve activation of the normally silent copy of a growth promotion gene or silencing of a normally expressed imprinted gene allele
- LOI at the *IGF2* gene in Wilms tumor could result from loss of *H19* expression
  - Rhabdomyosarcoma is associated with imprinting alterations of maternal chromosome 11p15.5
  - Sporadic osteosarcoma is associated with imprinting alteration of maternal chromosome 13
  - Imprinted genes can be involved in carcinogenesis in several ways
    - LOH at an imprinted region may result in deletion of the only functional copy of a tumor suppressor gene
    - LOI of a gene may lead to inappropriate expression of an imprinted gene that promotes cell growth
    - Inactivation of an imprinting control center results in aberrant expression of multiple imprinted oncogenes or tumor suppressor genes present in an imprinted chromosomal region
- Since imprinted genes are functionally haploid, they are susceptible to tumorigenic factors causing inactivation of the one active allele that would eliminate tumor suppressor gene expression

3.7 MicroRNA

3.7.1 Overview

- What is microRNA (miRNA)?
  - MicroRNAs (miRNAs) are a family of single-stranded, noncoding small RNA molecules 2–25 bases long
  - miRNAs are posttranscriptional regulators that bind to complementary sequences in the 3′ untranslated regions (UTR) of target mRNA
  - miRNAs are encoded by DNA and transcribed from DNA but not translated into protein
- miRNA transcription and processing
  - miRNA genes are transcribed by RNA polymerase II
  - miRNA is first transcribed as a long primary transcript (pri-MIR)
  - Pri-MIR is subsequently processed into a 60–120-nucleotide precursor with a hairpin (stem loop) structure (pre-MIR)
  - Pri-MIR is processed in the cell nucleus by the multi-subunit microprocessor complex, consisting of a nuclease (Drosha) and a double-stranded RNA (dsRNA)-binding protein (Pasha)

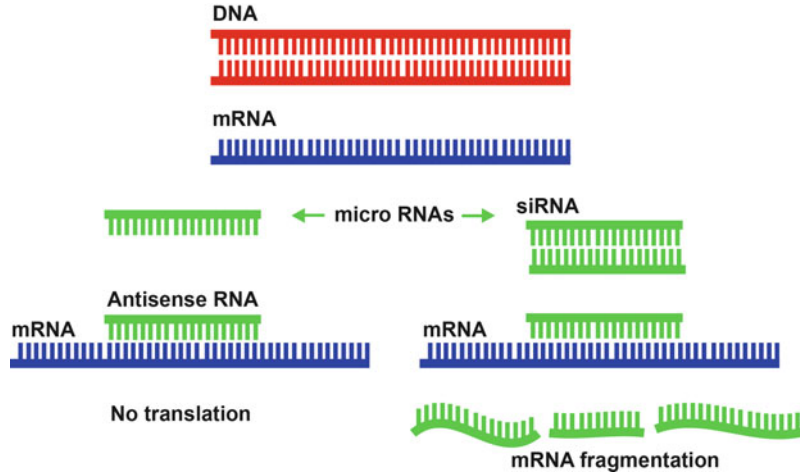


Fig. 3.12 Posttranscriptional gene repression is the basic function of miRNA. miRNA is transcribed from RNA genes and processed to miRNA. The miRNA molecules interact with the 3′ UTR of specific mRNA to suppress translation or to induce mRNA degradation. siRNA is dsRNA resembling viral RNA with similar effects toward posttranscriptional gene repression

  - A pri-miRNA sequence consists of five distinct components
    - Basal segment
    - Lower stem
    - Upper stem
    - Top stem
    - Terminal loop
  - Drosha and Dicer components of the microprocessor complex are members of ribonuclease III enzyme family that recognize structural regions of pri-MIR and pre-MIR molecules
- miRNA properties, recruitment, and function
  - Most pre-MIR molecules have an imperfect dsRNA structure topped by a terminal loop
  - Processed pre-MIRs are exported to the cytoplasm where they are further processed into mature miRNA
  - miRNA functions as a gene regulator by binding to the complementary sequence of one or more mRNAs (Fig. 3.12)
    - Animal miRNAs usually anneal to a site in the 3′ UTR
    - Plant miRNAs usually anneal to coding regions of mRNAs
    - Each miRNA may interact with multiple genes
    - Each mRNA is complementary to multiple miRNAs
    - miRNA-bound mRNA either remains untranslated or is degraded by RNAi effecter complex (RNA-induced silencing complex, RISC)
  - miRNA controls proliferation, differentiation, development, apoptosis, and stress response
- Posttranscriptional gene repression is the basic function of miRNA
  - The annealing of the miRNA to the mRNA inhibits protein translation and sometimes facilitates cleavage of the mRNA (Fig. 3.12)
  - miRNAs may also increase methylation of genomic sites corresponding to targeted mRNAs
  - miRNA function requires several ribonucleoproteins collectively termed the miRNP
  - miRNA activity may be blocked by antisense high-affinity locked nucleic acid oligo, morpholino oligo, or 2′-O-methyl RNA oligo
    - Locked nucleic acid RNA oligomers have modified ribose moieties with an extra bridge connecting the 2′ oxygen with the 4′ carbon by ether linkage
    - Morpholino RNA oligomers are nucleic acid analogs in which the ribose backbone is replaced by a chain of six-member morpholine rings joined by phosphorodiamidate linkages

    - 2′-O-methyl RNA oligomers are stabilized by adding a methyl group to the 2′ position of the ribose on the 5′ terminal residue
  - Antisense RNA oligomers synthesized to block-specific miRNAs are called antagomirs
- RNA genes
  - RNA genes encode RNA that is not translated into a protein
  - The human nuclear genome contains about 3,000 unique RNA genes (<10% of total gene number) to produce noncoding RNA for structural and catalytic functions in the cell
  - The human mitochondrial genome contains 24 RNA genes: two code for 23S and 16S rRNA subunits of mitochondrial ribosomes, the rest code for tRNAs

3.7.2 Clinical Implications

- miRNA-related diseases
  - miRNA regulates insulin secretion from β cells and may be involved in some cases of diabetes mellitus
  - miRNA encoded by viruses suggests a potential role for miRNA in the viral infection cycle
- miRNA and cancer
  - miRNA tumorigenesis cluster located on 13q31 is amplified in several lymphomas and other cancers
  - miRNAs, sometimes called oncomirs, may cause posttranscriptional downregulation of tumor suppressor gene expression and/or upregulation of oncogene expression
  - Mice engineered to produce excess lymphoma cell miRNA developed the disease within 50 days and died 2 weeks later
  - Two types of miRNA inhibit E2F1 protein translation, a critical regulator of cell proliferation
  - Expression patterns of 217 miRNA genes determined by bead-based flow cytometric analysis reveal gene activity fingerprints that can distinguish the developmental lineage and differentiation state of cancers
- Potential miRNA-based therapies
  - miRNA directed at oncogenes and tumor suppresser gene silencers may be therapeutic against cancer
    - miRNA let-7 represses *RAS* oncogene and significantly inhibits the growth of lung cancer
    - miRNA miR-15a and miR-16-1 repress *BCL2* oncogene and induce apoptosis in leukemic cell lines
  - Antagomirs (modified antisense miRNA), neutralizing miRNAs, are present in all tissues except brain
  - Antagomirs block and degrade target miRNA, allowing silenced gene reactivation

3.7.3 Methods for MicroRNA Discovery

- Computation-driven analysis identifies candidate miRNA sequences based on known structural features. The candidate sequences must be validated by experiment
- miRNA target analysis identifies "seed" nucleotide sequences that are complementary to known mRNA 3′ UTR sequences. Binding studies and functional analysis are essential to confirm true mRNA targets
- De novo identification of miRNA candidates involves sequencing of size-fractioned cDNA libraries to identify 22-nucleotide RNA molecules with hairpin sequences

3.8 RNA Interference

3.8.1 Overview

- RNA interference (RNAi) is the process of sequence-specific, posttranscriptional gene silencing in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene

- RNAi is a mechanism conserved in related eukaryotic organisms that preserves genomic integrity, regulates gene expression, and guards against exogenous virus infection
  - miRNA and small interfering RNA (siRNA) play a central role in RNA interference
    - siRNA is exogenous double-stranded RNA that is taken up by cells or enters via vectors
    - miRNA is endogenous noncoding RNA transcribed from somatic genes or mitochondrial DNA
- siRNA binds to target mRNA in a sequence-specific fashion and causes mRNA degradation
- A specific siRNA can transiently knock down expression of a targeted gene with high specificity and selectivity
  - Since siRNA must be efficiency delivered, the sequence is often introduced by a vector and replicated into a long dsRNA inside the cells, then the dsRNA is processed into siRNA by enzyme cleavage in the cytosol
  - siRNA is a dsRNA with 2-nucleotide 3′ overhangs on either end
  - Each strand has a 5′ phosphate group and a 3′ hydroxyl (−OH) group
- Biologic function of siRNA
  - Antiviral defense
  - Silencing mRNAs that are overproduced or translationally aborted
  - Guarding the genome from mutation induced by relocation of mobile genetic elements (transposons)
  - RNAi machinery is involved in miRNA processing and the resulting posttranslational repression
- siRNA-induced posttranscriptional gene silencing
  - siRNA interferes with the expression of a particular gene in many eukaryotes sharing a common homologous sequence
  - RNAi is mediated by the same cellular machinery that processes miRNA
  - siRNA molecules are involved in large-scale gene regulation in the cell
  - Before RNAi was well characterized, the phenomenon was known by other names, including *posttranscriptional gene silencing, transgene silencing*, and *quelling* (Fig. 3.13)
  - siRNA is bound in the RNA-induced silencing complex (RISC) and uses one of the strands as a guide to target complementary mRNA
- See Table 3.3 for differences between miRNA and siRNA

3.8.2 Biologic and Clinical Implications of Small Interfering RNA

- Applications of siRNA
  - siRNA's dramatic and selective reduction of an individual protein expression makes it a valuable research tool, both in cell culture and in vivo
  - Posttranscriptional silencing triggers assembly of a nuclease complex that targets homologous mRNAs for degradation
  - siRNA gene silencing may be amplified by priming the synthesis of additional dsRNA via RNA-directed RNA polymerase
  - siRNA can trigger transcriptional alterations at the genomic level by inducing methylation at sites of sequence homology
- Role of miRNA in medicine
  - siRNA has been used in clinical trials for treating macular degeneration and respiratory syncytial virus
  - RNAi can affect the complete reversal of virus-induced liver failure in mouse models
  - miRNA may inhibit viral gene expression in cancer cells
  - miRNA may stabilize neurodegenerative diseases, with particular attention to the polyglutamine diseases such as Huntington disease
- Potential risk of therapeutic siRNA in vivo
  - Introduction of too much siRNA can result in nonspecific activation of innate immune responses
  - Off-target effects may result in essential genes coincidentally similar to the targeted gene also being repressed

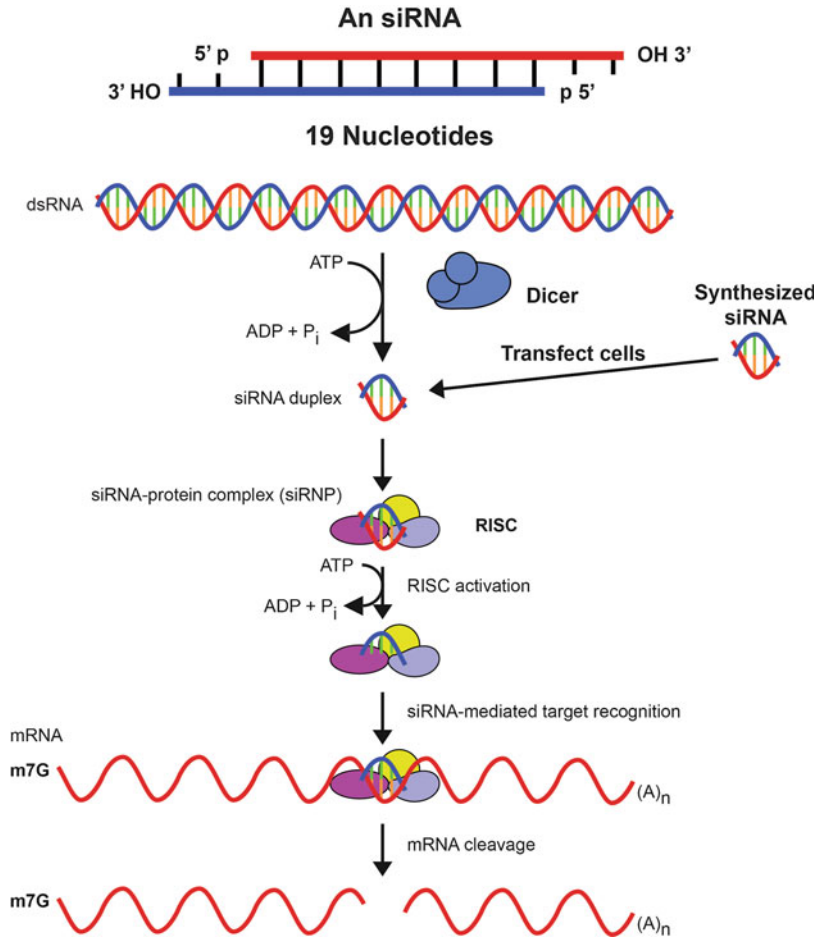


Fig. 3.13 siRNAs can be either endogenous or exogenous. siRNA is a 19-nucleotide dsRNA with 2-nucleotide 3′ overhangs on either end (upper panel). Each strand has a 5′ phosphate group (−P) and a 3′ hydroxyl group (−OH). The dsRNA chains are processed into siRNA molecules that form siRNA–protein complexes. The siRNA–protein complex can interact with mRNA to cause mRNA degradation. M7G indicates the cap site of the mRNA, and (A)n refers the poly A tail

Table 3.3 Differences between miRNA and siRNA

	miRNA	siRNA
Gene coding	Coded by RNA genes	Not coded by RNA genes
Structure	Single stranded	Double stranded
Processing	Posttranscriptional processing	Processing of long bimolecular RNA duplexes
Main function	Regulate gene expression	Posttranscriptional repression
Sequence conservation	miRNA sequences are conserved	siRNA sequences are not really conserved

3.9 Telomere

3.9.1 Overview

- What is a telomere?
  - The telomere is a ribonucleoprotein complex composed of at least seven proteins and an RNA primer sequence bound to a repetitive DNA sequence at the ends of the p and q arms (Fig. 3.14)
  - Telomeres are repeating simple sequences that belong to the minisatellite family (TTAGGG)
  - The repeat block extends between 3 and 20 kbp in length and is associated with

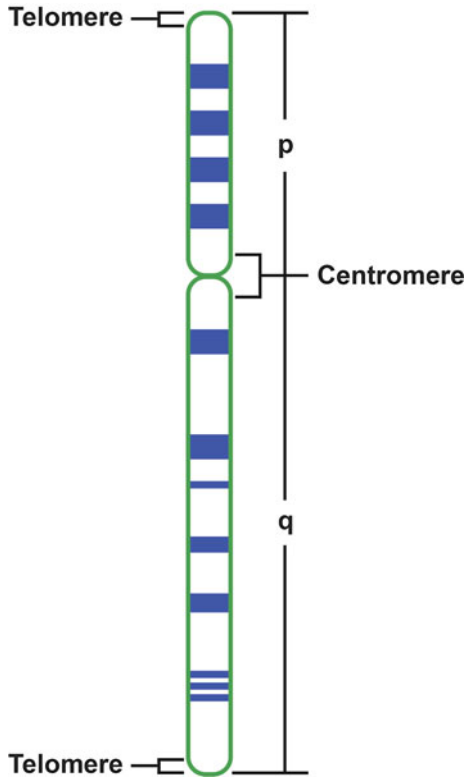


Fig. 3.14 The telomeres are located on both ends of the chromosome arms and function as a disposable buffer. Telomeres protect the chromosome from gene loss by ring chromosome formation or gain by nonreciprocal translocation during cell division

a 100–300-kbp segment of telomere-associated repeats between the telomere and the rest of the p or q arm

- The biologic function of telomeres
  - Telomeric DNA and telomere-specific binding proteins have an essential role in stabilizing chromosome ends by forming a cap structure that protects chromosome ends from degradation and terminal fusions
    - The basic function of telomeres is to seal the end of chromosome arms to protect them from exonuclease degradation
    - The telomere acts to protect the ends of chromosomes from fraying or ligation to unprotected DNA
    - Telomeres function as a disposable buffer to protect chromosomes from gene loss during division
    - Telomeres also associate with other telomeres at the nuclear envelope during interphase to maintain an ordered structure of the chromatin
  - Count the number of cell divisions
    - Telomeres judge the number of cell divisions that have occurred
    - Telomeres determine the cellular life span and dictate when replicative senescence will occur
  - Provide a mechanism for complete replication of DNA at the ends of chromosomes
    - Discontinuous replication of the lagging strand involves Okazaki fragments and loss of bases at the 3′ end
    - Telomerase adds hexamer repeats to 3′ ends, allowing DNA polymerase to complete synthesis of the lagging strand without losing bases from coding genes near the end of the chromosome

3.9.2 Telomere Structure and Maintenance

- In most organisms, telomere length is maintained within a limited range as a characteristic inherited trait
- Telomere length homeostasis requires that telomere degradation and extension are balanced, a process that is regulated by telomere proteins
- Telomeres normally shorten with each cell division and lengthen only in germ cells
- Telomeres are composed of both repeated DNA and specific DNA-binding proteins
- Telomeres contain double-stranded DNA in a closed loop with the 3′ single-strand overhang annealing to a complementary segment of the second strand, which is essential for telomere maintenance and capping
- The telomere loop (T-loop) is formed by this annealing of the DNA end back on itself
  - T-loop formation is catalyzed by a specific enzyme
  - The loop is stabilized by a protein complex

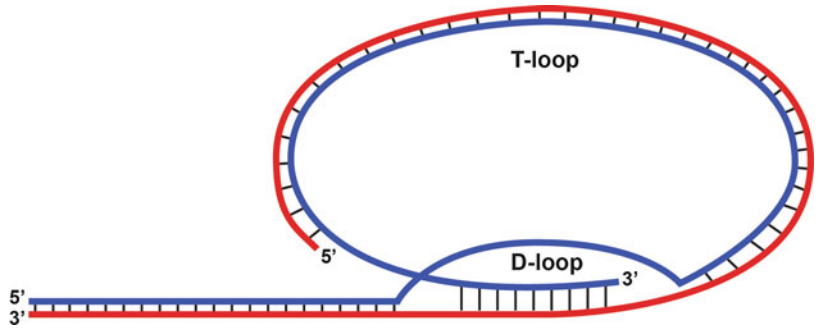


Fig. 3.15 The structures of D-loop and T-loop are essential components of the telomere. The T-loop is formed at the 3′ end of one DNA strand folding back to form a large loop. The D-loop is formed by the 3′ end binding to the 5′ end sequences of the telomere, displacing the normal complementary chain. Loop formation is catalyzed by specific enzymes and is stabilized by the protein complex. The loop structure protects the telomere

  - The overhang is buried inside the loop; the ends will not be recognized as a break in the double strand
  - When the T-loop structure is disturbed, the growth of the cell is arrested and cell cycle checkpoint proteins p53 and pRb are activated
    - If the cell cannot pass these checkpoints, the growth of the cell is arrested permanently and cell senescence occurs
    - If the cell cannot pass the p53 checkpoint, apoptosis occurs
    - If the checkpoints are bypassed, the cell grows continuously and indefinitely, resulting in genomic instability and risk of malignancy
- The displacement loop (D-loop) is formed by the 3′ G-rich strand extension overhang invading the duplex telomeric repeats (Fig. 3.15)
  - The D-loop is about 200 bp in length
- The stability of eukaryotic genomes is protected, at least in part, by the integrity of telomeres
- Catalytic component
  - Human telomere reverse transcriptase (hTERT) synthesizes DNA from an RNA template in the hTERT telomerase complex
  - hTERT is not expressed in most somatic cells
  - Reverse transcription by hTERT synthesizes telomeric sequences lost during cell division
  - hTERT activity is a critical factor in stabilizing telomeres through addition of TTAGGG repeats in human gametes

3.9.3 Cell Senescence and Telomere Shortening

- The number of cell cycles a cell may complete is limited by telomere length
- Telomeres shorten as cells repeatedly divide without primers of the lagging strand (Fig. 3.16)
- When eventually the shortening telomere reaches a senescence breakpoint, irreversible growth arrest or apoptosis will occur
- Shortened telomeres also increase the risk of chromosomal fusions
- Cell senescence arrests the growth of aged cells and prevents further mutation accumulation, thus reducing the risk of malignant transformation
- Shorter telomeres are associated with shorter organism life. Among people older than 60, those with shorter telomeres were three times more likely to die from heart disease and eight times more likely to die from infectious disease
  - Telomere shortening places a limit on replication capacity of cells
  - Telomere shortening can initiate CIN, which is a major driving force in malignant transformation

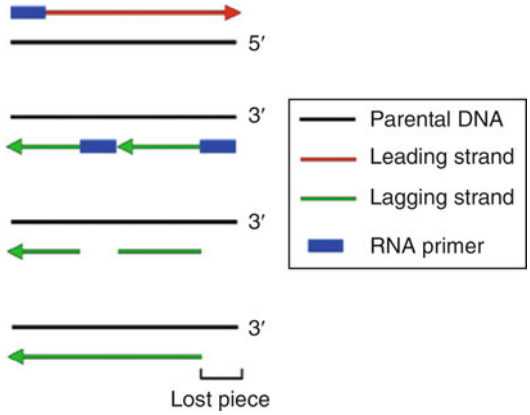


Fig. 3.16 Telomeres shorten during cell division. DNA polymerase requires an RNA primer to initiate synthesis in the 5′–3′ direction. DNA polymerase can synthesize the leading strand to the end of the chromosome. However, for the lagging strand, DNA synthesis produces a series of fragments, each requiring an RNA primer. When DNA ligase joins the fragments, it must have an adjacent 3′ DNA end to complete the conversion of the RNA primer. Since the last RNA primer has no joining DNA, a portion of the lagging strand of telomere is lost, resulting in telomere shortening

  - Long telomeres are a significant barrier to cancer formation
  - Telomere shortening associated with aging explains increasing cancer risk with age
- Telomere extension
  - Telomere extension is catalyzed by an enzyme, telomerase
  - Telomerase extends telomeres with its own octamer primer to add repeat TTAGGG sequences (Fig. 3.17)
  - Telomerase is active only in germ cells of multicellular eukaryotes to restore youthful telomere length in gametes
  - Telomerase-independent pathway: alternative lengthening of telomeres (ALT)
    - ALT is active in 15% of the telomerase negative neoplasias
    - The ALT pathway is preferentially active in mesenchymally derived cells, compared with those of epithelial origin
    - Some ALT-positive cells are associated with malfunction of promyelocytic leukemia nuclear body, a large deoxyribonucleoprotein complex essential for chromatin remodeling in resting cells
    - ALT involves homologous recombination between telomeres. Sequences are copied from one telomere to another by complementary annealing as a means of priming new telomeric DNA

3.9.4 Telomere Shortening and Cancer

- Shortened telomeres are found in many cancers, including malignancy of pancreas, bone, prostate, bladder, lung, kidney, larynx, and nasopharynx
- Telomere shortening is associated with genetic instability and increased cancer risk
  - The mean telomere length in cancer cells is about 5 kb
  - Telomere lengths between 7 and 9 kb are considered long
  - Telomere lengths between 3 and 5 kb are considered short
- Telomere shortening is also observed in many premalignant lesions
  - Marked telomere shortening is observed in 93% of prostatic intraepithelial neoplasia cases, a precursor to prostate cancer
  - Telomere shortening is prevalent in premalignant lesions from a variety of human epithelial tissues
    - Cancers with short telomeres often have high telomerase activity
- Cancers with short telomeres are likely to respond quickly to antitelomerase therapy

3.9.5 Telomerase Summary

- Telomerase is a reverse transcriptase (hTERT) carrying its own RNA template, repetitively coding the elongating telomere
- Telomerase is only transiently expressed during S phase in differentiated normal cells

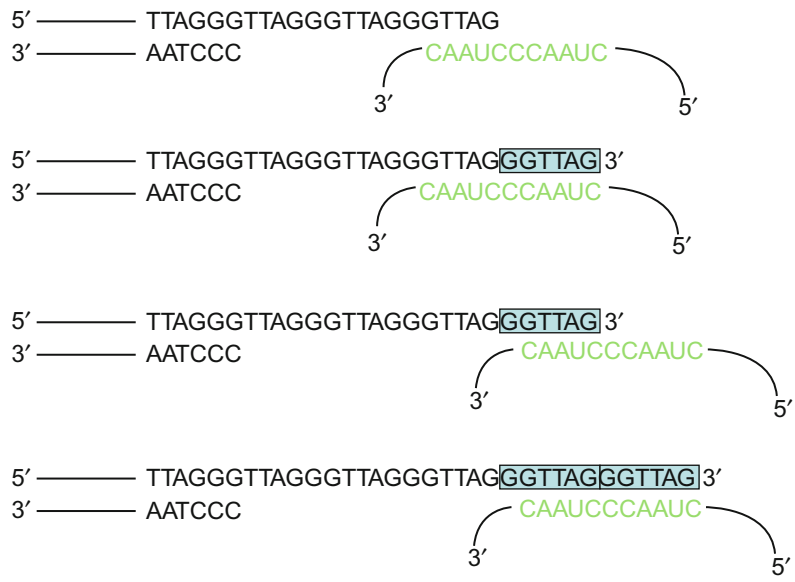


Fig. 3.17 Telomere extension is catalyzed by the enzyme telomerase. The leading strand is synthesized using an RNA molecule (in green) as a template to direct DNA synthesis. Each time the DNA is extended a block of TTAGGG toward the 3′ direction as shown in the box. The lagging strand is synthesized in fragments using the newly synthesized leading strand as a template and then ligated to form a new strand. Telomerase is thus a reverse transcriptase carrying its own RNA template, repetitively coding the elongating telomere

- Telomerase adds specific DNA sequence repeats (TTAGGG) to the 3′ end of DNA strands in the telomere regions
- Inhibition of telomerase reduces cell proliferation and accelerates loss of the telomere 3′ overhang
- Cancer cells with shortened telomeres overcome the checkpoints by expression of telomerase to reestablish their telomeres
- Telomerase provides cancer cells with at least two critical functions
  - Suppresses CIN
  - Grants unlimited cell replication
- Telomerase is the key to cellular evasion of senescence and apoptosis in many cancer cell lines
  - Normally, only germline and stem cells have highly active telomerase
  - Telomerase restores the ability of senescent cells to divide
- A variety of premature aging syndromes are associated with short telomeres
- Regulation of telomerase
  - The telomerase gene is located on the distal short arm of chromosome 5p (5p15.33)
  - Telomerase is highly active in gametogenesis and early embryogenesis
  - Telomerase is downregulated during cell differentiation and has undetectable expression in differentiated somatic cells
  - Amplifications of 5p15 is detected in some cancers, suggesting that increased copy number may be one mechanism for increasing telomerase expression in human tumors
  - The telomerase gene promoter contains a number of regulatory sites including two *MYC* binding sites, which may contribute to upregulated expression of telomerase. The oncogene *MYC* is overexpressed in up to 93% of solid cancers, making it the most frequently upregulated oncogene
  - It has been suggested that the telomerase gene promoter might contain a *TP53*-binding site that negatively regulates

telomerase expression. Since the majority of human cancers are deficient in p53 protein, this might also contribute to telomerase overexpression

3.9.6 Telomerase and Cancer

- Telomerase is expressed in many cancers but not in normal differentiated cells
- Many cancer cells are considered "immortal" because telomerase activity allows them to evade senescence
  - The telomeres in cancer cells have lost a substantial number of subunits due to uncontrolled replication, but activated telomerase stabilizes the shortened telomeres, allowing the over-proliferative cells to become not only immortal but also genetically unstable
- Activated telomerase precludes death by chromosome instability or by senescence-activated induction of apoptosis pathways
- Telomerase also promotes survival by continued activation of proliferation pathways
- Telomerase activation is observed in 90% of all human tumors but not in normal somatic tissue, suggesting that the cellular imperishability conferred by this enzyme plays a key role in cancer development
- Some cancer cells have alternative lengthening of telomeres (ALT), a nonsynthetic telomere lengthening pathway involving the transfer of telomere tandem repeats between sister chromatids
- Alternatively, tumor cells can repair their shortened telomere ends by other means, such as DNA recombination, in which one chromosome obtains DNA from another

3.9.7 Diagnostic and Therapeutic Implications of Telomerase

- Telomerase might play a role for screening and diagnosis of cancer
  - Increased telomerase activity is detectable in stage 0–1 breast cancer
  - Telomerase is expressed in most urinary bladder cancers (90%), prostate cancers (80%), and kidney cancers (69%)
  - Application of a FISH telomere assay improved the sensitivity and specificity for detecting malignant cells in cytology specimens
  - Telomerase tests have also shown promising results for improving the sensitivity and specificity of malignant cell screening in urine, pleural and peritoneal effusions, and in bronchoalveolar lavage fluid
  - Telomerase also appears to be a prognostic indicator. In breast cancer, there is a direct correlation between telomerase activity and higher stage, size of tumor, and nodal status
  - Telomerase expression is low in normal peripheral blood lymphocytes and benign lymph nodes. But telomerase expression increases in malignant lymphocytes, which could be helpful for making a distinction between benign and malignant lymph tissue

3.9.7.1 Antitelomerase Therapy

- Telomerase-inhibiting compounds would inhibit malignant cell growth and permit other antineoplastic drugs to induce apoptosis in susceptible cancer cells
- Telomerase inhibition might be useful in the treatment of any cancer with telomerase expression
  - Telomerase inhibitors suppress telomerase activity and reduce the proliferation rate of lung, breast, liver, and prostate cancer cells
  - Induction of differentiation, inhibition of reverse transcriptase, telomerase promoter downregulation, telomerase primer antisense inhibition, and blockage of telomere/telomerase interactions are different approaches being attempted for telomerase-targeting treatment
  - siRNA antitelomerase treatment decreases telomerase activity and inhibits cancer cell growth in vitro

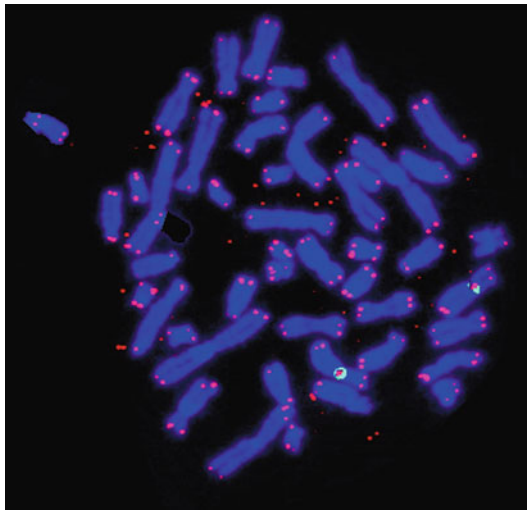


Fig. 3.18 FISH with telomere probes shows that each chromosome has four telomeres, seen on the tips of the p- and q-arms. The fluorescence intensity may be used to assay the telomere length

3.9.8 Methods for Telomere Analysis

- Terminal restriction fragment Southern blot
  - Radiolabeled $(AATCCC)_n$ antisense oligonucleotide probe against the TTAGGG telomere motif is allowed to hybridize to Hinf/Rsa I digested genomic DNA on a nylon membrane
  - The hybridized membrane is exposed to autoradiography film, and the intensity is converted into relative telomere length
- FISH-labeled telomeres in cytogenetic metaphase spreads
  - Fluorescence-conjugated telomere probes bind to the chromosome telomeres
  - The fluorescence intensity of hybridized bands is converted to the relative telomere length based on the ratio with control intensity (Fig. 3.18)
- The real-time PCR assay determines the telomere-to-single copy gene (T/S) ratio using separate primers and separate reaction wells to compare telomere hexamer to copy number of a standard housekeeping gene. The T/S ratio is proportional to the average telomere length in a cell
- Telomere length is generally expressed relative to a count of the average tumor cell chromosome number, the chromosome number of a normal cell from the host or a single copy gene in the test cell population

Further Reading

Ashley EA, Butte AJ, Wheeler MT, et al. Clinical assessment incorporating a personal genome. Lancet. 2010;375:1525–35.

Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. Cell. 2011;145:30–8.

Baker M. RNA interference: from tools to therapies. Nature. 2011;464:1225.

Bartel DP. MicroRNAs: genetics, biogenetics, mechanism, and function. Cell. 2004;116:281–97.

Bautch VL. Cancer: tumour stem cells switch sides. Nature. 2010;468:770–1.

Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 2011;11:726–34.

Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. Nat Med. 2011;17:320–9.

Brower V. Epigenetics: unravelling the cancer code. Nature. 2011;471:S12–3.

Cernilogar FM, Onorati MC, Kothe GO, et al. Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. Nature. 2011;480:391–5.

Cheng L, Sung MT, Cossu-Rocca P, et al. OCT4: biological functions and clinical applications as a marker of germ cell neoplasia. J Pathol. 2007;211:1–9.

Cheng L, Williamson SR, Zhang S, et al. Understanding the molecular genetics of renal cell neoplasia: implications for diagnosis, prognosis and therapy. Expert Rev Anticancer Ther. 2010a;10:843–64.

Cheng L, Davidson DD, MacLennan GT, et al. The origins of urothelial carcinoma. Expert Rev Anticancer Ther. 2010b;10:865–80.

Cheng L, Alexander RE, Zhang S, et al. The clinical and therapeutic implications of cancer stem cell biology. Expert Rev Anticancer Ther. 2011a;11:1133–45.

Cheng L, Zhang S, MacLennan GT, et al. Bladder cancer: translating molecular genetic insights into clinical practice. Hum Pathol. 2011b;42:455–81.

Cheng L, Zhang S, Alexander R, et al. Sarcomatoid carcinoma of the urinary bladder: the final common pathway of urothelial carcinoma dedifferentiation. Am J Surg Pathol. 2011c;35:e34–46.

Cheng L, Alexander RE, MacLennan GT, et al. Molecular pathology of lung cancer: key to personalized medicine. Mod Pathol. 2012;25:347–69.

Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. Nat Med. 2011a;17:297–303.
Chin L, Hahn WC, Getz G, Meyerson M. Making sense of cancer genomic data. Genes Dev. 2011b;25:534–55.
Chitwood DH, Timmermans MC. Small RNAs are on the move. Nature. 2011;467:415–9.
Clark MF, Fuller M. Stem cells and cancer: two faces of Eve. Cell. 2006;124:1111–5.
de Lange T. How telomeres solve the end-protection problem. Science. 2009;326:948–52.
Delude C. Tumorigenesis: testing ground for cancer stem cells. Nature. 2011;480:S43–5.
Frohling S, Dohner H. Chromosomal abnormalities in cancer. N Engl J Med. 2008;359:722–34.
Goffin J, Eisenhauer E. DNA methyltransferase inhibitors-state of the art. Ann Oncol. 2002;13: 1699–716.
Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481:306–13.
Green ED, Guyer MS. National human genome research I. Charting a course for genomic medicine from base pairs to bedside. Nature. 2011;470:204–13.
Haber DA, Gray NS, Baselga J. The evolving war on cancer. Cell. 2011;145:19–24.
Hamburg MA, Collins FS. The path to personalized medicine. N Engl J Med. 2010;363:301–4.
Hamilton A, Baulcombe D. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science. 2006;286:950–2.
Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100:57–70.
Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74.
Hannon G. RNA interference. Nature. 2002;418:244–51.
Harris TJ, McCormick F. The molecular pathology of cancer. Nat Rev Clin Oncol. 2010;7:251–65.
Herman J, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med. 2003;349:2042–54.
Ionov Y, Peinado MA, Malkhosyan S, et al. Ubiquitous somatic mutations in simple repeated sequences reveals a new mechanism for colonic carcinogenesis. Nature. 1993;363:558–61.
Joeng KS, Song EJ, Lee KJ. Long lifespan in worms with long telomeric DNA. Nature. 2004;36:607–11.
Jones PA, Gonzalgo ML. Altered DNA methylation and genome instability: a new pathway to cancer? Proc Natl Acad Sci USA. 1997;94:2103–5.
Jordan CT, Guzman ML, Noble M. Cancer stem cells. N Engl J Med. 2006;355:1253–61.
Keverne EB. Neuroscience: a mine of imprinted genes. Nature. 2011;466:823–4.
Kong A, Steinthorsdottir V, Masson G, et al. Parental origin of sequence variants associated with complex diseases. Nature. 2009;462:868–74.
Kouzminova N, Lu T, Lin AY. Molecular basis of colorectal cancer. N Engl J Med. 2011;362:1245–6.
Lander ES. Initial impact of the sequencing of the human genome. Nature. 2011;470:187–97.
Ledford H. Gene reading steps up a gear. Nature. 2011;470:155.
Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancer. Nature. 1998;396:634–9.
Leslie M. Cell biology. Are telomere tests ready for prime time? Science. 2011;332:414–5.
Loeb LA. Human cancers express mutator phenotypes: origin, consequences and targeting. Nat Rev Cancer. 2011;11:450–7.
Lowden MR, Flibotte S, Moerman DG. DNA synthesis generates terminal duplications that seal end-to-end chromosome fusions. Science. 2011;332:468–71.
Macconaill LE, Garraway LA. Clinical implications of the cancer genome. J Clin Oncol. 2010;28:5219–28.
Manolio TA, Green ED. Genomics reaches the clinic: from basic discoveries to clinical impact. Cell. 2011;147:14–6.
McDermott U, Downing JR, Stratton MR. Genomics and the continuum of cancer care. N Engl J Med. 2011;364:340–50.
Medema JP, Vermeulen L. Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. Nature. 2011;474:318–26.
Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010;11:685–96.
Pearson H. Epidemiology: study of a lifetime. Nature. 2011;471:20–4.
Polyak K, Hahn W. Roots and stem: stem cells in cancer. Nat Med. 2006;12:296–300.
Potti A, Schilsky RL, Nevins JR. Refocusing the war on cancer: the critical role of personalized treatment. Sci Transl Med. 2010;2:28cm13.
Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature. 2006;444:444–54.
Rini BI, Atkins MB. Resistance to targeted therapy in renal-cell carcinoma. Lancet Oncol. 2009;10: 992–1000.
Roychowdhury S, Iyer MK, Robinson DR, et al. Personalized oncology through integrative high-throughput sequencing: a pilot study. Sci Transl Med. 2011;3:111ra21.
Shackleton M, Quintana E, Fearon ER, et al. Heterogeneity in cancer: cancer stem cells versus clonal evolution. Cell. 2009;138:822–9.
Stratton MR. Exploring the genomes of cancer cells: progress and promise. Science. 2011;331:1553–8.
Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458:719–24.
Velculescu VE, Diaz LA, Jr. Understanding the enemy. Sci Transl Med. 2011;3:98ps37.
Visvader JE. Cells of origin in cancer. Nature. 2011;469:314–22.
Weinstein IB, Joe A. Oncogene addiction. Cancer Res. 2008;68:3077–80; discussion 80.

Genetic Inheritance and Population Genetics

4

Tatiana Foroud and Daniel L. Koller

Contents

4.1 Polymorphisms 111
4.2 Hardy-Weinberg Law 112
4.3 Autosomal Recessive Inheritance 113
4.4 Autosomal Dominant Inheritance 116
4.5 X-Linked Dominant Disorders 117
4.6 X-Linked Recessive Disorders 118
4.7 Bayes Theorem and Risk Calculation 118
4.8 Gene Mapping and Recombination 118
4.9 Disorders with Complex Genetic Inheritance 123
4.10 Genome-wide Association Studies 124
4.11 Next-Generation Sequencing 125
Further Reading 127

4.1 Polymorphisms

- Variation in DNA sequence produces genetic polymorphisms (Greek: many forms). A polymorphism is any DNA sequence in which the less common allele occurs at a frequency of 1% or greater. The first polymorphisms identified were those defining the serologic blood groups. Subsequently, many more polymorphisms have been identified
- Simple sequence repeats are a widely studied type of polymorphism. This type of polymorphism is often termed a microsatellite marker and consists of alleles defined by a variable number of two, three, or four tandem nucleotide sequences. These polymorphisms are commonly used to map disease genes
- Single nucleotide polymorphisms (SNPs) are a single base pair change in the DNA sequence that can be found in either the coding or non-coding region of the DNA. Many of the originally identified polymorphisms were SNPs; however, most were in the coding region and resulted in a phenotypic effect. Many of the SNPs used today do not result in any noticeable phenotypic effect. SNPs are estimated to occur, on average, once per every 1 in 100–300 bp, suggesting that there are over 10 million SNPs in the human genome
- Insertion/deletions (indels) and copy number variants (CNVs) are additional types of polymorphism that are being studied using recently developed molecular techniques

T. Foroud, PhD (✉) • D.L. Koller, PhD
Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_5,

4.2 Hardy-Weinberg Law

- Population genetics is defined as the study of genes in populations as distinguished from the study of the behavior of genes in families. However, this is a rather narrow definition since this area of genetics encompasses, for example, inbreeding and prediction of recurrence of a disorder in families
- The fundamental law in population genetics is the Hardy-Weinberg law. It is named after the English mathematician G. H. Hardy and the German physician W. Weinberg who independently published their versions of the law in 1908
- The Hardy-Weinberg law has important implications in the field of population genetics and for risk assessment for members of families with genetic disorders. The importance of the Hardy-Weinberg law lies in its ability to take information about the frequency of alleles in the population and then make predictions about the frequency of genotypes in the population
- This law allows us to estimate the frequency of gene carriers in a population. Then, this risk or frequency can be used to estimate the probability that another family member will be affected with a particular disease. The Hardy-Weinberg law defines the frequencies of genotypes in a randomly mating population based on the frequencies of the alleles at a locus. By random mating, we mean that matings occur without regard to the genotypes of the individuals
- The following illustrates the Hardy-Weinberg law with random mating for a locus with two alleles: A and a. The frequency of the two alleles is p and q, respectively, and the sum of the two allele frequencies is 1. There are three possible genotypes at this gene. An individual can have two copies of allele A (i.e., homozygous) and thus have genotype AA. The individual could have one copy of the A allele and one copy of the a allele (i.e., heterozygous) and thus have genotype Aa. The individual could have two copies of allele a (i.e., homozygous) and have genotype aa

Box 4.1: Assumptions of the Hardy-Weinberg law

- Random mating
  - Individuals select their mate at random and do not select individuals with particular disease or clinical traits
- Constant mutation rate
  - The frequency at which a normal allele mutates to an abnormal allele is the same as the rate at which an abnormal allele sustains another mutation
- No selection
  - Individuals of all genotypes reproduce at similar rates (i.e., there is no selection against the reproduction of a particular genotype)
- No fluctuation of gene frequencies due to migration or random causes
  - The composition of the population genotypes does not alter due to the entrance or departure of individuals with certain genotypes, or due to chance, in small populations (genetic drift)

Table 4.1 Transmission of gametes

			Male gametes	
			Frequency of A allele, p	Frequency of a allele, q
Female gametes	A allele	Frequency of A allele, p	AA (p x p)	Aa (p x q)
	a allele	Frequency of a allele, q	Aa (p x q)	aa (q x q)

- If the assumptions of the Hardy-Weinberg law are met (Box 4.1), then the frequency of these three genotypes (AA, Aa, and aa) in the population can be estimated using the allele frequency estimates of A and a obtained from the same population. The relationship between the estimated allele frequencies and the genotypic frequencies is the basis of the Hardy-Weinberg law and is shown below in Table 4.1. Therefore, the Hardy-Weinberg law allows us to estimate the frequency of

Table 4.2 Using the Hardy Weinberg law to estimate genotypic frequencies. Consider the β-globin locus to have two alleles. The normal allele is termed β^A and an abnormal allele, β^S. Individuals who have two copies of the β^A allele are considered normal. Individuals who have one normal allele and one abnormal allele, with genotype $\beta^A\beta^S$, have sickle trait. Those individuals who have two abnormal copies of the beta-globin gene, genotype $\beta^S\beta^S$, have sickle cell disease. Using the counting method, the frequency of the β^A allele is estimated to be ([2{884} + 1{112}]/[2 × 1000]) = 0.94 while the frequency of the β^S allele is 0.06. Therefore, in a sample of 1,000 individuals, we would expect the three genotypes to occur in the following frequencies:

Phenotype	Genotype	Frequency of genotype	Number of individuals
Normal	$\beta^A\beta^A$	$p^2 = (0.94)^2 = 0.884$	884
Sickle trait	$\beta^A\beta^S$	$2pq = 2(.94)(.06) = 0.112$	112
Sickle cell anemia	$\beta^S\beta^S$	$q^2 = (0.06)^2 = 0.004$	4

the three genotypes at this locus in the population to be frequency of AA = p^2, frequency of Aa = $2pq$, and frequency of aa = q^2 (Table 4.2)

- The Hardy-Weinberg law also holds for loci with more than two alleles. The frequency of a given homozygote, A_iA_i, is p_i^2 and for a heterozygote, A_iA_j, is $2p_ip_j$

4.3 Autosomal Recessive Inheritance

- There are many disorders with developmental disabilities that are inherited with an autosomal recessive pattern of inheritance. Individuals are only affected with the condition if they have inherited two abnormal alleles at the disease gene. Typically, the abnormal allele for an autosomal recessive disorder is denoted by a and the normal allele is indicated by A. A characteristic pattern of autosomal recessive inheritance is horizontal transmission. This means that affected individuals are only observed in a single generation and typically are siblings
- When studying autosomal recessive disorders, the only individuals whose genotypes are known with certainty are typically those who are affected with the disease and therefore must have genotype aa. Individuals who are unaffected may be either homozygous for the normal allele or might be heterozygous. Those individuals who have one normal and one abnormal allele at the disease gene are often termed "carriers"
- In populations meeting the assumptions of the Hardy-Weinberg law, we can estimate the frequency of a heterozygous individual to be $2pq$. If the frequency of individuals affected with a particular autosomal recessive disease is 1 in 2,500, then, using the Hardy-Weinberg law, we would assume the frequency of affected individuals = q^2 = 1/2,500. From this formula, we would estimate the frequency of the disease allele (q) to be the square root of 1 in 2,500, which is 1 in 50. We can estimate that the frequency of the normal allele must be 49/50. Using the frequency of the two alleles, we can now estimate the frequency of a disease allele carrier for this particular disease to be frequency of disease allele carrier = $2pq$ = 2(0.02) (0.98) = 0.0392. However, since the disease is relatively rare, the frequency of the normal allele (p) is very close to 1 and the frequency of the disease allele (q) is very small. Then, the frequency of a carrier = $2pq$ is approximately 2(1) $q = 2q$. For this particular disease, estimating the frequency of heterozygotes as $2q$ results in an estimated carrier frequency of 1 in 25, which is 0.04. Thus, the frequency of a gene carrier in a population meeting the assumptions of the Hardy-Weinberg law is simply twice the frequency of the disease allele (Box 4.2; Table 4.3)
- With the diagnosis of an individual with an autosomal recessive disorder, many

Box 4.2: Phenylketonuria (PKU)

- PKU has a frequency of 1/10,000. Then, the frequency of $q2 = 1/10{,}000$ and the frequency of the disease allele is $q = \sqrt{1/10{,}000} = 1/100 = 0.01$. From this, we can estimate the frequency of heterozygous disease allele carriers to be twice the disease allele frequency ($2q$), which is 1/50

Table 4.3 Frequencies of genotypes, genes, and carriers for an autosomal recessive disorder

Disease frequency (q^2)	Gene frequency (q)	Carrier frequency ($2q$)	Carrier frequency/ disease frequency
1/1,600	1/40	1/20	80
1/10,000	1/100	1/50	200
1/40,000	1/200	1/100	400
1/250,000	1/500	1/250	1,000

An important point to note is that carriers are much more common than individuals with the disorder; the rarer the disorder, the greater is the relative frequency of carriers (*last column*)

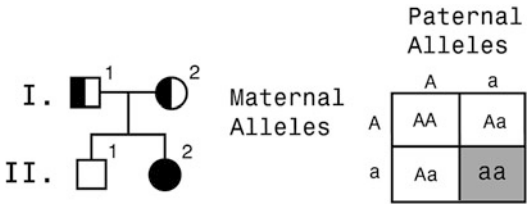


Fig. 4.1 (*Left*) Pedigree segregating an autosomal recessive disorder. (*Right*) Transmission of alleles in autosomal recessive inheritance

unaffected family members will be interested in determining whether or not they have also inherited a copy of the mutant disease allele and what the risk of disease is for their children

- In Fig. 4.1, a common clinical scenario is shown. In the left portion of the figure, a family is shown consisting of two parents and their two children. Their daughter has an autosomal recessive disorder. The affected daughter is fully shaded to denote that she is affected with the disease and must have two abnormal alleles at the disease gene. Her unaffected parents are half shaded to indicate that they must be carriers
- The right panel of Fig. 4.1 illustrates the potential transmission of alleles from the parents to their children. Since each heterozygous parent will, on average, transmit the abnormal allele 50% of the time, we would expect that 25% of their children will inherit two normal alleles (genotype AA) and be unaffected. We would also expect, on average, that 50% of their children will inherit one normal and one abnormal allele (genotype Aa) and also be unaffected. On average, 25% of their children will inherit from each parent an abnormal allele, resulting in genotype aa and an affected child. Thus, in this hypothetical scenario, the genotype of three individuals is known with certainty. The parents both have genotype Aa and the affected child has genotype aa
- The remaining individual in the family, the unaffected son, could have two possible genotypes. He could have inherited two normal alleles and have genotype AA or he could have inherited one normal and one abnormal allele and have genotype Aa. Both genotypes would result in an individual without symptoms of the disease; therefore, we cannot determine an individual's genotype through a careful physical examination or clinical evaluation
- We can estimate the likelihood that the son has genotype AA or Aa. Assuming that he cannot have genotype aa, we must recompute the risk that he is heterozygous at the disease locus by only considering the two genotypes that he can be, which are AA and Aa. Using the right portion of Fig. 4.1, we see that each of the boxed genotypes is equally likely. Therefore, if we eliminate the possibility that he is aa, the likelihood that he is AA is one of the three remaining boxes and is 1/3. The possibility that he is Aa is two of the remaining three equally likely boxes, which is 2/3. This can

be shown mathematically using the following formula
  - Probability the son has genotype AA, since he is unaffected, = P(AA)/(P[AA] + P[Aa]) = (1/4)/(1/4 + 1/2) = (1/4)/(3/4) = 1/3. Probability the son has genotype Aa, since he is unaffected, = P(Aa)/(P[AA] + P[Aa]) = (1/2)/(1/4 + 1/2) = (1/2)/(3/4) = 2/3
- Families having a member with an autosomal recessive disorder often wish to know the risk of disease to other family members. Figure 4.2 illustrates a common scenario. In this family, a couple has come for more information regarding their risk to have a child with a disorder similar to that of the wife's sister
- To address this question, we need to address three questions
  - What is the probability the father in the second generation is a carrier?
  - What is the probability the mother in the second generation is a carrier?
  - If they are both carriers, what is the probability their child will have the autosomal recessive disease?
- If we apply the questions listed above sequentially, we first have to estimate the probability that the father in the second generation is heterozygous at the disease gene. Using the Hardy-Weinberg law, we would use the estimate of the frequency of the disease allele in the population (q), and from this information, estimate the probability that he is a carrier for $2q$ (**Step 1**)
- His wife has an affected sister. We know that she does not have symptoms of the disease; therefore, we know that her genotype is not aa and must be AA or Aa. We have shown that the probability that she would have genotype AA is 1/3 and the probability that she would have genotype Aa is 2/3 (**Step 2**)
- Even if both parents are carriers, it is still only 25% of their children who will have the disease (**Step 3**)
- Therefore, the probability that this couple would have an affected child is the product of each of these probabilities. Thus, the answer would be $(2q)\ (2/3)\ (1/4) = q/3$
- See Box 4.3 for risk counseling example
- If a disease is rare, then the probability that two individuals who are heterozygous for the disease gene will marry becomes exceedingly rare. The probability that two individuals might both be carriers for the

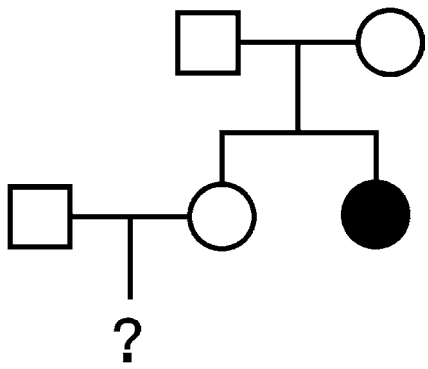


Fig. 4.2 Risk assessment for a family segregating an autosomal recessive disorder

Box 4.3: Counseling a family with PKU

- The family shown in Fig. 4.3 comes to learn more about their risk to have a child with PKU. The wife has in a previous union had a child with PKU. She now has a new partner and wants to know the probability that another child will have PKU
- Since the woman has already had a child with PKU, we know that she must be heterozygous at the PKU locus. Her new partner does not have any family history of PKU. Therefore, the probability that he is a carrier of a PKU mutation is the same as from the general population (i.e., $2q$). To estimate the frequency of a gene carrier, we use the information that the frequency of PKU is 1/10,000, so the frequency of heterozygotes in the population is 1/50. Therefore, the probability that the woman would have a child with PKU with her new partner is (1/50) (1) (1/4) = 1/200

mutant allele increases if the two individuals have both inherited the allele from a common ancestor (i.e., consanguinity). Consanguinity is more common in certain communities or countries. As a result, the descendents of a founder, who was heterozygous for recessive mutant allele, may have a higher frequency of a rare disease. In general, there is a higher frequency of consanguinity among parents of children with rare autosomal recessive disorders. When consanguinity occurs, it is not appropriate to utilize the Hardy-Weinberg law to estimate the probability that an individual in the pedigree is the carrier for an autosomal recessive disorder

- Consider the pedigree shown in Fig. 4.4. The parents of individual III-2 are both carriers of the autosomal recessive disorder. Individual II-2 inherited his disease allele from one of his parents. Therefore, if we arbitrarily assign the disease gene to his father, I-1, then his father is a heterozygote. The probability that his daughter (II-3) inherited the disease allele is 1/2. If II-3 inherited the disease allele, the probability that she transmitted the disease allele to her daughter (III-4) is 1/2. So, the risk of III-4 being a carrier is (1/2) (1/2) = 1/4. The probability of III-3 being a gene carrier is 2/3, since both his parents are carriers and he is not affected. Then, using the three steps to calculate risk, the probability that this couple would have a child with the same disorder as III-2 is (1/4) (2/3) (1/4) = 1/24
- Consanguinity results in a higher frequency of recessive disorders. One cause of consanguinity is incest between a parent and child. The result of incest among such first-degree relatives is a 40% risk for a handicapped child. The most common handicaps associated with incest are mental retardation and seizures

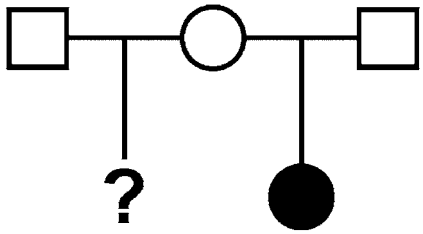


Fig. 4.3 Family risk counseling for PKU (see Box 4.3)

4.4 Autosomal Dominant Inheritance

- Unlike autosomal recessive inheritance where the genotype of the affected person is unambiguous, in the case of autosomal dominant inheritance, it is the affected individual who

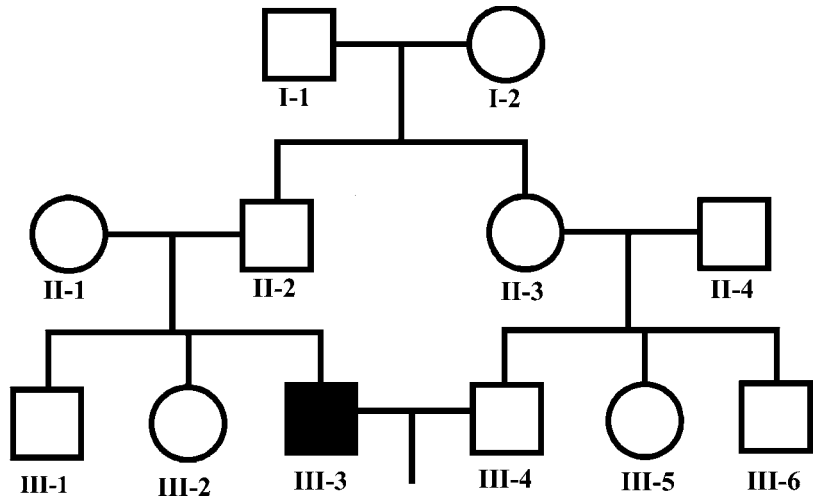


Fig. 4.4 Pedigree with consanguinity segregating an autosomal recessive disorder

Table 4.4 Hardy-Weinberg law for an autosomal dominant disorder

Genotype		
AA	Aa	aa
p^2 ↑	$2pq$ ↑	q^2 ↑
Affected	Affected	Unaffected

Note that p is the disease allele

may have an ambiguous genotype. They could have either the AA or Aa genotype (Table 4.4)

- Since individuals who are homozygous for the disease allele are very rare, the vast majority of affected individuals have the genotype Aa. Therefore, we can estimate the frequency of the disease allele by assuming that all affected individuals are Aa. Thus, the frequency of affected individuals is simply the frequency of heterozygotes, as estimated using the Hardy-Weinberg law. While the exact estimate of heterozygotes would be $2pq$, we note that the frequency of the normal allele (p) is close to 1, as we did for the autosomal recessive situation, so the frequency of the (heterozygous) affected individuals can be approximated as $2p$
- For example, the disorder dentinogenesis imperfecta is an autosomal dominant disorder. The frequency of affected individuals is 1/8,000. Therefore, $2p = 1/8{,}000$ and we estimate the frequency of the disease allele (p) to be 1/16,000. In other words, the frequency of an autosomal dominant disorder is twice the gene frequency
- In some instances, an individual may have an autosomal dominant disorder and there may not be any family history of the disorder. In this instance, it is likely that the individual who is affected has had a mutation in one of the alleles of the disease gene. This mutation occurred spontaneously in either the egg or sperm, which produced this individual. As a result, neither of the parents has the mutation in the other cells of their body, and hence they do not have the disease phenotype
- New mutations are more likely in disorders that are very severe, especially those that are lethal in childhood or prevent reproduction. Diastrophic dysplasia and osteogenesis imperfecta, type II, both of which are perinatal lethal conditions leading to severe bone and cartilage abnormalities, are virtually all the result of new mutations. New mutations are more likely in disorders that limit the ability to reproduce (i.e., reduce fitness of the individual). Specifically, if the disease remains at a constant frequency in the population, then those disease alleles that are lost as a result of the affected individual's decreased fitness must be replenished through new mutations
- As an example, individuals with achondroplasia have only 20% as many children as individuals of normal height. It is estimated that >80% of cases of achondroplasia are the result of new mutations. It appears that the frequency of mutation increases with paternal age

4.5 X-Linked Dominant Disorders

- When a disorder has X-linked dominant inheritance, both males and females may be affected with the disorder. Importantly, affected males would have a mutation only in their X chromosome while an affected female would typically have a mutation in at least one of their two X chromosomes
- Using the principles of the Hardy-Weinberg law, we can estimate the frequency of each genotypic group. Since males have only one X chromosome and females have two X chromosomes, we have to consider the frequency of the genotypes in the two sexes separately, as shown in Table 4.5
- In the left panel of Table 4.5, the frequency of affected males is simply the frequency of the mutant allele (p). The frequency of unaffected males is simply the frequency of the normal allele in the population (q). On the other hand, the frequency of unaffected females is the square of the frequency of the normal allele (q^2), as was true for an autosomal dominant disorder. Similar to the autosomal dominant situation, affected females can have one of two possible genotypes (AA or Aa), although

Table 4.5 Hardy-Weinberg law for an X-linked dominant disorder

Males		Females		
A	a	AA	Aa	aa
p ↑ Affected	q	p^2 ↑ Affected	$2pq$ ↑ Affected	q^2

Table 4.6 Hardy-Weinberg law for an X-linked recessive disorder

Males		Females		
A	a	AA	Aa	aa
p	q ↑ Affected	p^2	$2pq$	q^2 ↑ Affected

the vast majority will be heterozygous. Since most of the affected females will be heterozygotes, the frequency of affected females can be estimated as the frequency of heterozygotes, which is $2pq$, which is approximately $2p$. Using this approximation, for a rare, X-linked dominant disorder, the ratio of affected female to affected male will be about 2:1 (i.e., $2p$:p)

4.6 X-Linked Recessive Disorders

- When a disorder has X-linked recessive inheritance, the vast majority of the affected individuals will be male (Table 4.6). Similar to X-linked dominant inheritance, since males have only one X chromosome, the frequency of an affected male is simply the frequency of the disease allele in the population (q) (left panel, Table 4.6). Using the principles of the Hardy-Weinberg law, we can estimate the frequency of each genotypic group in the females. Females are only affected if both of their alleles are mutant; therefore, the frequency of affected females is q^2. Similar to autosomal recessive inheritance, females who are homozygous for the normal allele or heterozygous for the normal allele will be unaffected
- When considering a rare, recessive X-linked disorder, the frequency of homozygous females is the square of the frequency in males (i.e., q^2 vs q), and therefore homozygous affected females are extremely rare. Another important principle is that the frequency of heterozygous carrier females is approximately $2q$, or twice the frequency of affected males. Lastly, it should be noted that one-third of the mutant alleles are in males while two-thirds of the mutant alleles are in females. Therefore, there are twice as many carrier females as affected males

4.7 Bayes Theorem and Risk Calculation

- Bayes theorem is widely used to calculate the probability of an event, specifically when there is other information that can be used to modify or alter the initial probability
- The prior probability is the initial probability of the event, before considering any additional information that might modify or alter that estimate
- The conditional information is any information that can be used to modify or alter the initial or prior probability
- Table 4.7 presents the general format applied when utilizing Bayes theorem to modify the probability of an event
- See Boxes 4.4–4.6 for risk counseling examples

4.8 Gene Mapping and Recombination

- During the past two decades, there has been substantial progress made in elucidating the genes that contribute to many genetic disorders. This has been accomplished through the use of genetic linkage analysis, an experimental method in which many families having affected members are studied so as to identify a chromosomal region inherited by all affected family members in a family and which is not inherited by unaffected family members

Table 4.7 General format for applying Bayes theorem

	Probability (C occurs)	Probability (C does not occur)
Prior probability	P(C)	P(NC)
Conditional probability If O occurs	P(O/C)	P(O/NC)
Joint probability	P(C) × P(O/C)	P(NC) × P(O/NC)
Posterior probability	P(C) × P(O/C)	P(NC) × P(O/NC)
	[P(C) × P(O/C)] + [P(NC) x P(O/NC)]	[P(C) × P(O/C)] + [P(NC) x P(O/NC)]

C – event (i.e., individual is affected)
NC – event C does not occur (i.e., individual is not affected)
O – observation that if it occurs could affect the probability of the event (i.e., child born with disease, individual 40 years old and not affected)

Box 4.4: Risk Counseling Autosomal Dominant Disorder

- Consider the family shown in Fig. 4.5. The father has an autosomal dominant disorder that is fully penetrant
- The probability that a future offspring will be affected is 50%

Box 4.5: Risk Counseling Autosomal Dominant Disorder, Reduced Penetrance

- Consider the family shown in Fig. 4.5. The father has an autosomal dominant disorder
- This disorder is only 70% penetrant. This means that 70% of those who inherit the mutant allele will develop the disease. This also means that 30% of those who inherit the mutant allele will not develop the disease
- The probability that a future offspring will be affected is computed by multiplying the probability that the offspring will inherit the disease allele from the father (50%) by the probability that the offspring will be affected if they inherit the mutation (70%)
- Thus, the probability that a future offspring will be affected is 35%

Box 4.6: Risk Counseling Autosomal Dominant Disorder, Reduced Penetrance

- Consider the family shown in Fig. 4.6. The father has an autosomal dominant disorder
- His son (II-1) is not affected. This could be because he has inherited a normal allele from his father or it could be that he inherited the mutant allele, but he is among the 30% of individuals who despite inheriting the mutation are not affected
- We seek to estimate the probability that a future offspring of II-1 will be affected
- Table 4.8 summarizes the approach to calculate the probability that the offspring will be affected. To do this, we must first estimate the probability that II-1 has inherited the mutation even though he is unaffected
- As shown in Table 4.8, the prior probability before we consider the information that II-1 is not affected is 50% that II-1 would inherit the mutation from his affected father and 50% that he would not
- We then condition on the information that II-1 is not affected. It should be clear that this information will make it less likely that II-1 has inherited the mutation and more likely that II-1 has not inherited the mutation

(continued)

- The joint probability is simply the product of the two previous rows
- The posterior probability provides us with the updated probability that II-1 has or has not inherited the mutation after conditioning on the information that II-1 is unaffected
- As shown in the last row of Table 4.8, the probability that II-1 has inherited the mutation has decreased from the initial (prior) probability of 50–23%
- Similarly, the probability that II-1 has not inherited the mutation from his father has increased from the initial (prior) probability of 50% to 77%
- However, we have not yet answered the question what is the probability that II-1 will have an affected child. To address this, we must perform one more calculation
- The probability that II-1 will have an affected child requires the following events to each occur: II-1 must have inherited the mutant allele from his father, and II-1 must pass the mutant allele to his offspring, and then the offspring must be within the proportion of individuals who are affected when they have the mutation
- Mathematically, this is calculated by multiplying each of these probabilities (e.g., $0.23 \times 0.50 \times 0.70 = 0.0805$)
- This means that the probability that the offspring of II-1 will be affected is only 8.05%

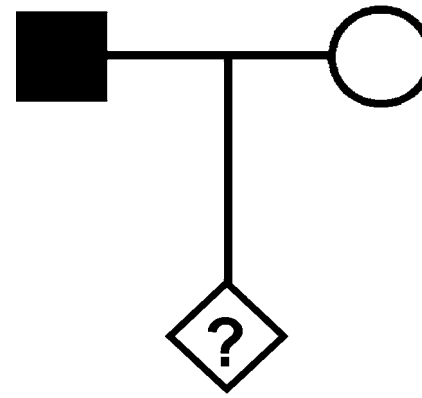


Fig. 4.5 Family risk counseling for autosomal dominant disorder (see Boxes 4.4, 4.5)

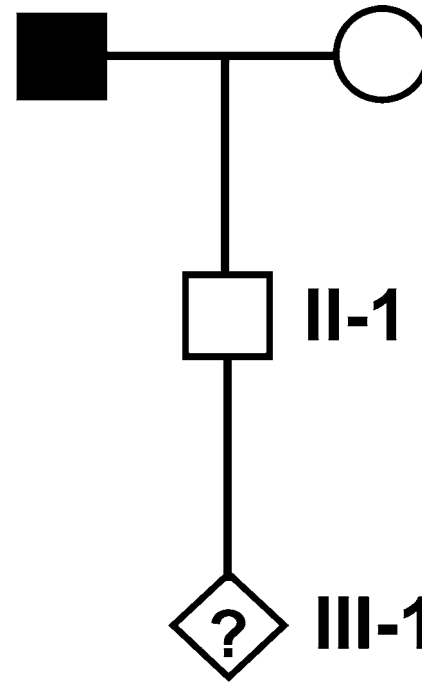


Fig. 4.6 Family risk counseling for autosomal dominant disorder (see Box 4.6)

- The basic principle underlying genetic linkage analysis is the study of the segregation of homologous chromosomes in meiosis. A gene, typically a disease gene, is considered linked to a chromosome if the gene of interest cosegregates with a gene or marker known to be on that chromosome. During meiosis I, homologous chromosomes pair together and crossover occurs (Fig. 4.7). On average, about 30–40 crossovers (about 1–2/chromosome) occur during a meiotic division. Genes that are close together on a chromosome will tend to be inherited together, violating Mendel's second law of independent segregation. Genes that are far apart on a chromosome will be more likely to have a crossover

Table 4.8 Using Bayes theorem to solve Box 4.6

	P (II-1 inherited the mutation)	P (II-1 did not inherit the mutation)
Prior probability	1/2	1/2
Conditional probability II-1 is not affected	P (not affected/inherited the mutation) = 0.30	P (not affected/did not inherit the mutation) = 1
Joint probability	(1/2) × (0.30) = 0.15	(1/2) × (1) = 0.50
Posterior probability	0.15/[0.15 + 0.50] = 0.23	0.50/[0.15 + 0.50] = 0.77

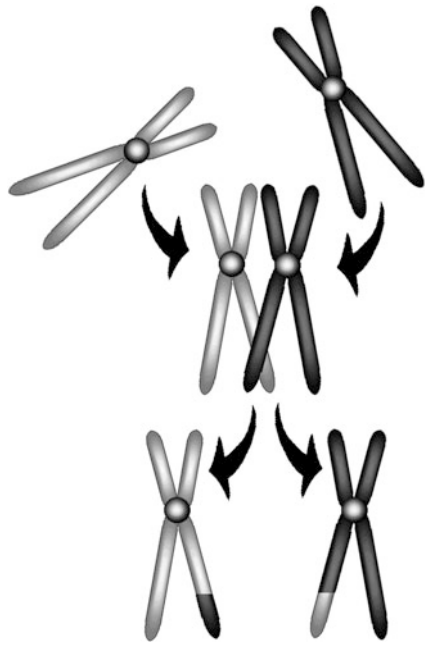


Fig. 4.7 Crossing over of chromosomes during meiosis

occur between them. Genes on different chromosomes will segregate independently of each other

- The observed result of crossing over is recombination, often denoted by the symbol theta (θ). The recombination fraction (θ) can take values between 0 and 0.5. When the recombination fraction is 0.0, this means the two loci are so close together on the chromosome that recombination never occurs between them. When the recombination fraction is 0.50, this indicates the two loci are segregating independently and either are far apart on the same chromosome or are on different chromosomes
- The estimate of recombination between two loci can be used to approximate the physical distance between the two loci. A recombination fraction of 1% corresponds to roughly 1 cM or one million bp of DNA
- The identification of common DNA sequence variation has rapidly altered the study of genetics. The identification of these sequence variants results in a polymorphism (Greek: many forms) of the DNA sequence. These sequence variations are commonly termed genetic markers and have become an indispensable tool in the study of genetics
- A disease gene can be mapped to a chromosome by studying the cosegregation of the disease gene with a marker of interest in families having individuals affected with the disease. This approach is termed the logarithm of odds (LOD) score method. Using the LOD score method, it is possible to determine a small region on a chromosome where a disease gene is located and sequentially refine its location by analyzing more markers in that region. The LOD score method will also determine how far a disease gene is from the marker being tested by estimating θ, the percentage of recombination between the disease gene and the marker. Thus, if a disease gene is estimated to be 1 cM from a marker, then there is typically recombination between the disease gene and the marker on 1% of the chromosomes during meiosis
- For all autosomes, that is, chromosomes 1–22, an individual has a pair of chromosomes, with one member of the pair maternally inherited and the other paternally inherited. When studying the segregation of marker and disease alleles, we are interested in determining which allele, at each of the two loci, is on the same chromosomal strand. The orientation of the four

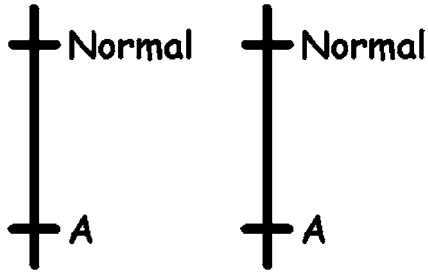


Fig. 4.8 Phase for a double homozygote

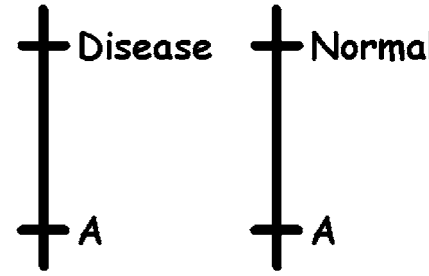


Fig. 4.9 Phase for a single heterozygote

alleles (two at the disease locus and two at the marker locus) on the pair of chromosomes is termed "phase" (see Fig. 4.8 for an illustration)

- Phase is an important concept in the LOD score method and refers to the alleles at each locus, which are on the same chromosome. It is essential that phase be known, or at least estimated, in order to determine whether or not recombination has occurred during meiosis. If a marker and disease gene are physically near each other on the same chromosome, recombination is less likely to occur between them. Through the examination of many meioses, it is possible to perform a statistical test, which will determine whether a disease gene and the marker being tested are close together on the same chromosome. This provides the first piece of information needed to determine the location of a disease gene within the human genome
- For individuals who are homozygous at both loci, it is easy to set phase, unambiguously. In fact, the two chromosomes are identical at these loci
- In some instances, the individual is heterozygous at only one of the two loci. For example, the individual could be heterozygous at the disease locus, which would correspond to an affected individual for an autosomal dominant disorder or a carrier for an autosomal recessive disorder. At the marker locus, the individual would then be homozygous (see Fig. 4.9). In a second example, the individual could be homozygous at the disease locus and heterozygous at the marker locus. For individuals who are heterozygous at only one of the two loci, phase can still be set unambiguously. Importantly, the two chromosomes will not be identical
- If the individual is heterozygous at the disease locus and also heterozygous at the marker locus, there are two possible ways to arrange the disease and marker genotypes (Fig. 4.10). These two possibilities are mutually exclusive
- The pedigree (Fig. 4.11) is segregating an autosomal dominant disorder. The first thing to consider in a pedigree of this type is what each person's genotype is at the disease locus. The affected individuals are all heterozygous at the disease locus (i.e., D/d) since this is an autosomal dominant disorder and the unaffected individuals are all homozygous (d/d). Once we know this, we can begin to determine the genetic phase of each individual
- It is typically easiest to start with those individuals whose phase is unambiguous. For an autosomal dominant disease, first assign the phase to the individuals who are unaffected or who are affected and are homozygous at the marker. This allows the phase of I:1, I:2, II:2, III:2, III:4, and III:7 to be determined
- Then, since you know that II:1 inherited his disease allele from his father, I:1, we can determine the phase for II:1. Similarly, since we know that II:2 must have given each of her children a normal allele at the disease locus and an A allele at the marker, the phase of all the affected children can now be determined
- Reviewing the pedigree, we see that individuals III:1, III:2, III:3, III:4, III:6, and III:7 all are nonrecombinant while individual III:5 has

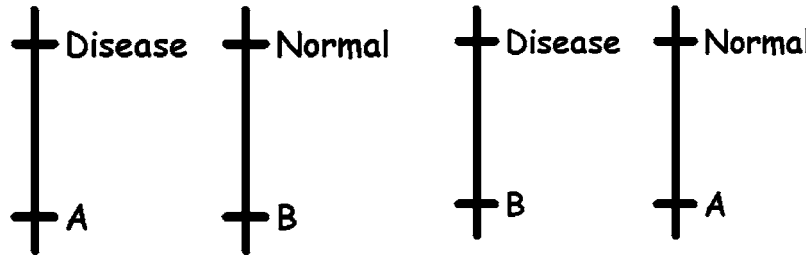


Fig. 4.10 Possible phase for a double heterozygote

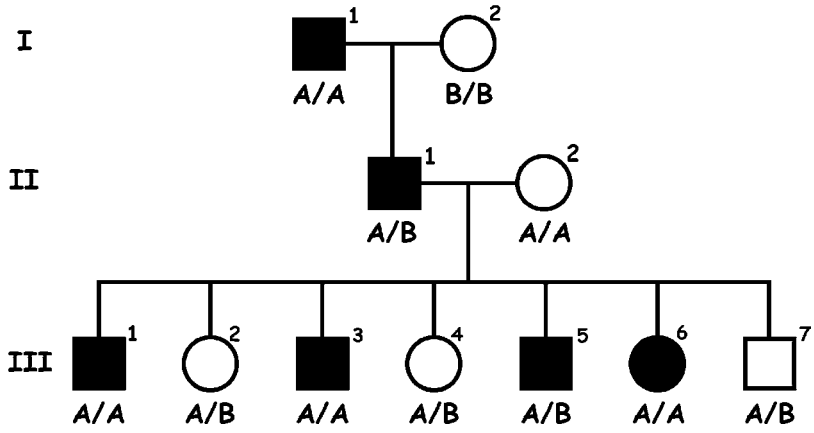


Fig. 4.11 Example 1 of an autosomal dominant disorder with marker data

had a recombination between the disease and the marker on the chromosome he inherited from his affected father

- When there is data on only two generations of individuals, the phase of the marker and disease locus cannot be determined unambiguously. For example, consider the pedigree shown in example 2, which is the same as the previous pedigree, but without generation I (Fig. 4.12)
- When we try to determine the phase of the unaffected mother, I:2, her phase is still known unambiguously, since she is not a double heterozygote. However, the phase of individual I:1 is now unclear. He can be either of the two phases shown in Fig. 4.10
- Now, when we consider the offspring in generation II, it is ambiguous whether they are recombinant or nonrecombinant. This can only be determined if we set the phase of the father, I:1. So, what we must do is determine for each child their status (recombinant or nonrecombinant) twice, once with the father's phase set as shown on the left in Fig. 4.10 and then again when the father's phase is set as shown on the right panel of Fig. 4.10

4.9 Disorders with Complex Genetic Inheritance

- There are a number of common disorders that clearly have a genetic contribution, such as autism, epilepsy, cardiovascular disease, and dementia, but which are not inherited in a simple Mendelian fashion
- These disorders are often considered to have complex genetic inheritance. It is hypothesized that multiple genes contribute to disease susceptibility and that the effect of a gene or genotype may vary depending on the environment in which the individual is placed
- The identification of the susceptibility genes or allele that contributes to these disorders will lead to new challenges for the field of genetics

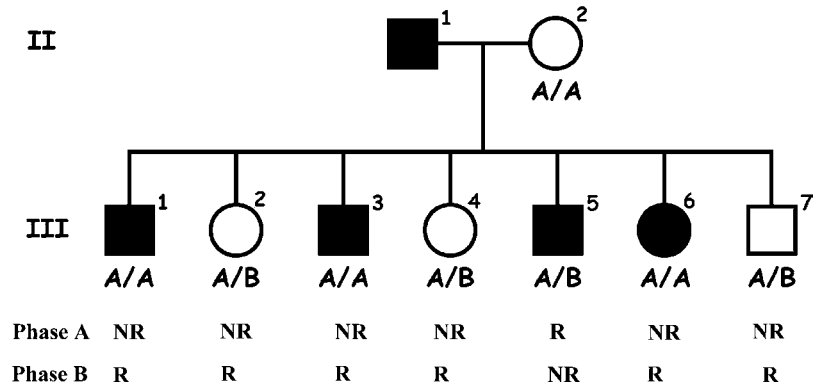


Fig. 4.12 Example 2 of an autosomal dominant disorder with marker data

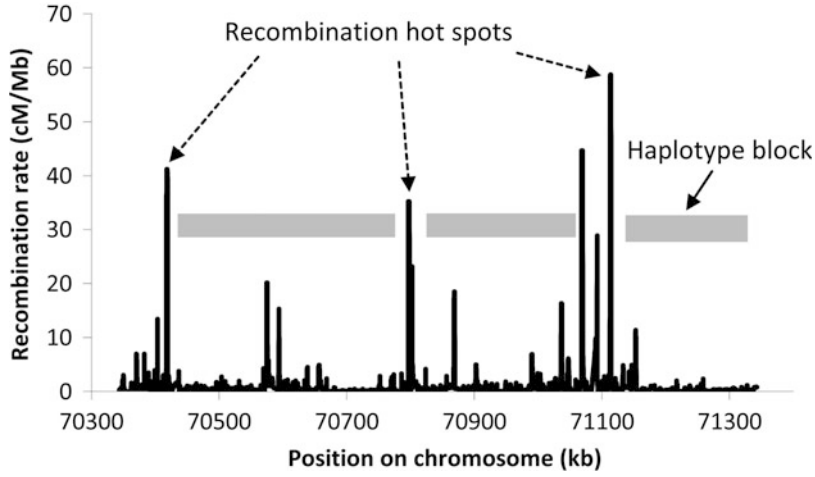


Fig. 4.13 Particular sites within the genome in which recombination is more likely (recombination hot spots)

and genetic counseling. Individuals who inherit a susceptibility allele(s) may be at an increased risk to develop the disorder, but it will be difficult to quantify the exact risk to that individual or their offspring

4.10 Genome-wide Association Studies

- With the sequencing of the human genome along with very detailed studies exploring variation in the genome across populations, researchers have learned a great deal about genomic architecture
- There are particular sites within the genome in which recombination is more likely (recombination hot spots), as well as regions in which recombination is less likely (cold spots). Little is currently known regarding prediction of whether a region will be likely to undergo more or less recombination based on sequence (Fig. 4.13)
- If recombination has occurred in a region over many generations, two markers within the region will be in linkage equilibrium. If recombination has occurred sparingly over many generations, then the two markers will not have achieved equilibrium and instead will be considered in linkage disequilibrium
- Linkage equilibrium can be assessed using statistical methods (Box 4.7)
- Recent developments in and automation of molecular methods have made it possible to genotype hundreds of thousands or even millions of SNPs simultaneously

Box 4.7: Estimating Equilibrium

- Assume there are two loci (locus A and locus B). They are near each other physically
- Locus A has two alleles, A1 and A2, with frequency 0.20 and 0.80, respectively
- Locus B has two alleles, B1 and B2, with frequency 0.60 and 0.40, respectively
- If the two loci are in equilibrium, the frequency of each haplotype (one allele at locus A and one allele at locus B) is simply the product of the frequency of the two alleles (see Table 4.9)
- If the two loci are not in equilibrium (linkage disequilibrium), then the frequency of the haplotypes will differ significantly from that expected if in equilibrium. This can be assessed statistically

Table 4.9 Estimation of frequency of haplotypes if the two loci are in linkage equilibrium

Haplotypes (allele at locus A and allele at locus B)	Frequency if in linkage equilibrium
A1B1	$0.20 \times 0.60 = 0.12$ (12%)
A1B2	$0.20 \times 0.40 = 0.08$ (8%)
A2B1	$0.80 \times 0.60 = 0.48$ (48%)
A2B2	$0.80 \times 0.40 = 0.32$ (32%)

- Most studies employing these genome-wide SNP arrays utilize a case control design
- A case control design compares SNP allele or genotype frequencies in a group of related cases with disease and a group of controls without disease (Fig. 4.14)
- The underlying hypothesis is that an SNP being genotyped will be in linkage disequilibrium with a disease risk factor (Fig. 4.15)
- Using this approach, the entire genome can be interrogated using 1–2.5 million SNPs

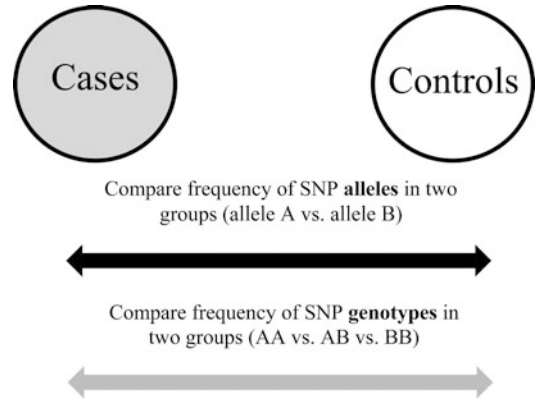


Fig. 4.14 Single nucleotide polymorphism (*SNP*) allele frequency and genotype frequency case control design comparison

- Most complex disease genome-wide association studies have required large samples of thousands of cases and controls to detect risk factors with relatively small odds ratios in the range of 1.1–1.5
- The genome-wide association study design has reasonable power to identify common risk factors of small effect in large samples, but relatively low power to detect rare risk factors, even those of moderate to large effect on disease risk (Table 4.10)
- There is great interest in developing SNP profiles that might be used to effectively screen patients to identify those disorders for which they are at greatest risk
- Efforts are already underway to implement screening for SNPs found to confer increased susceptibility to toxic drug side effects

4.11 Next-Generation Sequencing

- Another recent molecular development is the ability to perform in-depth sequencing using next-generation technologies
- Currently, sequencing of the entire exome (all exons in the genome) is becoming widely available, and it is anticipated that sequencing of the entire genome will become affordable shortly

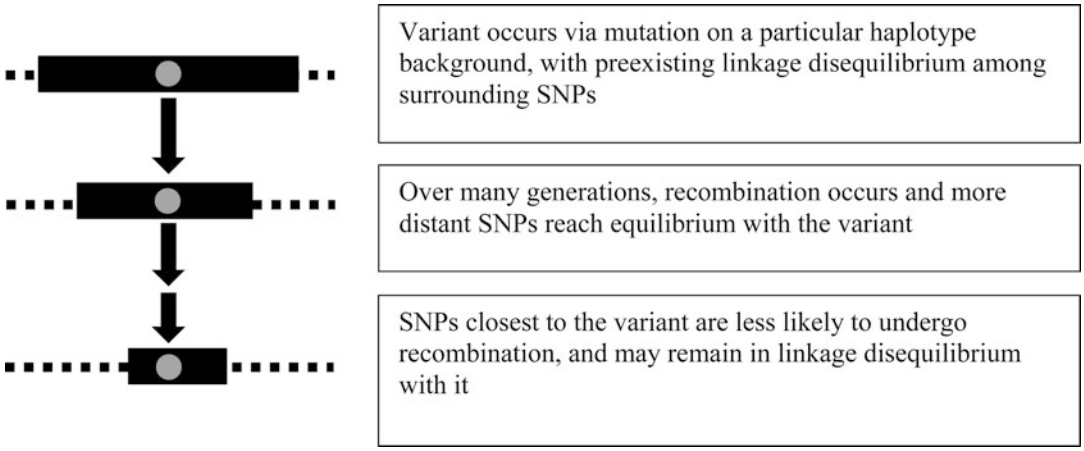


Fig. 4.15 Genotyped single nucleotide polymorphisms (*SNPs*) and linkage disequilibrium

Table 4.10 Comparison of linkage and association approaches

	Linkage	Association
Study design	Requires families with multiple affected members	Typically employs a series of unrelated cases and controls
Marker/ density	Microsatellites (5–10 million base pairs part) or SNPs every 500 K base pairs	SNPs with typically 1–2.5 million tested
Marker – risk factor	Need to be linked, which means they can be 20 million base pairs apart	Need to be in linkage disequilibrium, which means they should be within a few thousand base pairs to each other
Type of effect	Identify variants of large effect typically requiring variants in the same gene in many families to detect	Requires a common risk factor that has occurred on only a few haplotypes

- Whole-exome sequencing is currently being used to identify rare variants that result in Mendelian disorders
- This approach can be contrasted to that of genome-wide association studies that seek to identify common variants occurring on a common haplotype (Fig. 4.16)
- The greatest success thus far has been in the identification of novel genes resulting in disorders with recessive inheritance
- This is primarily because whole-exome sequencing identifies thousands of variants, and it is a bioinformatics challenge to narrow the variants to those that are novel and likely to be disease producing (Fig. 4.17)
- By focusing initially on recessive disorders, it is possible to require two variants in a gene as a means to rapidly narrow the number of variants and genes that must be further studied
- It is anticipated that next-generation sequencing will be widely applied to identify novel variants contributing to many different disorder

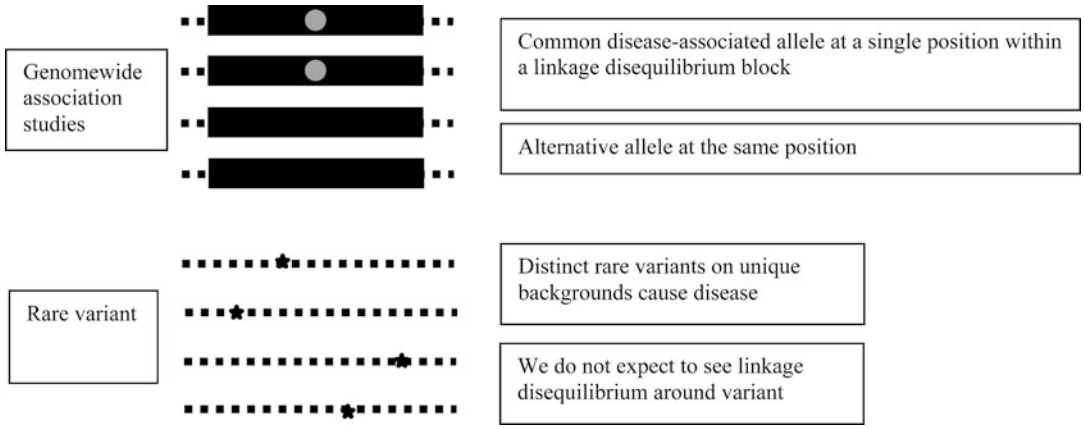


Fig. 4.16 Whole-exome sequencing for rare variant identification and genome-wide-association studies (*GWAS*) for common variant identification

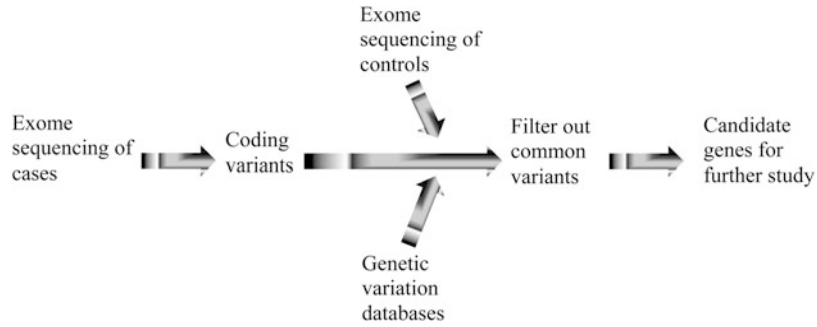


Fig. 4.17 Whole-exome sequencing bioinformatics challenge

Further Reading

Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science. 2008;322:881–8.

Bansal V, Libiger O, Torkamani A, Schork NJ. Statistical analysis strategies for association studies involving rare variants. Nat Rev Genet. 2010;11:773–85.

Cantor RM, Lange K, Sinsheimer JS. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. Am J Hum Genet. 2010;86:6–22.

Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet. 2010;11:415–25.

Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. Nat Rev Genet. 2010;10:241–51.

Gilad Y, Pritchard JK, Thornton K. Characterizing natural variation using next-generation sequencing technologies. Trends Genet. 2009;25:463–71.

Ioannidis JP, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. Nat Rev Genet. 2009;10:318–29.

Lambert SWJ, Uitterlinden AG. Genetic testing in clinical practice. Annu Rev Med. 2009;60:431–42.

Manolio TA. Genomewide association studies and assessment of the risk of disease. N Engl J Med. 2010;363:166–76.

Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010;11:685–96.

Nussbaum RL, McInnes RR, Willard HF. Thompson and Thompson genetics in medicine. 7th ed. Philadelphia: WB Saunders; 2007.

Clonality Analysis and Tumor of Unknown Primary: Applications in Modern Oncology and Surgical Pathology

5

Liang Cheng, Shaobo Zhang, Federico A. Monzon, Timothy D. Jones, and John N. Eble

Contents

5.1 Clonal Expansion Is the Hallmark of Neoplasia 130
5.1.1 Overview 130
5.1.2 Tumorigenesis Models 131
5.1.3 Patch Phenomenon 135

5.2 X Chromosome-Linked Clonality Analysis 136
5.2.1 Principle and Implication of X Chromosome Inactivation 136
5.2.2 X Chromosome Inactivation Control Mechanisms 138
5.2.3 Human Androgen Receptor Gene X Chromosome Inactivation Analysis 138
5.2.4 Other X Chromosome-Linked Clonality Analyses 144
5.2.5 Selected Applications 145

5.3 Loss of Heterozygosity (LOH) as a Clonal Marker 147
5.3.1 Overview 147
5.3.2 Evaluation and Interpretation of the LOH Analysis Results 147
5.3.3 Advantages and Limitations of LOH Analysis 149
5.3.4 Technical Considerations 149
5.3.5 Methods for LOH Analysis 150
5.3.6 Selected Applications 151

5.4 Other Methods of Clonality Analysis 153
5.4.1 Somatic Mutation 153
5.4.2 Gene Rearrangement Analysis 154
5.4.3 Cytogenetic Analysis and FISH 155
5.4.4 DNA Methylation as a Clonal Marker 156
5.4.5 Microsatellite Instability 157
5.4.6 Viral Integration Analysis 158
5.4.7 Comparative Genomic Hybridization (CGH) 159
5.4.8 Gene Expression Profiling/Array-Based Clonality Analysis 159
5.4.9 MicroRNA Signatures 159
5.4.10 Protein-Based Clonality Analysis 160

5.5 Tissue Contamination and Patient Identity Mismatch Testing 160
5.5.1 Overview 160
5.5.2 Technical Approaches 160
5.5.3 Caveats 163

5.6 Identification of Donor Origin in Transplantation Patients 164
5.6.1 Overview 164
5.6.2 Technical Approaches (Fig. 5.22) 164
5.6.3 Caveats 164

5.7 Bone Marrow Engraftment Testing 166
5.7.1 Overview 166
5.7.2 Technical Approaches 166
5.7.3 Caveats 166

5.8 Molecular Diagnosis of Hydatidiform Mole 167
5.8.1 Overview 167
5.8.2 Technical Approaches 168
5.8.3 Caveats 169

L. Cheng, MD (✉) • S. Zhang, MD • J.N. Eble, MD, MBA
Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

F.A. Monzon, MD
Departments of Pathology & Immunology and Molecular & Human Genetics, Baylor College of Medicine, Houston, TX, USA

T.D. Jones, MD
East Louisville Pathologists, PSC, Louisville, KY, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_6,

5.9 Cancer of Unknown Primary Origin (CUP) 169
5.9.1 Overview 169
5.9.2 Diagnostic Approaches 171
5.10 Summary 175
Further Reading 175

5.1 Clonal Expansion Is the Hallmark of Neoplasia

5.1.1 Overview

- Clonality refers to the principle that a group of cells are descended from and genetically or epigenetically identical to a single common ancestor, a stem cell
- Clonal proliferation is a fundamental characteristic of all human neoplasia
  - Tumors arise as a result of a series of genetic/epigenetic alterations occurring in a stem cell or progenitor cell
  - One cancer stem cell gives rise to daughter cells, which exhibit the same genetic/epigenetic alterations that initially provided a growth advantage to the stem cell
  - Further genetic alterations in subsequent daughter cells provide additional growth advantages
- Recent revolutionary progress in human genomics is reshaping the approach to cancer therapy and diagnosis
  - Molecular markers aid in detecting tumors at their earliest, most treatable stages
  - Patients are managed in accordance with genetic profiling
  - Novel approaches deepen the understanding of the biology of cancer and translate the knowledge into more effective and less toxic therapies
- The major principle for clonality analysis was proposed in the late 1970s, and the basis for which was the discovery of differential X chromosomal inactivation patterns in female patients and of a variety of somatic mutations
- The principle of clonality is evolving over time and so are the methods for clonality analysis
  - Traditional cancer models suggest that all tumor cells can form new tumors and are equally tumorigenic
    - Tumor is caused by a disruption of the regulatory mechanism of proliferation in general cells
      - Any cell type could be the target of carcinogenesis

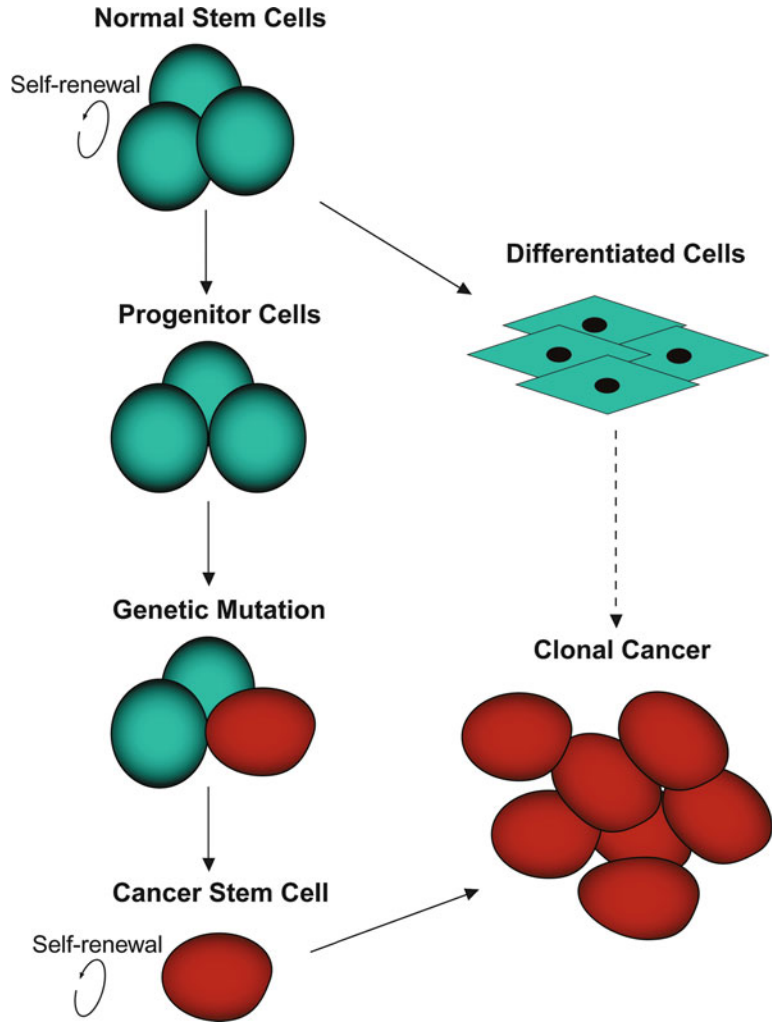


Fig. 5.1 Clonal proliferation is the hallmark of cancer. Normal stem cells renew themselves through asymmetric division to maintain the stem cell population and generate more differentiated cell populations. Carcinogenetic factors can affect stem cells in its self-renewal process, causing epigenetic and genetic alterations, and transform the stem cells into cancer stem cells. Cancer stem cells are the cellular source of cancer. These cells divide to expand the cancer cell population, forming a clonal cancer. The modern carcinogenesis theory suggests that cancer develops from a cancer stem cell, whereas the traditional carcinogenesis theory suggests that cancer can develop from any cell type, even differentiated cells as indicated by the *dashed line*

- The tumor is a clonal proliferation
- Modern cancer models propose that only a minority of tumor stem cells can form new tumors (Fig. 5.1)
  - Unregulated cell growth in tumors results from disruption of the regulatory mechanism of stem cell self-renewal
  - Thus, cancer is a regulatory disorder in stem cells and not simply an augmentation of proliferation signals
  - Only stem cells or progenitor cells are targets of carcinogenesis and possess the capacity for self-renewal (see also Chap. 3)
  - The genetic change is inheritable and shared by the entire population of cancer cells that are derived from the same stem cell origin
  - With prolonged carcinogen insults, part or all cancer stem cells may acquire new molecular alterations and form subclone

5.1.2 Tumorigenesis Models

5.1.2.1 Monoclonal Tumor Model

- Genetic damage or epigenetic alterations play critical roles in carcinogenesis
- Carcinogenesis is a multistep process at both the phenotypic and genetic levels
- Many tumors have been reported to result from the clonal expansion of a single stem

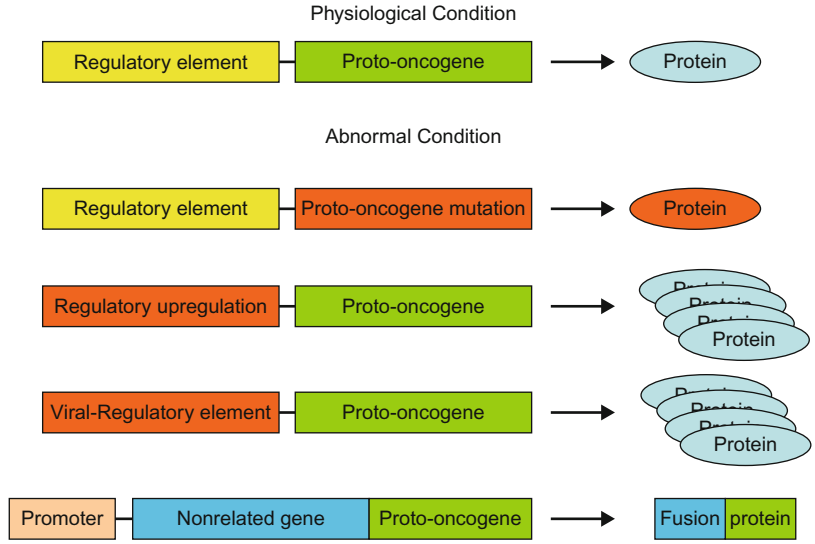


Fig. 5.2 Proto-oncogenes are functional genes in physiologic conditions. Normal proto-oncogenes are expressed in a well-controlled way and produce normal amounts of protein product (aqua-coded oval). An oncogene is a modified proto-oncogene, which produces either an increased amount of normal protein (*aqua*) or a mutated form of the protein (*orange*), which is involved in cancer development. Upregulation of oncogene expression can be induced by mutation of the regulatory fragment or by incorporation of viral DNA, which may provide a stronger promoter and increased expression of a normal form of oncoprotein (*aqua*). Also if the coding sequence is linked to another gene sequence by a chromosomal translocation, a fusion protein may result (*blue* and *green*)

cell founder (cancer stem cell); thus, the tumor is composed of a clonal population of tumor cells
- The monoclonal cancer theory is appealing in that it agrees with many known molecular events in tumorigenesis
- Somatic mutations and/or epigenetic alterations acquired by cancer stem cells can accumulate and be transferred from the cancer stem cell into daughter cells to form a tumor clone with the same genotype and phenotype among all cells
- Because the mutations acquired by the cancer stem cell are stably passed to its progeny, the presence of these mutations could be used as clonal markers
- The clonal tumor cells exhibit a concordant pattern of somatic mutations and/or epigenetic alterations
- Activation of an oncogene and deletion/inactivation of a tumor suppressor gene in a cancer stem cell and its progeny are common pathways in carcinogenesis (Fig. 5.2)
- Activation of antiapoptosis genes and inactivation/loss of apoptosis genes are also characteristics shared by clonal cancers
  - Simultaneous upregulation of antiapoptosis genes and downregulation of apoptosis genes result in a selective kinetic advantage and the expansion of a dominant clone
- Epigenetic disorders alter the expression levels of related genes during carcinogenesis
  - The epigenetic status is often consistent among a clonal population of cells and can, thus, be used in clonality analysis
  - The most important epigenetic process is mediated through DNA methylation, which silences genes

5.1.2.2 Multistep Carcinogenesis

- Multistage carcinogenesis is a stepwise accumulation of gene/epigenetic aberrations in the genome of susceptible stem cells
- In 1997, Kinzler and Vogelstein proposed the concept of two different types of

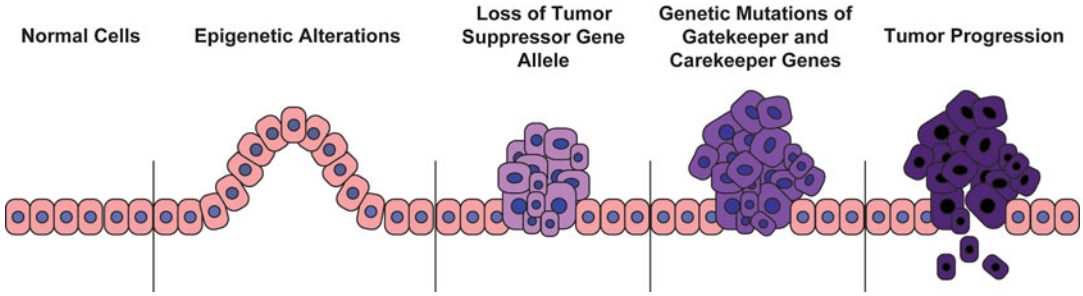


Fig. 5.3 Multistep carcinogenesis theory involves several epigenetic and genetic alterations. Epigenetic alterations and loss of tumor suppressor gene alleles are thought to be important initial steps in carcinogenesis. The accumulation of genetic mutations involving gatekeeper and caretaker genes will eventually initiate cancer development and eventually lead to a more advanced malignancy

carcinogenetic genes, gatekeepers and caretakers (Fig. 5.3)

- Gatekeepers are a class of genes that directly regulate the growth of tumors by inhibiting cell growth or promoting cell death. Known gatekeeper genes include *APC*, *TP53*, β-*catenin, Rb, NF1, VHL*, and others
  - The tumorigenesis is initiated by disruption of a tissue-specific gatekeeper pathway
  - Inactivation of specific gatekeeper in a specific cell type leads to a specific cancer
  - In the susceptible individual with one defect allele in a gate keeper gene, the gatekeeper pathway only need one additional mutation, which inactivates the remaining allele to initiate a cancer
- Caretakers are genes that maintain the integrity of the genome
  - Inactivated caretaker gene does not directly promote tumors but results in an increased mutation rate affecting all genes including gatekeeper genes
  - Caretakers include nucleotide excision repair genes, mismatch repair genes, *ATM* genes, *BRCA1* and *BRCA2* genes

- Three subsequent somatic mutations are required to initiate a cancer: mutation of the remaining caretaker allele and mutations of both gatekeeper alleles
- Cancer incidence increases with age, supporting the theory of multistep carcinogenesis
- Multistep carcinogenesis is a progressive event in which the sequential accumulation of somatic mutations transforms the stem cell into a cancer stem cell, which clonally expands into a cancer
  - Along with the accumulations of mutations, the affected cells also experience a morphologic transition from normal histology to hyperplastic/dysplastic lesions and finally to cancer
- The genetic changes are inheritable and shared by the entire population of cancer cells that are derived from the same stem cell origin, referred as clonal

5.1.2.3 Polyclonal Tumor Model

Field Cancerization Theory

- The concept of field cancerization was first introduced by Slaughter et al. in 1953
  - Slaughter proposed that the multicentric origin and high probability of cancer recurrence in the oral mucosa suggests a microscopically invisible "field" where genetically altered precancerous cells exist
- The theory further suggested an increased susceptibility of entire affected "field" to carcinogenesis within which the cells tend to develop cancers
  - Environment insults or carcinogens involve stem cells within a larger area of organ surface leading to damage of the cells in multiple sites
  - Field cancerization described the forming of premalignant states of clonal cell patches within the entire exposed surface

Table 5.1 Monoclonal versus polyclonal origin theories of tumorigenesis

	Monoclonal origin	Polyclonal origin
Founder	Single transformed cancer stem cell	Multiple transformed cells in the field
Tumor clone	Single	Multiple
Somatic mutations	Similar mutations in all cancer cells	Different mutations
Multifocal development	Implantation, migration	Locally transformed
Mechanism of carcinogenesis	Monoclonal proliferation	Field cancerization

  - Multiple cellular patches of premalignant lesions formed in the field lead to a higher prevalence of multiple local second primary tumors in an anatomical location synchronously or metachronously
  - Ultimately, each of the affected stem cells in a proliferating field will expand, gradually replace the normal epithelium, and develop into clonal cancer
  - Distinct X chromosomal inactivation patterns and discordant patterns of genomic mutations in the same tumor (or in separate tumors in the same patient) provide supporting evidence for the hypothesis
- The theory explains the development of multiple primary tumors and of locally recurrent tumors
  - Field cancerization generates multifocal areas of cancer development from multiple genetically distinct clones due to carcinogenic events (Table 5.1)
  - Each cancer clone originates from a different cancer stem cell and bears different genetic alterations
  - The cancer clone can form tumors synchronously or metachronously within the field
- Many organ systems have been studied including oral cavity, lung, esophagus, breast, skin, urinary bladder, vulva, and colon
  - Molecular findings have supported this field cancerization model in that genetic alterations can be detected in many fields in which the cells acquire genetic alterations and grow to form a patch (see Sect. 5.1.3)
  - External carcinogens cause independent genetic alterations at different sites leading to the development of multiple, genetically unrelated tumors
- The important implication is that some effects of the carcinogens in the field remain after the primary cancer is removed and may lead to new cancers
- Field carcinogenesis assumes a multistep process in which neoplastic changes evolve over a period of time due to the accumulation of somatic alterations in a single-cell lineage
  - One of the important evidence supporting field cancerization is that genetic alterations are often found within the morphologically normal cells surrounding the tumor in the same field
  - These morphologically normal cells adjacent to the tumor occupy an intermediate step in the transformation from phenotypically normal to a dysplastic epithelium
- Separate origination of multiple tumors suggests a concept of generalized carcinogenesis, through which a larger area or even the whole organ is affected by the carcinogens
- Although multiple tumor clones may exist during the carcinogenetic process, one clone may become dominant due to an additional growth advantage resulting in a pseudomonoclonal appearance on clonality analyses
- The genetic alterations may be found long before histologic evidence of cancer development, and the cells bearing the alterations may or may not themselves be precursors of cancer

5.1.2.4 Clonal Evolution in Cancer

- Clonal evolution theory introduced the genetic dynamic over time frame
- Carcinogenesis involve one or a few transformed stem cells
- Carcinogenesis is a process that involves clonal expansion, genetic diversification, and clonal selection
  - Successive subclones formed after acquiring additional genetic–epigenetic alterations with the genetic background inherited from their founder cell

  - The selective pressures from environment allow some subclones to expand while others become extinct or remain dormant
    - A cancer is clonally heterogeneous, and the genetic makeup of cells in a cancer may not be completely identical to their founder cell or subclonal counterparts
    - The cancer is formed by complex of cells with genetic heterogeneity among cancer cells that leads to heterogeneity in phenotype, function, and response to therapy
    - Cancers evolve over time and represent different subclones that occupy in different histologic and anatomic compartment

5.1.2.5 Histologically Distinct Colocalized Tumors

- The exact histogenesis of histologically distinct colocalized tumors remains a matter of controversy
- Two opposing theories, based on clonality, have been proposed to explain the origin of these morphologically diverse, biphasic tumors
  - Some investigators believe that it represents a collision tumor of two independent, monoclonal neoplasms occurring simultaneously
- Others suggest that these tumors have a common clonal origin with divergent differentiation into histologically divergent tumors

5.1.2.6 Random Collision Theory

- The collision theory is primarily based on morphologic analysis. Such tumors frequently have abrupt transition between divergent components without a gradual transition
- Collision tumors have been reported in various organs, and they represent a coexistence of two adjacent but histologically distinct tumors (without histologic admixture) in an organ
- Random collision theory proposes the possibility that two distinct tumor types initiated in a close proximity can result in a polyclonal neoplasm that may be recognized clinically as a single tumor
- Collision tumors have been reported in the stomach, lung, esophagus, ovary, and intestine
- Such tumors are polyclonal and display different genetic alterations in each component

5.1.2.7 Divergent Differentiation Theory

- Divergent differentiation theory proposes that colocalized tumors with distinct histology has a common clonal origin, for example, sarcomatoid carcinoma
  - The supporting evidence for this hypothesis is that both the components retain concordant genetic alterations
  - One tumor type may progress through multistep carcinogenesis with the accumulation of genetic alterations, genetic instability, and generation of multiple subclones
- Tumor stem cells depend on both genetic and microenvironmental alterations for clonal expansion and phenotypic differentiation
  - One tumor type may represent transdifferentiation from one phenotype to another secondary to molecular programs, induced by either internal or external environment

5.1.3 Patch Phenomenon

- Patch phenomenon is a concept introduced by Schmidt and Mead in the 1990s through the use of X chromosome inactivation analysis
  - A patch is generally regarded as a group of cells which are derived from a common stem cell, sharing common genetic characteristics and having inactivation of the same X chromosome in female individuals (Fig. 5.4)
  - Patch size varies. The patch size in bladder and stomach is up to 1 cm^2 and in hair follicle is 0.2–1.0 cm^2. In other tissues, the patch size is about 0.2–0.3 cm^2
  - Some patches correspond to anatomic boundaries such as an intestinal crypt or breast duct–lobular unit; however, some do not follow anatomic boundaries, such as patches in the liver
  - Because of potentially large patch sizes, cells isolated for X chromosome inactivation analysis may come from a single patch.

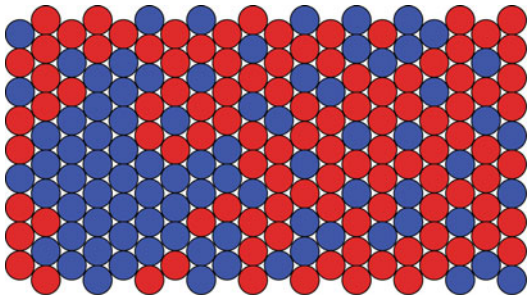


Fig. 5.4 A patch is a cluster of cells derived from the same founder cell. Cells in female tissues demonstrate a mosaic pattern of methylated X chromosomes (from either maternal or paternal origin). In this illustration, the *red* and *blue* cells represent cells with inactivated X chromosomes of maternal and paternal origin, respectively. Sometimes the cells can grow in clusters, remaining adjacent to one another as a result of clonal growth in a discrete territory referred to as a patch. The cells in a patch share identical genetic characteristics. The cluster of blue cells in the *left* lower corner constitutes a patch, all of which show inactivation of the paternally derived X chromosome

A multiple site cell isolation technique, in which much larger areas can be sampled, may be helpful in achieving more accurate results in these clonality analyses
    - The terminal lobuloductular unit of breast often lies entirely within one patch; therefore, a polyclonal origin of breast tumors may be difficult to demonstrate unless a proper normal control is used
    - Because of the large patch sizes in colon epithelium, X chromosome inactivation studies are heavily biased toward monoclonal results
- Patch phenomenon in field carcinogenesis
  – In field carcinogenesis, many stem cells in the field acquire somatic alterations and divide to form genetically altered cell clusters as patches
  – These altered cells in the patch may or may not be morphologically recognizable, but all have identical genomic changes
    - These patches can be recognized on the basis of genetic mutations such as *TP53* mutations or abnormal epigenetic processes such as nonrandom X chromosome inactivation
    - The cells in a patch may be more susceptible to carcinogens and to undergoing malignant transformation as somatic mutations accumulate
  – A single patch may be formed from the progeny of one or several stem cells that show the same genetic alteration pattern
    - The cells in these patches are monophenotypic but may or may not be monoclonal
- The patch proliferation–malignant transformation model is one of the basis for multifocal, multiclonal carcinogenesis (i.e., more than one genetically unrelated primary tumors)
- To truly demonstrate monoclonality, postulated somatic genetic changes have to be directly demonstrated rather than inferred on the basis of apparent X chromosome inactivation data

5.2 X Chromosome-Linked Clonality Analysis

5.2.1 Principle and Implication of X Chromosome Inactivation

- X chromosome inactivation (also called lyonization) is a process proposed by Mary Lyon in 1961. It is a process of chromosome-wide epigenetic gene silencing by which one of the two copies of the X chromosome present in female mammals is randomly and permanently inactivated (for purposes of dosage compensation)
- The inactive X chromosome is methylated, followed by histone deacetylation, resulting in compaction of the chromatin into repressive heterochromatin, forming a Barr body (also called sex chromatin)
- X chromosome inactivation occurs in early embryogenesis (blastocyst stage) and is permanent (Fig. 5.5). The process is random, and either the maternally or paternally derived X chromosome is inactivated
- Once established, the same inactivated X chromosome is stably maintained and passed to daughter cells through all subsequent cell divisions

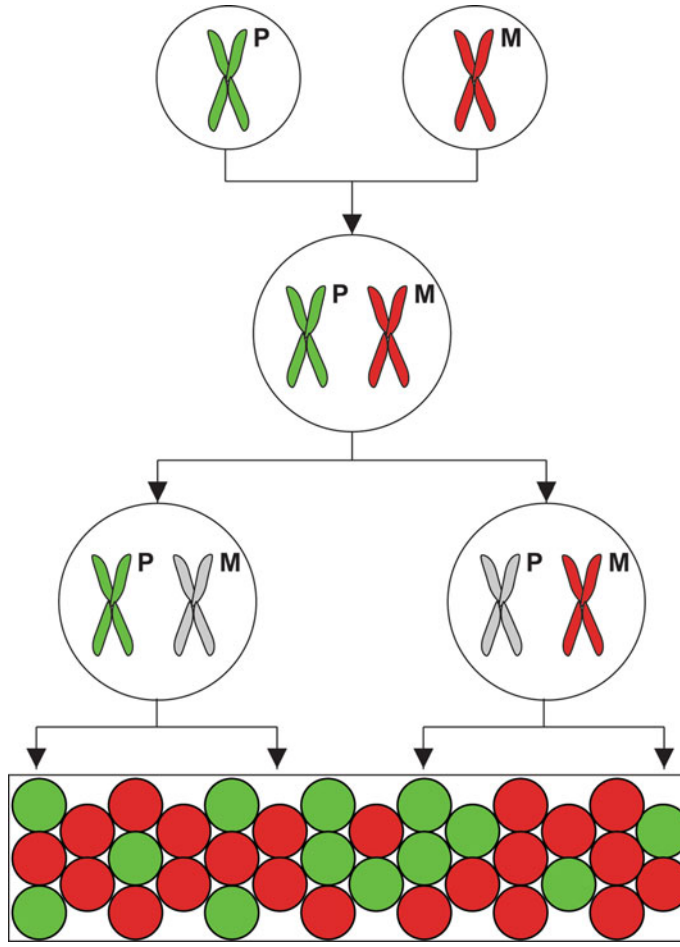


Fig. 5.5 X chromosomes of female cells are of paternal (P, *green*) and maternal (M, *red*) origin. In early embryogenesis, the paternally and maternally derived X chromosomes are both in an activated state. One of the X chromosomes in each cell becomes randomly inactivated after early embryogenesis (inactivated X chromosomes are *gray*). Normal female tissues are composed of a mosaic of cells having either the maternal or paternal X chromosome inactivated (*red* and *green* cells in lower box represent cells having an active maternal or paternal X chromosome, respectively). Theoretically, 50% of cells will have an inactivated paternal X chromosome and 50% of cells will have an inactivated maternal X chromosome. The X chromosome that is inactivated in a cell will be stably passed to its progeny, allowing this feature to be used as a clonal marker

- Analysis of the differential methylation (inactivation) of X chromosomes forms the basis of clonality analysis in women
- Cells derived from a common progenitor share the same inactivated X chromosome in all progeny
- X chromosome inactivation is normally a random process with approximately equal numbers of maternally and paternally derived X chromosomes being inactivated in the female
- As a consequence, normal tissues are composed of cellular mosaics with random X chromosome inactivation patterns. In contrast, a clonal expansion of cells, such as that present in tumors, exhibits a nonrandom pattern of X chromosome inactivation in all cells. Tumors in a female arising from a single progenitor cell have the same inactive X chromosomes (Fig. 5.6)

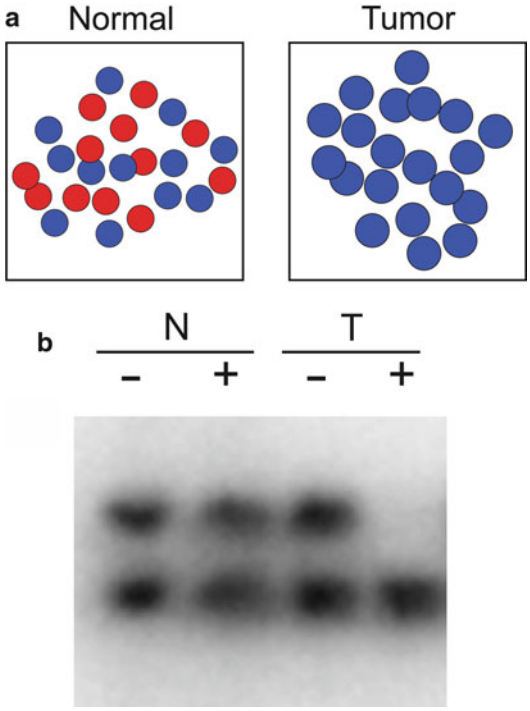


Fig. 5.6 Tumor cells are derived from a common progenitor cell and, therefore, share the same X chromosome inactivation pattern. Panel A is a schematic illustration showing random (normal) and nonrandom (tumor) X chromosome inactivation. Panel B shows a gel picture of X chromosome inactivation analysis. N designates normal control and T designates tumor. + and − indicate with or without HhaI methylation-sensitive restriction enzyme digestion. The *upper* allele of tumor could not be amplified after HhaI digestion, indicating a nonrandom X chromosome inactivation pattern. Two bands are seen in normal tissue after HhaI digestion, consistent with randomly inactivated X chromosomes in the constituent cells

5.2.2 X Chromosome Inactivation Control Mechanisms

- The X inactivation center (XIC) is located at Xq12–13 on the X chromosome, and it is necessary and sufficient to cause X chromosome inactivation
  - The XIC contains four X chromosome inactivation involved nontranslated RNA genes including *XIST*, *TSIX*, *JPX*, and *FTX*
  - The XIC also contains binding sites for both known and unknown regulatory proteins
- The X chromosome-inactive specific transcript gene (*XIST*) encodes a large nontranslation RNA that mediates the specific silencing of the X chromosome from which it is transcribed
  - The inactive X chromosome (Xi) is progressively coated by *XIST* RNA, but the active X chromosome (Xa) is not
    - The silencing of genes along the Xi occurs soon after coating by *XIST* RNA
  - X chromosomes lacking of *XIST* gene cannot be inactivated
- *XIST* antisense RNA gene (*TSIX*) encodes a large RNA which is antisense to *XIST*
  - *TSIX* is a negative regulator of *XIST*
  - *The TSIX* gene transcripts the opposite strand of DNA from the *XIST*; thus, the transcripts are antisense to *XIST*
  - X chromosome lacking *TSIX* expression is inactivated. Abnormally express the *XIST* gene on another chromosome leads to silencing of that chromosome
  - The process is initiated by transcription and cis localization of the noncoding *XIST* RNA, which then recruits much of the epigenetic machinery associated with maintenance of constitutive heterochromatin and silencing of genes on the inactivated X chromosome (e.g., histone modifications and DNA methylation)
- Inactivated X chromosome has high levels of DNA methylation, low levels of acetylated histone H4, low levels of histone H3 lysine-4 methylation, and high levels of histone H3 lysine-9 methylation

5.2.3 Human Androgen Receptor Gene X Chromosome Inactivation Analysis

- The human androgen receptor gene (*HUMARA*) X chromosome inactivation analysis is based on methylation-sensitive restriction enzyme that can only cut nonmethylated restriction site (Fig. 5.7)
  - The *HUMARA* is located at Xq11.2–12
  - A highly variable region of CAG trinucleotide repeats is located within the first exon of the gene

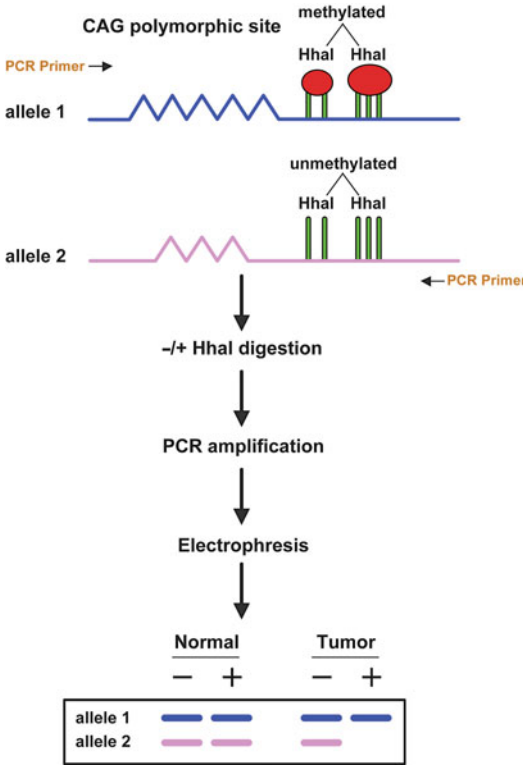


Fig. 5.7 Schematic illustration showing the HUMARA X chromosome inactivation assay. Genomic DNA is isolated from normal tissue and from tumor and is subjected to methylation-sensitive restriction enzyme HhaI digestion. Allele 1 (*blue*) and allele 2 (*purple*) are from different parental origins. Allele 1 is methylated and cannot be digested; allele 2 is nonmethylated and, thus, can be digested. Digested (+) and nondigested (−) DNA is PCR amplified and separated by gel electrophoresis. The digested normal tissue shows a double-band pattern since its X chromosomes are randomly inactivated. The digested tumor DNA shows only one band since its X chromosomes are nonrandomly inactivated. The X chromosome inactivation pattern is shared among the cells in the clonal population

  - These repeats reside close to 5′ promoter region of the *HUMARA* gene which contains proximity to differential methylation sites and to methylation-sensitive restriction endonuclease (*Hha*I or *Hpa*II) cutting sites
  - Studies demonstrated that methylation of *HpaII* and *HhaI* sites near these polymorphic repeats are correlated with X chromosome inactivation
  - Alternately unmethylated cytosines can be modified by sodium bisulfite, and the cytosine is changed to uracil, whereas methylated cytosines remain unaffected
  - PCR (polymerase chain reaction) using methylation-specific primers could differentiate the methylated DNA from nonmethylated
- Brief procedures
  - Genomic DNA is extracted from microdissected normal and tumor tissues
  - The genomic DNA is digested using methylation-sensitive restriction enzymes *Hha*I or *Hpa*II
  - The methylated (inactivated) X chromosome cannot be digested by methylation-sensitive restriction enzymes *Hha*I or *Hpa*II and thus can be amplified
  - The unmethylated (active) X chromosome was cut and thus cannot be amplified
  - The PCR products are subjected to electrophoresis to visualize the DNA fragments
- Interpretation of the X chromosome inactivation analysis results
  - Clonality analysis of a cell population can be interpreted correctly only in relation to the clonality of surrounding normal tissue of the same embryologic origin
  - Informative case: two allelic bands were present in normal DNA after PCR amplification
  - Noninformative case: only one allelic band was present in normal DNA after PCR amplification
  - In informative cases, the PCR products from digested normal control tissue show a double band on electrophoresis (random X chromosome inactivation), while digested clonal tumor DNA shows only one band (nonrandom X chromosome inactivation)
  - A nonrandom X chromosome inactivation pattern indicates clonal proliferation; a random X chromosome inactivation pattern indicates a polyclonal process (mosaic cell populations)
  - Identical X inactivation patterns in two distinct tumors (with either upper or lower

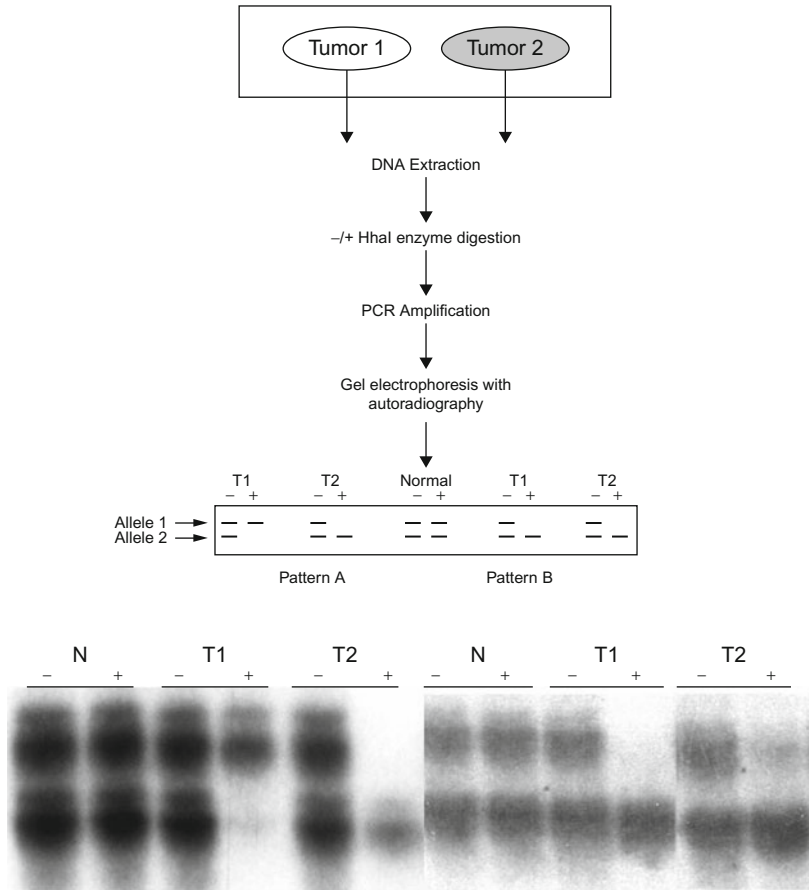


Fig. 5.8 X chromosome inactivation analysis can be used to analyze clonal relationships of separate tumors. Genomic DNAs from a normal control and from two separate tumors are isolated and subjected to a methylation-sensitive restriction enzyme digestion followed by PCR amplification and gel electrophoresis. In pattern A the two tumors (T1 and T2) display an opposite patterns of nonrandom X chromosome inactivation, indicating different clonal origins. Pattern B shows identical patterns of nonrandom X chromosome inactivation in each tumor, indicating a common clonal origin. As expected, normal control tissue (N) displays a random pattern of X chromosome inactivation with two bands present on gel electrophoresis before and after methylation-sensitive enzyme digestion

bands diminished) suggests a possible common monoclonal origin (Fig. 5.8)
  - Opposite X inactivation patterns, such as loss of upper band in one tumor and loss of lower band in another tumor, support an independent origin for each tumor

5.2.3.1 Advantages and Limitations of HUMARA X Chromosome Inactivation Clonality Analysis

- Advantages
  - X chromosome inactivation is the most consistently informative marker of clonal proliferation in women
  - 90% of females are heterozygous and suitable for X chromosome inactivation analysis
  - The proximity of differentially methylated restriction sites to the loci of genes allows for the use of these techniques on fragmented DNA (the size of PCR products is about 200–280 bp)
  - The method is simple and the result is stable
- Limitations
  - X chromosome inactivation is only applicable to women
  - Inherited or acquired unbalanced methylation (skewed, nonrandom pattern of X chromosome inactivation) and abnormal patterns of DNA methylation in malignancies can complicate the interpretation of results
  - Strict dosage compensation may not be necessary for all genes in the X chromosome
    - 15% of X-linked genes escape inactivation and an additional 10% of X-linked genes show variable patterns of inactivation

  - X-linked clonality analysis only distinguishes random and nonrandom X chromosome inactivation
  - Not all clonal proliferations are neoplastic. Clonal processes are not equivalent to neoplastic processes

5.2.3.2 Technical Considerations

Skewed DNA Methylation

- The process of X chromosome inactivation is usually random, producing tissues with equal mixtures of cells having active X chromosomes of either maternal or paternal origin
- However, skewed X inactivation patterns can occur, and this skewing can be inherited or acquired and can result from complex mechanisms of action during the early phase of embryonal life (Table 5.2)
- Skewed methylation is an asymmetric distribution of inactivated maternally or paternally derived X chromosomes (Fig. 5.9)
- X chromosome inactivation with a 3:1 ratio of inactivated Xp:Xm or vice versa is accepted as skewed X chromosome inactivation
- Primary skewed methylation patterns are related to the limited number of stem cells present at the time of random X chromosome inactivation during embryogenesis
  - Asymmetric division of stem cells is another cause of skewed methylation due to loss of some stem cells through terminal differentiation
  - Decreased precursor pool size contributes to skewed X chromosome inactivation
  - Monozygotic twins have an excessive skewing rate which may be related to the twinning process. The twinning process reduces the number of cells contributing to the embryo
- Skewed methylation patterns can be influenced by hereditary factors and can be genetically transmitted
  - Family concentrations of highly skewed methylation patterns with preferential activation of one parental X chromosome have been reported

Table 5.2 Mechanisms of skewed X chromosome inactivation in normal tissue

Inherited
Dysregulation, mutation, or deletion of *XIC* at Xq12–13
Xq28 deletion
Xist mutation
XCE human equivalent of mutant murine Xce
Decreased precursor pool size
Monozygotic twinning
Confined placental mosaicism
Carrier state of X-linked diseases
Agammaglobulinemia
Severe combined immunodeficiency
Wiskott–Aldrich syndrome
Adrenoleukodystrophy
Incontinentia pigmenti type II
α-Thalassemia with mental retardation syndrome
Duchene muscular dystrophy
HPRT deficiency
Facial dermal hypoplasia
Dyskeratosis congenita
Acquired
Selection and growth advantage
Aging process
Tissue specific (e.g., hematopoietic cells)
Artifact
Artificial allelic dropout
Reduction of DNA template quantity
Damaged or salt-contaminated DNA
Different laboratory criteria for skewing

- Mutation or abnormal imprinting of the *XIST* gene can result in skewed inactivation
- Dysregulation of XIC or physical deletion of Xq28 (pseudoautosomal region) could cause skewing of X chromosome inactivation
- *XIST* mutations are also directly related to skewing of X chromosome inactivation
- Several X-linked disorders are known to cause skewing of X chromosome inactivation (Table 5.2)
- Acquired skewing of X chromosome inactivation may also be related to selection and the aging processing
  - Age has been shown to have a direct influence on inactivation. A much higher

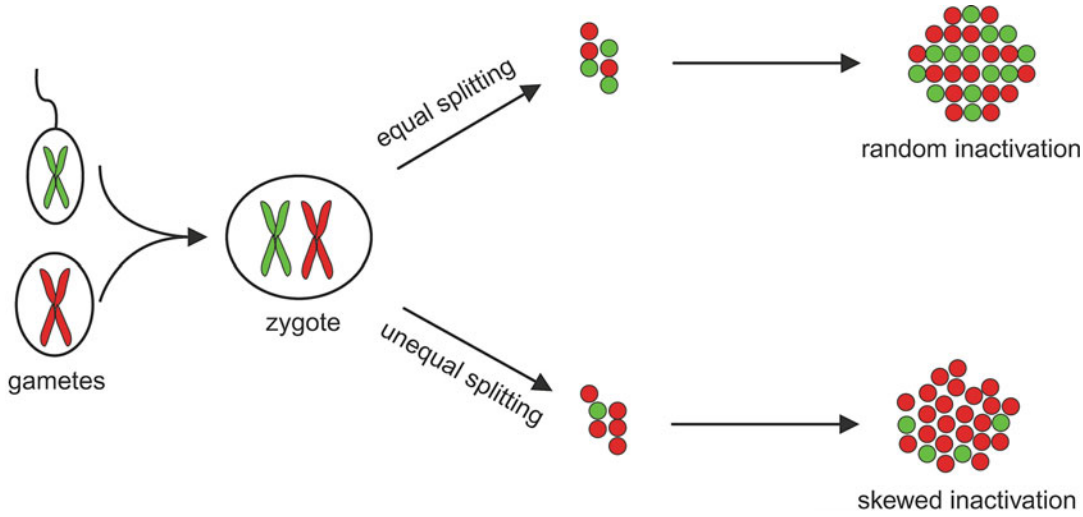


Fig. 5.9 Schematic illustration of a mechanism for a skewed methylation pattern. The random inactivation of an X chromosome in each cell of a female occurs during early embryogenesis. If equal splitting of cells with an inactivated maternal or paternal X chromosome occurs during the early X chromosome inactivation process, the tissues will be composed of a cellular mosaic. If there is unequal splitting of stem cells at this time, the stem cells with inactivated paternal (*green*) and maternal (*red*) X chromosomes will be in unequal number, resulting in a skewed pattern of X chromosome inactivation

incidence of extremely unbalanced X inactivation patterns is seen in elderly women. There is depletion of stem cells through the aging process
  - Somatic selection occurs in the aging process. A small growth advantage may preferentially affect certain cell populations or clones. Such selection is variable among different tissues
- Study on 5–10-year-old monozygotic and dizygotic female twin pairs found that the skewed XCI may already be established in early childhood
  - Acquired skewing in XCI after establishment is primarily mediated by stochastic mechanisms
- X chromosome inactivation patterns in different tissues in the same female may vary, a phenomenon which is referred to as tissue-specific X inactivation pattern
  - Skewed methylation is low in gastrointestinal mucosa and thyroid but significantly higher in blood cells
  - Hematopoietic cells are particularly subject to the selection process due to short life spans and high turnover. Depending upon the definition and the quantitative accuracy of the measurement, up to 20% of such specimens may have skewed X chromosome inactivation
  - Before a clonal population of cells can be demonstrated with X chromosome inactivation analysis, the pattern of X chromosome inactivation observed in a tissue sample must be interpreted with reference to that seen in normal tissue of the same lineage. If a hematologic malignancy is being studied, an alternative method of clonality analysis should be considered because of the high incidence of skewed methylation in blood cells
- Methodology divergence may also cause skewed results on X chromosome inactivation analysis, for example, artificial allelic dropout due to insufficient PCR amplification and reduction of DNA template quantity due to tissue preservation and processing, damaged or salt-contaminated DNA, and different laboratory criteria for skewing

Patch Phenomenon

- Patch size can be large or small in normal tissues (see previous discussion). A patch can contain 200 or more cells and have a diameter of 2–3 mm or greater in some tissues
  - For example, there have been reports that the patch size in bladder epithelium could be about 120 mm^2 and composed of 2×10^6 cells
  - Monoclonal patch size of normal human thyroid tissue is between 48 and 128 mm^2, containing 4×10^5 cells. In fact, when 20 normal thyroids were microdissected and subjected to X chromosome inactivation analysis, 70% demonstrated monoclonality, a likely reflection of large patch size
- X chromosome inactivation analysis cannot discriminate between a monoclonal proliferation and a patch
- Without considering the patch factor, X chromosome inactivation studies in human tissue, especially when applied to epithelial neoplasms, cannot readily answer questions about clonality
  - A microdissection area of >2 mm^2 or multiple site cell harvesting could be helpful in differentiating between a normal patch and an abnormal clonal proliferation

PCR Bias

- PCR bias is the phenomenon in which PCR can preferentially amplify one of two heterozygous alleles
- Differential methylation of androgen receptor gene, HUMARA, permits identification of nonrandom X inactivation in a monoclonal tumor
- Coamplification of two alleles in a heterozygote generates PCR products in different sizes
- Under optimized conditions the amplification efficiency of two alleles is equivalent yielding equal band intensities
- Highly imbalanced PCR products of heterozygous alleles may be present with preferential amplification of lower molecular weight alleles
  - PCR bias can be caused by different factors
  - Biased amplification consistently favors the lower allele
  - Regional secondary structure of DNA is another factor leading to PCR bias. Titrating the melting temperature is necessary to solve this issue
  - Adequate genomic DNA quantities are essential for a consistent allelic amplification. Five nanogram or more of genomic DNA can generate consistent amplifications
  - The quality of genomic DNA is critical when DNA is extracted from paraffin-embedded tissues. A report suggests that using 7-deaza-2′-dGTP could adjust the upper band coefficient by fourfold

Persistence of Biallelic Bands in Tumor Samples (Table 5.3)

- Up to 40% of cancers may have a random X chromosome inactivation patterns
- The loss of X chromosome inactivation may be related to loss of the XIC located at Xq13
- Contamination with normal tissue is one of the major causes of loss of sensitivity. Precise microdissection may be required for accurate interpretation
- Incomplete DNA digestion leads to the amplification of nonmethylated alleles, which will greatly reduce the sensitivity by showing a pseudorandom inactivation
- Methodology divergence, including different laboratory criteria for allelic loss, PCR conditions, and assay method selection, is also among the reasons for persistence of biallelic patterns
- X chromosome aneuploidy
  - X chromosome inactivation mechanisms result in only one active X chromosome. The other is subject to inactivation even in the setting of X chromosome aneuploidy such as XXX, XXY, or XXXX
  - Multiple X chromosomes from paternal and maternal origin may be inactivated and show falsely random inactivation in clonality analysis. Pseudorandom

Table 5.3 Persistence of biallelic bands in tumor samples

Possible explanations
Deletion of XIC
Contamination with normal tissue
Incomplete DNA digestion
X chromosome aneuploidy
Coexistence of multiple tumor subclones of independent origin
Variable methylation patterns at the HUMARA locus
Reactivation of inactive X chromosome
Others

Table 5.4 Commonly used clonality analysis techniques

X chromosome-linked methods
DNA based
DNA methylation
Human androgen receptor locus (HUMARA)
M27β probe for DXS255 locus
Restriction fragment length polymorphism (RFLP)
Glucose-6-phosphate dehydrogenase (G6PD) locus
Hypoxanthine phosphoribosyl transferase (HPRT)
Phosphoglycerate kinase (PGK) locus
RNA based
Palmitoylated membrane protein *p55* gene locus
Iduronate-2-sulfatase (*IDS*) gene locus
Protein based
Glucose-6-phosphate dehydrogenase (G6PD) isoenzyme
Non-X chromosome-linked method
Loss of heterozygosity
Somatic mutation (e.g., *TP53*)
Gene rearrangement (e.g., T cell receptor and immunoglobulin rearrangement for lymphoma workup)
Restriction fragment length polymorphism (RFLP)
Cytogenetics and fluorescence in situ hybridization (e.g., FISH for il2p)
DNA methylation
Microsatellite instability
Viral integration analysis (e.g., EBV, HBV, HCV, HPV)
Comparative genomic hybridization (CGH)
Microarray-based clonality analysis
MicroRNA fingerprint
Protein-based analysis (e.g., OCT4, TTF-1)

EBV Epstein–Barr virus, *HBV* hepatitis B virus, *HCV* hepatitis C virus, *HPV* human papillomavirus

X chromosome inactivation should be excluded through other techniques, such as fluorescence in situ hybridization (FISH)

- Coexistence of multiple tumor clones of independent origin may show false random X chromosome inactivation representing more than one clonal tumor. Precise small area microdissection of tumor cells may be helpful in solving this problem
- Variable methylation patterns at the HUMARA locus can be seen in neoplastic and nonneoplastic cells
  - About 15% of X-linked genes escape inactivation to some degree, and the proportion of genes escaping inactivation differs dramatically between different regions of the X chromosome
  - The incidence of variable X chromosome inactivation in healthy females varies from 4% to 33%, which may be related to tissue-specific X chromosome inactivation patterns
- Age-related reactivation of inactivated X chromosomes may be related to the loss of critical methylation sites on the X-linked genes

5.2.4 Other X Chromosome-Linked Clonality Analyses

- Various methods have been used in the past, essentially all of which have been gradually replaced by HUMARA X chromosome inactivation analysis. The reader may refer to specific articles for more detailed discussions (Table 5.4)
- M27β probe for DXS255 locus
  - Locus DSX255 on the X chromosome is consistently differentially methylated and can be analyzed directly to assess X chromosome inactivation status
  - The M27β probe is hybridized to electrophoretically separated DNA
  - The polyclonal tissue shows two equal bands, and the monoclonal tumor shows a pattern of unequal bands

- Glucose-6-phosphate dehydrogenase (G6PD), hypoxanthine phosphoribosyl transferase (HPRT), and phosphoglycerate kinase genes
  - These are restriction fragment length polymorphism (RFLP)-based analyses
  - DNA from tumor and control is extracted and purified and then PCR amplified
  - The amplified DNA is cut into restriction fragments using suitable endonucleases, which only cut the DNA molecule at a specific recognition sequence
  - The restriction fragments are then separated by agarose gel electrophoresis
  - Clonal and nonclonal populations show different band patterns accordingly
- Palmitoylated membrane protein *p55* gene locus
  - Human *p55*, located at Xq28, encodes an abundantly palmitoylated phosphoprotein of the erythroid membrane
  - The allele on the inactivated X chromosome is silenced
  - mRNA is extracted and reverse transcribed using random primers. cDNAs are amplified with specific primer sets for two rounds
  - Two bands are seen when X chromosome inactivation is random, but cells with nonrandom X chromosome inactivation show a single band
- Iduronate-2-sulfatase (*IDS*) gene locus
  - *IDS* clonality analysis is an mRNA analysis for a functional silencing of the gene
  - mRNA is extracted and reverse transcribed by random primers. PCR and primer extension analysis are carried out for *IDS*
  - Two bands are seen when X chromosome inactivation is random
- G6PD isoenzyme analysis
  - This is a classic tool that has be used to study X chromosome inactivation status and clonality since the 1970s
  - Protein is extracted from cells and the G6PD fraction is collected and then separated by electrophoresis
  - The protein bands are compared with known heterozygous blood cells
  - A clonal cell population will show only one isoenzyme band

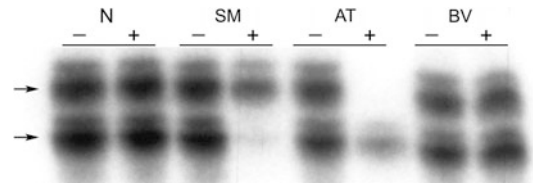


Fig. 5.10 A representative gel picture of X chromosome inactivation analysis of the various components of a renal angiomyolipoma. *N* normal, *SM* smooth muscle, *AT* adipose tissue, *BV* blood vessel. – and + indicate that DNA is nondigested or digested with methylation-sensitive restriction enzyme. In SM and AT, opposite patterns of nonrandom X chromosome inactivation indicate that the lesions are clonal proliferations with each component having a different clonal origin. BV shows a random X chromosome inactivation pattern (two allelic bands)

5.2.5 Selected Applications

5.2.5.1 Defining the Monoclonal Nature of the Lesion

- Renal angiomyolipoma is a benign neoplasm composed of blood vessels, smooth muscle, and adipose tissue
  - Whether renal angiomyolipoma is a hamartoma or a neoplastic process has long been debated
  - It is also uncertain which components of angiomyolipoma represent clonal growths and if various components share the same clonal origin
- Cheng et al. (2001b) found the smooth muscle cells and the adipose tissue to have differing patterns of nonrandom X chromosome inactivation, indicating that both are monoclonal and probably originate from independent clones (Fig. 5.10)

5.2.5.2 Defining the Clonal Relationship of Separate Tumors

- Two proliferative cell populations (e.g., two separate tumors) that share the same nonrandom X chromosome inactivation patterns have a 50% probability of being derived from a common progenitor cell (monoclonal) since the chance of either paternal or maternal origin of the inactivated X chromosome is 50%

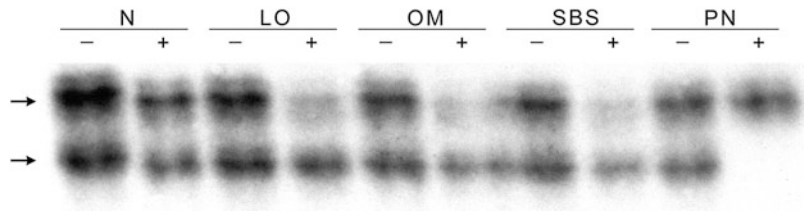


Fig. 5.11 X chromosome inactivation analysis of multifocal ovarian papillary serous tumors of LMP. Tumor locations are as follows: *LO* left ovary, *OM* right ovary, *SBS* peritoneum, and *PN* pelvic lymph node. *N* represents normal. – and + indicate that DNA is nondigested or digested with methylation-sensitive restriction enzyme. The tumors from the LO and SBS share the same X chromosome inactivation pattern, both showing loss of the upper alleles after digestion. This is suggestive of a common clonal origin. However, tumor from the PN shows a different clonal origin. The allele from the *right* ovary shows a reduced upper allele but does not reach the cutoff value and, thus, is interpreted as negative

- Different patterns of nonrandom X chromosome inactivation in separate tumors support independent origin
- With a larger number of cell populations analyzed, results of clonality assessments become more meaningful
  - The probability of different cell populations with the same pattern of X chromosome inactivation representing a polyclonal process (i.e., unique genetic origin) decreases as the sample number (n) of cell populations increases [probability = $(0.5)^n$]
  - For example, patients with ovarian papillary serous tumor of low malignant potential (LMP) may have peritoneal "implants" of histologically identical tumors
    - Gu et al. (2001) studied a group of women with advanced stage ovarian papillary serous tumors of LMP using X chromosome inactivation
    - Most of the patients with peritoneal and ovarian tumors showed different X inactivation patterns, suggesting that peritoneal tumors associated with ovarian LMP tumors may arise independently from their own primary tumor clones rather than through an "implantation" process
    - Some patients with bilateral ovarian tumors of LMP showed different X chromosome inactivation patterns in tumors of each ovary, indicating that the patients had bilateral primary tumors instead of one ovarian tumor with metastasis to the opposite ovary (Fig. 5.11)

5.2.5.3 Defining the Clonal Relationship of Different Components of the Same Tumor

- The identification of components of different biologic aggressiveness within a single neoplasm is a common finding in pathology
- These variable components are thought to result from tumor cell dedifferentiation or transformation, with the subsequent evolution of different subpopulations of tumor cells, a concept that is exemplified by the coexistence of small cell and urothelial carcinoma of the bladder
  - Cheng et al. (2005) found a concordant pattern of nonrandom X chromosome inactivation between small cell cancer and coexisting urothelial carcinoma, suggesting that both tumor components originate from the same progenitor cells in the urothelium. These findings may provide new insights into the treatment of small cell carcinoma of the urinary bladder

5.2.5.4 Defining the Clonal Relationship Between Precursors (Such as Intraepithelial Neoplasia) and Cancer

- Guo et al. (2000) studied cervical intraepithelial neoplasia with coexisting invasive cancers using X chromosome

inactivation. The authors found that precancerous lesions and coexisting cervical cancers were clonally related. Carcinogenesis of cervical cancers involves selection of subclones from originally polyclonal precursors

5.2.5.5 Defining the Clonal Relationships Between Primary and Metastatic Tumors

- Evidence of genetic heterogeneity and tumor subclones within urothelial carcinoma of the bladder has raised questions about the clonal origin of urothelial carcinoma and its metastases
  - Jones et al. (2005a) investigated female patients who underwent radical cystectomy for urothelial carcinoma
  - The X chromosome inactivation analysis showed identical nonrandom inactivation patterns in primary bladder cancer and pelvic lymph node metastases, suggesting that the capacity for metastasis arises in only a single clonal population in the primary tumor

5.3 Loss of Heterozygosity (LOH) as a Clonal Marker

5.3.1 Overview

- Microsatellites are polymorphic loci that consist of repeating units of 2–6 bp that repeat 10–100 times without interruption. There are approximately 200,000 microsatellite loci in the human genome
- The polymorphism of microsatellites is the basis for microsatellite analysis
- Slipped strand mispairing is the primary mechanism for polymorphism, which leads to deletion or insertion of the microsatellite repeat unit
  - Slipping is caused by regional nonpairing, which forms a "bubble" containing one or more repeat units. If the slippage involves a newly synthesized strand, it is called "backward slippage"; if it involves the parental strand, it is called "forward slippage"
  - 5′–3′ slipping (backward slippage) causes insertion of a repeat unit, and 3′–5′ slipping (forward slippage) causes deletion of a repeat unit
- In a heterozygote, PCR amplification using microsatellite locus-specific primers will result in two distinct bands, representing maternally and paternally derived alleles
- LOH represents the loss of one parent's contribution to the cellular genome (Fig. 5.12)
- LOH often indicates the presence of a tumor suppressor gene in the lost region
  - If LOH occurs at a selected polymorphic region that is related to a known tumor suppressor gene, it is highly suggestive of deletion of the corresponding gene as this also explains the growth advantage of the tumor cells carrying that genetic alteration
  - LOH can result in haploinsufficiency of a specific allele
  - LOH has been considered an early event during carcinogenesis. Often, the remaining copy of the tumor suppressor gene will be inactivated by mutations during cancer progression
  - LOH is inheritable and can be passed to the cellular derivatives. The presence of a uniform and nonrandom alteration of a tumor suppressor gene as demonstrated by LOH analysis in all cells of a tumor confirms a clonal origin
  - LOH is related to chromosomal instability. A cell with chromosomal instability carries mutations that result in an increased rate of LOH
- A number of mechanisms can lead to LOH, including local deletion, nondisjunction of chromosomes, mitotic recombination, gene conversion, double-strand break resulting in loss of a chromosome arm, or whole chromosome loss
- The principle of LOH clonality analysis is that the clonal expanding cells share a common set of allelic losses

5.3.2 Evaluation and Interpretation of the LOH Analysis Results

- LOH can be identified by detecting the presence of heterozygosity in germline DNA and

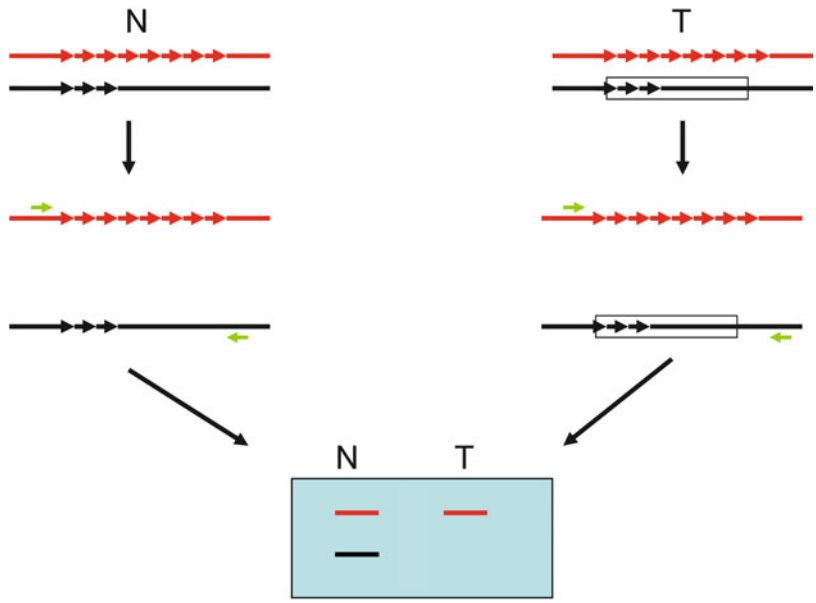


Fig. 5.12 LOH in a cell represents the loss of one parent's contribution of microsatellite DNA to a cell's genome. It often indicates the presence of tumor suppressor gene loss adjacent to the microsatellite locus. An informative locus is the one that the maternal and paternal alleles are different in repeat numbers. Specific primers (*green*) are used to amplify the alleles. For normal (N) cells, both alleles are intact and, thus, can be amplified. When the tumor cells (T) have lost a segment of its DNA, the lost allele (*boxed*) cannot be amplified. Two allelic bands are seen in normal control DNA, whereas tumor DNA shows loss of the lower allele (LOH)

the absence of heterozygosity at the same locus in the tumor DNA

- LOH can be used as a clonal marker. Tumor cells derived from the same progenitor cells will share the same LOH patterns at multiple microsatellite loci
- Brief procedures (Fig. 5.13)
  - Genomic DNA is extracted from microdissected normal and cancer tissues
  - PCR amplification of microsatellite loci from tumors and normal control DNA is performed
  - The PCR products are subjected to electrophoresis. Microsatellite patterns of normal and tumor samples are compared
- Informative case
  - Two alleles are present after PCR amplification in normal control samples
  - Informative is a synonym for the heterozygous state
  - Microsatellite loci with a relatively high heterozygous rate should be selected for the study
- Noninformative case
  - It is due to the cells posses two alleles with identical size at the locus, which could not be distinguished, as seen in the homozygous state
  - The noninformative case shows only one allelic band after PCR amplification in normal control DNA
- LOH clonality analysis results are based on the comparison of band patterns of tumor DNA to the normal control DNA
  - A noninformative case cannot be analyzed by this method
- Identical allelic loss patterns shared among separate tumors suggests a common clonal origin
- Different or opposite allelic loss patterns in two separate tumors suggests a different clonal origin (Fig. 5.13)
- Copy neutral LOH
  - Copy neutral also referred as uniparental disomy is a common mechanism leading to LOH by duplication of a maternal or

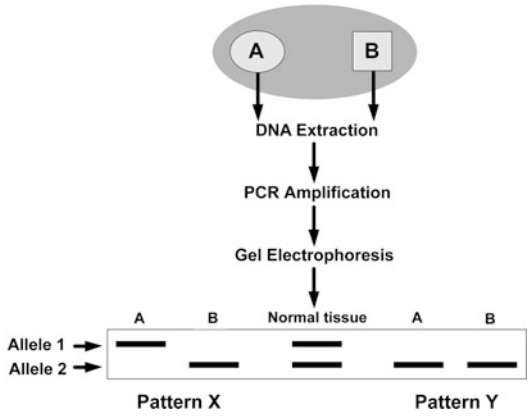


Fig. 5.13 An LOH-based clonality test compares the allelic banding patterns between tumors and normal tissue. Pattern X shows an opposite allelic loss pattern in two separate tumors (A and B), suggesting a different clonal origin. Pattern Y shows the same allelic loss pattern in each tumor, suggesting a common clonal origin. Normal tissue shows double alleles indicating that the patient is heterozygous (informative) at the locus

paternal (uniparental) allele and concurrent loss of the other allele
- No net gain/loss in the copy number occurs in the affected individual
- Copy neutral LOH can be classified as acquired, clonal derivation or a constitutional, nonclonal derivation
- Copy neutral LOH can be the result of an early embryonic mitotic event or truly constitutional
- The copy neutral LOH can arise from autozygosity or meiotic events, including trisomic rescue, gamete complementation, duplication of a monosomic chromosome, or nonhomologous recombination
- It can be segmental or involve an entire chromosome
- Copy neutral LOH can be biologically equivalent to the second hit in the Knudson hypothesis
- SNP-based arrays can detect the copy neutral LOH

5.3.3 Advantages and Limitations of LOH Analysis

- Advantages
  - Unlike X chromosome inactivation, this technique is applicable to both men and women
  - LOH is a sensitive marker for morphologically normal precancerous cells
  - Many microsatellite loci have high LOH rates in different malignancies
  - In analysis of multiple tumors by LOH analysis, multiple microsatellite markers generally required to decrease the likelihood of random matches (see discussion below)
  - The clonal population of cells may be recognizable even during subclone evolution in tumorigenesis
- Limitations
  - Accurate interpretation of results requires a pure population of target cells. Tissue microdissection is often required to avoid contamination
  - Selection of appropriate microsatellite loci for specific applications may be difficult
  - For formalin-fixed, paraffin-embedded tissue (FFPE), a long PCR microsatellite fragment is difficult to amplify (PCR product of $<$200 base pair is preferred)
  - LOH may occur in normal tissue (see discussion below)
  - The occurrence of homozygous deletion and microsatellite instability (MSI) may complicate interpretation of results

5.3.4 Technical Considerations

5.3.4.1 Allele Drop-Off

- Allele drop-off refers to PCR bias against one allele, especially the upper one, resulting in preferential amplification of the lower allele
- Allele drop-off should not be misinterpreted as noninformative or LOH
- To avoid this artifact, appropriate DNA extraction methods should be used. PCR conditions, especially the cycle number, need to be precisely controlled

5.3.4.2 Loss of Sensitivity

- Loss of sensitivity is a common reason for failure of the test
  - Most commonly due to amplifying the locus for too many cycles
    - When genomic DNA is in low concentration or in lower quality, there is a tendency to amplify for more cycles, which equalizes the alleles
    - The PCR amplification cycle number should be controlled within the range of log phase
  - The genomic DNA from FFPE tissue is fragmented
    - Selection of microsatellite markers with PCR products <200 bp may enhance the sensitivity
    - A different method (such as FISH) should be used to exclude homozygous deletion of a specific allele

5.3.4.3 LOH in "Normal" Cells

- The background level of LOH in cells with normal morphology has been reported to be from 4% to 20%, regardless of the detection system used
  - It suggested that the earliest genetic alterations that fit into the multistep model of tumorigenesis might be detected in cells that appear morphologically normal
  - The LOH is more likely to be found around a tumor
  - Histologically normal epithelium and/or simple hyperplasia associated with cancer is often characterized by genetic alterations
  - Genetic alterations, such as LOH leading to tumor development, occur earlier than morphologic alterations in precancerous cells
  - The genetic alteration may represent generalized carcinogenesis and might be the basis for the development of multifocal tumors or recurrence of tumors (see 5.1.2.3.1, "Field Cancerization" section above)
  - These genetic alterations may represent an intermediate step in carcinogenesis
- LOH could also be randomly acquired and irrelevant to tumor development
  - The existence of genomic alterations in morphologically normal cells is not an indication of malignancy
  - Random genomic alteration in morphologically normal cells has fewer tendencies to involve multiple loci
  - Incidences of LOH are significantly higher in precursor and cancer cells
- Careful selection of normal controls is critical for accurate interpretation of results

5.3.5 Methods for LOH Analysis

5.3.5.1 Radioisotope PCR Incorporation Gel Electrophoresis

- Genomic DNA is prepared from the pure cell population, generally from microdissected paraffin sections
- Radioisotope PCR incorporation gel electrophoresis uses locus-specific primer pairs to amplify the heterozygous loci from the normal and tumor genomic DNA in the presence of radioisotope-labeled dNTPs
- The PCR products are separated by electrophoresis and visualized by autoradiography
- The intensity of gel bands from normal control and tumor DNA is compared to decide if there is a diminished allele at the locus
- LOH is present when one band from either allele has disappeared or is greatly decreased when compared with the other allele
- The method is stable, simple, and inexpensive, but radioisotope is needed

5.3.5.2 High-Resolution Fluorescent Microsatellite Analysis

- Genomic DNAs are extracted from tissue sections and amplified using primers labeled with fluorescent compounds
- The PCR products are denatured and loaded onto a sequencer. The mobility of each sample is processed using computer software to convert amplified DNA to fluorescence band peaks
- The height of each allele peak is compared with the normal control to decide if LOH is present
- This method is sensitive and standardized but needs special equipment and software such as

ABI Genetic Analyzers (Applied Biosystem, Foster City, CA)

5.3.5.3 High-Performance Liquid Chromatography (HPLC)

- Genomic DNAs are amplified using primers labeled with fluorescent compounds
- Denatured amplicon is then gradually reannealed by decreasing sample temperature and injecting it into the HPLC analyzer
- DNA is eluted from the column
- DNA fragments of different sizes can be separated by elution over time
- The method is sensitive and highly repeatable, but a specific analyzer is needed (e.g., WAVE^{R} Fragment Analysis System [Transgenomic, Omaha, NE])

5.3.5.4 Single-Nucleotide Polymorphism (SNP) Array

- Single-nucleotide polymorphisms refer variable positions in the genome with two different allelic types. SNP does not cause disease
- SNP array identifies the LOH in whole genome
- SNP array karyotyping allows for the detection of uniparental disomy and cryptic chromosomal abnormalities
- Whole genome screening for LOH could provide allele profiling for each tumor

5.3.6 Selected Applications

5.3.6.1 Defining Clonal Relationships of Separate Tumors

- Concordant LOH patterns involving multiple loci in separate tumors support a common clonal origin: discordant LOH patterns over multiple loci suggest an independent clonal origin (Fig. 5.13) (Table 5.4)
- Prostate gland usually contains two or more widely separate tumors
  - A critical issue is whether the multiple tumors are independent in origin
  - Discordant pattern of LOH in the majority of multifocal prostate cancers has been found
  - LOH supported that the multiple loci of prostate tumors arise independently from a field effect
  - Independent expansion of cancer stem cells within the field leads to the development of multiple tumors

5.3.6.2 Defining Clonal Relationships of Different Histologic Components of the Same Tumor

- Teratomas of the testis in postpubertal patients are histologically diverse tumors that often coexist with other types of germ cell tumors
  - Clonality relationship of the tumor components in a mature teratoma and its relationship to other components of malignant mixed germ cell tumors were investigated using LOH by Kernek et al. (2004)
  - Complete concordant LOH patterns among mature teratoma and other germ cell tumor components were seen in over 70% of patients
  - LOH analysis suggested the histogenetic relationships of various germ cell tumor components and testicular tumorigenesis

5.3.6.3 Defining the Precursor Nature of Lesions

- LOH could be used to determine the precursor nature of a lesion
  - Atypical adenomatous hyperplasia (AAH) of the prostate has been proposed as a possible precursor lesion of prostate cancer
  - Cheng et al. (1998a) identified a high frequency of LOH, similar to that commonly seen in high-grade prostatic intraepithelial neoplasia (PIN) and prostate cancer, in AAH
  - The result provides evidence of a genetic link between some cases of AAH and prostate carcinoma

5.3.6.4 Defining Clonal Relationships Between Precursor Lesions and Cancer

- Concordant LOH patterns involve multiple microsatellite loci suggest common clonal origin between precursor lesion and cancer
  - Prostate carcinoma is usually heterogeneous and multifocal, with diverse clinical and morphologic manifestations

- Understanding the molecular basis for this heterogeneity is limited, particularly for the putative precursor, high-grade prostate intraepithelial neoplasia (PIN)
- Analysis of LOH in multiple foci of PIN and matched foci of carcinoma demonstrated strong genetic similarities between PIN and prostate carcinoma, suggesting that evolution and clonal expansion of PIN may account for the multifocal etiology of prostate cancer

5.3.6.5 Defining Clonal Relationships Between Primary and Metastatic Tumors

- Epidermotropic metastases of melanoma are unique in their varied histopathologic appearances, which can mimic a primary lesion
  - Primary tumors demonstrated a concordant LOH pattern with corresponding epidermotropic metastases
  - The data suggest that these metastases are clonally related to the primary lesion and that some cases diagnosed as epidermotropic metastatic melanoma might be divergent clones or new primaries rather than metastatic disease

5.3.6.6 Defining the Genetic Basis of Tumorigenesis and Cancer Progression

- Carcinogenesis is a process in that normal cells are transformed into cancer cells after acquiring genetic–epigenetic alterations
- Tumor progression involves accumulation of changes on cellular and genetic level
- The dynamic LOH represents the evolution of the cells from benign to malignant and progression
  - Boland et al. (1995) analyzed the temporal sequence of allelic losses on 5q, 17p, and 18q during neoplastic progression of colorectal cancer
    - No LOH was found in normal tissues surrounding colorectal neoplasms
    - The LOH occurred abruptly on 5q at the transition from normal colonic epithelium to the benign adenoma
    - LOH occurred on 17p at the transition from adenoma to carcinoma
    - The allelic losses on 5q and 17p are associated with abrupt waves of clonal neoplastic expansion

5.3.6.7 Defining Stromal–Epithelial Interaction During Tumorigenesis

- During tumorigenesis the prevailing model suggests a process whereby precancerous epithelial cells acquire multiple genetic mutations and the associated stroma becomes activated
  - Recent evidences demonstrated that epithelial cells without specifically engineered mutations can be induced to form carcinomas by association with genetically altered fibroblasts
  - The tissue specificity of stromal–epithelial interactions probably accounts for a tissue type-specific and cell type-specific role of the microenvironment in carcinoma development
  - In the study of 11 female patients with ductal carcinoma in situ and five cases of invasive ductal carcinoma, all but one of the polymorphic DNA markers showed LOH in the stroma with a frequency of 10–66.55% but none in the control stroma
  - The stroma in the tumor represents a neoplastic interactive component rather than a reactive response to the carcinoma

5.3.6.8 Defining the Evolution of Neoplastic Cell Lineages in Carcinogenesis

- Cancer progression develops as a consequence of an acquired genetic instability and the subsequent evolution of clonal populations with accumulated genetic errors. Barrett esophagus is a precursor of esophageal cancer
  - Barrett et al. (1999) found that the clonal evolution of neoplastic progeny frequently involved LOH at 5q, 13q, and 18q that occurred in no obligate order relative to each other
  - These findings suggest a progenitor cell was undergoing clonal expansion

  - With increasing genetic instability, these lesions will enter a phase of clonal evolution that begins in premalignant cells and continues after the development of invasive cancer

5.3.6.9 LOH Patterns Provide Fingerprints for Genetic Heterogeneity in Multistep Cancer Progression

- The multistep cancer carcinogenesis and progression in cancers are mediated by dysregulation of a series of functional networks, which may presented as LOH in related chromosomal loci
  - Maley et al. (2006) found that clonal diversity based on LOH can predict tumor progression from precursor lesions such as Barrett esophagus
- Some clear cell renal cell carcinomas contain components with sarcomatoid morphology
  - It has been suggested that the sarcomatoid elements arise from the clear cell tumors as a consequence of clonal expansions of neoplastic cells with increasing genetic instability
  - Jones et al. (2005) found different patterns of allelic loss in multiple chromosomal regions in clear cell and sarcomatoid components, possibly evolutes from the patterns of clear cell components
  - The genetic heterogeneity indicates genetic divergence during the clonal evolution of renal cell carcinoma

5.4 Other Methods of Clonality Analysis

5.4.1 Somatic Mutation

- Somatic mutations have been found in many human malignancies
  - Mutations are acquired by the tumor founder cell and can be stably passed to all of the clonally related daughter cell populations; thus, they could be used as clonal markers
- Somatic mutations are a part of the carcinogenesis process and are related to the biologic behavior of tumors
  - Somatic mutations most often affect oncogenes and tumor suppressor genes through different pathways (see Sect. 5.1.2.2)
- Methods for somatic mutation-related clonality analysis
  - DNA sequencing
    - DNA sequencing identifies a point mutation directly from the tumor DNA sequence
    - DNA sequencing is also used to analyze positive cases after a qualitative test such as single-strand conformation polymorphism (SSCP)
    - A similar mutation in key genes, such as an identical *TP53* mutation in the same codon in multiple samples from separate tumors, suggests a common clonal origin
    - Different *TP53* mutations, either at different codon or different nucleotide change, are compatible with independent clonal origins
  - Single-strand conformation polymorphism (SSCP)
    - The basic principle of SSCP is that the DNA sequence change will alter the three-dimensional conformation of single-stranded DNA, which alters the migration properties of DNA on gel electrophoresis. Differing band patterns indicate base changes
    - SSCP is a widely used PCR-based method to detect mutations in a large number of samples for clinical diagnostics, for population genetics, and for molecular epidemiologic study
    - A single-nucleotide change in a sequence of double-stranded DNA cannot be distinguished by electrophoresis since the physical properties of the double strands are almost identical for both alleles
    - SSCP detects a conformational difference from single-stranded nucleotide sequences of identical length. The

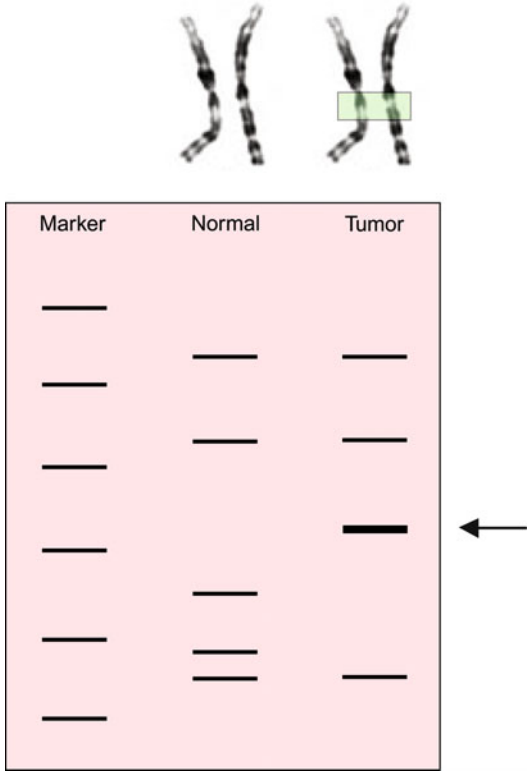


Fig. 5.14 RFLP analysis. DNA from normal tissue and tumor are extracted followed by PCR amplification. The amplified DNA is digested into restriction fragments using a suitable endonuclease, which only cuts the DNA molecule at a specific recognition sequence. The restriction fragments are then separated by gel electrophoresis. Normal DNA shows a specific restriction fragment panel. The restriction site of tumor DNA was destroyed by a mutation (*boxed*). Therefore, the tumor DNA cannot be cut and shows a changed band pattern (*arrow*)

detected abnormal conformation needs to be verified by sequencing
    - SSCP is tend to be replaced by sequencing with the development of sequencing technology
- Restriction fragment length polymorphism (RFLP)
    - The basic principle of RFLP is that mutation will destroy or create a restriction cutting site, leading to an altered restriction fragment banding profile (Fig. 5.14)
    - Regions of some genes encompass restriction enzyme cutting sites in their coding region, which will yield a specific restriction fragment pattern
    - Genomic DNA is extracted and purified, followed by PCR amplification for the target fragment
    - The amplified DNA fragment is then cut with selected restriction endonucleases, which only cut the DNA sequence where specific restriction sites exist
    - A gene fragment with a single-base mutation may destroy or create a restriction enzyme cutting site and therefore alter the restriction cutting pattern
    - The method is of low cost but only suitable for certain specific mutations
    - Tumors with the same altered RFLP pattern suggested a common clonal origin
- Denaturing high-performance liquid chromatography (DHPLC)
    - Denaturing HPLC is a method of separating one chemical from another based upon their absorption of ultraviolet light and group interaction with the analytic column substrate
    - When DNA containing a nucleotide differs from normal DNA, a mismatched heteroduplexes will form with a different melting temperature
    - A gene fragment is amplified from genomic DNA extracted from microdissected tumor tissues
    - The PCR product is denatured and injected into the HPLC instrument to be column separated, and the fractions are analyzed
    - The method is fast with high specificity and with ability to systemically detect all of the mutation sites in a gene fragment
    - This method requires expensive instrumentation

5.4.2 Gene Rearrangement Analysis

- Clonality analysis using T cell receptor and B cell immunoglobulin gene rearrangement has been routinely employed in lymphoma workups (see Chaps. 30–33 for details)

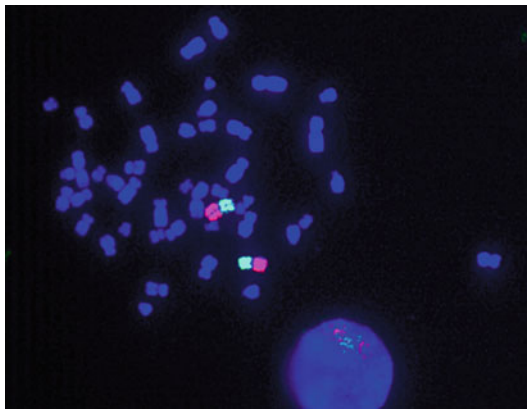


Fig. 5.15 Arm painting of chromosomes. The fluorescence-labeled DNA probes are hybridized to specific chromosome arms, "painting" the arms into different colors. Chromosome arm painting is a useful tool in the identification of chromosomal translocations, isochromosomes, and arm losses

5.4.3 Cytogenetic Analysis and FISH

- Cytogenetic analysis is an important tool for clonality analysis
  - Cytogenetic alterations are inheritable to the subsequent cell lineages during clonal expansion and can serve as clonal markers
- Karyotyping determines numerical and structural changes of chromosomes
- More recently, chromosomal painting techniques have been widely used in cellular karyotyping
  - Chromosome painting can be used to label the whole chromosome or an arm
  - It is also useful in identifying chromosomal translocations (Fig. 5.15)
- FISH has been routinely employed for aiding in diagnosis and identification of tumor origin
  - FISH uses fluorescent-labeled probes to hybridize to part of a chromosome with a high degree of sequence similarity
  - It can be used to identify numerical and structural anomalies of chromosomes from interphase or metaphase cell nuclei (Fig. 5.16a)
  - FISH can be used to map sequences to a specific position on a chromosome, for example, centromeres, arms, and telomeres showing loss and/or gain of DNA
  - Characteristic chromosome alterations could be used in aid of tumor diagnosis
    - Isochromosome 12p is characteristic of tumors of germ cell origin (Fig. 5.16b). Cheng et al. (2007b) found a high incidence of chromosome 12p anomalies in "fibrosis" from residual retroperitoneal fibrous masses after chemotherapy for testicular germ cell tumors. These findings suggest that the fibrous cells in the residual mass are derived from the same tumor progenitor cells as the preexisting metastatic germ cell tumor
    - Many tumors are characterized by specific types of chromosomal translocations, such as the t(9;22)(q34;q11) translocation (bcr-abl) in chronic myelogenous leukemia, which can be demonstrated by FISH using a specific probe set (Fig. 5.16c)
    - Consistent chromosomal changes suggest a clonal origin such as 1p36/19q13 deletion in oligodendroglioma (Fig. 5.16d)
    - FISH can also be used for chromosome painting to demonstrate entire chromosomes or chromosomal arm anomalies (Fig. 5.16e)
    - Multiprobe FISH, such as UroVysion™ by Vysis, uses probes labeled with different fluorescent dyes and is especially helpful in identifying specific colocalized abnormalities (Fig. 5.16f)
    - In patients with papillary renal cell carcinoma, it is not uncommon to find two or more anatomically distinct and histologically similar tumors at radical nephrectomy. Whether these multiple papillary lesions result from intrarenal metastasis or arise independently is unknown. Using a FISH approach, Jones et al. (2005) reported multiple papillary renal cell carcinomas arise independently. Intrarenal metastasis does not seem to play a major role in the spread of papillary renal cell carcinoma

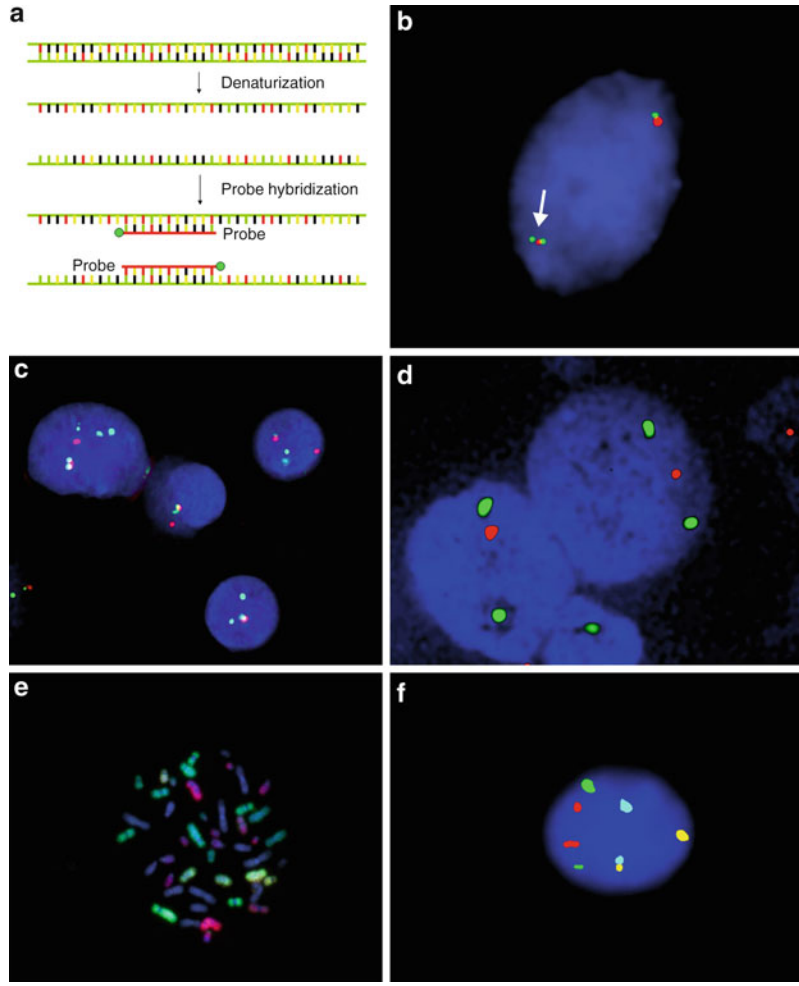


Fig. 5.16 FISH is a widely used technique in clonality analysis. Panel A shows an illustration of the principle of FISH. A fluorescence dye-labeled DNA probe is hybridized to a complementary sequence on a chromosome. Panel B shows an example of isochromosome 12p (*arrow*), a well-recognized marker for tumors of germ cell origin. *Green* signals show 12p and *red* signals show the centromere of chromosome 12. Note there are two chromosome 12 centromere signals and three 12p signals with specific patterns of signal aggregation. Panel C shows a typical bcr-abl translocation as seen in chronic myelogenous leukemia. The DNA fragments normally located on chromosome 9 (*red*) are transferred to chromosome 22 (*green*). Panel D shows a 1p36 deletion, a typical finding in about 80% of oligodendrogliomas. Most of the cells show two *green* (1q) and one *red* (1p36) signal pattern, indicating a clonal origin. Panel E demonstrates chromosomal painting with whole chromosome arms labeled with a specific color. Panel F shows FISH performed on normal urothelium using UroVysion probes specific for centromeres 3 (*red*), 7 (*green*), 17 (*aqua*), and 9p21 (*gold*). Gains of chromosomes 3, 7, and 17 and loss of 9p21 have been observed in more than 80% of urothelial carcinomas

5.4.4 DNA Methylation as a Clonal Marker

- DNA methylation is an epigenetic process involving a chemical modification of DNA without changing the DNA sequence (see Chap. 3 for more details)
- Epigenetic alterations, such as DNA methylation, can be inherited or acquired
  - Patterns of DNA methylation have been consistent findings in some cancer cells
  - Epigenetic characters of a tumor may provide a highly specific and sensitive molecular markers for tumor clonality

  - Catto et al. (2006) analyzed DNA methylation of seven gene promoters in bladder urothelial carcinoma and found that these epigenetic events are useful in defining the clonal origin of the tumor
- Brief procedures for methylation analysis (Fig. 5.17)
  - The genomic DNAs are extracted from microdissected normal and cancer tissues
  - The genomic DNAs are chemically modified by sodium bisulfite
    - The unmethylated cytosine in the CpG site is converted into uracil, but the methylated cytosine could not be converted
    - The primers designed according to methylated sequence (nonconvertible C) or unmethylated sequence (convertible C) could only recognize and amplify the methylated or unmethylated sequences, respectively
    - Methylated alleles can be amplified by methylation-specific primer pairs; unmethylated alleles can be amplified by a different set of primer pairs
    - The PCR products are subjected to gel electrophoresis, and microsatellite banding patterns are compared

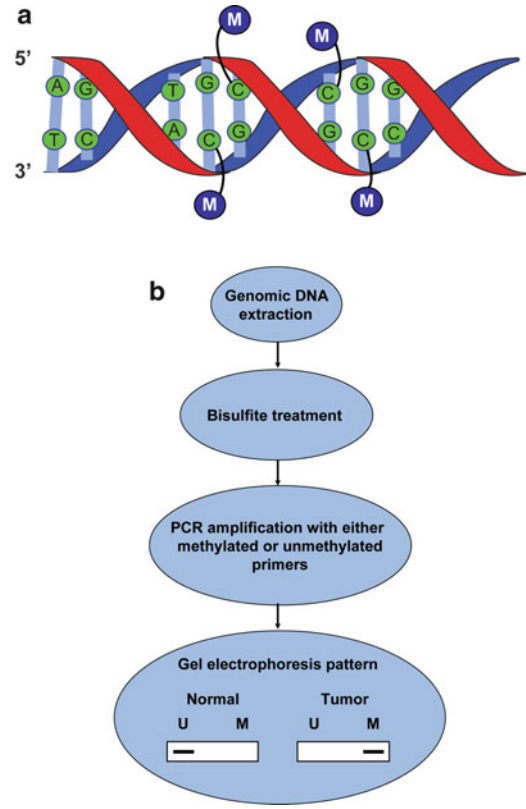


Fig. 5.17 The most important epigenetic process is mediated through DNA methylation, which silences a gene. Methylation occurs at the cytosine of a CpG dinucleotide (**a**). Clonal cells show an identical DNA methylation pattern. Methylation-specific PCR is a commonly used method for DNA methylation analysis (**b**). Genomic DNA is extracted from tumor and control tissue. Hydroquinone and sodium bisulfite convert any unmethylated cytosine to uracil (U), which allows all nonmethylated cytosine bases to be identified. Methylation-specific primers are designed according to the converted (nonmethylated) or nonconverted (methylated) sequences. Specific primers for nonmethylated or methylated sequences can only amplify the nonconverted or converted sequences, respectively. The resulting gel electrophoresis shows the different methylation states of the tissue. Normal cells exhibit a nonmethylated (U) DNA band, whereas the tumor tissue shows a methylated PCR product (M)

5.4.5 Microsatellite Instability

- MSI is a condition of having longer or shorter microsatellite repeats than germline DNA of normal cells
- MSI is a manifestation of DNA damage due to defects in the normal mismatch repair process
  - This is due to failure of the mismatch repair system to correct errors in the replication of microsatellite short sequence repeats (see Chap. 3 for more detail)
  - Instead of losing one heterozygous allele as seen in LOH analyses, MSI shows retention of both alleles in tumor DNA
    - Changes in the length of microsatellite loci are due to deletion or insertion of microsatellite repeating units
    - Like LOH, MSI could also be used as a clonal marker in clonality analysis (Fig. 5.18)
    - Tumor cells derived from the same clonal origin share the same pattern of MSI, and these genetic alterations can be transmitted through subsequent cell generations

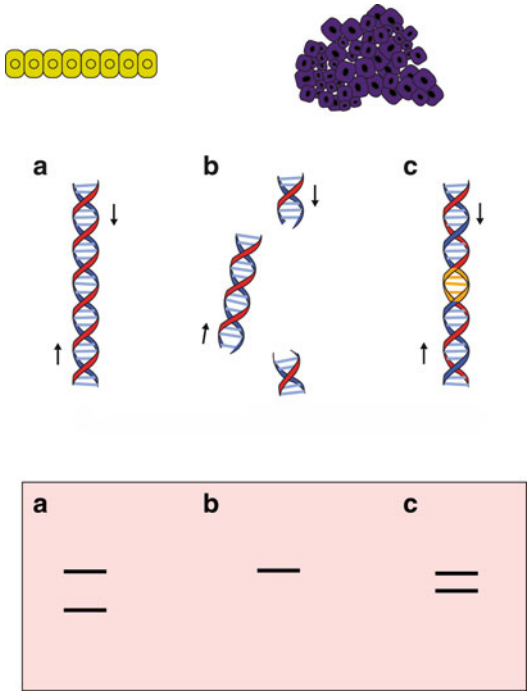


Fig. 5.18 Schematic illustration for microsatellite alterations. Normal informative cells (*yellow*, **a**) show both alleles after PCR amplification of the microsatellite locus. Cancer cells (*magenta*, **b** and **c**) show alteration of the microsatellite locus either through deletion of a DNA fragment to form LOH (**b**) or by altering the microsatellite repeat numbers to form MSI (**c**). The lower box shows the gel patterns. Normal cells (**a**) show both alleles, but tumor cells show LOH (**b**) and MSI (**c**). LOH generally represents the loss of an adjacent tumor suppressor gene. MSI represents a condition of impaired mismatch repair ability

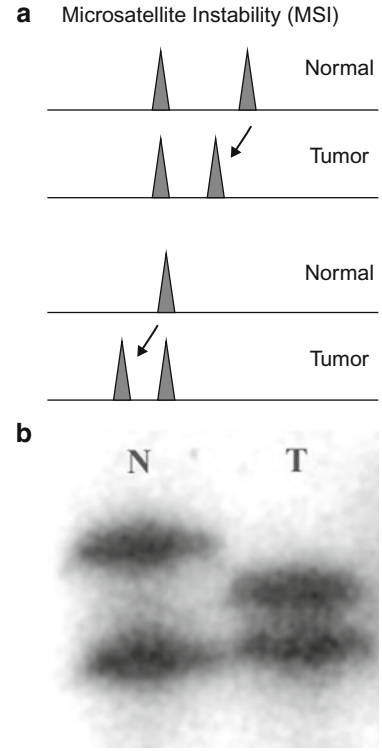


Fig. 5.19 MSI is a condition in which a group of cells acquires longer or shorter microsatellite alleles than normal cells. Alteration in the length of a microsatellite allele is due to deletion or insertion of single nucleotides or repeating units. MSI is different from LOH in that the tumor retains both alleles. Panel A is a schematic MSI pattern showing that one allele in tumor DNA has shifted its location (*arrow*) relative to that which is seen in normal control DNA. Alternatively, a shift from a noninformative control (one peak in normal) to a double-peaked (*arrow*) pattern in tumor DNA indicates the presence of MSI. Panel B is a gel photograph of MSI demonstrating that tumor (T) has an upper allele with reduced length

- The same technology can detect both MSI and LOH at a specific locus, but MSI is less common than LOH in solid tumors (Fig. 5.19)
- Accurate classification of two or more synchronous and metachronous tumors as independent tumors or as a primary tumor with metastatic foci has important clinical implications
  - Kaneki et al. (2004) analyzed a series of synchronous endometrial and ovarian adenocarcinomas
  - Based on MSI pattern comparison, the authors concluded that most patients (82%) could be diagnosed as having either single or double clonal tumors
  - MSI analysis can be used to distinguish primary versus metastatic origin of tumor

5.4.6 Viral Integration Analysis

- Oncoviruses are viruses associated with malignances and include viruses with DNA and RNA genomes
- Infection of oncoviruses plays an important role in the carcinogenesis of some cancers, including the associations of Epstein–Barr virus with lymphoma and nasopharyngeal carcinoma, of human papilloma virus with cervical cancer, and of hepatitis B virus with hepatocellular carcinoma

- The patterns of viral DNA incorporations are often concordant among tumor cells of the same clonal origin. Southern blot and PCR are commonly used methods for virus clonality analysis

5.4.7 Comparative Genomic Hybridization (CGH)

- CGH is a technique that permits the detection of chromosomal copy number changes in FFPE. CGH gives a global overview of chromosomal gains and losses in the whole genome of a tumor
  - CGH hybridizes green fluorochrome-labeled tumor DNA and red fluorochrome-labeled normal DNA to normal human metaphase preparations
  - Images of fluorescent signals are captured, and the green to red signal ratio is quantified for each chromosome locus along the chromosomal axis
    - If the signal from a locus is skewed to red, it indicates a loss of that chromosome locus in green-labeled tumor DNA
    - If the signal is skewed to green, it suggests a gain of that chromosome locus in tumor DNA
- The characteristic gain or loss of chromosomal material in a clonal tumor will be shared by the whole clonal population and derivatives, which can be used as a clonal marker
  - Jiang et al. (2005) investigated paired samples of primary cancer and pulmonary metastases obtained from patients who had undergone two consecutive surgeries
  - The overall CGH profiles were similar between primary carcinomas and their pulmonary metastases, indicating a common clonal origin of primary and metastatic tumors

5.4.7.1 Array Comparative Genomic Hybridization

- Array CGH combings the CGH with microarray technology to detect copy number changes at a level of 5–10 kb, even detect the change at a resolution of 200 bp by high resolution CGH
  - DNA from a test sample and normal reference sample are differentially labeled with fluorophores and hybridized to thousands probes printed on a glass slide
  - The fluorescence intensity of the sample and reference DNA is measured to calculate the ratio between them and subsequently the copy number changes for a particular location in the genome

5.4.8 Gene Expression Profiling/ Array-Based Clonality Analysis

- See Sect. 5.9 and Chap. 10 for more details

5.4.9 MicroRNA Signatures

- MicroRNA (miRNA) is a single-stranded RNA molecule encoded by genes. It is transcribed from encoding gene but not translated into protein (see Chap. 3 for more details)
- miRNA profiles in tumors reflect the developmental lineage and differentiation state of the tumors
  - A tissue-specific miRNA-expression signature is highly specific and reproducible and will be shared among cells with a common clonal origin
  - miRNA-expression profiling of human tumors can aid in diagnosis, staging, prognosis, and prediction of treatment response and determination of clonal origin
  - microRNA signature represents the tissue specificity of the tumor, which can be used to determine the tissue origin of a tumor
- Technical approaches
  - A set of oligoprobes corresponding to human miRNA are made into a microarray
  - The miRNA is extracted and reverse transcribed into cDNA
  - The cDNA is then PCR amplified, labeled, and hybridized to the array
  - The data is analyzed by an array analyzer

5.4.10 Protein-Based Clonality Analysis

- Isoenzyme-based clonality analysis, such as comparison of different isoforms of G6PD isoenzymes (see Sect. 5.9, X chromosome inactivation for more details)
- Proteomics (see Chap. 11 for more details)
- Immunohistochemistry (IHC)-based clonality analysis
  - Tumors from a common clonal origin express a common set of protein markers
  - Tissue-specific antigens, such as prostate-specific antigen (PSA), thyroid transcription factor-1 (*TTF-1*), CK7, CK20, CDX2, *OCT4*, and so on, are helpful in defining tumor origins and clonal relationships of tumors
  - Overexpression of oncogenes, such as *ras* and *myc*, can be readily detected in many cancers and in their metastases, and detection of these gene products has been explored as a means of clonality assessment

5.5 Tissue Contamination and Patient Identity Mismatch Testing

5.5.1 Overview

- Patient identity mismatches may happen before or after a specimen arrives in the laboratory. Possible sources of intralaboratory contamination include the dissection bench, instruments, ink marking of tissue margins, sectioning, and the water bath used in histologic sectioning
- Clonality analysis is one of the most commonly used approaches for identity testing. Molecular identity testing is based on the principle that cells from the same individual share a common set of genetic characteristics (Fig. 5.20)
- Commonly used methods include microsatellite profiling (Combined DNA Index System [CODIS]), FISH for sex chromosome determination, and human leukocyte antigen system (HLA)-related genotyping

5.5.2 Technical Approaches

- FISH (Fig. 5.21)
  - FISH is usually the first step in identity testing to exclude sex-mismatched contamination
    - Sex-mismatched tissue contaminations can be readily resolved by FISH through the detection of differing sex chromosome patterns (XX and XY) in a FFPE tissue section
    - Both host tissue (known to be from patient's sample) and suspected contaminating tissue are processed and hybridized with fluorescent-labeled DNA probes for X and Y chromosomes. The slides then are observed under fluorescence microscopy for the X and Y signals
    - The cells with XX signal pattern in their nuclei are from a female; the cells containing XY in their nuclei are from a male individual. A mismatched sex chromosome pattern indicates that the tissues in a sample are derived from different individuals
    - If the same sex chromosome pattern is observed in the known patient's tissue and the suspected contaminating tissue, further testing is needed to determine whether the tissues are derived from one or more than one individual
- The marker amelogenin, encoded by *AMELX*, is also frequently used in sex determination
  - The *AMELX* gene intron 1 on X chromosome is 6-bp shorter that its Y chromosome counterparts
    - PCR amplification of intron 1 of the AMELX gives rise to a 106-bp PCR product, but the Y chromosome gene, the AMELY, gives rise to a 112-bp PCR product
    - After gel electrophoresis, a male patient (XY) will show two bands (106-bp and 112-bp PCR product); a female patient (XX) will show single band of 106-bp PCR product

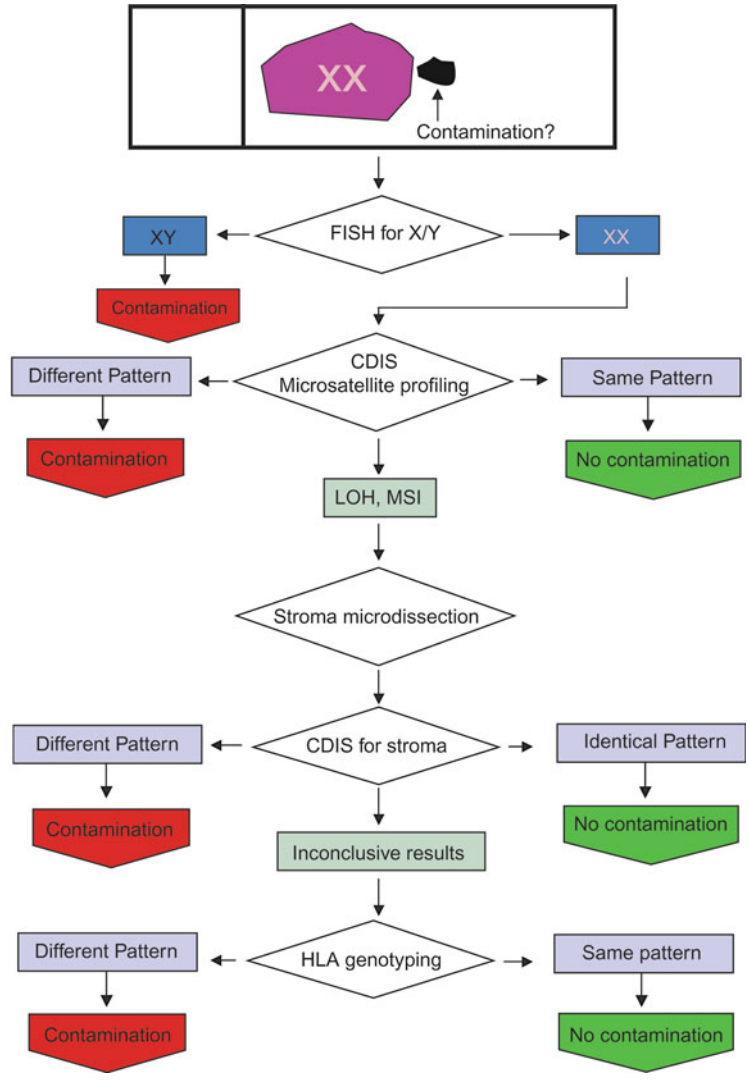


Fig. 5.20 Diagnostic strategies for identifying tissue contamination or patient identity mismatches. FISH for sex chromosomes is usually the first test used to distinguish sex-mismatched samples. When the suspected contaminating tissue displays the same pattern of sex chromosomes as the patient control tissue, microsatellite markers may be employed to assess for contamination (identity mismatch). CODIS loci are usually used for microsatellite profiling. There are three possible outcomes: the suspected contaminating tissue may display a different microsatellite profile, consistent with specimen contamination; the suspected contaminant may display an identical microsatellite profile, ruling out specimen contamination; or the test may yield inconclusive results due to LOH or MSI in tumor samples. In this latter situation, stromal cell DNA can be tested to resolve the issue. HLA genotyping may also be employed in these cases

- Microsatellite profiling and CODIS for identity testing
  - Microsatellite profiling is a technique which uses polymorphisms of microsatellite loci, which are unique for each individual. Polymorphism patterns of multiple microsatellite loci are shared by every cell from the same individual, and determining this pattern is referred to as "DNA fingerprinting" or "DNA typing." Alleles from different individuals can be unambiguously defined by the differing numbers of microsatellite repeats, which are expressed as lengths of amplified fragments (Fig. 5.21)
  - Microsatellite markers were used for human identity testing in the early 1990s. In 1997, 13 DNA polymorphic markers were identified as core markers for the US national database known as the CODIS (http://www.fbi.gov/hq/lab/codis/index1.htm). This system covers microsatellite loci on chromosomes 2, 3, 4, 5, 7, 8, 11, 12, 13, 16, 18, and 21 (Table 5.5)
    - In the United States and United Kingdom >5 million criminal justice DNA

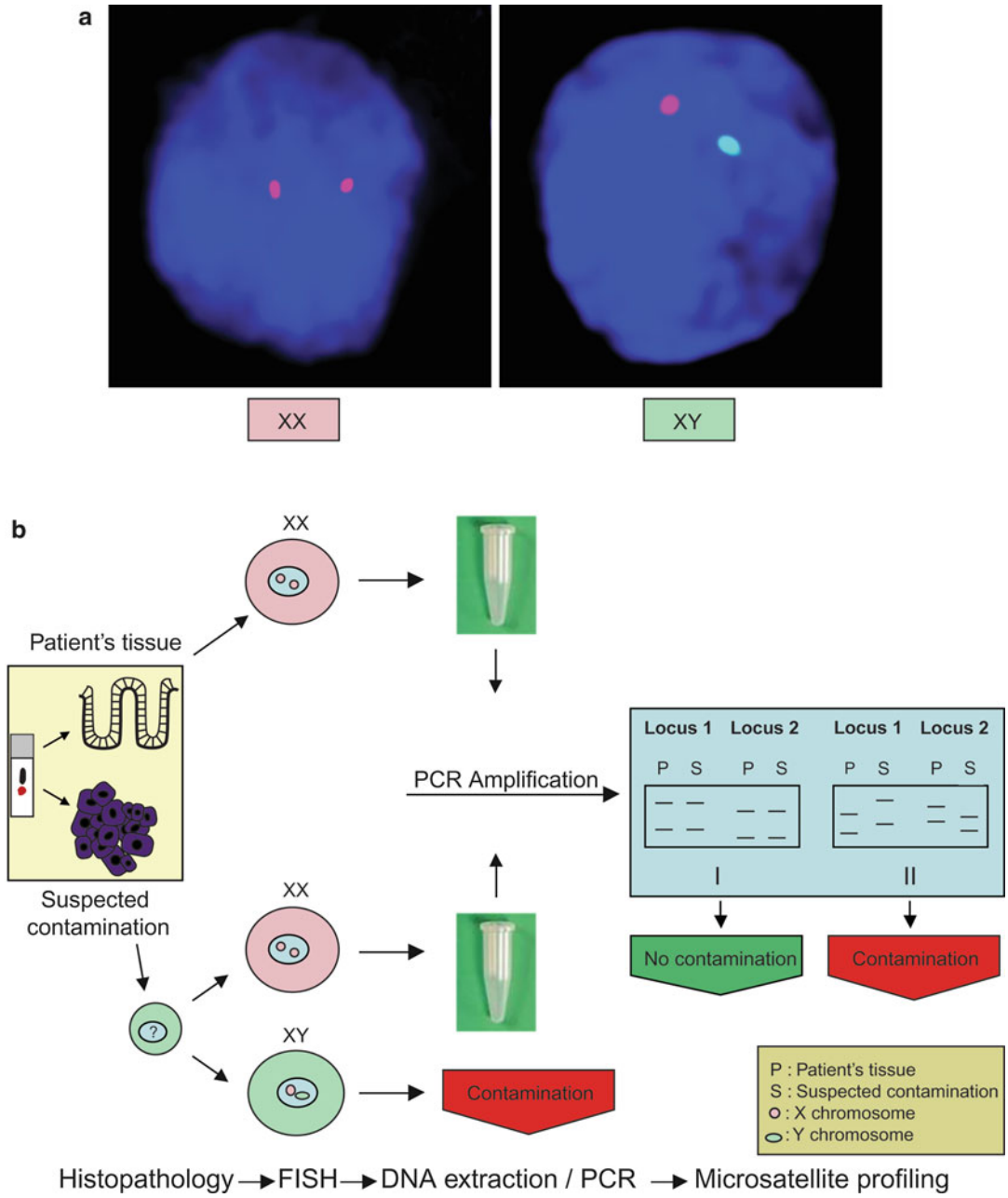


Fig. 5.21 Molecular identity testing is based on the principle that cells from the same individual share a common set of genetic characteristics. FISH testing for sex chromosomes will distinguish sex-mismatched cases (**a**). However, there is a 50% chance that the suspected contaminating tissue is from a same gender individual. Microsatellite profiling can be used to differentiate the tissues (**b**). Genomic DNA is extracted from the tissues and PCR amplified followed by gel electrophoresis. Schematic allelic patterns are shown in the *box* on the *right*. P and S designate patient and suspected contaminating tissues, respectively. (I) Concordant allelic patterns indicating the tissue is from the same individual; (II) discordant allelic patterns indicating that the tissue is from different individuals. Only informative loci are shown here for illustrative purposes

Table 5.5 The combined DNA index system

Locus	Repeats	Location	Fragment size
TPOX	5–14	2p23-pter	102–138
D3S1358	8–20	3p21	99–147
FGA	12–51	4q28	308–464
D5S818	7–16	5q21–31	133–169
CSF1PO	6–16	5q33–34	281–321
D7S820	5–15	7q	205–234
D8S1179	7–19	8q24.1–24.2	157–205
TH01	3–14	11p15	146–190
vWA	10–25	12p-pter	122–182
D13S317	5–16	13q22–31	157–201
D16S539	5–15	16q22–24	141–173
D18S51	7–27	18q21.3	262–342
D21S11	24–38	21q21.1	186–244

profiles are in the database. The random match probability for each allele locus ranges from 8×10^{-4} to 7.2×10^{-19} depending on the polymorphism and number of repeats at each locus

- Commercially available microsatellite panel kits and automated systems such as the ABI3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) allow the screening of multiple short tandem repeat (STR) loci in one run. With proven high specificity and sensitivity, CODIS markers have become the most commonly used loci in identity testing. Small screening panels can confirm an identity at a specified level of confidence
- Brief procedures for microsatellite identity analysis (Fig. 5.21)
  - The tissues are harvested separately from known host tissue or suspected contaminating tissue using tissue microdissection; genomic DNA is isolated. Patient blood DNA, if available, is also extracted as a reference
  - The DNA from different tissue origins is amplified using microsatellite locus-specific primers. The PCR products are separated by polyacrylamide gel electrophoresis or capillary electrophoresis
  - The allelic patterns from the known patient DNA sample and from the suspected contamination DNA are compared
  - If the tissues are derived from the same patient, both DNA samples will show identical band patterns (also identical to that of the patient blood DNA reference); simultaneous detection of different allele patterns over multiple loci strongly suggests that the tissues are from different individuals
- HLA genotyping is useful in tests of identity mismatch or tissue contamination
  - The HLA system is a group of genes that reside on chromosome 6, which encode cell surface antigen-presenting proteins and many other proteins. The chance of unrelated individuals having identical HLA genotypes is very low; therefore, the HLA genotype can be used as a clonal marker
  - A commercial kit can be used to amplify and distinguish 21 different genotypes at a polymorphic HLA locus
  - Brief procedures
    - DNA is extracted from a tissue section, then PCR is amplified against the HLA DQ-α locus (6p21.3), followed by hybridization to nylon membrane strips with allele-specific HLA-DQ-α probes
    - The pattern of signal dots indicates the homology to the HLA-DQ-α alleles, which reflects the genotype
    - Different genotypes indicate that tissue samples are from different individuals

5.5.3 Caveats

- The ability of a single microsatellite marker to distinguish between individuals depends on the degree of polymorphism the marker exhibits. The reported rate ranges from 59% to 91%, but absolute matching requires several polymorphic microsatellite markers
- Since most of the analyzed materials are from FFPE tissue, the rate of detectable amplification products declines with increasing amplicon length. Shorter fragments, preferably <200 bp, should be used to obtain a reasonable sensitivity for the microdissected samples

- The standard, or optimal, number of DNA loci that needs to be examined for tests of tissue contamination and patient identity mismatch has yet to be determined
  - Many studies have shown that the use of five to eight polymorphic DNA loci could well achieve a high specificity for mismatch detection
  - The use of nine markers achieved a power of exclusion of 99.7% in a paternity test (a power of exclusion of 99% is legally required)
  - The available commercial kits use 3–16 DNA markers. The amelogenin gender-determining marker is often included
  - The most frequently used markers are derived from the CODIS panel, such as FGA, TH01, TPOX, VWA, D3S1358, D8S1179, D18S51, and D21S11 (Table 5.5)
  - The probability of random matching of unrelated individuals at 3, 4, and 8 markers is about 1.0×10^{-3}, 7.8×10^{-4}, and 7.4×10^{-10}, respectively. The probability of two unrelated Caucasian individuals matching at 12 loci is about 1.12×10^{-12} based on published heterozygosity values for the known markers
- The high LOH and MSI rate in tumor tissues may interfere with the microsatellite profiling tests, especially for tissues consisting predominantly of tumor cells. Alternative tissues such as stroma could be analyzed. HLA genotyping may also be used in difficult cases

5.6 Identification of Donor Origin in Transplantation Patients

5.6.1 Overview

- The incidence of cancer in solid organ transplantation recipients is reported to be 4–18%. Tumors in transplant recipients may be transmitted inadvertently through the transplanted organ
- Most recipients of solid organ transplantation receive immunosuppressive therapy for prolonged periods, which greatly increases the risk for cancer development. Tumors may also arise de novo in the transplanted organ
- Cancer incidence has been reported to be 40–50%, 20 years after renal transplantation
- The identification of tumor origin is of biologic, clinical, therapeutic, and possible legal importance

5.6.2 Technical Approaches (Fig. 5.22)

- Similar approaches to those used in patient identity mismatch testing are used (see previous discussion)
- FISH
  - The tumor origin of sex-mismatched organ allograft can be differentiated through sex chromosome patterns
- Microsatellite profiling
  - The CODIS is often used as a microsatellite profile panel, which can distinguish tissues of donor or recipient origin
  - HLA genotyping could also be used in the analysis of tumor origin after organ transplantation (see previous discussion)

5.6.3 Caveats

- During microsatellite profiling, donor and recipient alleles sometimes could be presented simultaneously due to the presence of recipient stroma and blood cells in the transplanted organ. Precise microdissection could help to avoid contamination with recipient cells
- Microsatellite alleles may undergo alterations such as LOH and MSI. Multiple microsatellite loci are needed in creating a reliable microsatellite profile
- Ideally, DNA samples from donor (blood or paraffin-embedded tissue), recipient (blood or paraffin-embedded tissue), and tumor are analyzed in parallel

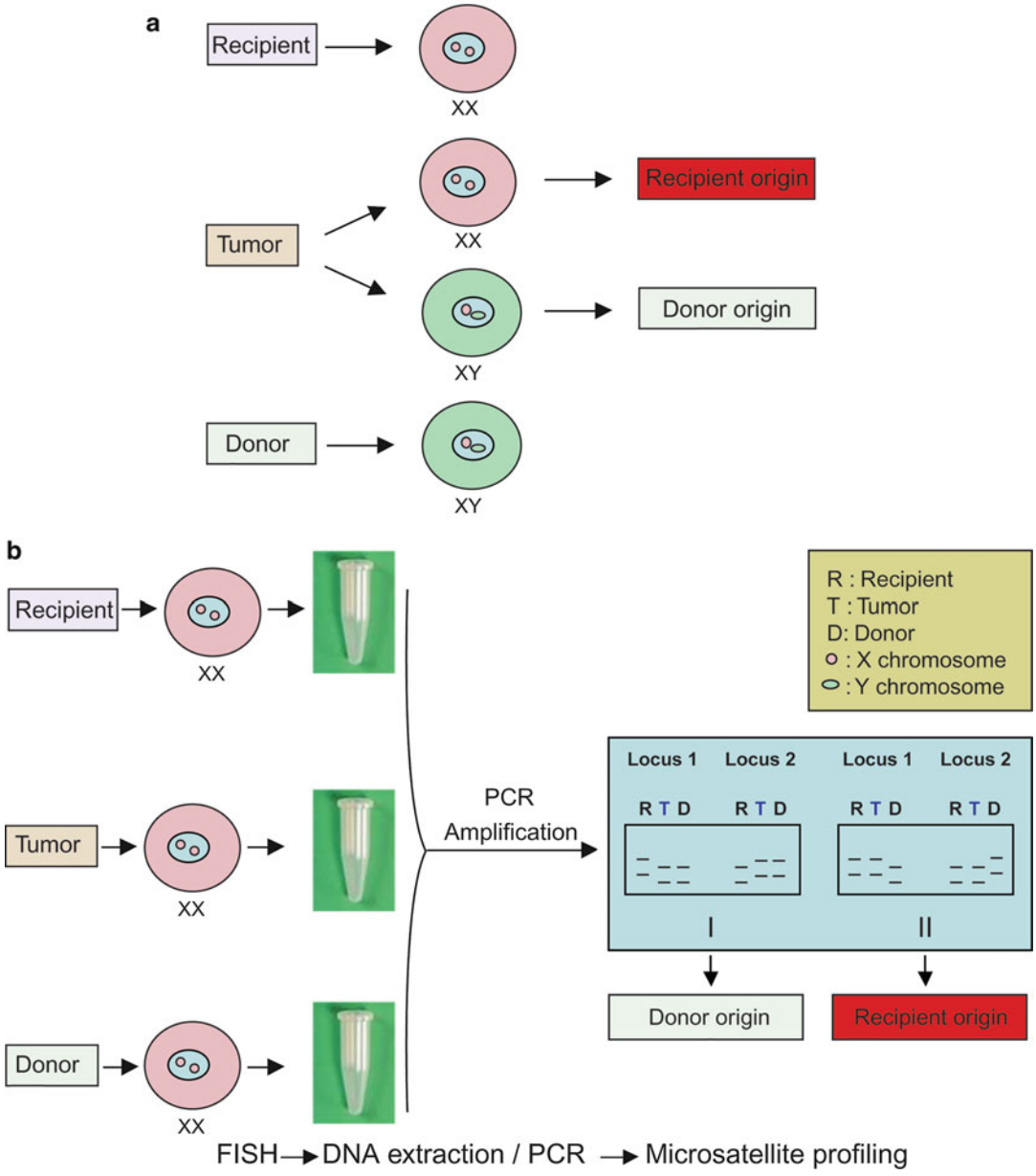


Fig. 5.22 Posttransplantation tumors can originate from the recipient, can be transmitted with the transplanted organ, or can arise de novo from the donor organ. FISH can be used to differentiate sex-mismatched cases by showing different sex chromosome patterns (panel A). Microsatellite profiling can distinguish donor and recipient by comparing the allelic patterns at microsatellite loci (panel B). Genomic DNA is extracted from donor (D), recipient (R), and tumor (T) and is PCR amplified. As shown in the green box, a tumor of donor origin will show an allele pattern identical to that of the donor (I). A tumor of recipient origin will show an allele pattern identical to that of the recipient (II). Only informative loci are shown here for illustrative purposes

– In many incidences, donor DNA is not available. Accurate results could still be obtained in these circumstances if the experiments are carefully controlled with meticulous attention to the selection of markers

5.7 Bone Marrow Engraftment Testing

5.7.1 Overview

- Hematopoietic stem cell transplantation, also known as bone marrow transplantation, is a procedure in which hematopoietic stem cells are transferred from a donor to a recipient. Allogeneic bone marrow transplantation is primarily used for the treatment of hematopoietic malignancies and hereditary disorders
- Engraftment is the process by which transplanted donor stem cells begin to proliferate and produce blood cells within the recipient. Engraftment of bone marrow transplants usually takes between 10 and 20 days to occur
- A low percentage of hematopoietic cells from the recipient may still exist after successful stem cell transplantation. This condition is called mixed chimerism (MC) and can be detected in the peripheral blood of the recipients after transplantation
- Engraftment testing is used to monitor the efficiency of engraftment, to detect graft rejection, and to assess the risk for relapse and the effectiveness of therapy

5.7.2 Technical Approaches

- FISH
  - This method is only suitable for recipients with a sex-mismatched donor
  - Posttransplantation blood cells are prepared and hybridized with fluorescently labeled probes for chromosomes X and Y
  - The cells with donor and recipient chromosome profiles are counted and calculated for the percentage of recipient cells
- Microsatellite profiling
  - Engraftment testing is performed using gene amplification of microsatellite markers in pretransplant and posttransplant samples of donor and recipient DNA
  - CODIS markers are among the most frequently used markers
  - Pretransplantation blood from both donor and recipient are needed to determine the informative microsatellite loci to be used in the analysis. The posttransplantation analysis establishes relative amounts of recipient and donor cells in the recipient's blood (Fig. 5.23)
  - Brief procedures
    - Genomic DNA is extracted from the donor and recipient pretransplantation bone marrow or blood specimens and from the recipient posttransplantation blood or bone marrow specimens
    - Informative microsatellite loci are selected according to the results of donor and recipient pretransplantation DNA testing
    - The DNAs are amplified by PCR, and the DNA products are separated by capillary electrophoresis or gel electrophoresis
    - The donor and recipient patterns are compared, and the ratio of recipient/donor cells is calculated
      - The relative intensity of recipient to donor is calculated by summation of two recipient allele intensities divided by summation of two recipient and two donor allele intensities
      - $MC = (R1 + R2)/(R1 + R2 + D1 + D2) \times 100\%$ (R1, R2: recipient allele 1 and recipient allele 2; D1, D2: donor allele 1 and donor allele 2)
    - Alternatively, the relative percentage can also be inferred and calculated by comparing the recipient allele intensity to the control sample containing 5% mixed recipient DNA

5.7.3 Caveats

- Selection of informative microsatellite markers is critical for a successful engraftment

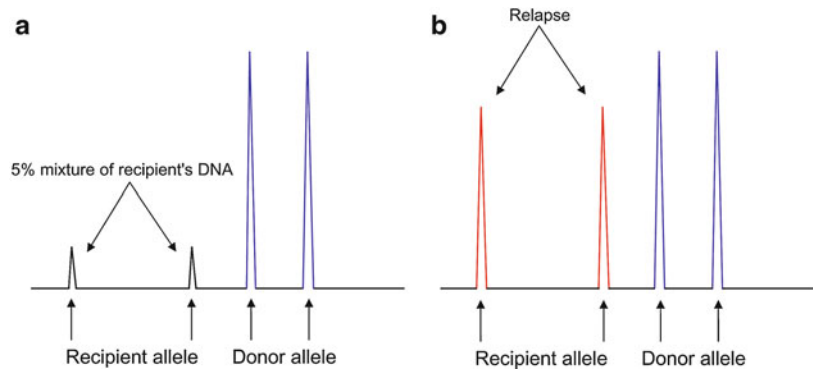


Fig. 5.23 Bone marrow engraftment analysis. The schematic illustration shows microsatellite locus amplification a short time after bone marrow transplantation (**a**) and several months after bone marrow transplantation (**b**). The main peaks in panel (**a**) represent donor alleles (reconstitution of the hematopoietic pool). The lower peaks on the *left* are from a 5% mixture of recipient DNA. There are no residual recipient cells detectable in the blood at this time. Several months after bone marrow transplantation, the test shows increased recipient alleles (*red*) from recipient blood cells (B), consistent with disease relapse

test; the donor and recipient pretransplantation blood should be tested to ensure that loci are informative (two alleles are present) and nonhomologous (the band sizes are different between donor and recipient)
- Many laboratories use 10 or more microsatellite loci to ensure the accuracy of the test
- A higher percentage or increasing level of recipient cells carries an increased risk for relapse. However, the exact cutoff value for the level of MC is yet to be established. Some studies suggest that detection of >1% of recipient cells correlates with an increased risk of relapse
- Measurement of chimerism in a particular subpopulation can be achieved by selective enrichment of target cells using flow cytometry (see Chap. 9 for more details)

5.8 Molecular Diagnosis of Hydatidiform Mole

5.8.1 Overview

- Hydatidiform mole is a trophoblastic proliferation disease characterized by marked enlargement of villi. About one in every 1,000 pregnancies is affected with much higher incidences in Asia
- This condition is characterized by abnormal development of both embryonic and extraembryonic tissues and is associated with abnormal chromosomes
- Hydatidiform moles are divided into two types: complete and partial moles (Table 5.6)
  - About 20% of patients with complete moles and 5% of patients with partial moles may develop persistent gestational trophoblastic disease
    - Partial moles are mostly triploid (69 XXX, 69 XXY, or 69 YYX) with a genome that is almost always composed of two sets of chromosomes of paternal origin (diandric) and of a haploid maternal set
    - Partial moles result either from the fertilization of a single ovum by two different haploid sperms or from fertilization by a single sperm with a duplicated genome
    - An extrapaternal set of chromosomes results in trophoblast overgrowth with underdevelopment of the embryo
    - The risk for progression to choriocarcinoma is lower compared with a complete mole (Table 5.6, Figs. 5.24 and 5.25)

Table 5.6 Characteristics of complete and partial mole

	Chromosome origin	Proposed mechanism	Karyotyping	Fetal parts	FISH	Microsatellite[a]	Risk of choriocarcinoma
Partial mole	Maternal and paternal	Two sperm fertilized one normal ovum	69XXY (70%) 69XXX (27%) 69XYY (3%)	Underdeveloped embryo	Trisomy	3 alleles (PPM)	Lower risk
Complete mole	Paternal	Empty ovum fertilized by 2 sperms or fertilized by 1 sperm with duplicated genome	46XX (90%) 46XY (10%)	Do not form	Disomy	2 alleles (PP)	2%
Villus from normal pregnancy	Maternal and paternal	–	46XX, 46XY	Embryo formed	Disomy	2 alleles (MP)	Exceedingly low

[a]Only informative alleles are considered here for illustrative purpose. *M* maternal allele, *P* paternal allele

- A complete mole is diploid
  - About 90% of complete moles are 46XX and 10% are 46XY
  - The genomes of complete moles are purely paternal and fetal parts do not form
  - Complete moles result from androgenesis, in which an empty ovum is fertilized by two sperms or by one sperm that has duplicated its genome
  - The risk for progression to choriocarcinoma is about 2%

5.8.2 Technical Approaches

- The basis for the molecular diagnosis of moles is the distinctive genome compositions of partial and complete moles
  - FISH can differentiate partial and complete moles by detection of a different numbers of sex chromosomes
  - Microsatellite profiles can distinguish partial and complete moles by demonstration of the allelotype and dosage differences of paternal and maternal alleles (Fig. 5.25)
- FISH detection of sex chromosomes (Figs. 5.24 and 5.25)
  - Partial moles are composed of trisomic cells with one maternal X and two paternal sex chromosomes
    - FISH shows an XXY, XYY, or XXX pattern
  - Complete moles are composed of cells with diploid genomes that are made up entirely of paternal chromosomes
    - The FISH pattern could be XX or XY
    - FISH could not distinguish complete mole from normal pregnancy
- Microsatellite profiling
  - CODIS loci have been used in many studies (see Sect. 5.5 for more details)
  - The microsatellite loci from villus DNA and from a maternal control are amplified, and the allele patterns are compared (Fig. 5.25)
  - Partial moles show two paternal (P) alleles and one maternal (M) allele as a three allele pattern at each locus (MPP pattern)
  - Among the three alleles demonstrated, one is of maternal origin, matching a maternal allele position (Fig. 5.25, Table 5.6)

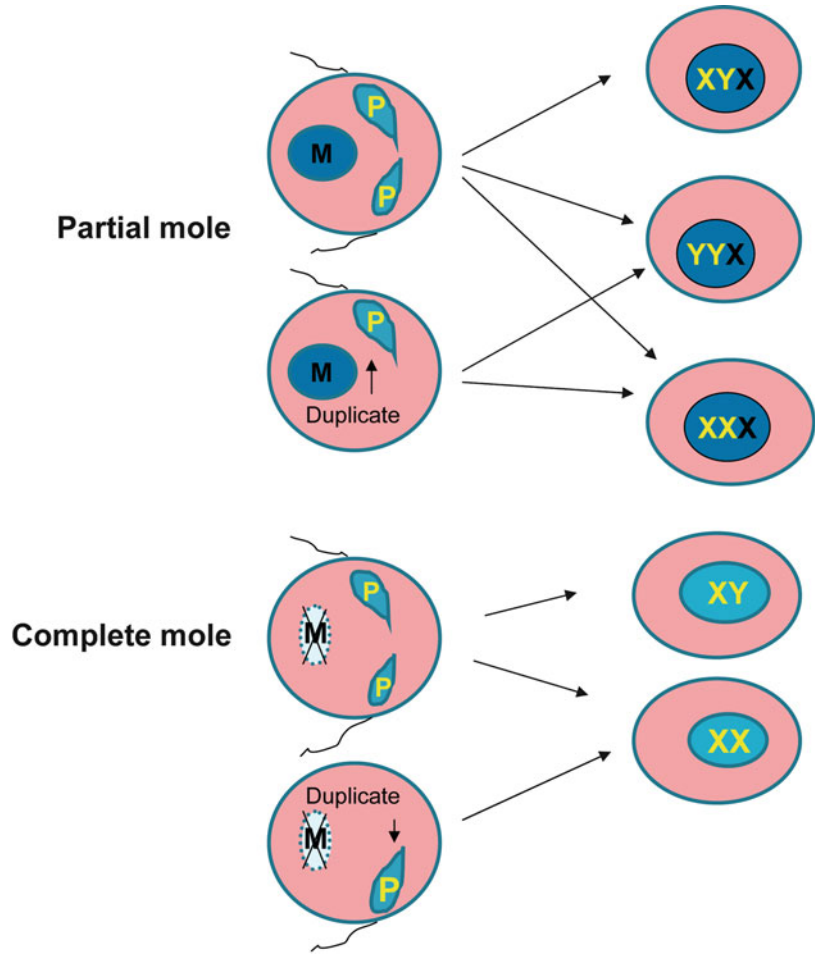


Fig. 5.24 Mechanisms of partial and complete mole formation. Partial moles are derived from both maternal and paternal chromosomes and are mostly triploid (69 XXX, 69 XXY, and 69XYY) with two sets of paternal chromosomes (P, *yellow*) and a set of chromosomes of maternal origin (M, *black*). Partial moles may result from two sperms fertilizing a haploid ovum or from one sperm fertilizing the ovum and duplicating its chromosomes. Complete moles are diploid and their chromosomes are purely paternal. Complete moles result from two sperms fertilizing an empty ovum or from one sperm fertilizing the empty ovum and duplicating its chromosomes. Ninety percent of complete moles are 46XX and 10% are 46XY

- Partial mole may also show an MP pattern if two paternal alleles are homozygous
- Complete moles show an allelic pattern identical to the paternal allele and may be either homozygous (P) or heterozygous (PP) (Fig. 5.25). A maternal allele will be absent
- Normal villi will show one paternal and one maternal allele (MP pattern)

5.8.3 Caveats

- FISH testing cannot distinguish a normal villus from a complete mole. Therefore, microsatellite profiling is preferred in many laboratories
- For microsatellite polymorphism testing, the genomic DNA that is used for a maternal control can be obtained either from blood or from maternal tissue microdissected from paraffin sections (such as endometrium)
- Because some microsatellite markers are noninformative, it is critical that multiple markers be tested to allow for accurate interpretation

5.9 Cancer of Unknown Primary Origin (CUP)

5.9.1 Overview

- CUP is defined as a metastatic malignant tumor whose primary site of origin is not identified after routine diagnostic procedures

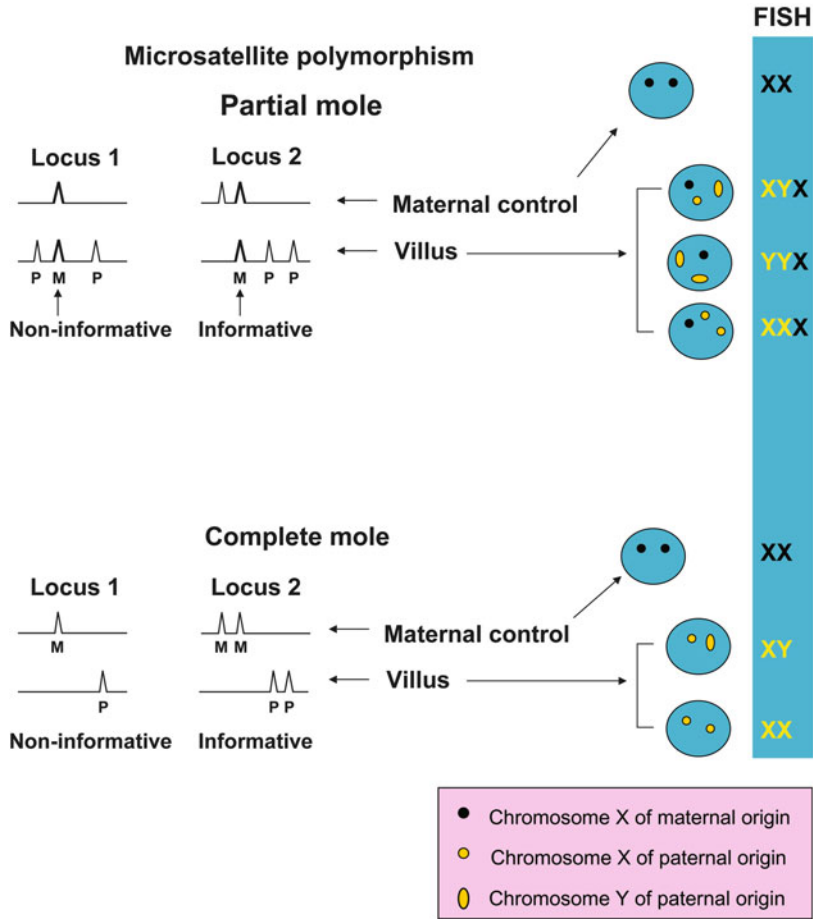


Fig. 5.25 Molecular diagnosis of molar pregnancy using microsatellite allelotyping and FISH. DNA samples are extracted from a maternal control (blood) and from the mole (villus). Partial moles (*upper* panel) show one maternal allele (bold, M and *arrow*) and two paternal alleles (P). Complete moles (*lower* panel) show no maternal allele but may show one or two paternal alleles. Locus 1 represents a noninformative DNA marker as the maternal control is homozygous at this locus. Locus 2 represents an informative DNA marker as the maternal control is heterozygous at this locus. Multiple polymorphic microsatellite loci should be employed in the analyses. FISH can also be used for mole identification. As shown in the *right* panel, partial moles show an XXY, XXX, or XYY pattern (maternal, *black*; paternal, *yellow*) in contrast to the XX pattern in the maternal control. Complete moles are diploid with either an XX or XY complement of sex chromosomes, both of paternal origin (*yellow*)

- CUPs comprise about 3–5% of all human cancers
  - One of the 10 most frequently diagnosed cancers in developed countries
  - Approximately 30,000 cases of CUP are diagnosed in the United States annually
- CUP represents a heterogeneous collection of tumor types and clinical presentations
- Diagnosis requires extensive diagnostic workup, including a thorough clinical history and physical examination, detailed cytohistopathologic morphological studies, immunohistochemistry (IHC) panels, serum tumor markers, and modern imaging technology to help identify the site of primary tumor
  - There is no consensus on an IHC panel and timing of diagnostic tests for all patients at clinical presentation
  - Detection of a primary site is not always accomplished before the patient's demise, but the clinically occult primaries have been found at autopsy in about 75% of patients

- Lung and pancreas are the most common primary sites of origin in CUP autopsy series
- Approximately 60% of CUPs are classified as well to moderately differentiated adenocarcinomas, 30–35% are poorly differentiated or undifferentiated neoplasms (including carcinomas), and 5% are squamous cell carcinomas
- Identification of a primary tumor can help to predict tumor behavior and to determine appropriate therapy
  - Therapies for the majority of CUP patients have been empiric and relatively ineffective
  - Patients with a CUP diagnosis have a dismal prognosis with a median survival of less than a year
- Mechanisms of CUP
  - The primary tumor has involuted or regressed and is no longer detectable by conventional methods when the metastasis becomes widespread
  - Alternatively, CUP may represent a unique tumor type in which the primary tumor has acquired a special metastatic phenotype and genotype soon after tumor initiation
    - These tumors have a preference for metastatic spread over local tumor growth

5.9.2 Diagnostic Approaches

- Principle: cancer is a clonal process. The cells in a metastatic tumor share a similar phenotype and genotype with the tumor cells from the primary site (Table 5.7)

Table 5.7 Diagnostic strategy for discovering an unknown primary tumor

Clinical
Medical history and physical examination
Age and sex
Race
Location of metastasis
Prior medical history
Family history of cancer
Cancer serum markers such as PSA and α-fetoprotein
Imaging analysis
Chest X-ray
Mammography
Computed tomography
Positron emission tomography (imaging)
Magnetic resonance imaging
Pathology
Light microscopic examination
IHC using tissue-specific markers such as PSA, TTF1, and OCT4
Ultrastructural (electron microscopy) examination
Molecular
Cytogenetic and FISH analysis
Isochromosome 12p for germ cell tumor
Selected translocations
Others
Mutational screening
Oncogenes (*EGFR*, *c-kit*, or *PDGFR*)
Tumor suppressor gene (*TP53*)
Metastasis suppressor gene (*Kiss-1*)
Gene expression profiling/array-based analysis
CGH (Array CGH)
SAGE
Viral integration analysis (such as EBV, HPV, and HBV genome detection)
MicroRNA profiling
Proteomic profiling

5.9.2.1 Imaging, Clinical, and Morphologic Approach

- The diagnostic workup for patients with CUP requires
  - Epidemiologic data (including age, sex, and race), prior medical history, location of metastasis, and careful physical examination
  - Imaging studies such as computer-assisted tomography (CAT), magnetic resonance imaging (MRI), and positron emission tomography (PET) scans
  - Routine pathological studies with immunohistochemistry (IHC) panels targeted to the primary site differential
  - Serum tumor markers for tumors of suspected endocrine origin
- Pathology
  - Histologic evaluation using light microscopy
  - Hematoxylin and eosin evaluation of the metastatic lesion remains the cornerstone

for the identification of a primary tumor in CUP patients
- Tissue-specific protein markers can be identified using a selected IHC panel of antibodies
  - The most common algorithm for IHC panels include the use of cytokeratin (CK) subtypes (e.g., CK7 and CK20) that allow broad categorization of the tumor based on their pattern of expression (e.g., upper or lower gastrointestinal tract)
  - Cytokeratin studies are often supplemented with IHC stains for antigens with restricted expression in specific tissue types (e.g., thyroid transcription factor-1, thyroglobulin, CDX2, prostate-specific antigen, mammaglobulin)
  - Numerous antibodies are now commercially available in the routine workup for CUP (such as PSA, HMB45, TTF1, CDX2, CK7, CK20, and OCT4)
- Poorly differentiated tumors often lose the expression of such differentiation markers
- Validation of the sensitivity and specificity of IHC markers is usually performed with normal tissues and known primaries and is usually not validated with undifferentiated neoplasms, and thus, the sensitivity and specificity with CUP cases is unknown
- A meta-analysis study revealed that histopathologic evaluation with IHC studies was successful in identifying the origin for metastatic tumors in 65.6% of cases
- Fine needle aspiration of body fluid sample is essential to establish a diagnosis of CUP since it helps inform the focus of imaging studies to follow
- Electron microscopy
  - Ultrastructural analysis may be useful in the diagnosis of a primary tumor in CUP patients. Characteristic ultrastructural features of tumor cells can be recognized and aid in the differential diagnosis
- Serum markers are helpful when hormone-secreting neoplasms are considered in the differential (e.g., β-HCG, AFP, and PSA)
  - Some commonly tested epithelial serum markers (e.g., CA19–9, CA15–3, CEA, CA125) have low specificity for specific tissue types

5.9.2.2 Molecular Diagnostics Approach

- The ability to distinguish cell types based on gene expression profiles that reflect their tissue of origin and function has allowed the development of assays that measure mRNA or miRNA transcripts in tissue specimens from patients with unknown primary tumors
- Three different molecular gene expression assays for determination of tissue of origin are available for clinical use in the United States
  - A real-time quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) assay that evaluates 92 mRNA transcripts (bioTheranostics, CancerTYPE ID®)
  - A microarray-based assay to evaluate mRNA transcriptomic patterns comprising 2,000 genes (Pathwork Diagnostics, Tissue of Origin Test)
  - A qRT-PCR assay that evaluates 48 microRNA (miRNA) mature transcripts (Rosetta Genomics, miRview® mets2)
- Although these assays vary in performance across specific tissue types, the overall accuracy in identifying tissue of origin in studies using known primary samples appears to be in the range of 75–89%
- All of these assays can be successfully performed using RNA from FFPE tissues, thus allowing their use with specimens from routine diagnostic procedures
- Cytogenetics and FISH
  - Characteristic chromosomal alterations may suggest a specific tumor type and organ site as these cytogenetic abnormalities are found in high frequency in certain tumors, such as t(X;18) in synovial sarcoma and trisomy 7/17 for papillary renal cell carcinoma (refer to Table 2 in Chaps. 1 and 2 for details)

  - Isochromosome 12p (i[12p]) is a marker for tumors of germ cell origin
    - Testicular cancer patients may have metastatic tumors of diverse histologic types that lack features of germ cell tumors
    - Kernek et al. (2004) found FISH for 12p amplification in routinely processed surgical specimens to be a useful adjuvant diagnostic tool in confirming the germ cell origin of metastatic tumors having the histologic appearance of somatic-type neoplasms
- Mutational screening
  - Overexpression of oncogenes and mutations of tumor suppressor genes are common in solid tumors. Detection of oncogenes and tumor suppressor gene mutations can be helpful in the identification of tumor primary sites and in clinical outcome prediction
- Gene expression profiling/array-based analysis
  - DNA microarrays are used to analyze genome-wide gene expression (gene expression profiling) (see Chap. 10 for more details)
  - It can monitor expression levels of thousands of genes simultaneously and can discriminate different disease states through qualitative and quantitative measures
  - An expression profiling study begins with samples from well-defined tumors, from which a specific panel of genetic markers (molecular signature) is established. Gene expression profiles of metastatic tumors closely resemble primary tumors of the same origin. Therefore, the gene expression profile can be used to identify tumor origin. Comparison to the relative abundance of specific panels of genes known to be expressed in different tissues (tissue-specific signatures) can provide clues to the cellular origin and primary site of a metastatic tumor
- Comparative genomic hybridization array (CGH)
  - See Chaps. 8–15 for more details
- Serial analysis of gene expression (SAGE)
  - See Chaps. 8–15 for more details
  - SAGE is a sequence-based sampling method for comprehensive analysis of the gene expression patterns. This technology does not require a preexisting known mRNA sequence; therefore, it can be used to identify new genes and to quantify known genes
  - The underlying principle of SAGE technology is that a short representative sequence tag (10–14 bp) can be used to uniquely identify a transcript, and the tag numbers directly reflect the abundance of corresponding transcripts
  - The SAGE libraries are highly accurate, quantitative, and comprehensive representations of the samples from which they are derived
  - Dennis et al. (2005) used 15 publicly available SAGE data libraries and found that a site of origin can be determined for a CUP with an 88% accuracy
- Viral integration analysis
  - Identification of viral genomic sequences in a CUP is useful to establish a primary site (e.g., identification of Epstein–Barr virus DNA in a CUP is helpful for establishing a nasopharyngeal primary in the appropriate clinical setting)
- miRNA profiling
  - miRNAs are short 20–22 nucleotide RNA molecules that are negative regulators of gene expression (see Chap. 7 for more details)
  - miRNAs are transcribed from RNA genes and are involved in many cellular processes including differentiation, proliferation, and the stress response
  - A distinct signature based on miRNA-expression profiling can be used to identify tumor origin
- Proteomic profiling
  - Proteomic methods detect a unique panel of proteins, which can differentiate one tissue from another according to established algorithms. These methods may provide insights into specific expressions of

proteins which could serve as unique tumor markers
  - See Chap. 11 for more details

5.9.2.3 CancerTYPE ID® Test (bioTheranostics Inc., San Diego, CA)

- The CancerTYPE ID® assay is a laboratory-developed test (LDT) that evaluates 92 mRNA transcripts to classify specimens into 30 tumor types and 54 histological subtypes
- This test uses RNA extracted from FFPE tissue through microdissection
- The gene expression measurement of 87 transcripts (content genes) is normalized to the expression of five reference transcripts, and the resulting profile is compared to a database of 2,206 tumor profiles, and a tissue of origin is assigned according to their similarity to tissues in the database
- The currently available version of the CancerTYPE ID® test (expanded version) was evaluated with through a leave-one-out cross validation analysis using the entire 2,206-sample reference dataset and a separate 187 sample test set
- The overall sensitivity of the assay has been reported as 87% at the main type level (30 cancer types) and 85% at the histological subtype level (54 subtypes) with the cross validation sample and 83% at the main type level with the independent 187 sample set
- In the study that reported the expanded version, the independent test set represented 28 tumor classes, with six tumor classes being represented by three samples or less with some cases in which 100% concordance indicates the correct classification of a single sample for a specific tissue type

5.9.2.4 Pathwork Tissue of Origin Test (Pathwork Diagnostics, Redwood City, CA)

- The Pathwork Tissue of Origin Test is an FDA-cleared assay that is performed using microarray-based gene expression profiling using RNA from FFPE tissues including cytology specimens
- The test evaluates the gene expression of 2,000 mRNA transcripts
- The FFPE version of the Tissue of Origin Test was developed from gene expression profiles using a database of 2032 frozen and 104 FFPE specimens
- The test's proprietary algorithm reports a similarity score (SS) which is a value ranging from 0 (very low similarity) to 100 (very high similarity) for each of 15 tissue types reported by the test
  - The highest SS indicates the likely tissue of origin, and an SS of <5 rules out tissue types
- A blinded multicenter validation study using 462 FFPE tumor specimens demonstrated an overall sensitivity of 89% (95% CI: 85–91%) and an overall specificity of 99.8%
  - A 91% agreement was seen with metastatic tumor specimens (comprising 38% of the of the sample)
- In the study that reported the FFPE assay, the independent test set represented 15 tumor types with representation of 25–57 specimens per tumor type
- An independent study with 43 poorly differentiated and undifferentiated tumor specimens showed 97% (95% CI: 80.4–99.8%) agreement between the Pathwork TOO Test result and the diagnosis
  - They also evaluated off-panel specimens in which the test showed significantly decreased performance but utility to exclude tissue types from the differential diagnosis
- Given that some off-panel tissues can have expression profiles similar to on-panel tumors, Pathwork has released new tests (LDTs) that distinguish between ovarian and endometrial tumors (tissue of origin endometrial) and between squamous lung and head and neck carcinomas (tissue of origin head and neck)

5.9.2.5 miRview® mets Test (Rosetta Genomics, Ltd., Philadelphia, PA, and Rehovot, Israel)

- The miRview® mets test assay is a laboratory-developed test (LDT) that evaluates 48 miRNA biomarkers by qRT-PCR and

classifies 25 different tumor types corresponding to 17 distinct tissues and organs
- The test uses two independent algorithms, a binary decision tree and a *k*-nearest neighbors (k-NN], to predict the tissue of origin for a given specimen
  - A high-confidence prediction is given when both algorithms agree and a low-confidence prediction when each predicts a different site of origin
- The test was developed through an initial selection miRNA markers from microarray data from 504 tumor specimens, which was then shortened to 48 genes from qRT-PCR gene expression data from a 356 FFPE training tumor sample
- Performance of the test was evaluated in an independent set of 204 specimens (188 yielded results) including nearly 100 metastatic tumors
- The study defined a correct result when any one of the two algorithms correctly identified the tissue of origin and reported an 85% accuracy using that definition
  - In 66% of the specimens, the algorithms were concordant (high-confidence classification); 90% agreed with the reference diagnosis
  - When using single prediction, the accuracy of the assay was 59%
- Tissue type representation for this study ranged from 1 to 38 specimens per tissue type
- In a prospective study, the test's results were consistent or compatible with the clinicopathologic features in 84% of cases processed successfully
- The miRview® mets^2 is a second-generation version of this assay which is reported to identify 42 primary origins
- Rosetta Genomics claims this version of the assay has an overall sensitivity of 85%, but this performance data has not yet been published in the peer review literature

5.10 Summary

- CUP represents a group of biologically and clinically heterogeneous tumors
- Management of patients with CUP requires a truly multidisciplinary approach and close communications between the pathologists and oncologists to ensure an effective workup
- A variety of clonality analysis methodologies can be employed to determine a site of origin for the metastatic lesion
- Gene and protein expression profiling technologies are potentially of great benefit in the identification of a primary tumor and may aid in the accurate assessment of prognosis
- Identification of CUP-specific molecular signatures may offer new avenues for diagnosis and treatment

Further Reading

Abbosh PH, Wang M, Eble JN, et al. Hypermethylation of tumor suppressor gene CpG Islands in small cell carcinoma of the urinary bladder. Mod Pathol. 2008a; 21:355–62.

Abbosh PH, Zhang S, MacLennan GT, et al. Germ cell origin of testicular carcinoid tumors. Clin Cancer Res. 2008b;14:1393–6.

Allen RC, Zoghbi HY, Moseley AB, et al. Methylation of *Hpa*II and *Hha*I sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. Am J Hum Genet. 1992;51:1229–39.

Anderson GG, Weiss LM. Determining tissue of origin for metastatic cancers: meta-analysis and literature review of immunohistochemistry performance. Appl Immunohistochem Mol Morphol. 2010;18:3–8.

Andreoiu M, Cheng L. Multifocal prostate cancer: biological, prognostic, and therapeutic implications. Hum Pathol. 2010;41:781–93.

Armstrong AB, Wang M, Eble JN, et al. TP53 mutational analysis supports monoclonal origin of biphasic sarcomatoid urothelial carcinoma (carcinosarcoma) of the urinary bladder. Mod Pathol. 2009;22:113–8.

Ashley EA, Butte AJ, Wheeler MT, et al. Clinical assessment incorporating a personal genome. Lancet. 2010;375:1525–35.

Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. Cell. 2011;145:30–8.

Bahrami S, Cheng L, Wang M, et al. Clonal relationships between epidermotropic metastatic melanomas and their primary lesions: a loss of heterozygosity and X-chromosome inactivation-based analysis. Mod Pathol. 2007;20:821–7.

Barrett MT, Sanchez CA, Prevo LJ, et al. Evolution of neoplastic cell lineages in Barrett oesophagus. Nat Genet. 1999;22:106–9.

Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 2011;11:726–34.

Begum S, Gillison ML, Nicol TL, et al. Detection of human papillomavirus-16 in fine-needle aspirates to determine tumor origin in patients with metastatic squamous cell carcinoma of the head and neck. Clin Cancer Res. 2007;13:1186–91.

Belmont JW. Genetic control of X inactivation and processes leading to X-inactivation skewing. Am J Hum Genet. 1996;58:1101–8.

Bhowmick NA, Chytil A, Plieth D, et al. TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. Science. 2004;303:848–51.

Bishop E, Eble JN, Cheng L, et al. Adrenal myelolipomas show nonrandom X-chromosome inactivation in hematopoietic elements and fat: support for a clonal origin of myelolipomas. Am J Surg Pathol. 2006; 30:838–43.

Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. Nat Med. 2011;17:320–9.

Boland CR, Sato J, Appelman HD, et al. Microallelotyping defines the sequence and tempo of allelic losses at tumour suppressor gene loci during colorectal cancer progression. Nat Med. 1995;1:902–9.

Bostwick DG, Shan A, Qian J, et al. Independent origin of multiple foci of prostate intraepithelial neoplasia (PIN): comparison with matched foci of prostate cancer. Cancer. 1998;83:1995–2002.

Brandli DW, Ulbright TM, Foster RS, et al. Stroma adjacent to metastatic mature teratoma after chemotherapy for testicular germ cell tumors is derived from the same progenitor cells as the teratoma. Cancer Res. 2003;63:6063–8.

Brunelli M, Eble JN, Zhang S, et al. Gains of chromosomes 7, 17, 12, 16, and 20 and loss of Y occur early in the evolution of papillary renal cell neoplasia: a fluorescent in situ hybridization study. Mod Pathol. 2003a; 16:1053–9.

Brunelli M, Eble JN, Zhang S, et al. Metanephric adenoma lacks the gains of chromosomes 7 and 17 and loss of Y that are typical of papillary renal cell carcinoma and papillary adenoma. Mod Pathol. 2003b;16:1060–3.

Brunelli M, Eble JN, Zhang S, et al. Eosinophilic and classic chromophobe renal cell carcinomas have similar frequent losses of multiple chromosomes from among chromosomes 1, 2, 6, 10, and 17, and this pattern of genetic abnormality is not present in renal oncocytoma. Mod Pathol. 2005;18:161–9.

Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature. 2005;434:400–4.

Catto JW, Hartmann A, Stoehr R, et al. Multifocal urothelial cancers with the mutator phenotype are of monoclonal origin and require panurothelial treatment for tumor clearance. J Urol. 2006;175:2323–30.

Cheng L. Establishing a germ cell origin for metastatic tumors using OCT4 immunohistochemistry. Cancer. 2005;101:2006–10.

Cheng L, Song SY, Pretlow TG, et al. Evidence of independent origin of multiple tumors from prostate cancer patients. J Natl Cancer Inst. 1997;90:233–7.

Cheng L, Shan A, Cheville JC, et al. Atypical adenomatous hyperplasia of the prostate: a premalignant lesion? Cancer Res. 1998a;58:389–91.

Cheng L, Song SY, Pretlow TG, et al. Evidence of independent origin of multiple tumors from patients with prostate cancer. J Natl Cancer Inst. 1998b;90:233–7.

Cheng L, Leibovich BC, Bergstralh EJ, et al. p53 alteration in regional lymph node metastases from prostate carcinoma: a marker for progression? Cancer. 1999;85:2455–9.

Cheng L, Bostwick DG, Li G, et al. Conserved genetic findings in metastatic bladder cancer: a possible utility of allelic loss of chromosomes 9p21 and 17p13 in diagnosis. Arch Pathol Lab Med. 2001a;125:1197–9.

Cheng L, Gu J, Eble JN, et al. Molecular genetic evidence for different clonal origin of components of human renal angiomyolipomas. Am J Surg Pathol. 2001b;25:1231–6.

Cheng L, Gu J, Ulbright TM, et al. Precise microdissection of human bladder carcinomas reveals divergent tumor subclones in the same tumor. Cancer. 2002;94:104–10.

Cheng L, MacLennan GT, Pan CX, et al. Allelic loss of the active X chromosome during bladder carcinogenesis. Arch Pathol Lab Med. 2004a;128:187–90.

Cheng L, MacLennan GT, Zhang S, et al. Laser capture microdissection analysis reveals frequent allelic losses in papillary urothelial neoplasm of low malignant potential of the urinary bladder. Cancer. 2004b;101:183–8.

Cheng L, Thomas A, Roth LM, et al. OCT4: a novel biomarker for dysgerminoma of the ovary. Am J Surg Pathol. 2004c;28:1341–6.

Cheng L, Jones TD, McCarthy RP, et al. Molecular genetic evidence for a common clonal origin of urinary bladder small cell carcinoma and coexisting urothelial carcinoma. Am J Pathol. 2005;166:1533–9.

Cheng L, Zhang S, MacLennan GT, et al. Interphase fluorescence in situ hybridization analysis of chromosome 12p abnormalities is useful for distinguishing epidermoid cysts of the testis from pure mature teratoma. Clin Cancer Res. 2006;12:5668–72.

Cheng L, Sung MT, Cossu-Rocca P, et al. OCT4: biological functions and clinical applications as a marker of germ cell neoplasia. J Pathol. 2007a;211:1–9.

Cheng L, Zhang S, Wang M, et al. Molecular genetic evidence supporting the neoplastic nature of stromal cells in "fibrosis" after chemotherapy for testicular germ cell tumors. J Pathol. 2007b;213:65–71.

Cheng L, MacLennan GT, Zhang S, et al. Evidence for polyclonal origin of multifocal clear cell renal cell carcinoma. Clin Cancer Res. 2008;14:8087–93.

Cheng L, Zhang S, Davidson DD, et al. Molecular determinants of tumor recurrence in the urinary bladder. Future Oncol. 2009a;5:843–57.

Cheng L, Zhang S, MacLennan GT, et al. Molecular and cytogenetic insights into the pathogenesis, classification, differential diagnosis, and prognosis of renal epithelial neoplasms. Hum Pathol. 2009b;40:10–29.

Cheng L, Zhang S, Alexander R, et al. Sarcomatoid carcinoma of the urinary bladder: the final common pathway of urothelial carcinoma dedifferentiation. Am J Surg Pathol. 2011a;35:e34–46.

Cheng L, Zhang S, Alexander R, et al. The landscape of EGFR pathways and personalized management of nonsmall cell lung cancer. Future Oncol. 2011b;7:519–41.

Cheng L, Alexander RE, MacLennan GT, et al. Molecular pathology of lung cancer: key to personalized medicine. Mod Pathol. 2012a;25:347–69.

Cheng L, Davidson DD, MacLennan GT, et al. The origins of urothelial carcinoma. Expert Rev Anticancer Ther. 2012b;10:865–80.

Cheng L, Williamson SR, Zhang S, et al. Understanding the molecular genetics of renal cell neoplasia: implications for diagnosis, prognosis and therapy. Expert Rev Anticancer Ther. 2012c;10:843–64.

Cheng L, Zhang S, Eble JN, et al. Molecular genetic evidence supporting the neoplastic nature of fibrous stroma in testicular teratoma. Mod Pathol. 2012;25:1432–8.

Cheng L, Zhang S, MacLennan GT, et al. Laser-assisted microdissection in translational research: theory, technical considerations, and future applications. Appl Immunohistochem Mol Morphol. 2012, [Epub ahead of print].

Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. Nat Med. 2011a;17:297–303.

Chin L, Hahn WC, Getz G, Meyerson M. Making sense of cancer genomic data. Genes Dev. 2011b;25:534–55.

Cossu-Rocca P, Eble JN, Delahunt B, et al. Renal mucinous tubular and spindle carcinoma lacks the gains of chromosomes 7 and 17 and losses of chromosome Y that are prevalent in papillary renal cell carcinoma. Mod Pathol. 2006a;19:488–93.

Cossu-Rocca P, Eble JN, Zhang S, et al. Acquired cystic disease-associated renal tumors: an immunohistochemical and fluorescence in situ hybridization study. Mod Pathol. 2006b;19:780–7.

Cossu-Rocca P, Zhang S, Roth LM, et al. Chromosome 12p abnormalities in dysgerminoma of the ovary: a FISH analysis. Mod Pathol. 2006c;19:611–5.

Davidson DD, Cheng L. Field cancerization in the urothelium of the bladder. Anal Quant Cytol Histol. 2006;28:337–8.

Dennis JL, Hvidsten TR, Wit EC, et al. Markers of adenocarcinoma characteristic of the site of origin: development of a diagnostic algorithm. Clin Cancer Res. 2005;11:3766–72.

Dumur CI, Fuller CE, Blevins TL, et al. Clinical verification of the performance of the pathwork tissue of origin test. Utility and limitations. Am J Clin Pathol. 2011;136:924–33.

Emerson RE, Ulbright TM, Zhang S, et al. Nephroblastoma arising in a germ cell tumor of testicular origin. Am J Surg Pathol. 2004;28:687–92.

Emerson RE, Wang M, Liu L, et al. Molecular genetic evidence of an independent origin of serous low malignant potential implants and lymph node inclusions. Int J Gynecol Pathol. 2007a;26:387–94.

Emerson RE, Wang M, Roth LM, et al. Molecular genetic evidence supporting the neoplastic nature of the Leydig cell component of ovarian Sertoli-Leydig cell tumor. Int J Gynecol Pathol. 2007b;26:368–74.

Erlander MG, Ma XJ, Kesty NC, et al. Performance and clinical evaluation of the 92-gene real-time PCR assay for tumor classification. J Mol Diagn. 2011;13:493–503.

Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61:759–67.

Frohling S, Dohner H. Chromosomal abnormalities in cancer. N Engl J Med. 2008;359:722–34.

Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481:306–13.

Green ED, Guyer MS. National human genome research I. Charting a course for genomic medicine from base pairs to bedside. Nature. 2011;470:204–13.

Greco FA, Spigel DR, Yardley DA, et al. Molecular profiling in unknown primary cancer: accuracy of tissue of origin prediction. Oncologist. 2010;15:500–6.

Guo Z, Ponten F, Wilander E, et al. Clonality of precursors of cervical cancer and their genetical links to invasive cancer. Mod Pathol. 2000;13:606–13.

Haber DA, Gray NS, Baselga J. The evolving war on cancer. Cell. 2011;145:19–24.

Hainsworth JD, Spigel DR, Rubin MS, et al. Treatment of carcinoma of unknown primary site (CUP) directed by molecular profiling diagnosis: a prospective, phase II trial. J Clin Oncol. 2010;28(15 suppl):10540.

Halat S, Eble JN, Grignon DG, et al. Multilocular cystic renal cell carcinoma is a subtype of clear cell renal cell carcinoma. Mod Pathol. 2010;23:931–6.

Hamburg MA, Collins FS. The path to personalized medicine. N Engl J Med. 2010;363:301–4.

Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74.

Harris TJ, McCormick F. The molecular pathology of cancer. Nat Rev Clin Oncol. 2010;7:251–65.

Heim S, Mandahl N, Mitelman F. Genetic convergence and divergence in tumor progression. Cancer Res. 1988;48:5911–6.

Henley J, Summerlin DJ, Tomich C, et al. Molecular evidence supporting the neoplastic nature of odontogenic keratocyst: a laser capture microdissection study of 15 cases. Histopathology. 2005;47:582–6.

Hodges KB, Cummings OW, Saxena R, et al. Clonal origin of multifocal hepatocellular carcinoma. Cancer. 2010;16:4078–85.

Holiday R. X chromosome reactivation. Nature. 1987;327:661–2.

Hornberger JC, Amin M, Varadhachary GR, et al. Effect of a gene expression-based tissue of origin test's impact on patient management for difficult-to-diagnose primary cancers. J Clin Oncol. 2011; 29(4 suppl):abstr 459.

Ismael G, de Azambuja E, Awada A. Molecular profiling of a tumor of unknown origin. N Engl J Med. 2006;355:1071–2.

Jiang JK, Chen YJ, Lin CH, Yu IT, Lin JK. Genetic changes and clonality relationship between primary colorectal cancers and their pulmonary metastases – an analysis by comparative genomic hybridization. Genes Chromosomes Cancer. 2005;43:25–36.

Jones TD, Carr MD, Eble JN, et al. Clonal origin of lymph node metastases in bladder carcinoma. Cancer. 2005a;104:1901–10.

Jones TD, Eble JN, Cheng L. Application of molecular diagnostic techniques to renal epithelial neoplasms. Clin Lab Med. 2005b;25:279–303.

Jones TD, Eble JN, Wang M, et al. Clonal divergence and genetic heterogeneity in clear cell renal cell carcinomas with sarcomatoid transformation. Cancer. 2005c; 104:1195–203.

Jones TD, Eble JN, Wang M, et al. Molecular genetic evidence for the independent origin of multifocal papillary tumors in patients with papillary renal cell carcinomas. Clin Cancer Res. 2005d;11:7226–33.

Jones TD, Wang M, Eble JN, et al. Molecular evidence supporting field effect in urothelial carcinogenesis. Clin Cancer Res. 2005e;11:6512–9.

Jones TD, Wang M, Sung MT, et al. Clonal origin of metastatic testicular teratomas. Clin Cancer Res. 2006;12:5377–83.

Jones TD, Zhang S, Lopez-Beltran A, et al. Urothelial carcinoma with an inverted growth pattern can be distinguished from inverted papilloma by fluorescence in-situ hybridization, immunohistochemistry, and morphologic analysis. Am J Surg Pathol. 2007;31:1861–7.

Kaneki E, Oda Y, Ohishi Y, et al. Frequent microsatellite instability in synchronous ovarian and endometrial adenocarcinoma and its usefulness for differential diagnosis. Hum Pathol. 2004;35:1484–93.

Katona TM, Jones TD, Wang M, et al. Molecular evidence for independent origin of multifocal neuroendocrine tumors of the enteropancreatic axis. Cancer Res. 2006;66:4936–42.

Katona TM, Jones TD, Wang M, et al. Genetically heterogeneous and clonally unrelated metastases may arise in patients with cutaneous melanoma. Am J Surg Pathol. 2007a;31:1029–37.

Katona TM, O'Malley DP, Cheng L, et al. Loss of heterozygosity analysis identifies genetic abnormalities in mycosis fungoides and specific loci associated with disease progression. Am J Surg Pathol. 2007b; 31:1552–6.

Kernek KM, Ulbright TM, Zhang S, et al. Identical allelic losses in mature teratoma and other histologic components of malignant mixed germ cell tumors of the testis. Am J Pathol. 2003;163:2477–84.

Kernek KM, Brunelli M, Ulbright TM, et al. Fluorescence in situ hybridization analysis of chromosome 12p in paraffin-embedded tissue is useful for establishing germ cell origin of metastatic tumors. Mod Pathol. 2004;17:1309–13.

Kinzler KW, Vogelstein B. Cancer susceptibility genes gatekeepers and caretakers. Nature. 1997;386:761–3.

Knudson AG. Two genetic hits (more or less) to cancer. Nat Rev Cancer. 2001;1:157–62.

Kum JB, Grignon DJ, Wang M, et al. Mixed epithelial and stromal tumors of the kidney: evidence for a single cell of origin with capacity for epithelial and stromal differentiation. Am J Surg Pathol. 2011;35:1114–22.

Kuperwasser C, Chavarria T, Wu M, et al. Reconstruction of functionally normal and malignant human breast tissues in mice. Proc Natl Acad Sci USA. 2004;101:4966–71.

Lander ES. Initial impact of the sequencing of the human genome. Nature. 2011;470:187–97.

Loeb LA. Human cancers express mutator phenotypes: origin, consequences and targeting. Nat Rev Cancer. 2011;11:450–7.

Lott S, Wang M, Zhang S, et al. Fibroblast growth factor receptor 3 (*FGFR3*) and *TP53* mutation analysis in inverted urothelial papilloma: Incidence and etiological considerations. Mod Pathol. 2009;22:627–32.

Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. Nature. 2005; 435:834–8.

Lyon MF. Gene action in the X chromosome of the mouse. Nature. 1961;190:327–73.

Ma XJ, Patel R, Wang X, et al. Molecular classification of human cancer using a 92-gene real-time quantitative polymerase chain reaction assay. Arch Pathol Lab Med. 2006;130:465–73.

Macconaill LE, Garraway LA. Clinical implications of the cancer genome. J Clin Oncol. 2010;28:5219–28.

Maley CC, Galipeau PC, Finley JC, et al. Genetic clonal diversity predicts progression to esophageal adenocarcinoma. Nat Genet. 2006;38:468–73.

Manolio TA, Green ED. Genomics reaches the clinic: from basic discoveries to clinical impact. Cell. 2011;147:14–6.

McCarthy RP, Wang M, Jones TD, et al. Molecular evidence for the same clonal origin of multifocal papillary thyroid carcinomas. Clin Cancer Res. 2006; 12:2414–8.

McCarthy RP, Zhang S, Bostwick DG, et al. Molecular genetic evidence for different clonal origins of epithelial and stromal components of phyllodes tumor of the prostate. Am J Pathol. 2004;165: 1395–400.

McDermott U, Downing JR, Stratton MR. Genomics and the continuum of cancer care. N Engl J Med. 2011;364:340–50.

Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010;11:685–96.

Monzon FA, Koen TJ. Diagnosis of metastatic neoplasms: molecular approaches for identification of tissue of origin. Arch Pathol Lab Med. 2010;134:216–24.

Monzon FA, Lyons-Weiler M, Buturovic LJ, et al. Multicenter validation of a 1,550-gene expression profile for identification of tumor tissue of origin. J Clin Oncol. 2009;27:2503–8.

National Comprehensive Cancer Network (NCCN). NCCN clinical practice guidelines in oncology: occult primary (Cancer of Unknown Primary [CUP]). Version 2.2011

Novelli M, Cossu A, Oukrif D, et al. X-inactivation patch size in human female tissue confounds the assessment of tumor clonality. Proc Natl Acad Sci. 2003;100:3311–4.

Nowell PC. The clonal evolution of tumor cell populations. Science. 1976;194:23–8.

O'Malley DP, Orazi A, Wang M, et al. Analysis of loss of heterozygosity and X chromosome inactivation in spleens with myeloproliferative disorders and acute myeloid leukemia. Mod Pathol. 2005;18:1562–8.

O'Malley DP, Orazi A, Dunphy CH, et al. Loss of heterozygosity identifies genetic changes in chronic myeloid disorders, including myeloproliferative disorders, myelodysplastic syndromes and chronic myelomonocytic leukemia. Mod Pathol. 2007; 20:1166–71.

O'Shaughnessy JA. Molecular signatures predict outcomes of breast cancer. N Engl J Med. 2006; 355:615–7.

Ohno T, Stribley JA, Wu G, et al. Clonality in nodular lymphocyte-predominant Hodgkin's disease. N Engl J Med. 1997;337:459–65.

Oien KA. Pathologic evaluation of unknown primary cancer. Semin Oncol. 2009;36:8–37.

Paterson RF, Ulbright TM, MacLennan GT, et al. Molecular genetic alterations in the laser-capture-microdissected stroma adjacent to bladder carcinoma. Cancer. 2003;98:1830–6.

Pentheroudakis G, Golfinopoulos V, Pavlidis N. Switching benchmarks in cancer of unknown primary: from autopsy to microarray. Eur J Cancer. 2007; 43:2026–36.

Pillai R, Deeter R, Rigl CT, et al. Validation and reproducibility of a microarray-based gene expression test for tumor identification in formalin-fixed, paraffin-embedded specimens. J Mol Diagn. 2011;13:48–56.

Potti A, Schilsky RL, Nevins JR. Refocusing the war on cancer: the critical role of personalized treatment. Sci Transl Med. 2010;2:28cm13.

Poulos C, Cheng L, Zhang S, et al. Analysis of ovarian teratomas for isochromosome 12p: evidence supporting a dual histogenetic pathway for teratomatous elements. Mod Pathol. 2006;19:766–71.

Quackenbush J. Microarray analysis and tumor classification. N Engl J Med. 2006;354:2463–72.

Ramaswamy S, Ross KN, Lander ES, et al. A molecular signature of metastasis in primary solid tumors. Nat Genet. 2003;33:49–54.

Rini BI, Atkins MB. Resistance to targeted therapy in renal-cell carcinoma. Lancet Oncol. 2009;10: 992–1000.

Rosenfeld N, Aharonov R, Meiri E, et al. MicroRNAs accurately identify cancer tissue origin. Nat Biotechnol. 2008;26:462–9.

Rosenwald S, Gilad S, Benjamin S, et al. Validation of micro-RNA based qRT-PCR test for accurate identification of tumor tissue origin. Mod Pathol. 2010; 23:814–23.

Roychowdhury S, Iyer MK, Robinson DR, et al. Personalized oncology through integrative high-throughput sequencing: a pilot study. Sci Transl Med. 2011;3:111ra21.

Ruiz C, Lenkiewicz E, Evers L, et al. Advancing a clinically relevant perspective of the clonal nature of cancer. Proc Natl Acad Sci USA. 2011;108:12054–9.

Schalkwyk LC, Meaburn EL, Smith R, et al. Allelic skewing of DNA methylation is widespread across the genome. Am J Hum Genet. 2010;85:196–212.

Shackleton M, Quintana E, Fearon ER, et al. Heterogeneity in cancer: cancer stem cells versus clonal evolution. Cell. 2009;138:822–9.

Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer. 1953;6:963–8.

Slaughter DP. Multicentric origin of intraoral carcinoma. Surgery. 1946;20:133–46.

Stancel GA, Coffey D, Alvarez K, et al. Identification of tissue of origin in body fluid specimens with a gene expression microarray assay. Cancer Cytopathol. 2012;120:62–70.

Steeg PS. Tumor metastasis: mechanistic insights and clinical challenges. Nat Med. 2006;12:895–904.

Stratton MR. Exploring the genomes of cancer cells: progress and promise. Science. 2011;331:1553–8.

Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458:719–24.

Stokes A, Drozdov I, Guerra1 E, et al. Copy number and loss of heterozygosity detected by SNP array of formalin-fixed tissues using whole-genome amplification. PLoS One. 2011;6:e24503.

Sung MT, Eble JN, Wang M, et al. Inverted papilloma of the urinary bladder: a molecular genetic appraisal. Mod Pathol. 2006;19:1289–94.

Sung MT, Wang M, Maclennan G, et al. Histogenesis of sarcomatoid urothelial carcinoma of the urinary bladder: evidence for a common clonal origin with divergent differentiation. J Pathol. 2007;211: 420–30.

Sung MT, Zhang S, MacLennan GT, et al. Histogenesis of clear cell adenocarcinoma in the urinary tract: evidence of urothelial origin. Clin Cancer Res. 2008;14:1947–55.

Talmadge JE, Wolman SR, Fidler IJ. Evidence for the clonal origin of spontaneous metastases. Science. 1982;217:361–3.

Turhan AG, Humphries RK, Phillips GL, et al. Clonal hematopoiesis demonstrated by X-linked DNA polymorphisms after allogeneic bone marrow transplantation. N Engl J Med. 1989;320:1655–61.

Varadhachary GR, Raber MN, Matamoros A, et al. Carcinoma of unknown primary with a colon-cancer profile-changing paradigm and emerging definitions. Lancet Oncol. 2008;9:596–9.

Varadhachary GR, Spector Y, Abbruzzese JL, et al. Prospective gene signature study using microRNA to identify the tissue of origin in patients with carcinoma of unknown primary. Clin Cancer Res. 2011;17:4063–70.

Velculescu VE, Diaz LA, Jr. Understanding the enemy. Sci Transl Med. 2011;3:98ps37.

Visvader JE. Cells of origin in cancer. Nature. 2011;469:314–22.

Wang X, Jones TD, Zhang S, et al. Amplifications of EGFR gene and protein expression of EGFR, Her-2/neu, c-kit, and androgen receptor in phyllodes tumor of the prostate. Mod Pathol. 2007a;20:175–82.

Wang X, Zhang S, MacLennan GT, et al. Epidermal growth factor receptor protein expression and gene amplification in small cell carcinoma of the urinary bladder. Clin Cancer Res. 2007b;13:953–7.

Wang X, Wang M, MacLennan GT, et al. Evidence for common clonal origin of multifocal lung cancers. J Natl Cancer Inst. 2009;101:560–70.

Wareham K, Lyon M, Glenister P, et al. Age related reactivation of an X-linked gene. Nature. 1987;327:725–7.

Weber K, Thomaschewski M, Warlich MR, et al. GB marking facilitates multicolor clonal cell tracking. Nat Med. 2011;17:504–9.

Weinstein IB, Joe A. Oncogene addiction. Cancer Res. 2008;68:3077–80; discussion 80.

Widschwendter M, Fiegl H, Egle D, et al. Epigenetic stem cell signature in cancer. Nat Genet. 2007;39:157–8.

Williamson SR, Zhang S, Yao JL, et al. *ERG-TMPRSS2* rearrangement is shared by concurrent prostatic adenocarcinoma and prostatic small cell carcinoma and absent in small cell carcinoma of the urinary bladder: evidence supporting monoclonal origin. Mod Pathol. 2011;24:1120–7.

Wong CCY, Caspi A, Williams B, et al. A longitudinal twin study of skewed X chromosome-inactivation. PLoS One. 2011;6:e17873.

Modern Immunohistochemistry in Targeted Therapy

6

Allen M. Gown

Contents

6.1 Basic Principles of Immunohistochemistry 181
6.1.1 Antibodies 181
6.1.2 Detection Systems 182
6.1.3 Epitope (Antigen) Retrieval 182
6.1.4 Scoring of Results and Cutoffs for Positivity 183

6.2 Antibody Validation and Quality Control in Targeted Therapy Immunohistochemistry 183
6.2.1 Antibody and Method Validation 183
6.2.2 Quality Control in Immunohistochemistry 184

6.3 Immunohistochemistry of Specific Targets of Targeted Therapy 184
6.3.1 Breast Cancer: Estrogen and Progesterone Receptors 184
6.3.2 Breast Cancer: HER2 187
6.3.3 Gastrointestinal Stromal Tumors and c-kit 191
6.3.4 Rituximab (Anti-CD20) and B Cell Lymphoma 192

Further Reading 193

A.M. Gown, MD
PhenoPath Laboratories, Seattle, WA, USA

Department of Pathology, University of British Columbia, Vancouver, BC, Canada

6.1 Basic Principles of Immunohistochemistry

- Immunohistochemistry is the application of antibodies with predefined specificities to tissue coupled with the use of detection systems permitting visualization of the antibody target

6.1.1 Antibodies

- Successful use of immunohistochemistry (IHC) requires use of antibodies with maximum specificity for target molecule
- Antibody specificity is the property of antibodies that permit them to react with epitopes of antigens (which can be proteins, sugar moieties, etc.) and not with others. Antibodies can distinguish differences in primary structure of an antigen, isomeric forms of an antigen, or secondary or tertiary structure of an antigen
- Cross reactivity of antibodies may occur when an epitope is shared by the target molecule as well as a nontarget molecule which may or may not be related to the target molecule
- Antibodies may typically be rabbit (polyclonal), mouse (monoclonal), or rabbit (monoclonal)
- Advantages of monoclonal antibodies include absence of contamination by other antibodies of differing affinity and specificity, limitless supply of antibodies, and ability to obtain antibodies "customized" for use, e.g., in IHC

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_7,

- Disadvantages of polyclonal antibodies include differences in antisera raised in different animals, limited volumes of antibody that can be prepared, and presence of minor contaminating antibodies that can give unexpected cross reactivities
- Advantages of polyclonal antibodies are relative inexpense in production and generally robust nature of resulting antibodies
- Disadvantages of monoclonal antibodies include potential loss of single epitope to which they are directed by fixation, decalcification, etc.
- Existence of multiple vendors for antibodies requires selectivity and precision in determining the most appropriate reagent for antibodies
- Best to employ antibodies that have been used in published studies
- For some companion diagnostics for targeted therapy, there exist Food and Drug Administration (FDA)-cleared and non-FDA-cleared reagents. While some recommend the use of FDA-cleared reagents, they are not necessarily more accurate and precise than non-FDA-cleared reagents, and all antibodies or kits need to be validated in the laboratory prior to use
- Some FDA-cleared reagents (e.g., HercepTest™ for HER2 detection) do not employ the most robust antibodies and/or detection systems

6.1.2 Detection Systems

- Best to employ the most sensitive but reliable detection systems for IHC (Fig. 6.1)
- Major immunoperoxidase-based detection systems include
  - Avidin biotin-based detection systems
  - Polymer-based detection systems
  - Tyramide-based detection systems
- Advantages of avidin biotin-based detection system are low cost, long track record, and moderate to high sensitivity
- Disadvantages of avidin biotin-based detection systems are problems with endogenous biotin leading to potential false-positive signal, as well as relatively lower sensitivity compared with polymer-based detection systems
- Advantages of polymer-based detection systems include absence of problems with endogenous biotin, high to very high sensitivity
- Disadvantages of polymer-based detection systems include higher cost and shorter track record compared with avidin biotin-based systems
- Advantages of tyramide-based detection systems include very high sensitivity
- Disadvantages of tyramide-based detection systems include relative unreliability and instability compared with other detection systems
- Polymer-based detection systems best choice overall
- 3,3′-diaminobenzidine recommended for chromogen

6.1.3 Epitope (Antigen) Retrieval

- Most (but not all) antibodies require the prior use of "epitope retrieval" or "antigen retrieval" for maximum sensitivity on deparaffinized, formalin-fixed tissue sections
- Optimal use of epitope retrieval requires determination of the most appropriate buffer (salt composition, pH) and duration of retrieval process
- Most epitope retrieval can be performed with heat applied in microwave oven with slides immersed in buffer
- No single epitope retrieval solution is optimal for all antigens
- Optimal epitope retrieval method must be determined for each antibody, using the "test battery" approach with testing of several buffers across pH range
- Examples of buffers in "test battery" might include 10-mM citrate pH 6, 10-mM EDTA pH 8, 50-mM Tris/EDTA pH 9, 500-mM Tris pH 10
- Potential problems with epitope retrieval, both of which can result in false-negative studies, include
  - Failure to identify optimal buffer conditions

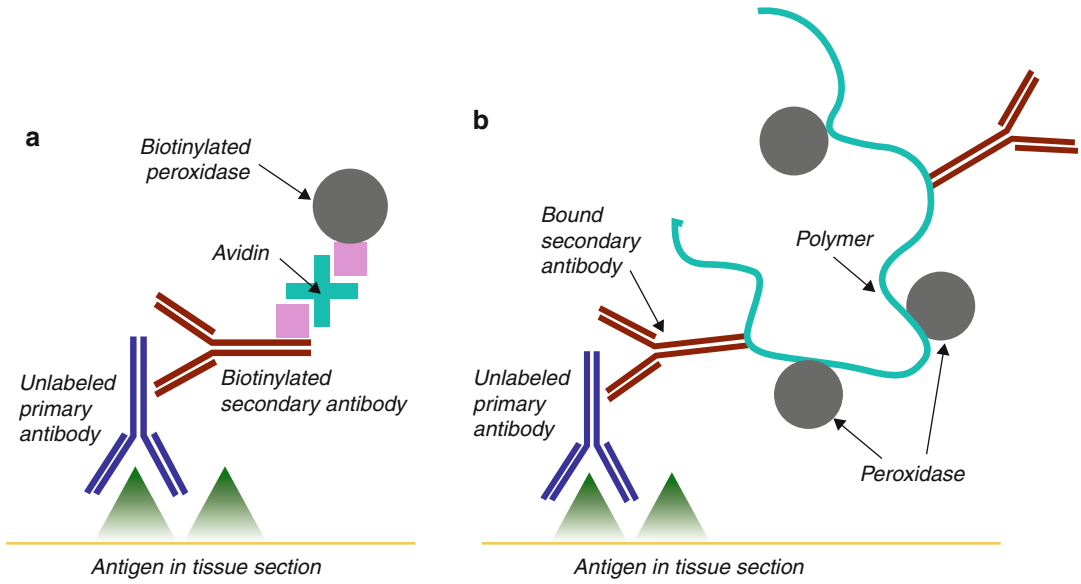


Fig. 6.1 Schematic diagrams of avidin biotin (**a**) and polymer-based (**b**) immunohistochemical detection systems

– "Overretrieval," which occurs when epitope retrieval duration is too long, which is most likely to occur in relatively underfixed tissues

6.1.4 Scoring of Results and Cutoffs for Positivity

- Scoring may be dichotomous (positive, negative), semiquantitative (e.g., low vs high positive, 1+ vs 2+ vs 3+), or quantitative (e.g., by direct counting of positive cells manually or by image analysis or by quantifying degree of positivity by image analysis)
- Scoring may be based on fraction of cells positive (e.g., 1%, 10%, 50%), with optimal cutoffs determined by prediction of clinical outcome
- Scoring may be based on intensity of immunostaining of the target cell (e.g., HER2, see below)
- Scoring may be based on a combination of fraction of cells positive plus intensity (e.g., H-score, Allred score for estrogen receptor [ER])

6.2 Antibody Validation and Quality Control in Targeted Therapy Immunohistochemistry

6.2.1 Antibody and Method Validation

- Before an antibody, and the immunohistochemical methodology employed with it, is introduced into the laboratory, it is necessary to perform validation studies
- Definition of validation: procedures designed to confirm the ability of the antibody to recognize its target antigen in normal and diseased tissues where it is expected to localize. Validation also includes procedures to calculate the sensitivity and specificity of the antibody in the tissues employed in the laboratory
- Antibody validation is required by the Clinical Laboratory Improvement Amendments (CLIA) (42CFR493.1253) as well as College of American Pathologists (CAP), as part of general lab checklist (GEN.41850, GEN.42020)

- There must be a written validation procedure for every antibody brought online in the laboratory
- There are two types of validation: method validation and clinical validation
  - Method validation
    - General principles of antibody validation, with respect to ensuring that antibody identifies target antigen in tissue section
      - Highly selective immunostaining pattern in tissue must be observed, in compartment of cell appropriate to location of antigen
      - When employed in Western blots, staining of protein of expected molecular weight
      - Disappearance of staining after preabsorption of antibody prior to immunostaining with purified epitope
      - Identical staining patterns of antibodies raised to different epitopes of same antigen
      - Correspondence between IHC staining patterns and results of other technique such as in situ hybridization in consecutive sections
      - Absence of staining of (mouse) tissues in animals genetically deficient for protein
    - There must be revalidation if a new antibody lot is introduced, if there is change to epitope retrieval methodology, if there is a change in the detection system employed, if a new application for an existing antibody is introduced, or if an existing antibody is displaying poor performance
    - Low-level as well as high-level positive cases must be included in the validation set. Special validation requirements for ER and progesterone receptor (PR) discussed below
    - Validation cases should be tested on multiple days to ensure between run and within run precision
    - No rules exist for number of cases required for validation, but there have been American Society of Clinical Oncology (ASCO)–CAP requirements (see below) for ER, PR, and HER2 validation
  - Clinical validation
    - Definition of clinical validation: studies to confirm that results of test predict clinical outcome or response to therapy
    - Requires access to large patient cohorts with long-term clinical outcome
    - Given impracticality for most laboratories of performing clinical validation, it is recommended that a laboratory cross-validate, i.e., show concordance between their antibody procedure against the procedure of a laboratory that has a clinically validated assay (see ASCO–CAP requirements for ER, PR and HER2, below)

6.2.2 Quality Control in Immunohistochemistry

- Revalidation of antibodies with changes in protocols
- Continuing education for technologists and pathologists
- Participation in proficiency testing through CAP or alternative testing modalities
  - CAP requires twice yearly proficiency testing with 90% success for ER and HER2 for lab accreditation (see below)

6.3 Immunohistochemistry of Specific Targets of Targeted Therapy

6.3.1 Breast Cancer: Estrogen and Progesterone Receptors

- Estrogen receptor is transcription factor expressed in nucleus of many normal breast epithelial cells and 70–80% of primary breast carcinomas (Fig. 6.2)
- Estrogen receptor is a weak prognostic marker of breast cancer (when controlled for other parameters such as tumor grade, stage, age, etc.)

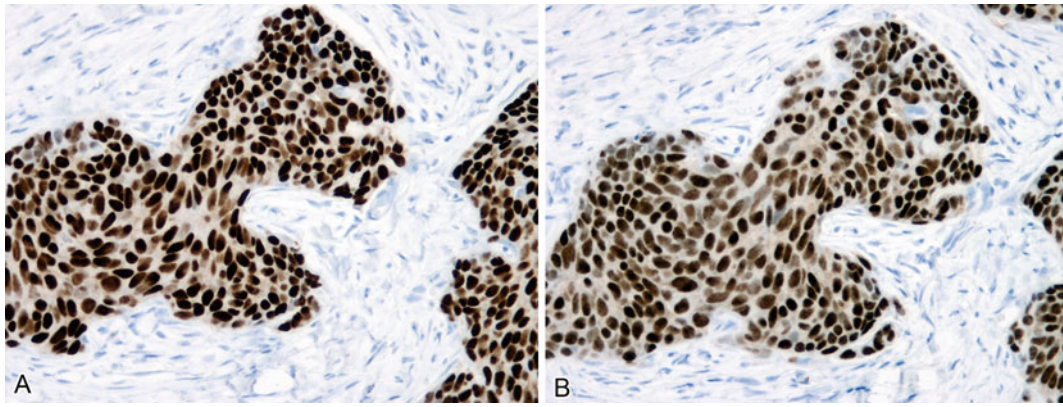


Fig. 6.2 Breast cancer showing uniform, high levels of expression of estrogen (**a**) and progesterone (**b**) receptor proteins

- Progesterone receptor status provides independent additional predictive/prognostic value
- ER and PR expression are strong predictive markers for response to hormonal-based therapy (e.g., tamoxifen, aromatase inhibitors)
- Immunohistochemistry for detection of ER has largely supplanted degradative techniques, such as enzyme immunoassay, dextran-coated charcoal method, etc., owing to stronger predictive power (largely owing to false-positives as a consequence of contamination by nonneoplastic breast tissue)
- Few published studies exist documenting accuracy and reproducibility of ER IHC, but those that do show poor sensitivity, particularly in cases showing low levels of ER expression
- Three potential sources of errors with ER and PR testing by IHC: preanalytical, analytical, and postanalytical
- Preanalytic sources include specimen type (core needle biopsy vs lumpectomy or mastectomy specimen), cold ischemia time (time between specimen removal and placement in formalin), fixative content (formalin vs other), duration of fixation, and additional treatment such as decalcification in the case of bone metastases
- Specimen type: More than 15 published studies show fairly high concordance between ER IHC results in parallel analysis of core needle biopsy and resection specimen of same tumor (discrepancy range from 0% to 38%), with 1% discrepancy rate cited by most recent large study
- As most discrepancies are false-negatives on resection specimen, core needle biopsy recommended
- Cold ischemia time: Defined as time interval between removal of specimen from patient until immersion into fixative
- Recommended by ASCO–CAP (see Table 6.1) to be less than 1 h
- Some published data documents adverse effects of prolonged cold ischemia time on breast cancer markers but others suggest no effect
- Fixative content: Only 10% neutral buffered formalin should be employed, and use of all nonformalin-based fixatives is not recommended. However, if a laboratory wishes to employ the latter, they need to validate tissues fixed in the alternative fixative using specimens fixed in both fixatives. ASCO–CAP strongly recommends use of formalin-fixed tissue (see Table 6.1)
- Fixative duration: ASCO–CAP requires fixation in 10% neutral buffered formalin for no less than 6 h and no more than 72 h. This is based in part on extrapolation from published work of Goldstein et al. (2003) demonstrating potential false-negative studies in tissues fixed less than 6–8 h in formalin

Table 6.1 Summary of ASCO–CAP requirements for ER and PR testing

Preanalytical factors
Cold ischemia time less than or equal to 1 h
Only 10% neutral buffered formalin
Fixation for no less than 6 h and no more than 72 h
Analytical factors
Acceptable antibody clones for ER include 1D5, 6F11, SP1, 1D5 + ER.2.123
Acceptable antibody clones for PR include 1294 and 312
No specific recommendation for epitope retrieval or detection system
Postanalytical factors
Results should be reported as either "positive" or "negative"
Positive interpretation requires at least 1% positive nuclear signal in tumor cells
Average intensity (strong, moderate or weak) should also be recorded but does not impact positive versus negative
ER and PR results should fit clinical profile, i.e., some cancer types (lobular, mucinous, tubular) should rarely be ER-negative
Additional
Validation of ER and PR IHC assays must be performed prior to reporting patient results, attaining at least 90% concordance with positive results and at least 95% concordance with negative results on at least 40 positive and 40 negative cases (with at least 10 positives weakly positive) cross tested against a clinically validated assay
Participation in proficiency testing required, with 90% score attained on graded challenges
Ongoing competency assessment required for all pathologists signing out ER and PR results
Monitoring of ER and PR positivity rate should be performed at least annually

ASCO–CAP American Society of Clinical Oncology-College of American Pathologists, *ER* estrogen receptors, *IHC* immunohistochemical, *PR* progesterone receptors

- However, more recent studies using polymer-based detection system fail to demonstrate any such effect of shorter or longer fixation times
- Analytical sources of error include choice of ER antibody, epitope retrieval method, and detection method
- ER antibody: A number of antibodies available, all of which have been clinically validated (6 F11, SP1, 1D5, 1D5 + ER.2.123; see Table 6.2)
  - SP1 rabbit monoclonal antibody demonstrated to be equivalent or superior to 1D5 in published studies
  - SP1 more robust reagent, less affected by preanalytical factors
- PR antibody: A number of antibodies have been clinically validated or technically validated with clinically validated assay (1294, 312, PgR 1A6, PgR 636, 1E2; see Table 6.3)
- To ensure optimal epitope retrieval and detection system performance, use lower level ER and PR-positive tissues (e.g., myometrium) as external positive controls

Table 6.2 Anti-ER antibodies in IHC assays for ER expression

Antibody clone	Type of antibody
6F11	Mouse monoclonal
1D5	Mouse monoclonal
SP1	Rabbit monoclonal
ER-2-23/1D5	Mouse monoclonals (cocktail)

ER estrogen receptors, *IHC* immunohistochemical

Table 6.3 Anti-PR antibodies in IHC assays for PR expression

Antibody clone	Type of antibody
1294	Mouse monoclonal
312	Mouse monoclonal
PgR636	Mouse monoclonal
PgR1A6	Mouse monoclonal
1E2	Rabbit monoclonal

IHC immunohistochemical, *PR* progesterone receptors

- Postanalytical sources of error: scoring system, cutoff for positivity
- Scoring of ER-positive or PR-positive cells: ER or PR results can be assessed by estimating or

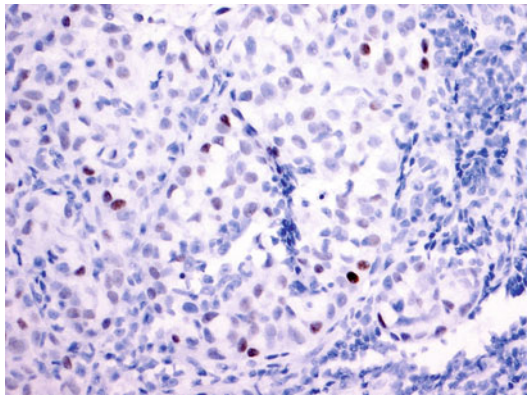


Fig. 6.3 Breast cancer showing low levels estrogen receptor expression; weak signal in more than 1% of tumor cells is still a positive result

quantifying the fraction of tumor cells positive, with or without the use of image analysis
  - Allred scoring system assigns two scores, one for percent-positive cells (0–5) and one for intensity (0–3), which are summed to yield "Allred score" of 0–8. Score of 3 or higher considered positive
  - 1% positivity score on tumor mandated by ASCO–CAP to be threshold for positivity; based on study of Harvey et al. (1999) demonstrating that clinical significance to use of Allred score 3 or higher for cutoff for positivity (Fig. 6.3)
  - No equivocal range for ER or PR positivity analogous to that of HER2 IHC (see below)
  - ASCO–CAP recommends recording of intensity in addition to percentage of positive cells
  - In the context of cytology specimens, a minimum of 100 cells should be counted according to ASCO–CAP
- Validation of ER and PR assays
  - Technical validation: confirms the ability of the antibody to identify its target antigen (Fig. 6.4)
  - Clinical validation: confirms that the IHC assay predicts clinical outcome
  - Recognizing inability of most laboratories to perform clinical validation, all ER and PR IHC assays need to be technically validated against an assay that has been clinically validated in another laboratory
    - Must show at least 90% concordance in ER-positive or PR-positive cases and at least 95% concordance in ER-negative or PR-negative cases, according to ASCO–CAP guidelines
- Limitations of ER assays by IHC
  - Do not speak to functionality of ER
  - Do not address activation of downstream pathways of ER activation
  - Do not identify the subset of ER-positive patients who will fail hormonal-based therapy

6.3.2 Breast Cancer: HER2

- HER2 is a member of the human epidermal growth factor receptor family of proteins [HER1 (EGFR), HER2, HER3, HER4] and the only member of the EGFR family without a known ligand. Normal cells have 20,000–50,000 receptors per cell, and HER2-overexpressing tumors (approximately 15–20% of breast cancers) can express 500,000–2,000,000 receptors per cell
- Activation of HER2 leads to signaling cascades, including the PI3K and MAPK pathways, leading to changes resulting in increased cell survival, increased cell proliferation increased cell mobility (e.g., invasion and metastasis), and angiogenesis
- HER2 status of breast cancer is a strong prognostic factor (predictor of adverse outcome) but also a predictor of negative response to traditional chemotherapy (e.g., cytoxan, methotrexate, fluorouracil), a predictor of positive response to anthracycline-based chemotherapy, a predictor of resistance to tamoxifen in ER-positive tumors, and a predictor of response to anti-HER2-targeted therapies such as trastuzumab and lapatinib
- Across several international clinical trials, ~50% improvement in disease-free survival mediated by adjuvant trastuzumab (NSABP B-31, HERA, and BCIRG)

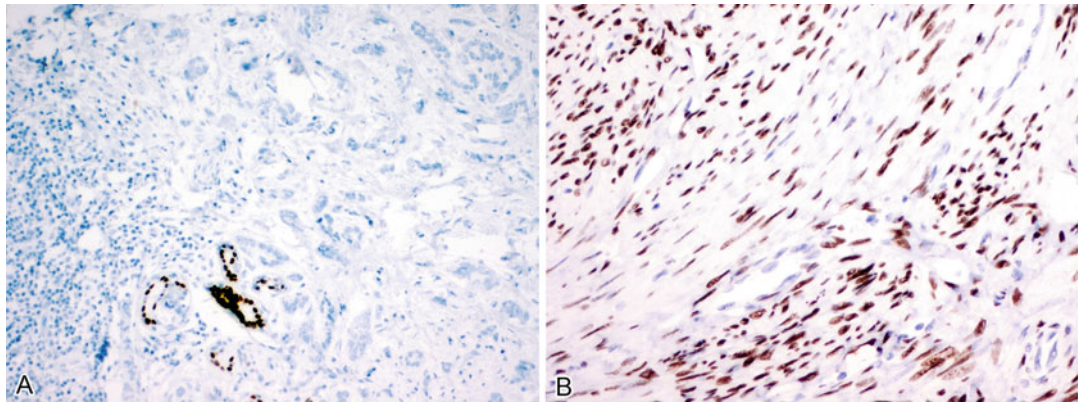


Fig. 6.4 (**a**) Breast cancer negative for estrogen receptor expression; note strong signal on nonneoplastic breast epithelium, serving as positive internal control. (**b**) Myometrium serves as an excellent external ER control

- HER2 testing should be routinely performed in patients with a new diagnosis of invasive breast cancer
- Overexpression of HER2 protein in >95% of cases associated with amplification of HER2 gene and concomitant overexpression of HER2 mRNA
- Many published studies document poor accuracy and reproducibility of HER2 IHC when comparing local laboratories with central laboratories, and comparing IHC versus fluorescence in situ hybridization (FISH) results on same specimens
- Three potential sources of errors with HER2 testing by IHC: preanalytical, analytical, and postanalytical
- Preanalytic sources include specimen type (core needle biopsy vs lumpectomy or mastectomy specimen), cold ischemia time (time between specimen removal and placement in formalin), fixative content (formalin vs other), duration of fixation, and additional treatment such as decalcification in the case of bone metastases
  - Specimen type
    - Very high concordance (>98%) has been reported between HER2 results in core needle biopsies and resection specimens of same breast cancers
    - ASCO–CAP recommends use of resection specimens owing to likelihood of artifacts in core needle biopsies which might confound interpretation (see Table 6.4); ASCO–CAP recommendations categorize core needle biopsies with "edge or retraction artifact involving entire core" or "crust artifact" as exclusion criteria to perform or interpret HER2 IHC assay
  - Cold ischemia time
    - While no recommendations in original ASCO–CAP document, ASCO–CAP document in 2011 recommends maximum of 1 h cold ischemia time for HER2 (as well as ER) testing
    - Published data show possible effects on FISH signals but no effect documented on HER2 IHC
- Fixation
  - ASCO–CAP guidelines require the use of 10% neutral buffered formalin, with fixation duration of no less than 6 h and no more than 48 h
  - Fixation requirements apply to resection specimens as well as core needle biopsies
  - Alternative fixatives may be employed but only if there has been validation of the performance characteristics of this alternative fixative to show that they are concordant with results using 10% NBF on the same specimens
- Analytic sources of error: choice of anti-HER2 antibody, epitope retrieval method, detection method

Table 6.4 Summary of ASCO–CAP requirements for HER2 testing

Preanalytical factors
Cold ischemia time less than or equal to 1 h
Only 10% neutral buffered formalin
Fixation for no less than 6 h and no more than 48 h
Analytical factors
No specific recommendation for anti-HER antibody clones
No specific recommendation for epitope retrieval or detection system
Scoring of slides as 0, 1+, 2+ or 3+
Postanalytical factors
Results should be reported as either "positive" or "negative" or "equivocal" for overexpression
HER2 results should fit clinical profile, i.e., some cancer types (lobular, mucinous, tubular) should rarely be HER2-positive
Additional
Validation of HER2 IHC assay on 25–100 cases must be performed prior to reporting patient results, attaining 95% positive and negative concordance with alternative method (e.g., fluorescence in situ hybridization) or validated IHC performed in another laboratory
Participation in proficiency testing required, with 90% score attained on graded challenges
Ongoing competency assessment required for all pathologists signing out HER2 results
Monitoring of HER2 positivity rate should be performed annually

ASCO–CAP American Society of Clinical Oncology–College of American Pathologists, *IHC* immunohistochemical

- HER2 antibody: a number of antibodies available, all of which have been technically validated against alternative techniques such as FISH (A0485, CB11, SP3, 4B5; see Table 6.5)
  - ASCO–CAP recommendations do not specify antibodies
  - Rabbit monoclonals SP3 and 4B5 most robust reagents for routine use, less affected by preanalytical factors
- To ensure optimal epitope retrieval and detection system performance, use both high-level (3+) positive controls as well as low-level (1+) controls
- Postanalytical sources of error: scoring system, cutoff for positivity
  - Scoring of HER2
    - 0: No signal
    - 1+: Weak incomplete membranous signal in any proportion of the tumor cells
    - 2+: Complete membranous signal that is either nonuniform or weak in intensity but with obvious circumferential signal in at least 10% of tumor cells
    - 3+: Intense, uniform membranous signal in greater than 30% of invasive tumor cells

Table 6.5 Anti-HER2 antibodies in IHC assays for HER2 overexpression

Antibody clone	Type of antibody
Dako A0485[a]	Rabbit polyclonal
SP3	Rabbit monoclonal
4B5	Rabbit monoclonal
CB11	Mouse monoclonal

[a]Same antibody that is part of HercepTest™
IHC immunohistochemical

  - Interpretation of results of HER2 (Fig. 6.5)
    - 0 or 1+: Considered negative for HER2 overexpression
    - 2+: Considered "equivocal" for HER2 overexpression
    - 3+: Considered "positive" for HER2 overexpression
- Tests resulting in equivocal (2+) results should retested by another method (e.g., fluorescence in situ hybridization, FISH) which may in turn yield positive, negative, or equivocal results
- Examples of exclusion criteria for HER2 IHC, according to ASCO–CAP
  - Fixation not corresponding to guidelines (although alternative fixatives can be validated)

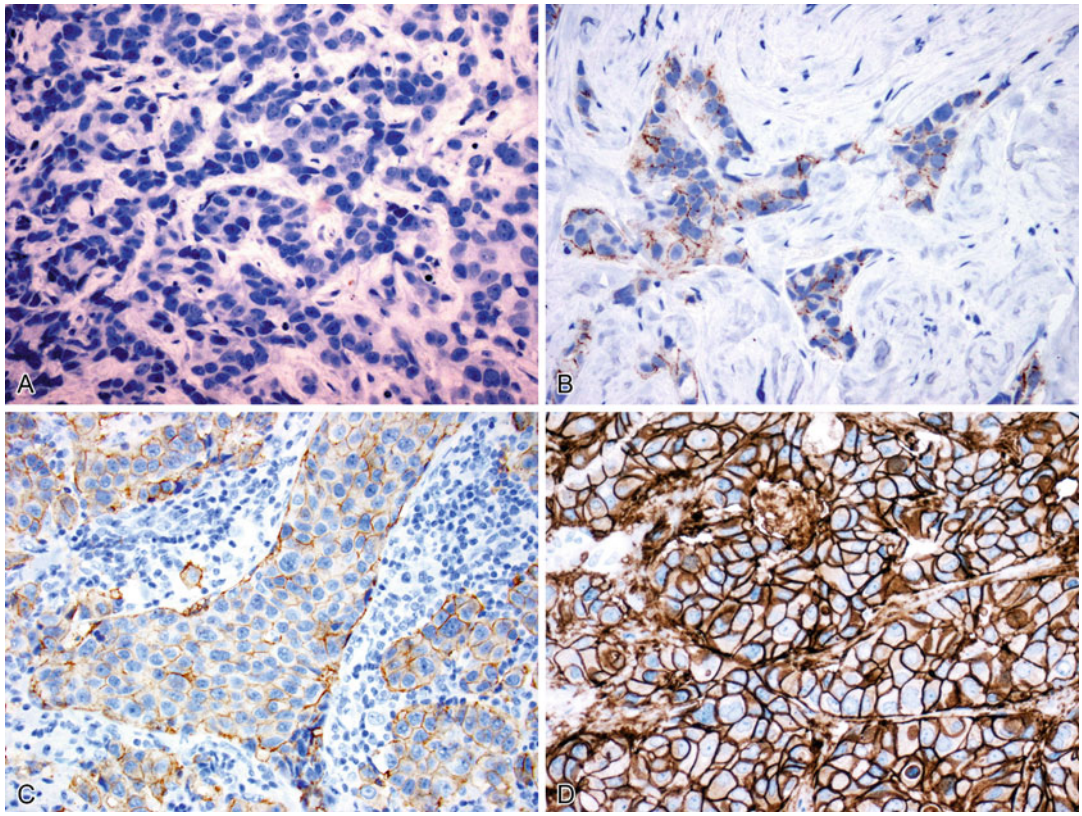


Fig. 6.5 Different breast cancers showing the spectrum of HER2 expression, from 0, 1+, 2+, and 3+. Only the 3+ case is considered positive for overexpression

  - Tissues with strong membranous signal in internal normal ducts
  - High background signal in cytoplasm (>10% of signal)
- Validation of HER2 immunohistochemical assay (ASCO–CAP 2007)
  - Validation of HER2 IHC should be performed prior to offering test
  - Initial test validation requires cross-testing of 25–100 samples initially tested by IHC by another method (e.g., FISH) in the same laboratory or by IHC or another method (e.g., FISH) in a different laboratory
    - 95% concordance required within categories of HER2-positive (e.g., 3+ IHC and FISH amplified) and HER2-negative (e.g., 0 or 1+ IHC and FISH nonamplified)
    - If multiple pathologists involved in HER2 IHC interpretation, each pathologist must demonstrate 95% concordance on positives and negatives
  - Revalidation required if IHC procedure (e.g., detection system) changed
- Optimization of HER2 IHC
  - Integrate morphology: HER2 overexpression very uncommon in low-grade breast carcinomas and especially rare in lobular, mucinous, and tubular variants
  - Consider using normalization: published studies demonstrate significant reduction in false-positive rate by "subtracting" IHC score of nonneoplastic ducts from that of tumor, yielding normalized HER2 score
  - Consider use of image analysis; some studies demonstrate significant improvement of interobserver and intraobserver agreement with computer-assisted reading of immunostained slides

- Ongoing competency assessment
  - Competency of laboratory personnel and pathologists must be continuously addressed
  - Proficiency testing at least twice per year, with satisfactory performance (90% or greater) on graded challenges

6.3.3 Gastrointestinal Stromal Tumors and c-kit

- KIT proto-oncogene (Hardy–Zuckerman 4 feline sarcoma viral oncogene homolog)
- Encodes c-kit, a growth factor and member of the type III receptor tyrosine kinase family that also includes platelet-derived growth factor (PDGFRA and PDGFRB)
- c-kit receptor synonyms: CD117, stem cell receptor
- c-kit molecule: 135 kd, contains extracellular domain, hydrophobic transmembrane domain, and intracellular tyrosine kinase domain
- Ligand for c-kit is known as kit ligand, steel factor, stem cell factor (SCF), mast cell growth factor
- Upon ligand binding c-kit undergoes dimerization and autophosphorylation, resulting in downstream signaling, including MAPK pathway and the PI3K–AKT pathology. Activation of these pathways results in increased cell survival and cell proliferation and/or differentiation
- c-kit found in many normal tissues, including tissue mast cells, melanocytes, germ cells, glandular epithelial cells of breast and salivary gland, and interstitial cells of Cajal (ICC) of intestine
- ICCs are pacemaker cells of the gastrointestinal tract first described by neuroanatomist Santiago Ramon y Cajal, thought to play important role in normal control of intestinal motility
- Gastrointestinal stromal tumors (GISTs) are rare subset of GI sarcoma (3,000–5,000 new cases/year in United States) but the most common mesenchymal tumor of the gastrointestinal tract

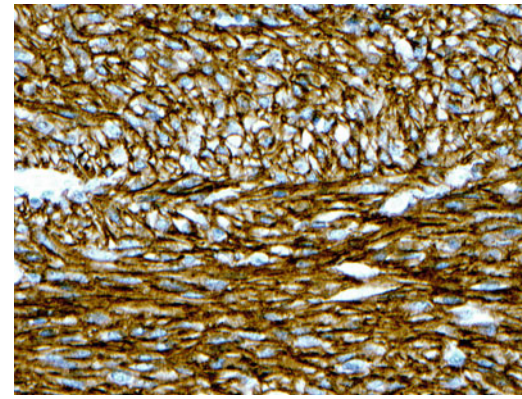


Fig. 6.6 Gastrointestinal stromal tumor showing uniformly positive expression of c-kit; note membranous and cytoplasmic distribution of antigen

- More than 50% of GISTs occur in stomach but can occur anywhere in GI tract from small intestine to rectum and may also rarely occur outside GI tract
- 70–80% of GISTs have mutated c-kit gene (two-thirds on exon 11)
- Mutation of c-kit leads to constitutive activation of tyrosine kinase
- Of GISTs lacking c-kit mutations, minority have mutation an activation of PDGFRA; c-kit and PDGFRA mutations are mutually exclusive in GISTs
- PDGFRA-mutant GISTs show clinical and pathologic features distinct from c-kit-mutant GISTs
  - Predilection for stomach
  - Variable or negative c-kit expression
  - Lower malignant potential
- 10–15% of GISTs have mutations neither in c-kit or PDGFRA
- Immunohistochemistry of c-kit in GISTs
  - c-kit positivity is the hallmark of GISTs (~95%) (Fig. 6.6)
  - GISTs also express CD34 (60–70%), smooth muscle actins (50–60%), but rarely S100 or desmin, helping to distinguish GISTs from nerve sheath tumors and leiomyoma/leiomyosarcoma
  - DOG1 an even more sensitive and specific marker of GISTs, including those negative for c-kit

  - c-kit expression is not specific for GISTs and is also expressed in chronic myelogenous leukemia and acute myelogenous leukemia; small cell carcinoma of the lung; melanoma; neuroblastoma; testicular cancer and seminoma; subset of breast cancer
  - c-kit detectable on cell surface of GIST tumor cells, though strong signal generally found in cytoplasm as well (occasionally with a "dot-like" pattern)
  - Recommended antibody for IHC is rabbit polyclonal (available from Dako)
- Immunohistochemistry and prediction of response to TKI therapy (e.g., imatinib)
  - Imatinib is tyrosine kinase inhibitor (TKI) developed for treatment of CML owing to ability to inhibit fusion oncoprotein BCR–ABL
  - Imatinib binds to ATP-binding site of c-kit and inhibits ATP binding and stabilizes the kinase in the inactive conformation
  - c-kit IHC serves as a surrogate cell type-specific marker which helps classify the tumor as GIST
  - c-kit IHC is not of and by itself predictive of response to imatinib
  - While c-kit is target of IHC, identifying the presence of immunoreactive protein does not imply the presence of mutated c-kit that is required for sensitivity to imatinib therapy
  - Molecular studies documenting presence of activating c-kit mutation required for prediction of response to imatinib
  - Primary and secondary resistance to imatinib can develop; vast majority of patients eventually develop disease progression following acquired mutations in c-kit or PDGFRA

6.3.4 Rituximab (Anti-CD20) and B Cell Lymphoma

- CD20 is 33-kd protein expressed on normal mature B lymphocytes and most malignant B cells. The vast majority of non-Hodgkin lymphoma arises from B cells (~30% diffuse large B cell lymphoma and ~20% follicular lymphoma) (Fig. 6.7)

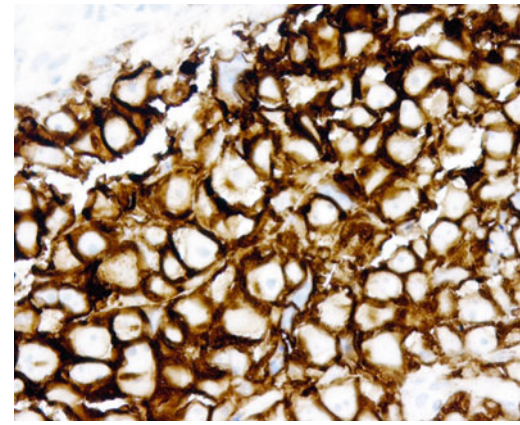


Fig. 6.7 Diffuse large B cell lymphoma positive for expression of CD20

- CD20 is transmembrane protein that plays a role in B cell proliferation, activation, and differentiation
- Over 95% of B cell non-Hodgkin lymphomas express CD20
- Chronic lymphocytic lymphoma (CLL) is also a B cell-derived malignancy expressing CD20
- CD20 is not normally expressed on pre-B cells or differentiated plasma cells
- Rituximab is genetically engineered chimeric monoclonal antibody to CD20
- Administration of rituximab depletes both malignant and normal CD20-positive B lymphocytes via mechanisms that include antibody-dependent cell-mediated cytotoxicity and complement dependent cytotoxicity
- Rituximab indicated as single agent therapy or in combination with chemotherapy for CD20-positive B cell non-Hodgkin lymphomas
- Principal disease applications: follicular and low-grade lymphoma, diffuse large B cell lymphoma (DLBCL), B cell chronic lymphocytic leukemia (BCLL), Burkitt lymphoma, and occasional B lymphoblastic leukemia/lymphoma or plasma cell neoplasms
- CD20 unique among the targets in this chapter, for as HER2 or ER it is a target of the therapy, and like c-kit it is also a marker of cell lineage

- Loss of immunostaining with antibodies to CD20 may reflect masking and inaccessibility of CD20 epitope by rituximab and/or downregulation of CD20 expression
- When evaluating specimen for B cell non-Hodgkin lymphoma, prudent to employ an additional B cell marker (e.g., PAX5 or CD79A) if there has been a history of rituximab therapy
- Most common antibody employed for IHC is L26 mouse monoclonal

Further Reading

Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Mod Pathol. 1998;11:155–68.

Allred D, Carlson R, Berry D, Burstein H, Edge S, Goldstein L, et al. NCCN task force report: estrogen receptor and progesterone receptor testing in breast cancer by immunohistochemistry. J Natl Compr Canc Netw. 2009;7(Suppl 6):S1–21.

Apple S, Pucci R, Lowe A, Shintaku I, Shapourifar-Tehrani S, Moatamed N. The effect of delay in fixation, different fixatives, and duration of fixation in estrogen and progesterone receptor results in breast carcinoma. Am J Clin Pathol. 2011;135:592–8.

Badve S, Baehner F, Gray R, Childs B, Maddala T, Liu M, et al. Estrogen- and progesterone-receptor status in ECOG 2197: comparison of immunohistochemistry by local and central laboratories and quantitative reverse transcription polymerase chain reaction by central laboratory. J Clin Oncol. 2008;26:2473–81.

Bartlett JMS, Brookes CL, Robson T, van de Velde CJH, Billingham LJ, Campbell FM, et al. Estrogen receptor and progesterone receptor as predictive biomarkers of response to endocrine therapy: a prospectively powered pathology study in the Tamoxifen and Exemestane adjuvant multinational trial. J Clin Oncol. 2011;29:1531–8.

Bilous M, Ades C, Armes J, Bishop J, Brown R, Cooke B, et al. Predicting the HER2 status of breast cancer from basic histopathology data: an analysis of 1500 breast cancers as part of the HER2000 International Study. Breast. 2003;12:92–8.

Bloom K, Harrington D. Enhanced accuracy and reliability of HER-2/neu immunohistochemical scoring using digital microscopy. Am J Clin Pathol. 2004;121:620–30.

Brock JE, Hornick JL, Richardson AL, Dillon DA, Lester SC. A comparison of estrogen receptor SP1 and 1D5 monoclonal antibodies in routine clinical use reveals similar staining results. Am J Clin Pathol. 2009; 132:396–401.

Burry RW. Immunocytochemistry: a practical guide for biomedical research. New York: Springer; 2009.

Carlson RW, Moench SJ, Hammond ME, Perez EA, Burstein HJ, Allred DC, et al. HER2 testing in breast cancer: NCCN task force report and recommendations. J Natl Compr Canc Netw. 2006;4(Suppl 3):S1–22.

Cheang MCU, Treaba DO, Speers CH, Olivotto IA, Bajdik CD, Chia SK, et al. Immunohistochemical detection using the new rabbit monoclonal antibody SP1 of estrogen receptor in breast cancer is superior to mouse monoclonal antibody 1D5 in predicting survival. J Clin Oncol. 2006;24:5637–44.

Chen LL, Trent JC, Wu EF, Fuller GN, Ramdas L, Zhang W, et al. A missense mutation in KIT kinase domain 1 correlates with imatinib resistance in gastrointestinal stromal tumors. Cancer Res. 2004;64:5913–9.

Cheson BD, Leonard JP. Monoclonal antibody therapy for B-cell non-Hodgkin's lymphoma. N Engl J Med. 2008;359:613–26.

Collins LC, Botero ML, Schnitt SJ. Bimodal frequency distribution of estrogen receptor immunohistochemical staining results in breast cancer: an analysis of 825 cases. Am J Clin Pathol. 2005;123:16–20.

Corless CL, Barnett CM, Heinrich MC. Gastrointestinal stromal tumours: origin and molecular oncology. Nat Rev Cancer. 2011;11:865–75.

Davis TA, Czerwinski DK, Levy R. Therapy of B-cell lymphoma with anti-CD20 antibodies can result in the loss of CD20 antigen expression. Clin Cancer Res. 1999;5:611–5.

Demetri GD, von Mehren M, Blanke CD, Van den Abbeele AD, Eisenberg B, Roberts PJ, et al. Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. N Engl J Med. 2002; 347:472–80.

Druker BJ, Tamura S, Buchdunger E, Ohno S, Segal GM, Fanning S, et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. Nat Med. 1996;2:561–6.

Fang C, Xu W, Li J-Y. A systematic review and meta-analysis of rituximab-based immunochemotherapy for subtypes of diffuse large B cell lymphoma. Ann Hematol. 2010;89:1107–13.

Fitzgibbons P, Murphy D, Hammond M, Allred D, Valenstein P. Recommendations for validating estrogen and progesterone receptor immunohistochemistry assays. Arch Pathol Lab Med. 2010;134:930–5.

Forstpointner R, Unterhalt M, Dreyling M, Böck H-P, Repp R, Wandt H, et al. Maintenance therapy with rituximab leads to a significant prolongation of response duration after salvage therapy with a combination of rituximab, fludarabine, cyclophosphamide, and mitoxantrone (R-FCM) in patients with recurring and refractory follicular and mantle cell lymphomas: Results of a prospective randomized study of the German Low Grade Lymphoma Study Group (GLSG). Blood. 2006;108:4003–8.

Gao G, Liang X, Jiang J, Zhou X, Huang R, Chu Z, et al. A systematic review and meta-analysis of immunochemotherapy with rituximab for B-cell non-Hodgkin's lymphoma. Acta Oncol. 2010;49:3–12.

Gavrielides MA, Gallas BD, Lenz P, Badano A, Hewitt SM. Observer variability in the interpretation of HER2/neu immunohistochemical expression with unaided and computer-aided digital microscopy. Arch Pathol Lab Med. 2011;135:233–42.

Glennie MJ, French RR, Cragg MS, Taylor RP. Mechanisms of killing by anti-CD20 monoclonal antibodies. Mol Immunol. 2007;44:3823–37.

Goldstein NS, Ferkowicz M, Odish E, Mani A, Hastah F. Minimum formalin fixation time for consistent estrogen receptor immunohistochemical staining of invasive breast carcinoma. Am J Clin Pathol. 2003; 120:86–92.

Gown A. Current issues in ER and HER2 testing by IHC in breast cancer. Mod Pathol. 2008;21(Suppl 2):S8–15.

Gown A, Goldstein L. The knowns and the unknowns in HER2 testing in breast cancer. Am J Clin Pathol. 2011;136:5–6.

Gown AM, Goldstein LC, Barry TS, Kussick SJ, Kandalaft PL, Kim PM, et al. High concordance between immunohistochemistry and fluorescence in situ hybridization testing for HER2 status in breast cancer requires a normalized IHC scoring system. Mod Pathol. 2008;21:1271–7.

Grillo-López AJ, Hedrick E, Rashford M, Benyunes M. Rituximab: ongoing and future clinical development. Semin Oncol. 2002;29:105–12.

Haines 3rd GK, Wiley E, Susnik B, Apple S, Frkovic-Grazio S, Reyes C, et al. HER2 in well differentiated breast cancer: is testing necessary? Breast Cancer Res Treat. 2008;112:551–6.

Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol. 2010;28:2784–95.

Harvey JM, Clark GM, Osborne CK, Allred DC. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. J Clin Oncol. 1999;17:1474–81.

Heinrich MC, Blanke CD, Druker BJ, Corless CL. Inhibition of KIT tyrosine kinase activity: a novel molecular approach to the treatment of KIT-positive malignancies. J Clin Oncol. 2002;20:1692–703.

Heinrich MC, Corless CL, Demetri GD, Blanke CD, von Mehren M, Joensuu H, et al. Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor. J Clin Oncol. 2003a;21:4342–9.

Heinrich MC, Corless CL, Duensing A, McGreevey L, Chen C-J, Joseph N, et al. PDGFRA activating mutations in gastrointestinal stromal tumors. Science. 2003b;299:708–10.

Hiddemann W, Kneba M, Dreyling M, Schmitz N, Lengfelder E, Schmits R, et al. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. Blood. 2005; 106:3725–32.

Hirota S, Isozaki K, Moriyama Y, Hashimoto K, Nishida T, Ishiguro S, et al. Gain-of-function mutations of c-kit in human gastrointestinal stromal tumors. Science. 1998;279:577–80.

Hodi Z, Chakrabarti J, Lee AHS, Ronan JE, Elston CW, Cheung KL, et al. The reliability of assessment of oestrogen receptor expression on needle core biopsy specimens of invasive carcinomas of the breast. J Clin Pathol. 2007;60:299–302.

Hsu SM, Raine L, Fanger H. A comparative study of the peroxidase-antiperoxidase method and an avidin-biotin complex method for studying polypeptide hormones with radioimmunoassay antibodies. Am J Clin Pathol. 1981a;75:734–8.

Hsu SM, Raine L, Fanger H. Use of avidin-biotin-peroxidase complex (ABC) in immunoperoxidase techniques: a comparison between ABC and unlabeled antibody (PAP) procedures. J Histochem Cytochem. 1981b;29:577–80.

Ibarra JA, Rogers LW. Fixation time does not affect expression of HER2/neu: a pilot study. Am J Clin Pathol. 2010;134:594–6.

Ibarra J, Rogers L, Kyshtoobayeva A, Bloom K. Fixation time does not affect the expression of estrogen receptor. Am J Clin Pathol. 2010;133:747–55.

Jazirehi AR, Bonavida B. Cellular and molecular signal transduction pathways modulated by rituximab (rituxan, anti-CD20 mAb) in non-Hodgkin's lymphoma: implications in chemosensitization and therapeutic intervention. Oncogene. 2005;24:2121–43.

Joensuu H, Roberts PJ, Sarlomo-Rikala M, Andersson LC, Tervahartiala P, Tuveson D, et al. Effect of the tyrosine kinase inhibitor STI571 in a patient with a metastatic gastrointestinal stromal tumor. N Engl J Med. 2001; 344:1052–6.

Khoury T, Sait S, Hwang H, Chandrasekhar R, Wilding G, Tan D, et al. Delay to formalin fixation effect on breast biomarkers. Mod Pathol. 2009;22:1457–67.

Khoury T, Kulkarni S, Morrison C, Bshara W, Zakharia Y, Edge S. HER2 testing: how to reach a high concordance rate between immunohistochemistry and fluorescence in situ hybridization. Appl Immunohistochem Mol Morphol. 2011;19:424–30.

Köhler A, Lauritzen B, Van Noorden CJ. Signal amplification in immunohistochemistry at the light microscopic level using biotinylated tyramide and nanogold-silver staining. J Histochem Cytochem. 2000;48:933–41.

Lee AHS, Key HP, Bell JA, Hodi Z, Ellis IO. Concordance of HER2 status assessed on needle core biopsy and surgical specimens of invasive carcinoma of the breast. Histopathology. 2012;60:880–4.

Lehr H-A, Mankoff DA, Corwin D, Santeusanio G, Gown AM. Application of photoshop-based image analysis

to quantification of hormone receptor expression in breast cancer. J Histochem Cytochem. 1997; 45:1559–65.

Mann GB, Fahey VD, Feleppa F, Buchanan MR. Reliance on hormone receptor assays of surgical specimens may compromise outcome in patients with breast cancer. J Clin Oncol. 2005;23:5148–54.

Mayer G, Bendayan M. Biotinyl-tyramide: a novel approach for electron microscopic immunocytochemistry. J Histochem Cytochem. 1997;45:1449–54.

Miwa S, Nakajima T, Murai Y, Takano Y, Sugiyama T. Mutation assay of the novel gene DOG1 in gastrointestinal stromal tumors (GISTs). J Gastroenterol. 2008;43:531–7.

Moatamed NA, Nanjangud G, Pucci R, Lowe A, Shintaku IP, Shapourifar-Tehrani S, et al. Effect of ischemic time, fixation time, and fixative type on HER2/neu immunohistochemical and fluorescence in situ hybridization results in breast cancer. Am J Clin Pathol. 2011;136:754–61.

Mohsin SK, Weiss H, Havighurst T, Clark GM, Berardo M, le Roanh D, et al. Progesterone receptor by immunohistochemistry and clinical outcome in breast cancer: a validation study. Mod Pathol. 2004;17:1545–54.

Motta G, Cea M, Moran E, Carbone F, Augusti V, Patrone F, et al. Monoclonal antibodies for non-Hodgkin's lymphoma: state of the art and perspectives. Clin Dev Immunol. 2010;2010:428253.

Onrust SV, Lamb HM, Balfour JA. Rituximab. Drugs. 1999;58:79–88.

Osamura R, Gown A. Personalized medicine-molecular targeted cancer therapy. J Cell Mol Med. 2009; 13:4285.

Pauletti G, Godolphin W, Press MF, Slamon DJ. Detection and quantitation of HER-2/neu gene amplification in human breast cancer archival material using fluorescence in situ hybridization. Oncogene. 1996;13:63–72.

Perez EA, Suman VJ, Davidson NE, Martino S, Kaufman PA, Lingle WL, et al. HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. J Clin Oncol. 2006;24:3032–8.

Pertschuk LP, Feldman JG, Kim YD, Braithwaite L, Schneider F, Braverman AS, et al. Estrogen receptor immunocytochemistry in paraffin embedded tissues with ER1D5 predicts breast cancer endocrine response more accurately than H222Sp gamma in frozen sections or cytosol-based ligand-binding assays. Cancer. 1996;77:2514–9.

Polak JM, Van Noorden S. Introduction to Immunocytochemistry. 3rd ed. Oxford: Bios; 2003.

Press MF, Slamon DJ, Flom KJ, Park J, Zhou J-Y, Bernstein L. Evaluation of HER-2/neu gene amplification and overexpression: comparison of frequently used assay methods in a molecularly characterized cohort of breast cancer specimens. J Clin Oncol. 2002;20:3095–105.

Press MF, Sauter G, Bernstein L, Villalobos IE, Mirlacher M, Zhou J-Y, et al. Diagnostic evaluation of HER-2 as a molecular target: an assessment of accuracy and reproducibility of laboratory testing in large, prospective, randomized clinical trials. Clin Cancer Res. 2005;11:6598–607.

Qiu J, Kulkarni S, Chandrasekhar R, Rees M, Hyde K, Wilding G, et al. Effect of delayed formalin fixation on estrogen and progesterone receptors in breast cancer: a study of three different clones. Am J Clin Pathol. 2010;134:813–9.

Reddy JC, Reimann JD, Anderson SM, Klein PM. Concordance between central and local laboratory HER2 testing from a community-based clinical study. Clin Breast Cancer. 2006;7:153–7.

Renshaw S. Immunohistochemistry. Oxfordshire: Scion; 2007.

Rhodes A, Jasani B, Balaton AJ, Barnes DM, Miller KD. Frequency of oestrogen and progesterone receptor positivity by immunohistochemical analysis in 7016 breast carcinomas: correlation with patient age, assay sensitivity, threshold value, and mammographic screening. J Clin Pathol. 2000a;53:688–96.

Rhodes A, Jasani B, Balaton AJ, Miller KD. Immunohistochemical demonstration of oestrogen and progesterone receptors: correlation of standards achieved on in house tumours with that achieved on external quality assessment material in over 150 laboratories from 26 countries. J Clin Pathol. 2000b;53:292–301.

Rhodes A, Jasani B, Barnes D, Bobrow L, Miller K. Reliability of immunohistochemical demonstration of oestrogen receptors in routine practice: interlaboratory variance in the sensitivity of detection and evaluation of scoring systems. J Clin Pathol. 2000c;53:125–30.

Rhodes A, Borthwick D, Sykes R, Al-Sam S, Paradiso A. The use of cell line standards to reduce HER-2/neu assay variation in multiple European cancer centers and the potential of automated image analysis to provide for more accurate cut points for predicting clinical response to trastuzumab. Am J Clin Pathol. 2004;122:51–60.

Ross JS, Fletcher JA. The HER-2/neu oncogene: prognostic factor, predictive factor and target for therapy. Semin Cancer Biol. 1999;9:125–38.

Sabattini E, Bisgaard K, Ascani S, Poggi S, Piccioli M, Ceccarelli C, et al. The EnVision++ system: a new immunohistochemical method for diagnostics and research. Critical comparison with the APAAP, ChemMate, CSA, LABC, and SABC techniques. J Clin Pathol. 1998;51:506–11.

Schulz H, Bohlius JF, Trelle S, Skoetz N, Reiser M, Kober T, et al. Immunochemotherapy with rituximab and overall survival in patients with indolent or mantle cell lymphoma: a systematic review and meta-analysis. J Natl Cancer Inst. 2007;99:706–14.

Shi SR, Key ME, Kalra KL. Antigen retrieval in formalin-fixed, paraffin-embedded tissues: an enhancement method for immunohistochemical staining based on microwave oven heating of tissue sections. J Histochem Cytochem. 1991;39:741–8.

Shi SR, Imam SA, Young L, Cote RJ, Taylor CR. Antigen retrieval immunohistochemistry under the influence of pH using monoclonal antibodies. J Histochem Cytochem. 1995;43:193–201.

Shi SR, Cote RJ, Yang C, Chen C, Xu HJ, Benedict WF, et al. Development of an optimal protocol for antigen retrieval: a "test battery" approach exemplified with reference to the staining of retinoblastoma protein (pRB) in formalin-fixed paraffin sections. J Pathol. 1996;179:347–52.

Shi SR, Cote RJ, Taylor CR. Antigen retrieval immunohistochemistry: past, present, and future. J Histochem Cytochem. 1997;45:327–43.

Shi SR, Shi Y, Taylor CR. Antigen retrieval immunohistochemistry: review and future prospects in research and diagnosis over two decades. J Histochem Cytochem. 2011;59:13–32.

Skaland I, Nordhus M, Gudlaugsson E, Klos J, Kjellevold KH, Janssen EAM, et al. Evaluation of 5 different labeled polymer immunohistochemical detection systems. Appl Immunohistochem Mol Morphol. 2010;18:90–6.

Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987;235:177–82.

Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science. 1989;244:707–12.

Smith MR. Rituximab (monoclonal anti-CD20 antibody): mechanisms of action and resistance. Oncogene. 2003;22:7359–68.

Tawfik OW, Kimler BF, Davis M, Donahue JK, Persons DL, Fan F, et al. Comparison of immunohistochemistry by automated cellular imaging system (ACIS) versus fluorescence in-situ hybridization in the evaluation of HER-2/neu expression in primary breast carcinoma. Histopathology. 2006;48:258–67.

Taylor CR. Paraffin section immunocytochemistry for estrogen receptor: the time has come. Cancer. 1996;77:2419–22.

Taylor CR, Cote RJ. Immunomicroscopy: a diagnostic tool for the surgical pathologist. 3rd ed. Philadelphia: Saunders Elsevier; 2006.

Taylor CR, Shi SR, Chen C, Young L, Yang C, Cote RJ. Comparative study of antigen retrieval heating methods: microwave, microwave and pressure cooker, autoclave, and steamer. Biotech Histochem. 1996;71:263–70.

Tedesco KL, Thor AD, Johnson DH, Shyr Y, Blum KA, Goldstein LJ, et al. Docetaxel combined with trastuzumab is an active regimen in HER-2 3+ overexpressing and fluorescent in situ hybridization-positive metastatic breast cancer: a multi-institutional phase II trial. J Clin Oncol. 2004;22:1071–7.

Turbin D, Leung S, Cheang M, Kennecke H, Montgomery K, McKinney S, et al. Automated quantitative analysis of estrogen receptor expression in breast carcinoma does not differ from expert pathologist scoring: a tissue microarray study of 3,484 cases. Breast Cancer Res Treat. 2007;110:417–26.

Tuveson DA, Willis NA, Jacks T, Griffin JD, Singer S, Fletcher CD, et al. STI571 inactivation of the gastrointestinal stromal tumor c-KIT oncoprotein: biological and clinical implications. Oncogene. 2001;20:5054–8.

Umemura S, Itoh J, Itoh H, Serizawa A, Saito Y, Suzuki Y, et al. Immunohistochemical evaluation of hormone receptors in breast cancer: which scoring system is suitable for highly sensitive procedures? Appl Immunohistochem Mol Morphol. 2004;12:8–13.

van Oers MHJ, Klasa R, Marcus RE, Wolf M, Kimby E, Gascoyne RD, et al. Rituximab maintenance improves clinical outcome of relapsed/resistant follicular non-Hodgkin lymphoma in patients both with and without rituximab during induction: results of a prospective randomized phase 3 intergroup trial. Blood. 2006;108:3295–301.

Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, et al. Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol. 2002;20:719–26.

von Wasielewski R, Mengel M, Gignac S, Wilkens L, Werner M, Georgii A. Tyramine amplification technique in routine immunohistochemistry. J Histochem Cytochem. 1997;45:1455–9.

Vosse BAH, Seelentag W, Bachmann A, Bosman FT, Yan P. Background staining of visualization systems in immunohistochemistry: comparison of the avidin-biotin complex system and the EnVision + system. Appl Immunohistochem Mol Morphol. 2007;15:103–7.

Werner M, Chott A, Fabiano A, Battifora H. Effect of formalin tissue fixation and processing on immunohistochemistry. Am J Surg Pathol. 2000;24:1016–9.

West RB, Corless CL, Chen X, Rubin BP, Subramanian S, Montgomery K, et al. The novel marker, DOG1, is expressed ubiquitously in gastrointestinal stromal tumors irrespective of KIT or PDGFRA mutation status. Am J Pathol. 2004;165:107–13.

Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. J Clin Oncol. 2007;25:118–45.

Yaziji H, Goldstein LC, Barry TS, Werling R, Hwang H, Ellis GK, et al. HER-2 testing in breast cancer using parallel tissue-based methods. JAMA. 2004;291:1972–7.

Systems Pathology

7

José Costa and Michael J. Donovan

Contents

7.1 Overview 197

7.2 Clinical Systems Pathology (CSP) 198

7.3 Data Analysis Approaches Used in CSP 198

7.4 The Tools Used in CSP and Some Common CSP Applications 199

7.4.1 Cytofluorometry 199

7.4.2 Proteomics 200

7.4.3 Functional Genomics 201

7.4.4 Structural Genomics of Somatic Cells (Mutations, CNV) 202

7.4.5 Epigenomics 203

7.4.6 Spatial Architecture 203

7.4.7 CSP Applications in Clinical Oncology 205

Further Reading 206

J. Costa, MD (✉)
Department of Pathology, Yale University School of Medicine, New Haven, CT, USA

M.J. Donovan, PhD, MD
Mount Sinai School of Medicine, New York, NY, USA

7.1 Overview

- Systems pathology strives to understand disturbed physiological systems by integrating all levels of functional and morphological information into a coherent model
  - Systems are composed of many parts that assemble at different scales of organization and work across different time scales
  - The behavior of each part of a complex system is variable and depends on the behavior of other parts. The measurement of dynamic interactions is an important component of systems pathology
- Integration of data obtained at different scales of organization (molecular, cellular, organ, organism, and population) is facilitated by computational power
  - High-density datasets obtained in the clinic or in the experimental laboratory are interpreted with the help of mathematical models and agent-based simulations
  - Refined heuristic models enable, then, the formulation of innovative hypotheses
- The scope of systems pathology is broad, and systems level thinking is fundamental to advance preclinical and clinical models
  - Clinical systems pathology attacks specific problems arising in a clinical case using the tools of complex sciences. Using these tools, it aims at providing patient-specific answers to diagnostic prognostic and predictive questions (Fig. 7.1)

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_8,

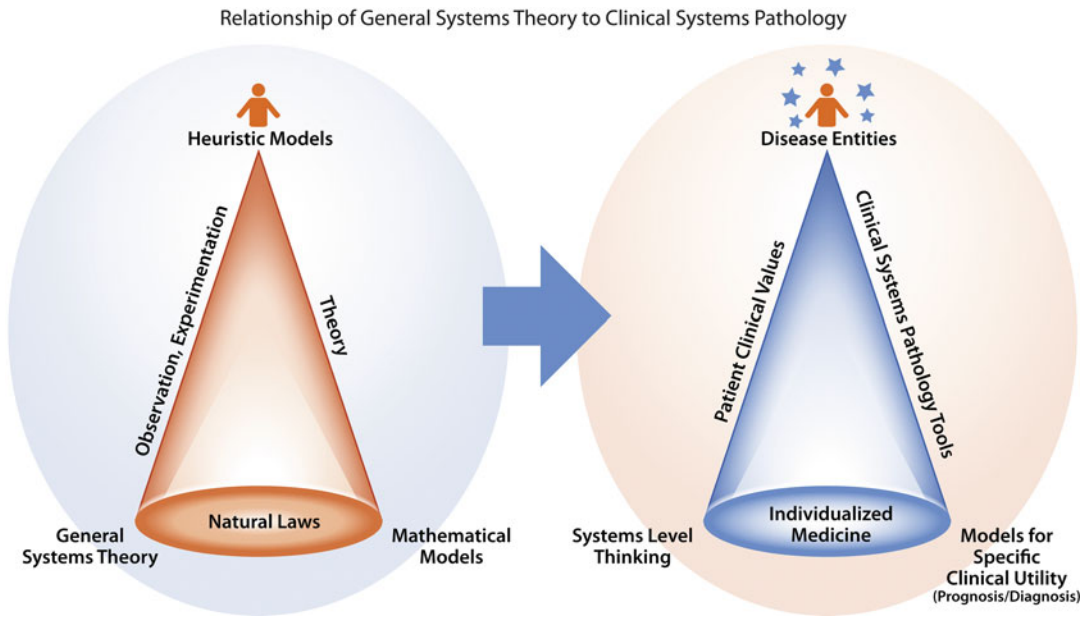


Fig. 7.1 Relationship of general systems theory to clinical systems pathology

7.2 Clinical Systems Pathology (CSP)

- Clinical systems pathology uses the conceptual tools and information streams traditionally used in clinical medicine to arrive at a diagnosis
  - A systems-based mechanistic diagnosis often provides a rational basis for therapeutic intervention
- CSP is highly integrative in nature as it makes use of all available data to arrive at the best answer to the question posed by a particular clinical case
- Relational databases link data derived from the clinical presentation laboratory parameters, macrophenotype and microphenotype (diagnostic imaging and histopathology), proteomics and functional genomics, and genomic sequencing data to generate a detailed model of a patient's state (Fig. 7.2)
- To generate a CSP model, it is necessary to assemble a series of cases representative of a given specific problem
  - These cases studied at the systems level generate a training dataset that is mined to generate a diagnostic or a predictive/prognostic algorithm
  - A successful algorithm, that is, one that provides an answer with sufficient clinical utility, is then tested on an independent set of cases. If the classifying or predictive powers are maintained, the model is fixed and can be used on individual patients
- As a result of this process, CSP is precise, predictive, objective, quantitative, integrative, and efficient and provides a good tool to practice evidence-based medicine
- A particular feature of CSP is the incorporation of spatial data at the macroscopic and microscopic scales
  - Modern technologies enable not only the quantitative capture of the phenotype (phenomics) but also the measure of molecular species (e.g., proteins and nucleic acids) in their spatial context

7.3 Data Analysis Approaches Used in CSP

- The approach to integrate and interpret the aggregate of data obtained with the tools of clinical systems pathology is based on

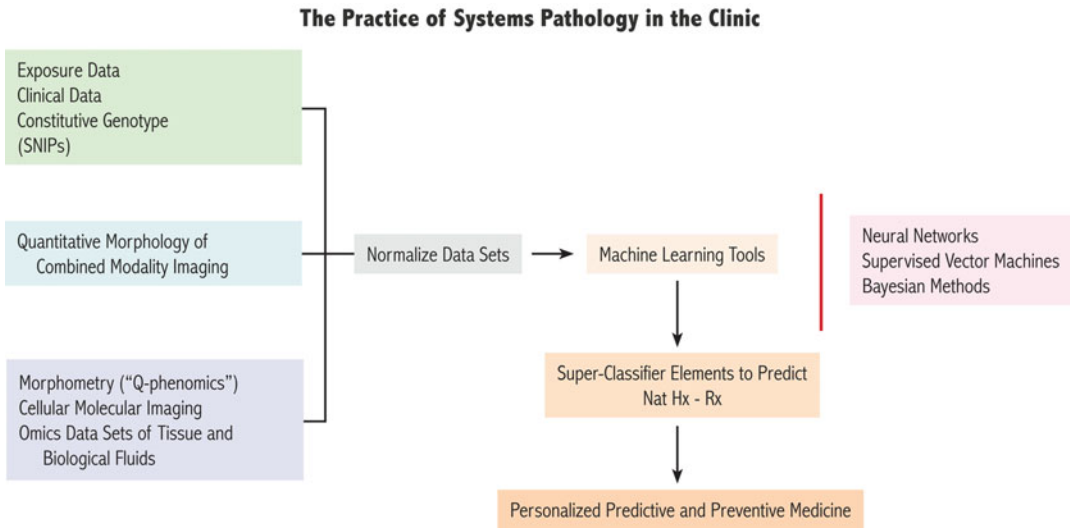


Fig. 7.2 Clinical systems pathology flowchart indicating the major sources of information used to arrive at a predictive diagnosis

(1) statistical learning and data mining and (2) modeling

- The fundamental principle of mining is to fit a model to data with minimal assumptions about what the model should be or how the variables in the data are related
- Several model architectures are available for the task
  - Neural networks
  - Classification and regression trees
  - Kernel machines (support vector machines)
- The major aim of data mining is prediction. For this a machine is built that given input(s) X predicts Y. Another common use is classification (diagnosis)
- When a predictive model is built, what we care about is not so much the error on the particular set of data used to build the model but the error we can expect on a new dataset
- Getting rid of the sampling noise in the mathematical model built on the original dataset leads to overfitting. Overfitting will generate error on new or out of sample datasets. The common strategies to avoid overfitting are cross validation, regularization, and capacity control
- The most common validation strategy is to divide the original data into a training set and a validation set
  - The training set of data is used to define the machine
  - Its performance is subsequently evaluated on the validation set

7.4 The Tools Used in CSP and Some Common CSP Applications

7.4.1 Cytofluorometry

- Cytofluorometry is a multichannel cell characterization technique designed to characterize (and sort) cell populations in suspension, commonly used to subclassify hematologic malignancies
- Stained cells traverse an excitation beam, and the fluorescence output provides a quantitative measure of various biochemical and biophysical properties of the cell
- Measurable optic parameters are also captured, including light absorption and light scattering, which is particularly useful for

measuring cell size, shape, density, granularity, and stain uptake
- Isolates rare cell populations including stem cells and circulating tumor cells, useful for patient stratification and treatment response
- CSP applications
  - DNA ploidy by cytofluorometry in various tumor types can help more accurately diagnose specific tumor types and predict outcome; for example, osteosarcomas can be appropriately characterized for future treatments
  - Cytofluorometry is used in conjunction with the circulating tumor cell platforms for pharmacodynamic assessment of the phenotypic characteristics of circulating tumor cells, posttreatment
  - Identify and monitor response/nonresponse between and within treatment groups. Examples include posttreatment evaluation of androgen receptor levels in castrate-resistant prostate cancer patients with antiandrogen, MDV3100, an oral androgen receptor antagonist
  - Assess proteosome activity function using fluorescent applications on isolated/sorted cells, bone marrow aspirates, etc.; determine response to proteasome inhibition with bortezomib (Velcade)
  - Quantitative flow can confirm and assess various measures of disease, for example, identification of collagen abnormalities in specific collagen myopathies. Approaches have the potential to be more accurate than proteomic methods such as immunohistochemistry
  - Applied to purify populations of engineered whole cell therapeutics (induced pluripotent stem cells) by removing residual embryonic stem cells, thereby removing potential for teratoma development
  - Microfluidic image cytometry has been applied to characterize disaggregated clinical brain tumor specimens and identify potentially relevant pathways, including PI3K/AKT/mTOR/PS6K signaling cascade – characterize an oncogenic driver for therapeutic targeting

7.4.2 Proteomics

- Study of the complete protein complement present within cells and tissues, particularly
  - Structure
  - Posttranslational modification (e.g., glycosylation, phosphorylation)
- Comprehensive techniques such as MALDI-ToF MS (matrix-assisted laser desorption/ionization time of flight mass spectrometry) offer a panoramic picture of the relative expression and modification of large numbers of proteins, but are not useful for differentiating proteins at the cellular level
- Novel approaches, including reverse phase lysate arrays and microwestern arrays,
  - Bridge the gap between large- and small-scale protein analysis approaches
  - Provide insight into the role that protein pathways play in several biologic processes
- Enzyme-linked, immunoabsorbent assays (ELISA) play a role in the management of some patients with newly diagnosed breast cancer, notably the levels of uPA and PAI1 antigen in breast tissue samples with respect to outcome
- Advances in immunofluorescence through multiplexing and image analysis have transformed in situ proteomic efforts to optimize cell characterization
- Nano-immunoassays can assess phosphorylation of proteins and therefore activation of selected signaling pathways in small numbers of cells, disaggregated biopsy specimens, or cell aspirates
- CSP applications
  - There already exist several clinically relevant diagnostic and prognostic tissue-based protein assays that routinely use immunohistochemistry to guide management decisions, including
    - In leukemia/lymphoma, by providing an accurate classification
    - In breast cancer HER2, estrogen and progesterone receptors (ER, PR)
    - The triple stain (CK5/6, p63, AMACR) used in the diagnosis of prostate cancer
  - Multiplex immunofluorescent assays merged with image analysis systems

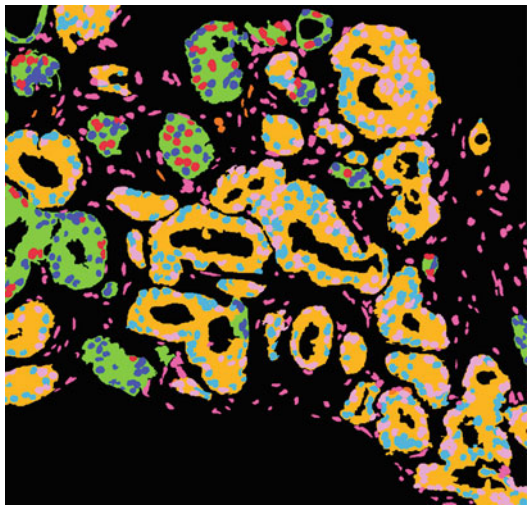


Fig. 7.3 Machine vision of a tissue section that has been stained using multiplexed antibodies and the presence of each antigenic determinant recognized on the tissue. Antigens are then mapped to the segmented morphology, generating a virtual functional histology. *Color key: orange*: AMACR(+) Epi (CK18+) gland, *green*: AMACR(–) Epi (CK18+) gland, *blue*: AR(–)/AMACR(–) nuclei, *red*: AR(+)/AMACR(–) nuclei, *cyan*: AR(–)/AMACR(+) nuclei, *light pink*: AR(+)/AMACR(+) nuclei, *magenta*: stromal nuclei

become quantitative tools for confirming cell-specific biomarker signatures with immediate relevance to assessing response to therapy and long-term outcome
  - Quantitative levels of both the androgen receptor (AR) and Ki67 have been shown to be associated with advanced disease and specifically for AR, with the durability of response to hormone therapy (Fig. 7.3)
- Immunohistochemistry of Ki67 in postneoadjuvant hormone breast cancer tissue samples from phase I/II randomized trials has demonstrated efficacy in supporting the decision to use chemotherapy
- Applying standard IHC to identify upregulated feedback loops from pretreatment and posttreatment samples has provided insight into the necessity for multimodal therapeutic regimens; for example, examination of upstream and downstream protein pathways for mTOR inhibitors (e.g., everolimus) for glioblastoma multiforme identified upregulation of the PI3K/PAKT feedback loop, promoting the use of dual PI3K/mTOR inhibition in clinical trials to assess impact on outcome endpoints
- The tissue-based ELISA that evaluates uPA and/or PAI1 has already achieved level one evidence (LOE1) for predicting the benefits of adjuvant chemotherapy in recently diagnosed breast cancer patients with node-negative disease
  - The transition of this assay to a more amenable nonfresh frozen tissue will remove, in the not-so-distant future, the need for a gene expression analysis to determine outcome in the same patient population
- The application of nanotechnologies to determine activation of pathways by protein phosphorylation will refine the current understanding of genomic variability/molecular phenotyping and the role signaling redundancy plays in driving disease
  - The NanoPro assays are able to quantify protein phosphorylation (e.g., EGFR, AKT, GSK, ERK1/2) in as few as 25 cells, thereby enabling analysis of functional heterogeneity in normal and diseased tissues
- Advances using traditional IHC have identified some cross-reactivity of the antibodies employed in the FDA-approved HER2 kits with other HER family members, raising questions regarding impact on patient response to standard of care therapy such as Herceptin in the HER2-positive patient

7.4.3 Functional Genomics

- Understanding the function of genes and other features of the genome, such as single nucleotide polymorphisms, repeat sequences, and noncoding regions of the DNA

- Systematic analysis of genes, their encoded proteins, and the relationship between gene activity and cell function
  - Enabled by manipulation of gene expression through gene silencing (siRNA) and assessment of phenotype (i.e., downregulation of proliferation, interruption of migration, induction of apoptosis)
  - Cloud computing through systems biology will continue to refine and clarify our assessment of the role of functional genomics in mechanisms of homeostasis, senescence, and the pathologic–biochemical manifestation of disease
- CSP applications
  - Identification of epithelial tumor translocations (see following section) has been driven almost entirely by the ability to assess the novel profiles of genes in specific subgroups of a particular cancer type
    - This approach led to the discovery of the TMPRSS2–ETS family members (ERG or ETV1) translocation in prostate cancer
  - Therapeutic use of gene silencing, for example, antisense inhibitor of the antiapoptosis gene clusterin in a phase II randomized clinical trial for patients with advanced, metastatic resistant prostate cancer
    - Biologic effects observed and improvement in overall survival noted

7.4.4 Structural Genomics of Somatic Cells (Mutations, CNV)

- Mutations of specific genes have been identified as critical for the evolution of disease and response to therapeutics. In tumors, notable examples include
  - *KRAS* mutations in colon, lung, and pancreatic cancer
  - *TP53* mutations in breast and ovarian cancer
  - *EGFR* mutations in colon, lung, and gliomas
  - *BRAF* V600E mutation in melanoma
- Gene copy number variation (CNV) produces structural genomic variants that arise from deletions or duplications of stretches of cellular DNA and result in a copy number change of the respective DNA segment
- The Human CNV Project identified CNVs across approximately 12% of the human genome, and association of specific CNVs with disease states raises the question of their potential role in causing disease
  - CNVs do play a clear role in drug-related genes by altering drug metabolism and drug response including toxicity
  - Some well-known CNVs include EGFR in glioblastoma multiforme and non-small cell lung cancer and HER2 in breast cancer
- Translocations also play an important role in the pathogenesis of tumors. They have been related mostly to tumors of mesenchymal origin (hematopoietic and sarcomas)
  - Leukemia and lymphoma are traditionally analyzed for a battery of translocations that allow for tumor classification and appropriate treatment
  - Translocations are also of diagnostic value in soft tissue tumors (synovial sarcoma, primitive neuroectodermal tumor, Ewing sarcoma, desmoplastic small cell tumor, alveolar rhabdomyosarcoma, inflammatory myofibroblastic tumor, among others)
- Relatively recent identification of the TMPRSS2–ERG translocation in prostate cancer and the EML4–ALK translocation in a subset of lung cancers has demonstrated that translocations are of pathogenetic importance in epithelial tumors and have found utility in the diagnosis and therapy of these tumors (e.g., crizotinib)
- CSP applications
  - Identification of patients with EGFR mutations in non-small cell lung cancer is now standard of care for deciding upon the appropriate use of EGFR inhibitors
    - Patients with activating mutations of exon 19–22 are more likely to respond to erlotinib or Tarceva
    - Patients with resistant mutations (T790M) should do as well with standard chemotherapy regimens

  - Identification of specific mutations is used to select appropriate biologic therapy; for example, EGFR inhibitors (cetuximab) for metastatic colon cancer are prescribed only in a setting of wild-type KRAS
  - Specific mutations (e.g., *BRAF* V600E) are currently used to guide treatment decisions in patients with metastatic melanoma
  - Detection of the EGFRviii mutation in brain tumor tissue and more recently in blood has been successfully applied to the early management of patients diagnosed with glioblastoma multiforme and pediatric medulloblastoma. Recent efforts have also expanded to include exosomes as a rich source of mRNA to use for assessing the presence of EGFRviii
  - In patients with prostate cancer, CNV has been proposed as a diagnostic/prognostic tool for deciding upon the use of adjuvant therapy postsurgery
    - The approach has also helped to define the meaning of molecular heterogeneity, evaluating not just one gene in a pathway but also multiple points of dysregulation as being "pathognomonic" for the disease in question, that is, activation of PI3K in combination with alterations in the expression patterns of several phosphatases in the same signaling cascade

7.4.5 Epigenomics

- Epigenetic mechanisms give rise to heritable changes in phenotype without altering the DNA sequence
  - A cell can dynamically alter its phenotype, ability to turn off and on gene activity through promoter methylation or modifying the state of chromatin
  - Epigenetics allows the cell to functionally adapt to its external–internal environment
- Epigenetic alterations including hypermethylation of gene promoters can be early events in neoplastic progression, and thus methylated genes may serve as biomarkers for the detection of cancer in clinical specimens
- Modifications through methylation are able to effectively transform the transcriptomic and proteomic output of the cell both in a normal and diseased state
- CSP applications
  - MGMT (O-6-methylguanine-DNA methyltransferase) is a DNA-repair enzyme that promotes resistance to alkylating agents such as temozolomide and is used in the treatment of malignant gliomas (e.g., glioblastoma [GBM])
    - Promoter methylation of MGMT reduces enzyme levels, thereby improving overall drug responsiveness
    - Clinical tests are currently used to evaluate the degree of MGMT methylation status in tumor samples and assess the likelihood of resistance to alkylating agents or the need to consider the use of novel alternative treatments (e.g., anti-integrin cilengitide)
  - The promoter methylation status of both GSTP1 (glutathione-S-transferase P1, an antioxidant enzyme that plays a significant role in allowing the organism to detoxify/remove free radicals) and APC (adenomatous polyposis coli, a tumor suppressor gene) has been incorporated into various biofluid and tissue assays to guide diagnostic treatment decisions for prostate cancer
  - Hypomethylation of Line-1 elements in squamous cell carcinoma of the head and neck has been linked to 5-FU response as well as prognosis

7.4.6 Spatial Architecture

- Quantification of the tissue architecture and/or cytological features is one of the tools used in CSP. A simple demonstration of how tissue architecture is integrated into CSP is provided by an exercise to test whether evidence of liver toxicity in a rat preclinical model could be identified with enough precision

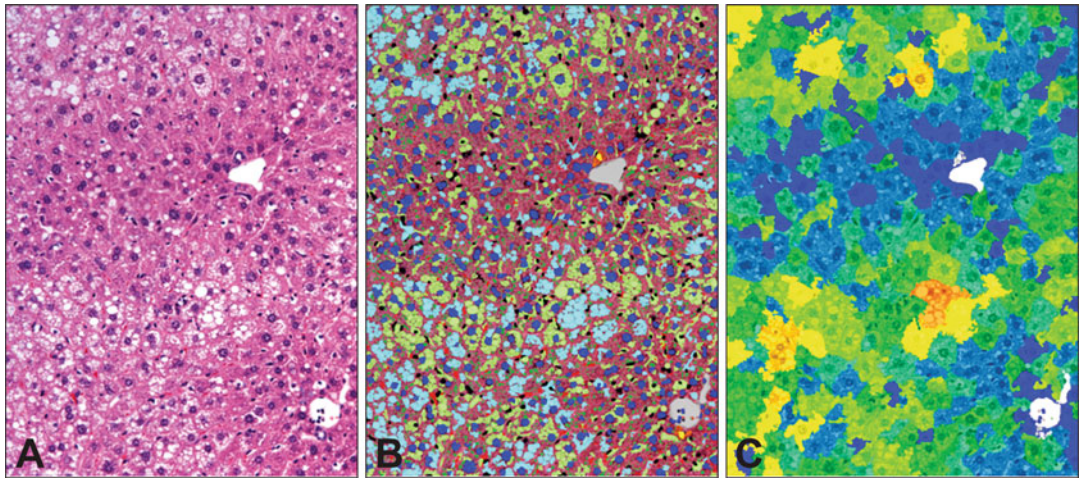


Fig. 7.4 Sections of liver (**a**) can be segmented and (**b**) subsequently analyzed/quantified; (**c**) the information derived from a group of normal and abnormal samples can be used to train a model to discriminate between the two groups with sufficient accuracy

  - H&E-stained rat liver sections from treated vs. nontreated animals were used to train a model for classifying animals as treated versus nontreated (i.e., exhibiting normal variation)
  - Rat liver H&E sections from treated animals demonstrated variable degrees of steatosis (fatty liver) as evidenced by hepatocyte vacuolization, which was distributed within the hepatic lobule
- Specific challenge: in all rats, regardless of treated vs. nontreated, there is a certain degree of hepatocytic vacuolization
  - Determine a threshold for the overall amount, volume, and distribution of steatosis that could reliably classify a liver as treated versus nontreated (Fig. 7.4)
- A classification algorithm was generated using a training set of 100 rats comprised of 50 animals that had been treated and 50 that had received no treatment
- The model was then applied to a de novo, never been seen cohort of 100 animals composed of 75% normal and 25% treated. It was found that the algorithm successfully classified this animal cohort as normal versus treated with an area under the curve (AUC) of 0.97. By comparison, a veterinary pathologist evaluation of the same set of animals yielded an accuracy metric (AUC) of 0.72
- This example demonstrates how spatial architecture at the microscopic scale can be captured and used to build a preclinical toxicity model that can be applied to multiple toxicity studies
  - The system can be automated for high throughput
  - The open architecture of the system enables the incorporation of additional predictors (i.e., serum analytes, genetic backgrounds, transcriptomic studies) to increase the classifying power and identify early signs and degree of cellular tissue toxicity
- A further example of revealing cytological changes in tissues, changes that can be quantified, is a study of nephrotoxicity
  - Formalin-fixed, paraffin-embedded tissue-based fluorescent in situ hybridization (PET-FISH) RNA detection assay successfully detected nephrotoxic changes at the gene expression level (in situ nascent RNA detection with PET-FISH) for proapoptotic biomarkers (e.g., clusterin) in kidney tubules prior to morphological change (e.g., cytoplasmic staining, features associated with necrosis)

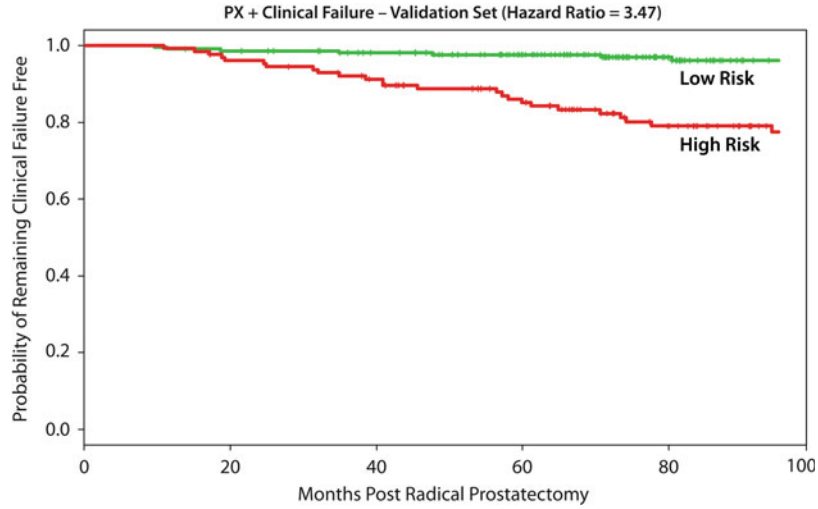


Fig. 7.5 Risk for fatal outcome can be predicted from the tissue analysis of prostate biopsies using the CSP methodology

7.4.7 CSP Applications in Clinical Oncology

- The CSP approach provides a novel framework in which very different types of information (including histologic morphometric features, clinical and pathology-based attributes, epidemiologic characteristics, and biomarker profiles) are integrated
  - Using artificial intelligence, algorithms are created that provide prognostic and predictive probabilities, which help to stratify patients into risk groups for an outcome of interest
- Some of these approaches can be based on prior knowledge. A given set of markers and clinical elements that play a role in a specific disease process can be tested on a CSP platform and so determine the relative importance of each of these features both individually as well as collectively
  - Usually a reduction in the number of features used in the model produces a clinically manageable signature that predicts a sought-after outcome (overall survival, response to a given therapy, or selection of therapy)
- Following the general scheme depicted in Fig. 7.2, the integration of quantitative multiplexed format biomarkers related to tissue and cytological architecture with clinical variables enabled the development of algorithms that could be applied to specific individual patients
  - For non-small cell lung cancer, a model was constructed in which increasing levels of cyclin D1, caspase-3, and activated and phosphorylated KDR (VEGFR2) along with advanced ECOG status predicted poor overall survival in patient's postmultimodal therapy including gefitinib
- A different study utilized the quantitative immunofluorescent platform of CSP to interrogate T and B lymphocytes on formalin-fixed, paraffin-embedded biopsy specimens and confirmed apoptotic resistance in steroid refractory and dependent patients with Crohn disease
  - By utilizing a multiplex immunofluorescent approach and features selected by their predictive performance in an agnostic fashion, even the findings obtained in endoscopic biopsy specimens were predictive of the patient class or status (refractory/dependent)
- A proof of concept of the effectiveness of the CSP approach has been provided by studies aiming at a more rational selection of the treatment of men with prostate cancer

  - The integration of clinical data with functional histologic attributes of the H&E-stained tissue section combined with quantitative immunofluorescent biomarker features has been successful in predicting disease recurrence both from the biopsy findings and from prostatectomy specimens (Fig. 7.5)
- The general principles of CSP indicate that integrating SNP-derived information that indicates susceptibility to prostate cancer would improve the predictive power of CSP. It is known to be the case for PSA at screening, but it is yet to be confirmed in the CSP setting

Further Reading

Aparicio SA, Huntsman DG. Does massively parallel DNA resequencing signify the end of histopathology as we know it? J Pathol. 2010;220:307–15.

Auffray C, Chen Z, Hood L. Systems medicine: the future of medical genomics and healthcare. Genome Med. 2009;1:2.

Azuaje F, Devaux Y, Wagner DR. Integrative pathway-centric modeling of ventricular dysfunction after myocardial infarction. PLoS One. 2010;5:e9661.

Azuaje FJ, Rodius S, Zhang L, Devaux Y, Wagner DR. Information encoded in a network of inflammation proteins predicts clinical outcome after myocardial infarction. BMC Med Genomics. 2011;4:59.

Capodieci P, Donovan M, Buchinsky H, Jeffers Y, Cordon-Cardo C, Gerald W, et al. Gene expression profiling in single cells within tissue. Nat Methods. 2005;2:663–5.

Capra F. The web of life: a new scientific understanding of living systems. 1st ed. New York: Anchor Books; 1996.

Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, BRIM-3 Study Group, et al. Improved survival with vermurafenib in melanoma with BRAF V600E mutation. N Engl J Med. 2011;364:2507–16.

Deisboeck TS, Zhang L, Yoon J, Costa J. In silico cancer modeling: is it ready for prime time? Nat Clin Pract Oncol. 2009;6:34–42.

Dietel M, Schafer R. Systems pathology–or how to solve the complex problem of predictive pathology. Virchows Arch. 2008;453:309–12.

Donovan MJ, Hamann S, Clayton M, Khan FM, Sapir M, Bayer-Zubek V, et al. Systems pathology approach for the prediction of prostate cancer progression after radical prostatectomy. J Clin Oncol. 2008;26:3923–9.

Donovan MJ, Costa J, Cordon-Cardo C. Systems pathology: a paradigm shift in the practice of diagnostic and predictive pathology. Cancer. 2009a;115(Suppl. 13):3078–84.

Donovan MJ, Khan FM, Fernandez G, Mesa-Tejada R, Sapir M, Zubek VB, et al. Personalized prediction of tumor response and cancer progression on prostate needle biopsy. J Urol. 2009b;182:125–32.

Donovan MJ, Kotstianti A, Bayer-Zubek V, et al. A systems pathology model for predicting overall survival in patients with refractory, advanced non-small-cell lung cancer treated with gefitinib. Eur J Cancer. 2009c;45:1518–26.

Donovan MJ, Osman I, Khan FM, Vengrenyuk Y, Capodieci P, Koscuiszka M, Anand A, et al. Androgen receptor expression is associated with prostate cancer-specific survival in castrate patients with metastatic disease. BJU Int. 2010;105:462–7.

Donovan MJ, Khan FM, Powell D, Bayer-Zubek V, Cordon-Cardo C, Costa J, et al. Postoperative systems models more accurately predict risk of significant disease progression than standard risk groups and a 10-year postoperative nomogram: potential impact on the receipt of adjuvant therapy after surgery. BJU Int. 2012;109:40–5.

Faratian D, Clyde RG, Crawford JW, Harrison DJ. Systems pathology–taking molecular pathology into a new dimension. Nat Rev Clin Oncol. 2009; 6:455–64.

Gaughan EM, Costa DB. Genotype-driven therapies for non-small cell lung cancer: focus on EGFR, KRAS and ALK gene abnormalities. Ther Adv Med Oncol. 2011;3:113–25.

Graversen JA, Suh LK, Mues AC, Korets R, Donovan MJ, Khan FM, et al. Independent diagnostic and post-treatment prognostic models for prostate cancer demonstrate significant correlation with disease progression endpoints. J Endourol. 2012;26:451–6.

Lei L, Sonabend AM, Guarnieri P, Soderquist C, Ludwig T, Rosenfeld S, et al. Glioblastoma models reveal the connection between adult glial progenitors and the proneural phenotype. PLoS One. 2011;6:e20041.

Noordermeer D, de Wit E, Klous P, van de Werken H, Simonis M, Lopez-Jones M, et al. Variegated gene expression caused by cell-specific long-range DNA interactions. Nat Cell Biol. 2011;13:944–51.

Oltvai ZN, Barabasi AL. Systems biology. Life's complexity pyramid. Science. 2002;298:763–4.

Peng Y, Jiang Y, Chuang ST, Yang XJ. Computer-aided detection of prostate cancer on tissue sections. Appl Immunohistochem Mol Morphol. 2009; 17:442–50.

Quintana E, Shackleton M, Foster HR, Fullen DR, Sabel MS, Johnson TM, et al. Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. Cancer Cell. 2010;18:510–23.

Saidi O, Cordon-Cardo C, Costa J. Technology insight: will systems pathology replace the pathologist? Nat Clin Pract Urol. 2007;4:39–45.

Santaolalla R, Mane J, Pedrosa E, Loren V, Fernandez-Barnares F, Mallolas J, et al. Apoptosis resistance of mucosal lymphocytes and IL-10 deficiency in patients with steroid-refractory of Crohn's disease. Inflamm Bowel Dis. 2011;17:1490–500.

Stricker T, Catenacci DV, Seiwert TY. Molecular profiling of cancer–the future of personalized cancer medicine: a primer on cancer biology and the tools necessary to bring molecular testing to the clinic. Semin Oncol. 2011;38:173–85.

Tanner NT, Pastis NJ, Sherman C, Simon GR, Lewin D, Silvestri GA. The role of molecular analyses in the era of personalized therapy for advanced NSCLC. Lung Cancer 2012;76:131–7.

Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling human prostate cancer. Cancer Cell. 2010;13:11–22.

Trock BJ, Brotzman MJ, Mangold LA, Bigley JW, Epstein JI, McLeod D, et al. Evaluation of GSTP1 and APC methylation as indicators for repeat biopsy in a high-risk cohort of men with negative initial prostate biopsies. BJU Int. 2012;110:56–62.

von Bertalanffy L. General system theory; foundations, development, applications. New York: G Braziller; 1969.

Walther Z, Sklar J. Molecular tumor profiling for prediction of response to anticancer therapies. Cancer J. 2011;17:71–9.

Section 2

Methodology in Molecular Genetic Pathology

Diagnostic Methodology and Technology in Molecular Genetic Pathology

8

Josephine Wu, Fei Ye, Miao Cui, Robert Shibata, Ruliang Xu, Liang Cheng, and David Y. Zhang

Contents

8.1 Sample Collection and Processing Methods 212
8.1.1 Sample Types 212
8.1.2 Sample Collections/Storage 212
8.1.3 DNA Extraction Methods 213
8.1.4 RNA Extraction Methods 215
8.1.5 Nucleic Acid Storage/Handling 219

8.2 Amplification Methods 219
8.2.1 Signal Amplification 220
8.2.2 Target-Based Amplification 227

8.3 Signal Detection Methods 244
8.3.1 DNA-Binding Dyes 244
8.3.2 Probe-Based Chemistries 246
8.3.3 Linear Probes 246
8.3.4 Structured Probes 248

8.4 Nonamplification Methods for the Detection of Nucleic Acid Targets 250
8.4.1 Hybrid Capture (HC) 250
8.4.2 In Situ Hybridization (ISH) 251
8.4.3 Southern Blot 254
8.4.4 Northern Blot 255
8.4.5 Allele-Specific Oligonucleotide (ASO) and Sequence-Specific Oligonucleotide (SSO) Hybridization 256
8.4.6 Reverse Hybridization 257

8.5 DNA Separation Methods 257
8.5.1 Gel Electrophoresis 257

8.6 Sequencing of Nucleic Acids 268
8.6.1 Sanger Sequencing 268
8.6.2 Dye Terminator Sequencing 270
8.6.3 The SNaPshot Method (Applied Biosystems, ABI) 271
8.6.4 Pyrosequencing 273
8.6.5 RNA Sequencing 276
8.6.6 Next-Generation Sequencing 276
8.6.7 Target Capture 279

8.7 Protein Detection Methods 280
8.7.1 Enzyme Immunoassay (EIA) 280
8.7.2 Protein Electrophoresis (SDS-PAGE) 283
8.7.3 Western Blotting (WB) 285
8.7.4 Technologies Used In Proteomics 287

Further Reading 299

J. Wu, DDS, CLSP(MB), CLDIR (✉) • F. Ye, PhD • M. Cui, MD • R. Shibata, MD • D.Y. Zhang, MD, PhD, MPH
Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

R. Xu, MD, PhD
Department of Pathology, New York University Langone Medical Center, New York, NY, USA

L. Cheng, MD
Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_9,

8.1 Sample Collection and Processing Methods

8.1.1 Sample Types

- General considerations
  - Ideal source of nucleic acid is fresh tissue, although paraffin-embedded tissue is acceptable
  - If extraction is not performed immediately, flash freezing of solid tissue or cells with liquid nitrogen preserves nucleic acids. This is particularly important with RNA, which is highly unstable and easily degraded by RNases
  - Fresh tissue may also be placed in commercially available reagents to preserve cellular RNA up to 1 week at room temperature
  - RNase and DNase are rapidly denatured by chaotropic agents like guanidinium isothiocyanate (GITC). For effective RNA stabilization, a minimum concentration of 5 mol/L is required. GITC-processed samples can be stored at room temperature

8.1.1.1 Whole Blood/Bone Marrow

- Most often the best available DNA source
- Requires the use of an anticoagulant to prevent clot formation

8.1.1.2 Buccal Cells

- Can be either collected and air-dried on cotton swabs or collected in a saline mouthwash and pelleted for immediate analysis

8.1.1.3 Cervical Cells

- Can be stored at room temperature for up to 2 weeks and longer when refrigerated in collection medium
- Fixed cytological specimen can also be used for nucleic acid isolation

8.1.1.4 Hair Root

- Useful in forensic testing, when other tissue is unavailable

8.1.1.5 Paraffin-Embedded Tissue

- The most common fixative is neutral buffered formalin. When exposed to nucleic acid, formalin causes the formulation of free nucleotide amino groups, methylene bridging of bases, and cross-linking of nucleic acid with protein, resulting in increased nucleic acid fragmentation
- With increasing fixation time, the amount of recoverable nucleic acid is progressively reduced. Tissue fixation in formalin for >24 h will significantly reduce the yield of nucleic acid
- Longer storage of the tissue block can lead to further degradation of nucleic acids
- Although RNA can be isolated, it is usually degraded; also, formalin may inhibit subsequent RT-PCR reactions
- Not suitable for Southern blot techniques

8.1.1.6 Bodily Fluids

- CSF: useful for analysis of CNS infection and lymphoma. Avoid blood contamination during spinal tap
- Urine: DNA can filter through the kidneys and is present in the urine of healthy individuals, usually as small fragments. Since DNA is very unstable in urine, samples must be processed within a few hours of collection
- Peritoneal and pleural fluid. Bloody specimens may need to remove red blood cells to avoid inhibition of PCR

8.1.2 Sample Collections/Storage

8.1.2.1 Whole Blood/Bone Marrow (Vacutainer Specimen Collection Tubes)

- Specimens should be transported within 24 h of draw and are best stored at 2–8 °C for up to 72 h. Storage at 22–25 °C is not recommended for >24 h. Freezing blood or bone marrow specimens without prior red blood cell lysis causes contamination with heme, which can inhibit PCR amplification. Leukocyte pellets can be stored for up to 1 year at −20 °C or for greater than 1 year at −80 °C

- Anticoagulants
  - EDTA (lavender-topped tubes)
    - Ethylenediaminetetraacetic acid
    - Preferred specimen collection type
  - Plasma Preparation Tube
    - BD Vacutainer® PPT™ Plasma Preparation Tube (BD PPT™ Tube) is a plastic evacuated tube for the collection of venous blood
    - Upon centrifugation, separates undiluted plasma for use in molecular diagnostic test methods
    - Use spray-dried K_2EDTA as anticoagulant for quantitative molecular diagnostic tests without dilution of plasma
    - Is a convenient, safe, single tube system for the collection of whole blood and the separation of plasma. It can be used for certain viral load testing, that is, HIV, HCV, and CMV
  - Citrate (yellow-topped tubes)
    - Acid citrate dextrose
    - Acceptable for molecular testing; provides a good yield of nucleic acids with greater than 70% of the original high molecular weight DNA
  - Heparin (Green-topped tubes)
    - Least preferred specimen collection tube
    - Heparin concentrations as low as 0.05 U per reaction volume may cause inhibition of enzymes, that is, DNA polymerase and Taq polymerase, and prevent amplification
    - Attempts to remove heparin activity, for example, ethanol precipitation, boiling and filtration, pH modification with gel filtration, or titration with protamine sulfate do not appear to eliminate inhibition. Serial washing of the buffy coat with saline prior to DNA extraction may be helpful
    - Heparinase treatment of extracted DNA may also be helpful but is expensive and not suitable for RNA due to its RNase activity. Alternatively, heparin-free RNA may be precipitated with lithium chloride
  - PAXgene blood RNA
    - A single tube for blood collection, RNA stabilization, and specimen transport and storage
    - Prefilled with intracellular RNA stabilization reagent to provide immediate RNA stabilization
    - Blood cell lysis in tube simplifies RNA purification
    - Suitable for downstream RNA analysis, such as RT-PCR
  - Cell Preparation Tube
    - BD Vacutainer® Cell Preparation Tube is for collection of whole blood and the separation of mononuclear cells
    - Cell separation medium is comprised of a polyester gel and density gradient liquid
    - Citrate is used for anticoagulation which is a better choice than EDTA for preserving cell markers and for PCR
- Fresh tissue
  - Should be collected in a sterile container (microcentrifuge tube) and frozen immediately in nitrogen at −80 °C. For travel of short distances, specimens should be placed sterile gauze prewet with sterile saline to prevent drying. If longer distance transportation is required, specimens should be frozen immediately and transported in dry ice
- Paraffin-embedded tissue
  - Can be stored at room temperature. However, the longer the duration of storage, the more DNA degradation occurs

8.1.3 DNA Extraction Methods

8.1.3.1 Manual

- QIAamp® DNA Blood Mini Kit (Qiagen, Hilden, Germany)
  - The QIAamp DNA Blood Mini Kit simplifies isolation of DNA from blood and related body fluids with fast-spin column or vacuum procedures. No phenol–chloroform extraction is required

  - The following samples can be processed: fresh and frozen whole blood (with common anticoagulants such as citrate, EDTA, and heparin), plasma, serum, buffy coat, bone marrow, lymphocytes, platelets, and body fluids
  - Typical yield from 200 ul of blood is 4–12 μg with a processing time of 20–40 min
  - DNA binds specifically to the QIAamp silica gel membrane while contaminants pass through. PCR inhibitors such as divalent cations and proteins are completely removed in two efficient wash steps, leaving pure nucleic acid to be eluted in either water or a buffer provided with the kit
  - Optimized buffers lyse samples, stabilize nucleic acids, and enhance selective DNA adsorption to the QIAamp membrane. Alcohol is added and lysates loaded onto the QIAamp spin column. Wash buffers are used to remove impurities, and pure, ready-to-use DNA is then eluted in water or low salt buffer
  - The complete process requires 20 min of handling time (lysis times differ according to the sample source)
  - With the QIAamp DNA Blood Mini Kit, blood can be processed via a vacuum manifold instead of centrifugation, for greater speed and convenience in DNA purification
  - Automated platforms are also available (Table 8.1)
- Gentra Puregene (Qiagen)
  - 150 μl to 20 ml of whole blood or bone marrow collected in common anticoagulants (EDTA, citrate, or heparin)
  - The PUREGENE kit works via alcohol and salt precipitation. The first step is to lyse cells with an anionic detergent in the presence of a DNA stabilizer that inhibits DNase activity, after which RNA is digested
  - The proteins are digested and removed along with other contaminants by salt precipitation. The DNA is then alcohol precipitated and dissolved in a DNA stabilizer
  - 25–60 min is required to process a sample
  - DNA yield is 35 μg/ml of blood
- AllPrep DNA/RNA/Protein Mini Kit (Qiagen)
  - Simultaneous purification of DNA, RNA, and protein from the same cell or tissue sample
  - It can be used in combination with Allprotect Tissue Reagent, which stabilizes DNA, RNA, and protein in tissue samples for long-term storage without freezing
  - Cells or tissues are first lysed and homogenized in lysis buffer, which immediately inactivates DNases and RNases as well as proteases to ensure isolation of intact DNA, RNA, and proteins
  - The lysate is then passed through an AllPrep DNA spin column. This column, in combination with the high salt buffer, allows selective and efficient binding of genomic DNA. The column is washed and pure, ready-to-use DNA is then eluted
  - Ethanol is added to the flow through from the AllPrep DNA spin column to provide appropriate binding conditions for RNA, and the sample is then applied to an RNeasy spin column, where total RNA binds to the membrane and contaminants are efficiently washed away. High-quality RNA is then eluted in RNase-free water
  - Protein precipitation solution is added to the flow through of the RNeasy spin column, and the precipitated proteins are pelleted by centrifugation. Intact total proteins are redissolved in an appropriate buffer and then ready to use in downstream applications
- QIAGEN Genomic-tips
  - Anion-exchange technology to purify high molecular weight DNA from a wide range of biological samples
  - DNA binds to the QIAGEN Resin in the column while other cell constituents such as proteins, carbohydrates, and metabolites flow through
  - Purified DNA is eluted in high salt buffer and precipitated with isopropanol

Table 8.1 Summary of automated Qiagen extraction platforms

Qiagen extraction platforms	
Instrument	Purification features
EZ1 advanced	DNA and RNA from 1 to 14 human and forensic samples or 1–6 samples in parallel
QIAcube	DNA from blood, plasmid DNA and RNA from animal cells; DNA/RNA cleanup from up to 12 samples; spin column format
QIAsymphony	DNA from 1 to 96 samples of blood, buffy coat, tissues, cells, or bacterial cultures; RNA from 1 to 96 samples of whole blood, cells, and tissues; 1–96 forensic and human identity samples
BioRobot Universal System	Genomic DNA, plasmid DNA, and DNA cleanup, plus downstream reaction setup; RNA whole blood, cells, or tissues, plus RT-PCR setup; 96-well format
BioRobot MDx workstation	Purification of genomic and viral DNA from human samples, bacterial DNA, and human RNA of whole blood; 96-well format
QIAxtractor	Purification of genomic DNA from virtually all sample types, 8–96 samples
Autopure LS	Purification of genomic DNA from a wide range of sample sources with scalable protocols, batches of 8–16 samples

  - DNA purified with QIAGEN Genomic-tips is sized up to 150 kb with an average length of 50–100 kb
- Agilent DNA Extraction Kit (Agilent, Santa Clara, CA)
  - The DNA Extraction Kit provides a simple, nontoxic method for efficiently isolating high molecular weight DNA from tissue, whole blood, and cultured cells
  - DNA isolated may be used directly for restriction digests, cloning, Southern blotting, PCR amplification, and other DNA analysis techniques
  - The DNA Extraction Kit is a modification of a procedure based on separating contaminating protein from DNA by salt precipitation. The procedure involves digestion of cellular proteins, subsequent removal of the proteins by "salting out" using standard sodium chloride, precipitation of the DNA with ethanol, and resuspension in the buffer of choice
    - 2 h are required to process a sample
    - DNA yield is 100–300 μg/5 ml of blood and 250 ug/g of tissue

8.1.3.2 Automated

- Various automated systems are available for the extraction and purification of nucleic acids (see Sect. 14.1.2)

8.1.4 RNA Extraction Methods

- General considerations
  - Adequate homogenization of cells or tissues is an essential step in RNA isolation to prevent RNA loss and degradation. The method of homogenization is best tailored to the particular cell or tissue type, that is, vortexing in a cell lysis solution for cultured cells or more rigorous disruption techniques such as enzymatic digestion for animal tissues, plant tissues, yeast, and bacteria for maximum recovery of RNA
  - Endogenous RNases must be inactivated immediately upon tissue harvesting to prevent RNA degradation. This can be effectively accomplished by
    - Homogenizing samples immediately after harvesting in a chaotropic-based cell lysis solution (e.g., containing guanidinium)
    - Flash freezing small tissue samples (homogenized) in liquid nitrogen
    - Utilization of an aqueous, nontoxic collection reagent (e.g., RNA*later*®, an RNA stabilization solution by Life Technologies) that stabilizes and protects cellular RNA in intact, unfrozen tissue and cell samples when samples cannot be immediately processed
      - Tissue samples must be in thin pieces (0.5 cm) so that the RNA*later* can

quickly permeate the tissue before RNases destroy the RNA. Cells or tissues can be harvested into RNA*later* and stored at room temperature for up to 1 week, at 4 °C for up to 1 month, or at −20 °C indefinitely
  - RNA is an unstable molecule and requires maintaining RNase-free laboratory conditions and diethylpyrocarbonate (DEPC)-treated glassware and water

8.1.4.1 Life Technologies (Grand Island, NY)

- Total RNA Isolation
  - LeukoLOCK® Total RNA Isolation System
    - Method for cellular fractionation of whole blood, total RNA stabilization, and extraction of RNA from the leukocytes (including T and B cells, neutrophils, eosinophils, basophils, monocytes, and other less abundant cell types)
    - It employs filter-based leukocyte depletion technology to isolate leukocytes from whole blood and stabilize the cells on a filter
    - Anticoagulated blood is passed through a LeukoLOCK filter, which captures the total leukocyte population while eliminating red blood cells (including reticulocytes), platelets, and plasma
    - The filter is flushed with RNA*later* to stabilize the RNA in the captured leukocytes. The RNA can be isolated immediately, or stabilized cells can be maintained for several days at room temperature, or for longer periods at −20 °C or −80 °C
    - 10–20 μg of RNA is isolated per 9–10 ml of whole blood
  - MagMAX™ AI/ND Viral RNA Isolation Kit, total RNA isolation
    - Ambion MagMAX system utilizes magnetic bead technology to isolate RNA from cells and viral RNA from cell-free samples, such as serum, plasma, swabs, and cell culture media
    - RNA is bound more efficiently than with glass fiber filter, resulting in higher and more consistent RNA yields. The MagMAX magnetic bead eliminates filter clogging from cellular particulates and allows the end user to concentrate RNA from large, dilute samples
    - As few as 10 copies of viral RNA from 100 to 400 μl sample can be recovered
    - Typical viral RNA recovery exceeds 50%
    - Ideal when working with low viral concentrations
    - Linear recovery of 50 transcripts
    - Extraction procedure takes approximately 30 min to complete
- mRNA Isolation
  - Poly(A) RNA (mRNA) makes up between 1% and 5% of total cellular RNA
  - mRNA isolation procedures are used in
    - Detection and quantitation of extremely rare mRNAs
    - Synthesis of probes for array analysis
    - The construction of random primed cDNA libraries, where the use of total RNA would generate rRNA templates that would significantly dilute out cDNAs of interest
  - Removal of ribosomal RNA (rRNA) and transfer RNA (tRNA) results in up to a 30-fold enrichment of a specific message
  - Poly(A) Purist® Kit (Ambion, Life Technologies)
    - Isolation of high-quality mRNA requires efficient removal of rRNA and specific recovery of poly(A) RNA
    - rRNA contamination occurs which is caused by nonspecific adsorption to the oligo(dT) matrix and binding to/copurifying with mRNA. Ambion Poly (A) purist kit minimizes this unwanted interaction while promoting efficient oligo(dT) selection
    - The Poly(A)Purist Kits are available in two formats, one that uses oligo(dT) cellulose-based selection (Poly(A) Purist Kit and MicroPoly(A)Purist Kit) and one that utilizes oligo(dT) magnetic bead-based purification (Poly(A)Purist MAG Kit)

  - Oligo(dT) cellulose-based selection – The oligo(dT) cellulose-based kits (Poly(A)Purist and MicroPoly(A) Purist) utilize batch binding of RNA to premeasured aliquots of oligo(dT) cellulose to avoid the problems of slow flow rates and clogged columns. A spin cartridge is used in the wash and elution steps
    - The Poly(A) Purist Kit contains reagents for 6 isolations, each for up to 2 mg total RNA, while the MicroPoly(A) Purist Kit contains reagents for 20 isolations, each for 2–400 μg total RNA
  - Magnetic bead-based selection employed in the Poly(A) Purist MAG Kit with less time to complete (<45 min)
    - The Poly(A) Purist MAG Kit contains reagents for up to 80 isolations, each from 100 μg of total RNA, or for eight large preps from as much as 1 mg of total RNA
- Total Nucleic Acid Isolation Kit
  - The MagMAX™ FFPE offers high-throughput purification of total RNA and DNA from formalin-fixed, paraffin-embedded (FFPE) tissue samples
  - Uses magnetic bead and no need for deparaffinization-free and xylene-free making for a simpler and safer process. Remove xylene and deparaffinization from your FFPE nucleic acid purifications
  - Purify total nucleic acids, including mRNA, miRNA, and DNA, from FFPE tissue samples
  - Isolate RNA or DNA in 3 h (automated) or 3.5 h (manual)

8.1.4.2 Purescript Total RNA Purification Kits (Gentra)

- These kits utilize a modified salt precipitation procedure in combination with highly effective inhibitors of RNase activity
- Samples may be whole blood or bone marrow, from 200 μl up to 30 ml, collected in common anticoagulants (EDTA, citrate, or heparin)
- Samples are processed in 60 min
- RNA yield is 2–7 μg/ml of blood

8.1.4.3 QIAamp® RNA Blood Mini Kit

- Simplifies isolation of RNA from blood with a fast-spin column procedure
- RNA binds specifically to the QIAamp silica gel membrane while contaminants pass through. PCR inhibitors such as divalent cations and proteins are completely removed in two efficient wash steps, leaving pure RNA to be eluted in either water or a buffer provided with the kit
- The typical yield is 1–5 μg RNA per ml healthy blood or up to 100 μg RNA from tissue in 30–100 μl RNase-free water elution
- Extracts total cellular RNA ready to use in downstream RT-PCR

8.1.4.4 Oligotex Direct mRNA Kits (Qiagen)

- Allow isolation of pure poly A^+ mRNA directly from cells or tissues in as little as 1 h
- Rigorous denaturing lysis conditions to generate an immediate RNase-free environment for the isolation of intact mRNA
- Use Oligotex resin for binding poly A^+ mRNA and spin columns for washing and eluting bound mRNA
- Greater than 90% recovery

8.1.4.5 NucliSENS® miniMAG (bioMérieux, Marcy l'Etoile, France)

- Generic extraction for DNA and RNA in various specimens using proprietary Boom technology with new magnetic silica particles. Samples lysed in buffer containing chaotropic agent that denatures and inactivates cells, bacteria, and viruses. The nucleic acids (DNA and RNA) are released, bound to silica particles, and eluted
- Standardized generic extraction protocol for DNA and RNA in various clinical samples
- Two platforms available
  - Manual magnetic bead-based system offering a throughput of 12 extractions in 40 min
  - Multiextraction automated platform offering a high throughput with 24 extractions in 40 min

8.1.4.6 PAXgene Tissue Kit (Qiagen)

- Tissue samples fixed and stored in PAXgene Tissue Containers can be paraffin-embedded and used for pathological studies as well as for subsequent purification of miRNA, RNA, and/or DNA
- The PAXgene Tissue miRNA Kit provides purification of total RNA, including RNA from approximately 18 nucleotides, from tissues fixed and stabilized in PAXgene Tissue Containers
- Purification is carried out using silica-based RNA purification technology in a spin column format
- Used with the containers, the kit provides a complete preanalytical solution for collection, fixation, and stabilization through to purification of high-quality mRNA and total RNA for molecular analysis

8.1.4.7 Quality and Quantity Assessment

- Analysis of the quality, quantity/concentration, and size of nucleic acid used is critical for the success of all aspects of molecular testing
- The following methods can be utilized to determine nucleic acid characteristics
  - Spectrophotometry
  - Fluorescent dyes
  - Electrophoresis

8.1.4.8 Spectrophotometry

- This is the simplest and most rapid method to evaluate purity, quantity, and quality of nucleic acids
- DNA and RNA demonstrate maximum absorption at approximately 260 nm. Protein absorbs at 280 nm, while background scatter absorbs at 320 nm
  - Protein absorption is primarily the result of the aromatic amino acids phenylalanine, tyrosine, and tryptophan
- DNA and RNA quality or purity can be measured by analysis of the optical density (OD) at 260 and 280 nm, $(OD_{260} - OD_{320})/(OD_{280} - OD_{320})$
  - A ratio of 1.7–2.0 is indicative of good quality nucleic acid. Less than 1.7 indicates too much protein or the presence of other contaminants, for example, organic solvents
- DNA and RNA quantity can be measured by OD_{260} reading. An OD_{260} reading of 1.0 corresponds to 50 μg/ml of double-stranded DNA, 40 μg/ml of single-stranded RNA, or 35 ug/ml single-stranded DNA. Concentrations are calculated as follows
  - dsDNA (ug/ml) = $(OD_{260} - OD_{320}) \times$ dilution factor $\times$ 50 ug/ml
  - ssRNA (ug/ml) = $(OD_{260} - OD_{320}) \times$ dilution factor $\times$ 40 ug/ml
  - ssDNA (ug/ml) = $(OD_{260} - OD_{320}) \times$ dilution factor $\times$ 35 ug/ml

8.1.4.9 Fluorescent Dye

- Alternative method to assess purity, quantity, and quality of nucleic acids
- The following fluorescent dyes bind nucleic acids: PicoGreen, ethidium bromide, acridine orange, diaminobenzoic acid (DABA), and propidium iodide. The most commonly used of these is PicoGreen
- PicoGreen can be used in quantitative assays. The intensity of fluorescence is dependent on the ratio of PicoGreen to nucleic acid
- Can reliably detect as little as 5–10 pg of DNA

8.1.4.10 Electrophoresis

- Small agarose gels ("minigels") offer an easy method to determine size and quantity of nucleic acid
- Molecular weight ladders (e.g., 100 bp marker) provide a reference standard for size determination
- Nucleic acid in samples can be quantified by extrapolation from a standard curve. The standard curve is produced by performing serial dilutions of samples with known nucleic acid concentrations
- Band smearing indicates DNA degradation or too much DNA loaded

8.1.4.11 Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA)

- This system is an alternative to traditional gel-based analysis that integrates the quantitation

of RNA samples with quality assessment in one quick and simple assay
- When used in coordination with the RNA 6000 LabChip®, as little as 1 μl of 10 ng/μl RNA is required per analysis
- In addition to assessing RNA integrity, this automated system also provides a good estimate of RNA concentration and purity (i.e., rRNA contamination in mRNA preparations) in a sample
- Concentration, integrity, and purity can be simultaneously analyzed in a single 5 ng sample and displayed as a gel-like image, an electropherogram, or tabular formats
  - RNA quality index (RQI) or RNA integrity number (RIN) determined based on 18S/28S rRNA ratio (measured on total RNA)

8.1.4.12 Troubleshooting

- No or low nucleic acid yield
  - Ensure that ample time is allowed for resuspension/rehydration of specimen
  - Concentrate dilute nucleic acid using ethanol precipitation
  - Repeat isolation from any residual specimen, modifying sample volume to compensate for possible low cell number or poor specimen handling
- Too much nucleic acid yield
  - Dilute sample and remeasure OD reading within instrument range
- Poor nucleic acid quality
  - For degradation, repeat isolation from residual specimen
  - For protein or other contamination, purify specimen by reisolation
- Band smearing
  - Due to poor DNA quality/DNA degradation
  - Due to poor quality of DNA synthesis

8.1.5 Nucleic Acid Storage/Handling

- General considerations
  - Nucleic acid samples are best stored as multiple aliquots in separate tubes in order to prevent degradation and damage from successive freeze–thaw events. Aliquotted tubes also minimize the potential for accidental contamination, for example, by DNase, RNase, specimen, or amplicon
  - To prevent amplicon contamination, exposed surfaces (bench space/work hoods, instruments, and floors) should be decontaminated with DNAase solution (DNase AWAY from Sigma) followed by DNase-free water or 10% bleach solution followed by 70% ethanol after use
  - To reduce the likelihood of exposure to ambient RNases, all laboratory surfaces, including pipettors, benchtops, glassware, and gel equipment should be decontaminated with a surface decontamination solution. RNase-free tips, tubes, and solutions should always be used, and gloves should be changed frequently
- DNA
  - Purified DNA should be stored in TE buffer at 4 °C for less than 1 week (short-term storage), –20 °C for longer-term storage, or −80 °C for indefinite storage (Fig. 8.1)
- RNA
  - For short-term storage, purified RNA should be stored at −20 °C
  - Purified RNA should be stored in RNase-free ultrapure water at −80 °C, for long-term storage
  - Although RNA resuspended in water or buffer can be stored at −80 °C, RNA is most stable in an NH_4OAc/ethanol precipitation mixture at −80 °C

8.2 Amplification Methods

- Molecular biological amplification technologies have undergone rapid evolution since the invention of the polymerase chain reaction (PCR) over 20 years ago. Most molecular diagnostic assays require amplification of the nucleic acid target, signal amplification, or both. Target-based amplification methods include PCR, strand displacement amplification (SDA), and transcription-mediated

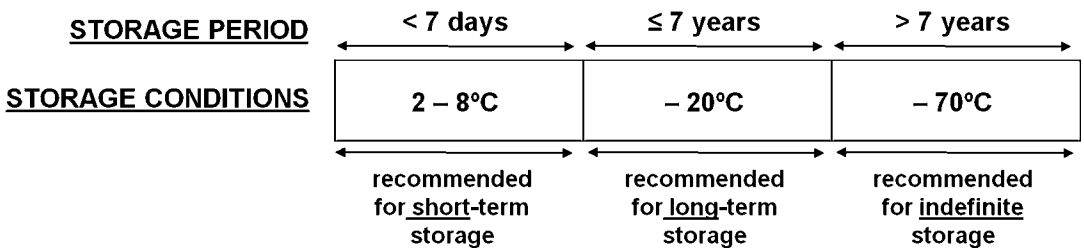


Fig. 8.1 Recommended conditions for DNA storage (Bogner and Killeen 2006)

amplification (TMA). Signal amplification methods include branched DNA (bDNA), rolling circle amplification (RCA), ramification amplification (RAM), and Invader® cleavase technology

- In general, amplification methods are rapid, specific, and extremely sensitive (in many situations, as few as 10 molecules of target nucleic acid can be detected per reaction)
- Owing to the exquisite sensitivity of amplification processes, special care must be taken to prevent contamination of samples with nucleic acid products from previous amplification reactions and other spurious material. Laboratory staff must strictly adhere to the principle of unidirectional workflow, in which physically separated areas of the laboratory are designated for preamplification, specimen preparation, and postamplification procedures, with one-way flow of samples, reagents, and amplification products between these areas. Other means of preventing contamination include cleaning work areas with 10% bleach at the end of each shift and the use of uracil-N-glycosylase (UNG) in PCR reactions

8.2.1 Signal Amplification

- Described below are probe-based methods of target detection which employ enhancement of a target-specific signal without increasing the copy number of the target as such. These methods have some intrinsic advantages over PCR-based techniques, including less risk of contamination, ease of operation, isothermal nature, and better linearity

8.2.1.1 Branched DNA (bDNA) (Bayer, Tarrytown, NY)

- General information
  - First described by Mickey S. Urdea, this technology was validated and developed by Chiron Diagnostics which is now owned by Bayer Inc.
  - The trade name is *Versant 3.0* in the United States and *Quantiplex 3.0* in Europe
  - Both DNA and RNA can be detected
  - Target: usually the entire genome
  - It is mainly used to detect infectious agents in clinical materials including HBV, HCV, HIV, and CMV
- Principle
  - Branched DNA (bDNA) is a signal amplification assay of nucleic acid probe
  - It uses dozens of probes which mediate the attachment of signal amplification molecules to viral nucleic acid targets
  - One end of the bDNA molecule is designed to bind to a specific target probe, while the other end of the bDNA molecule binds to multimers linked to alkaline phosphatase (ALP). ALP then catalyzes a chemiluminescence reaction
- Procedure (Fig. 8.2)
  - Target DNA or RNA is isolated; double-stranded DNA is denatured into single-stranded DNA. The RNA or single-stranded DNA is captured to a microwell by a first set of target probes which bind to capture probes coated on the microwell
  - A second set of target probes hybridizes to the DNA or RNA (the first set of target probes and the second set of target probes bind to different regions of the target DNA or RNA sequence)

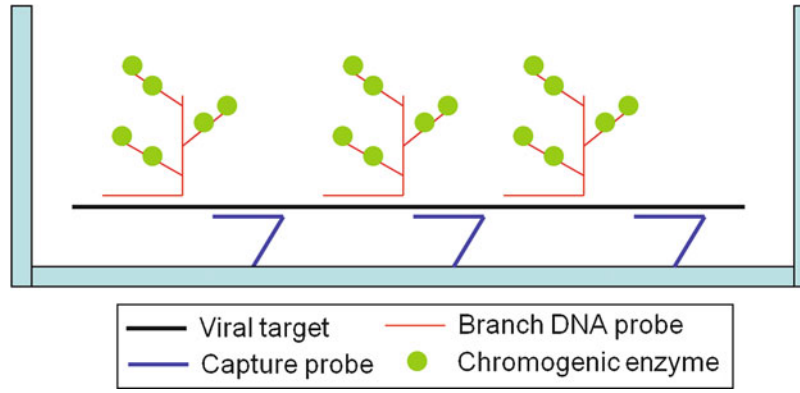


Fig. 8.2 Principle and procedure of branched DNA signal amplification (see text for description)

  - The second set of target probes hybridizes to the branched DNA molecules. Multiple copies of an ALP-labeled probe are then hybridized to the branched DNA molecules
  - Detection is achieved by incubating the ALP-bound complex with a chemiluminescent substrate. Light emission is directly related to the amount of target DNA or RNA present in each sample, and results are recorded as relative light units by the analyzer. A standard curve is defined by light emission from standards. Concentrations of DNA or RNA in specimens are determined from this standard curve
- Applications
  - bDNA technology can be used for Hepatitis C RNA viral load (VERSANT® version 3.0) detection. The assay has a broad dynamic range, accurately measures two-fold drops in viral load, predicts treatment nonresponders, and provides equal genotype quantitation
  - bDNA technology can also be used for HIV1 RNA viral loads (VERSANT® version 3.0) detection. The assay provides a broad linear range and a high level of precision and reproducibility and quantitates all major subtypes. It is also well suited to detect emerging subtypes
  - Other reported applications of the procedure include the detection and quantification of hepatitis B and cytomegalovirus
- Advantages
  - bDNA is very specific
  - The original amount of the target remains unmodified
  - bDNA allows broad detection of different genotypes in genetically diverse populations
  - bDNA is inherently quantitative
- Limitations
  - The bDNA assay has a narrower linear range for quantitation than quantitative PCR. Care must be taken when bDNA gives very low or very high results. In such cases, other methods, such as PCR-based quantitation, should be used for confirmation
  - bDNA requires multiple layers of probes to capture and signal the target molecule, which often produces high background
  - Although the bDNA technology format could be easily adapted to high-throughput screening, the assay's cost limits such applications
  - Because alkaline phosphatase is used in this assay, extreme care must be taken to avoid contamination with this ubiquitous enzyme
  - Sensitivity of bDNA is lower as compared to other target and signal amplification methods
  - A tedious procedure

8.2.1.2 Rolling Circle Amplification (RCA)

- General information
  - RCA is an isothermal signal amplification method

Fig. 8.3 Principle of RCA (see text for description)

  - The amplification mode of RCA is linear, and one primer is applied
  - The mechanism is based on the in vivo "rolling circle" replication of bacteriophages
- Principle and procedures (Fig. 8.3)
  - Uniquely designed circularizable probe (C probe or padlock probe) contains three regions: two target complementary sequences located at the 5′ and 3′ termini and an interposed generic linker region
  - Once the C probe hybridizes to its target, the 5′ and 3′ ends are juxtaposed. A closed circular molecule is then generated following incubation of the C probe-target complex with a DNA ligase
  - The resulting closed circular molecule is helically twisted around the target strand
  - The unique design of the C probe allows its amplification by a rolling circle (RCA) mechanism as observed in in vivo bacteriophage replication. In this scheme, a single forward primer complementary to the linker region of the C probe and a DNA polymerase-bearing strand displacement activity are employed
  - The polymerase extends the bound primer along the closed C probe for many revolutions and displaces upstream sequences, producing a long single-stranded DNA (ssDNA) of multiple repeats of the C probe sequence that can be as long as 0.5 megabase
  - This type of amplification results in linear growth of the products with up to several thousand-fold amplification
- Applications
  - When applied to a microarray system, RCA can detect 480 fM (150 molecules) of spotted primers, corresponding to an 8,000-fold increase in detection sensitivity over hybridization under the same conditions
  - The combination of RCA and DNA microarray technology allows for the real-time detection of multiple targets with great sensitivity and specificity
  - RCA can also be used to detect protein. Immuno-RCA has been developed. In this scheme, a primer is linked to an antibody and the signal is amplified by RCA
  - RCA has been used for in situ detection of mRNA
- Advantages
  - Since this method uses polymerases (such as phi 29 polymerase) that are capable of strand displacement, the reaction takes place under isothermal conditions
  - Because thermal cycling is not required, this is an ideal method for in situ amplification
  - RCA can make target capture, amplification reaction, and detection happen on the same solid support, such as a dipstick or latex beads. This makes the technology especially suitable for use in the field
- Limitations
  - Sensitivity is lower compared with RAM (see below), since only one primer is used
  - Since the procedure is technically complicated, future work will be required to evaluate its full potential

8.2.1.3 Ramification Amplification (RAM)

- General information
  - RAM is a novel isothermal DNA amplification method that amplifies a C probe exponentially through the mechanism of primer extension, strand displacement, and ramification
  - Also referred to as hyperbranched rolling circle amplification or cascade rolling circle amplification, this method represents an extension of the rolling circle amplification method described above

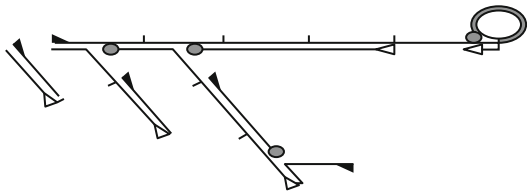


Fig. 8.4 Principle of RAM (see text for description)

- Principle and procedures (Fig. 8.4)
  - This method uses a specially designed circular probe (C probe) in which the 3′ and 5′ ends are brought together in juxtaposition by hybridization to a target, which forms an open loop
  - Before amplification begins, this open loop is covalently linked by a T4 DNA ligase in a target-dependent manner, producing a closed DNA circle
  - This circular DNA is the template for the forward primers to attach. Then a DNA polymerase extends the bound forward primer along the C probe and displaces the downstream strand, generating a multimeric single-stranded DNA (ssDNA)
  - This multimeric ssDNA then serves as a template for reverse primers to hybridize, extend, and displace downstream DNA, generating a large ramified (branching) DNA complex
  - This ramification process comes to an end when all ssDNAs become dsDNAs and no new primers are available
- Applications
  - The practical use of RAM has been shown in several studies for detecting target nucleic acids in clinical samples, such as *Chlamydia trachomatis* in cervical specimens collected in PreservCyt cytological solution
  - The RAM assay has also been used in the identification of *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli* (STEC) in food and human samples
- Advantages
  - Unlike PCR, the use of a thermocycler is not necessary. RAM employs the extension and displacement nature of some polymerases, such as phi 29 polymerase, making millionfold amplification feasible in a short period of time under isothermal conditions
  - The primer binding sequences of the C probe are identical, independent of the hybridization region. This feature facilitates multiplexing (simultaneous detection of multiple targets)
  - It is an ideal method for in situ amplification because RAM does not require thermal cycling
  - RAM can make target capture, amplification reactions, and detection happen on the same solid support, such as a dipstick or latex beads. This makes the technology especially suitable for use in the field
- Limitations
  - Specificity is somewhat low, and nonspecific background signals can be problematic
  - Since the procedure is technically complicated, future work will be required to evaluate its full potential

8.2.1.4 Invader Cleavase Technology (Hologic, Madison, WI)

- General information
  - The Invader® assay (formally Third Wave Technologies) is a homogeneous, isothermal DNA probe-based system for sensitive, quantitative detection of specific nucleic acid sequences
  - Invader reactions can be performed directly on either DNA or RNA, eliminating the need for target amplification and in the case of RNA, reverse transcription
  - The Invader system amplifies a target-specific signal but not the target itself
  - High specificity is achieved through a combination of sequence-specific oligonucleotide hybridization and structure-specific enzymatic cleavage
  - Cleavage is carried out using one of the Cleavase® (Third Wave Technologies) enzymes, a family of both naturally

occurring and engineered thermophilic structure-specific endonucleases
  - This technology can accurately detect single base changes, insertions, or deletions in DNA and RNA molecules
  - The Invader assay is also well suited for target quantitation over a broad dynamic range
- Principle and procedures (Fig. 8.5)
  - The basis for the Invader assay is the cleavage of a unique secondary structure formed by two partially overlapping oligonucleotides (an allele-specific primary probe and an invader probe) that hybridize to a target sequence to create a "flap"
  - Cleavase VIII (flap endonuclease I from *Archaeoglobus fulgidus*) recognizes this three-dimensional structure as a specific substrate and cleaves the 5′ flap of the primary probe
  - The flap initiates a secondary reaction in which the released 5′ flap serves as an invader probe on a fluorescence resonance energy transfer (FRET) cassette to create another overlapping tertiary structure that is, in turn, recognized and cleaved by the Cleavase enzyme
  - When the FRET cassette is cleaved, a fluorophore is separated from a closely adjoining quencher on the FRET cassette; the resulting free fluorophore emits a detectable fluorescence signal proportional in intensity to the concentration of the target sequence
- Applications
  - The Invader assay can be used to genotype DNA and has been applied in the detection of mutations in Factor V, Factor II (Prothrombin), ApoE, and methylenetetrahydrofolate reductase (MTHFR)
  - Recently, it is used for detecting HPV high-risk types and HPV genotype 16 and 18
  - The assay can be used directly on genomic DNA, total RNA or cell lysates without prior target amplification, or on samples previously amplified by PCR or RT-PCR
  - Invader assays can be used for manual, small-scale semiautomated analysis as well as fully automated, high-throughput studies
- Advantages
  - The Invader assay is an accurate, rapid, and cost effective tool, not only for SNP genotyping but also for the characterization of gene deletion and duplication events
  - Its suitability for both ultrahigh-throughput and low- to medium-throughput genotyping analyses has been well established
  - It has become the most widely used PCR-independent SNP and mutation detection technology
- Limitations
  - If DNA quality is poor and quantity is low, the assay may give invalid results
  - Insensitive detecting low copy number of target and preamplification may be necessary

8.2.1.5 Multiplex Ligation-Dependent Probe Amplification (MLPA) (MRC-Holland, Amsterdam, Netherlands)

- General information
  - Multiplex ligation-dependent probe amplification was first described in Nucleic Acid Research 30, e57 (2002) by Schouten et al.
  - It is commonly used for the detection of large deletion and copy number variations
  - Besides the genomic DNA, mRNA can be detected by reverse transcriptase MLPA (RT-MLPA) and the copy number quantification and methylation can be detected by methylation-specific MLPA (MS-MLPA)
  - AlleleID is the only software that designs synthetic probes for MLPA
- Principle
  - One MLPA probe consists of two probe oligonucleotides (Fig. 8.6)
  - Left probe oligo (LPO) which has a PCR primer sequence and left hybridizing sequence (LHS)

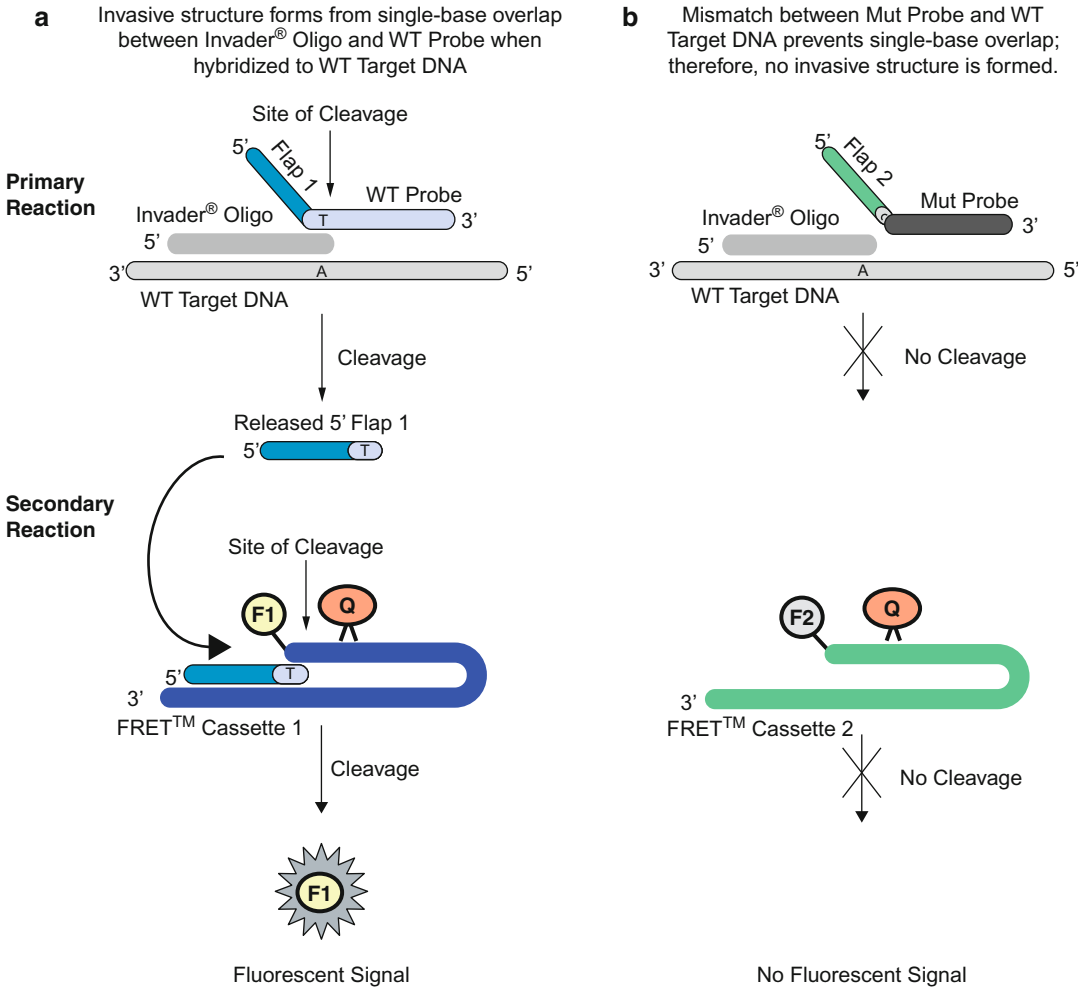


Fig. 8.5 Principle of Invader Cleavase technology (see text for description) (Source: www.twt.com)

- Right probe oligo (RPO) which has a PCR primer sequence, a stuffer sequence, and a right hybridizing sequence (RHS)
- LHS and RHS are directly adjacent but do not overlap when they hybridize to a target
- The length of LPO is 50–60 nt and RHS is 60–450 nt
- Tm of each hybridizing sequence (LHS or RHS): 70 °C (68 °C is absolute minimum)
- LPO primer recognition sequence: GGGTTCCCTAAGGGTTGGA, and RPO primer recognition sequence: TCTAGATTGGATCTTGCTGGCAC
- GC content: ~50% (if possible)
- LHS: preferably a maximum of 2 G/C nt in the 5 nt at its 3′ end, directly adjacent to the ligation site
- LHS/RHS: preferably a maximum of 3 G/C directly adjacent to the primer recognition sequence

• Procedure (Fig. 8.7)

- If the probe-target sequence is present in the sample after DNA denaturation, the two probes (LPO and RPO) oligos hybridize next to each other
- Hybridized probe oligos are joined by ligation

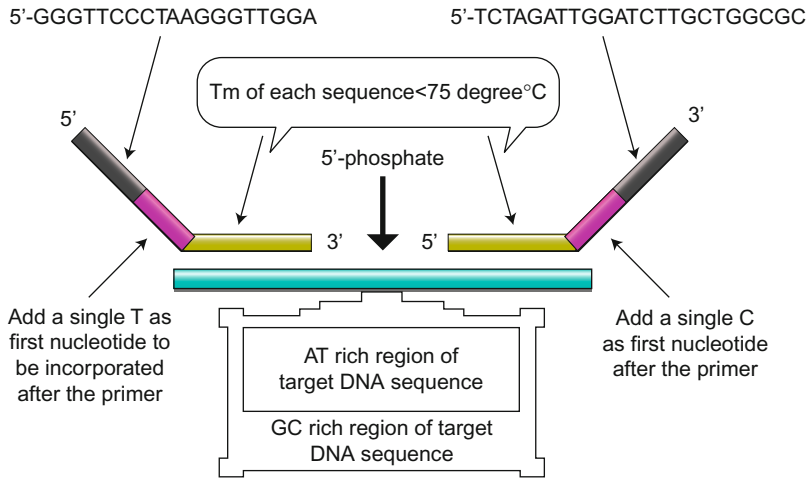


Fig. 8.6 Schematic presentation of the MLPA probe. One MLPA probe consists of two oligonucleotides: the left probe oligo (LPO) on the left and the right probe oligo (RPO) on the right. Both may contain an optional stuffer sequence in the middle, though not recommended for synthetic probes. LPO: 5′ end, recognition sequence of forward PCR primer (GGGTTCCCTAAGGGTTGGA); 3′ end, left hybridizing sequence (LHS). RPO: 5′end, right hybridizing sequence (RHS); 3′ end, recognition sequence reverse primer (TCTAGATTGGATCTTGCTGGCAC). The primer recognition sequences are used in the amplification of the probe during the PCR reaction

  - The ligated probes are amplified in a multiplex PCR. Not sample nucleic acids but the MLPA probes added to the sample are amplified by PCR (PCR reaction is very reproducible, as only one pair of PCR primers is used for amplification of all fragments)
  - Each probe generates an amplification product of unique length
  - Amplification products (130–490 nt) are analyzed by electrophoresis
- Applications
  - MLPA can be used for mutations and single-nucleotide polymorphisms scan: MLPA probe mixes for more than 100 genes are available in one reaction, such as BRCA1, MSH2, MSH6, MLH1, DMD, and APC
  - MLPA can be used for microdeletion syndromes, for example, MLPA kit P245 can detect 20 different microdeletion syndromes including Williams syndrome, Prader–Willi–Angelman, NF1, Cri-du-Chat, and DiGeorge (22q11 + 10p15)
  - Relative mRNAs quantification: MLPA probe mixes can be optimized for RNA samples derived from a specific tissue
  - Analysis of DNA methylation: detection of aberrant CpG island methylation
- Advantages
  - Detection of copy number of 40–50 DNA sequences in a single reaction. Only 20 ng human DNA required
  - Point mutation and SNP can be detected, and sensitivity and specificity were 0.997 and 0.989
  - An easy system: Only a thermocycler and a capillary electrophoresis system are required
  - High throughput: results available within 24 h
- Limitations
  - MLPA cannot be used on single cells. At least 3,000 cells are required
  - MLPA cannot detect balanced translocations. MLPA with telomere specific probe mixes can be used to detect most unbalanced translocations
  - MLPA cannot distinguish female triploid from diploid cells
  - MLPA assays for mRNA quantification have to be optimized for specific tissues
  - Dynamic range of MLPA is limited

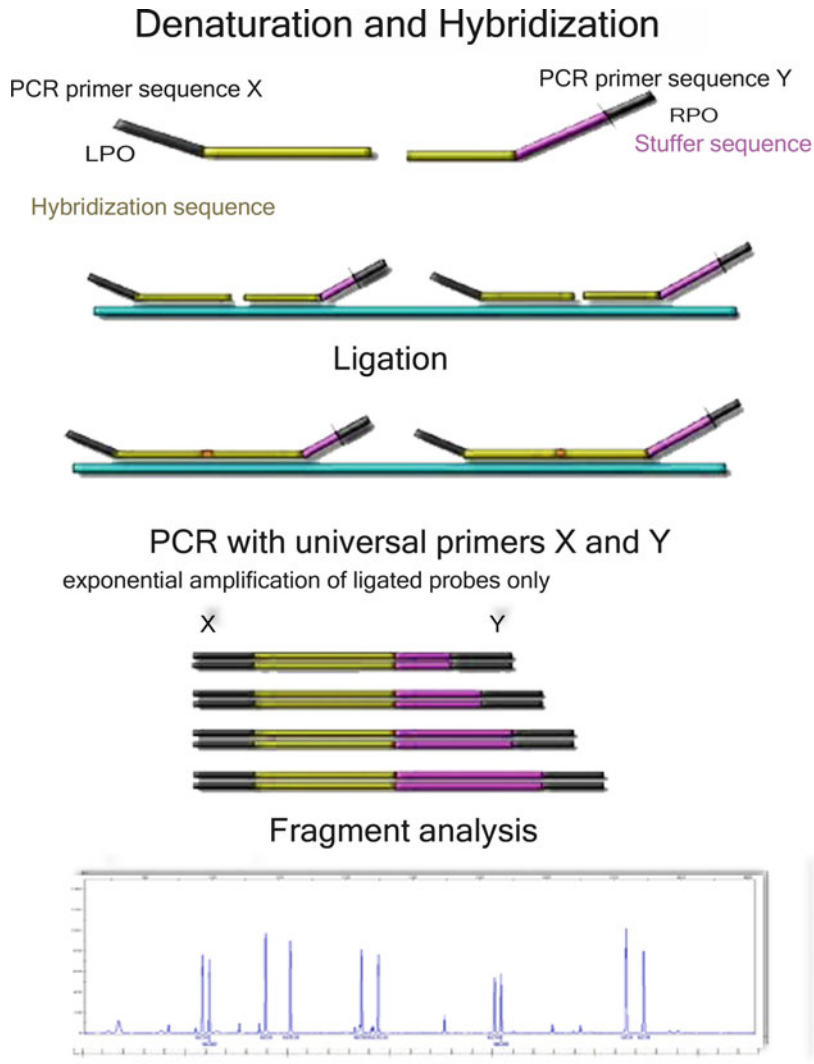


Fig. 8.7 Principle and procedure of MLPA. LPO and RPO are hybridized adjacent to target sequences, ligated and amplified. The amplification products are separated by capillary electrophoresis (see text or description)

8.2.2 Target-Based Amplification

- Target-based amplification methods are designed to detect and amplify the target gene of interest. Unlike probe-based amplification, the target is amplified over and over and the final signal is dependent on the amplified target

8.2.2.1 Polymerase Chain Reaction (PCR)

- General information
  - The polymerase chain reaction was invented by Kary B. Mullis in 1985
  - PCR technology has produced a revolution in molecular biology and is now the most widely used method of nucleic acid amplification in both clinical and research settings
  - PCR is a typical example of target-based amplification technology
- Principle and procedure (Fig. 8.8)
  - Oligonucleotide primers are designed to be complementary to the ends of the gene sequence of interest
  - Following heating to denature the DNA template and cooling to promote primer annealing, the oligonucleotide primers

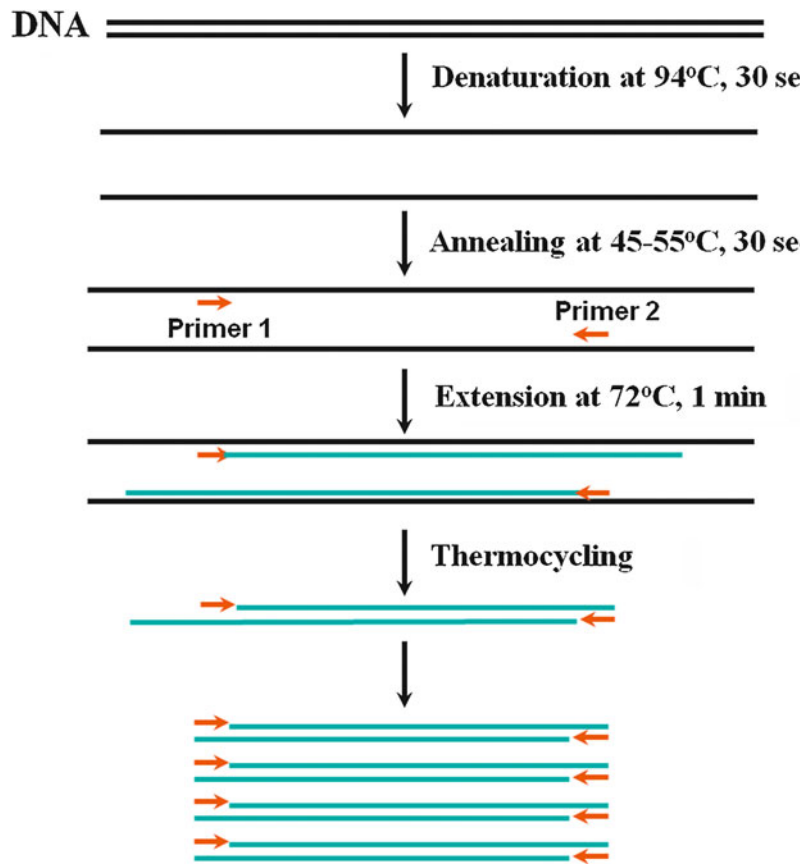


Fig. 8.8 Principle of PCR (see text for description)

each bind to the complementary strand of the target fragment
  - The primers are designed to anneal in positions such that when each is extended by a DNA polymerase, the newly synthesized strands will overlap the binding site of the opposite primer
  - As the steps of denaturation, annealing, and extension are repeated, the primers repeatedly bind to both the original DNA template and complementary sites in the newly synthesized strands and are extended to produce new copies of DNA
  - The end result is an exponential increase in the total number of target DNA copies
- Other considerations
  - Prevention of contamination
    - Wear a clean lab coat, not previously worn while handling amplified PCR products or used during sample preparation, and clean gloves
    - Change gloves whenever you suspect that they are contaminated
    - Maintain separate areas and dedicated equipment and supplies (see Fig. 8.9)
      - Reagent preparation
      - Sample preparation
      - PCR amplification
      - Analysis of PCR products
    - Never bring amplified PCR products into the PCR setup area (unidirectional workflow)
    - Open and close all sample tubes carefully. Try not to splash or spray PCR samples
    - Keep reaction tubes and tubes containing reaction components capped as much as possible

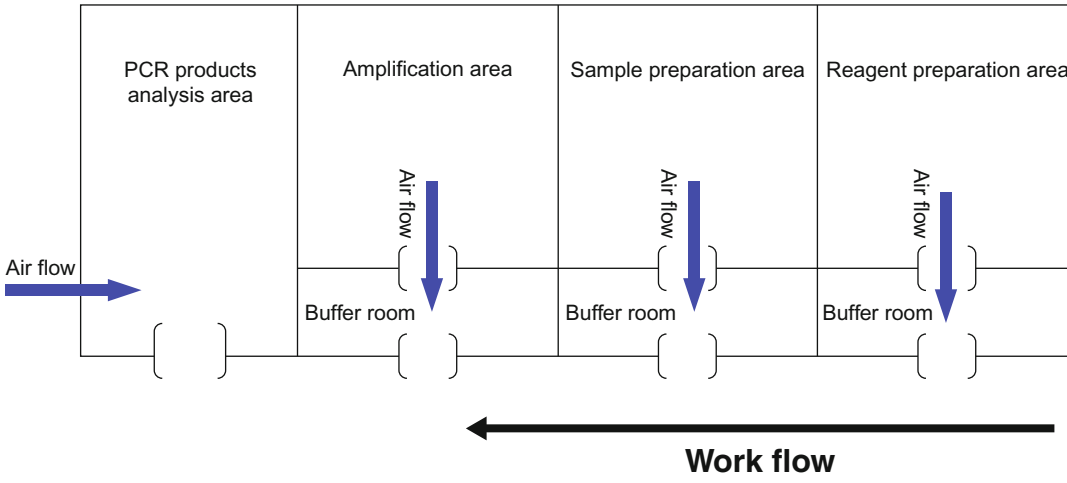


Fig. 8.9 Construction and floor map of a clinical PCR laboratory illustrating the principle of unidirectional workflow. Note that the PCR products analysis area is under negative air pressure; all other work areas are under positive pressure

- Use a positive displacement pipette or aerosol-resistant pipette tips
- Clean lab benches and equipment periodically with 10% bleach solution
- Use of real-time PCR can minimize amplicon contamination
- Recently introduced system includes all components (sample prep, reagents, PCR, and detection) in the same instruments, such as Roche COBAS AmpliPrep-COBAS TaqMan (CAP/TAQ). Therefore, eliminate the needs for separate areas

– Primer design
  - Primer design is a very important parameter for PCR. The primer sequence determines several things such as the length of the product, its melting temperature, and ultimately the yield. A poorly designed primer can result in little or no product due to nonspecific amplification and/or primer dimer formation, which can become competitive enough to suppress product formation
  - Several variables must be taken into consideration when designing primers, such as primer length, melting temperature (Tm), specificity, complementary primer sequence, G/C content and polypyrimidine (T, C) or polypurine (A, G) stretches, and 3′ end sequence

– Primer length: Since both specificity and the temperature and time of annealing are at least partly dependent on primer length, this parameter is critical for successful PCR
  - Oligonucleotides between 18 and 24 bases are extremely sequence specific, provided that the annealing temperature is optimal
  - Primer length is also proportional to annealing efficiency: In general, the longer the primer, the less efficient the annealing. With fewer templates primed at each step, this can result in a significant decrease in amplified product
  - The primers should not be too short, however, unless the application specifically calls for it
  - The goal should be to design a primer with an annealing temperature of at least 50 °C. The relationship between annealing temperature and melting temperature is one of the "black boxes" of PCR. A general rule of thumb is to use an annealing

temperature that is 5 °C lower than the melting temperature
    - The annealing temperature determined in this fashion will not be optimal and empirical experiments must be performed to determine the optimal temperature
- Melting temperature (Tm)
    - Both of the oligonucleotide primers should be designed such that they have similar melting temperatures
    - If primers are mismatched in terms of Tm, amplification will be less efficient or may not work at all since the primer with the higher Tm will misprime at lower temperatures and the primer with the lower Tm may not work at higher temperatures
    - The melting temperatures of oligos are most accurately calculated using nearest neighbor thermodynamic calculations with the formula: Tm (primer) = ΔH [ΔS + R ln (c/4)] – 273.15 °C + 16.6 log 10 [K+], where H is the enthalpy and S is the entropy for helix formation, R is the molar gas constant, and c is the concentration of primer. A good working approximation of this value (generally valid for oligos in the 18–24 base range) can be calculated using the formula: Tm = 2(A + T) + 4(G + C)
    - In addition to calculating the melting temperatures of the primers, care must be taken to ensure that the melting temperature of the product is low enough to ensure 100% melting at 92 °C
    - In general, products between 100 and 600 base pairs are efficiently amplified in many PCR reactions. For FFPE sample, it is recommended that the product should be less than 150 bp
    - The product Tm can be calculated using the formula: Tm = 59.9 + 0.41 (%G + C) − 675/length (under condition of 50 mM KCL)
- Specificity
    - Primer specificity is at least partly dependent on primer length
    - Primers must be chosen so that they have a unique sequence within the template DNA to be amplified
    - A primer designed with a highly repetitive sequence will result in a smear when amplifying genomic DNA. Because Taq polymerase is active over a broad range of temperatures, primer extension will occur at the lower temperatures of annealing. If the temperature is too low, nonspecific priming may occur which can be extended by the polymerase if there is a short homology at the 3′ end
- Complementary primer sequences
    - Primers need to be designed with absolutely no intraprimer homology beyond three base pairs. If a primer has such a region of self-homology, partially double-stranded structures can occur which will interfere with annealing to the template
    - Another concern is interprimer homology. Partial homology in the middle regions of two primers can interfere with hybridization. If the homology should occur at the 3′ end of either primer, primer dimer formation will occur, which, due to competition, will prevent the formation of the desired product
- G/C content, polypyrimidine (T, C) and polypurine (A, G) stretches
    - The base composition of primers should be between 45% and 55% GC

        - The primer sequence must be chosen such that there are no poly G or poly C stretches that can promote nonspecific annealing. Poly A poly T stretches are also to be avoided as these will open up stretches of the primer template complex. This can lower the efficiency of amplification
    - Polypyrimidine (T, C) and polypurine (A, G) stretches should also be avoided
        - Ideally the primer will have a near random mix of nucleotides, a 50% GC content, and be ~20 bases long. This will put the Tm in the range of 56–62 °C
    - 3′ end sequence
        - It is well established that the 3′ terminal position in PCR primers is essential for the control of mispriming
        - It is preferable for a G or C residue to occupy the 3′ terminal position of each primer. This "GC Clamp" helps to ensure correct binding at the 3′ end due to the stronger hydrogen bonding of G/C residues. It also helps to improve the efficiency of the reaction
  - When applied in clinical settings, adequate and strict QC/QA procedures must be followed
- Advantages
  - PCR is extremely sensitive and can be applied in a vast number of areas
  - Numerous variations on the basic PCR procedure further expand the utility of this methodology
- Limitations
  - Potential contamination is a major concern in PCR
    - The sources for contamination could be samples of high DNA concentration, DNA template controls or standards, and PCR carryover contamination
    - Following the guidelines discussed above will lessen the risk of contamination
    - Additionally, use of uracil-N-glycosylase (UNG, commercially known as AmpErase®) specifically mitigates the potential for contamination by PCR products from prior reactions
      - The AmpErase enzyme recognizes and catalyzes the destruction of DNA strands containing deoxyuridine, but not DNA containing deoxythymidine. Deoxyuridine is not present in naturally occurring DNA but is always present in amplicon due to the use of deoxyuridine triphosphate as one of the dNTPs in the Master Mix reagent; therefore, only amplicon contains deoxyuridine
      - Deoxyuridine renders contaminating amplicon susceptible to destruction by the AmpErase enzyme prior to amplification of the target DNA. The AmpErase catalyzes the cleavage of deoxyuridine-containing DNA at the deoxyuridine residues by opening the deoxyribose chain at the C1 position
      - When heated in the first thermal cycling step, the amplicon DNA chains break at the positions of the deoxyuridine, thereby rendering the DNA nonamplifiable
      - The AmpErase enzyme is inactivated by temperatures above 55 °C and therefore does not destroy target amplicon formed during amplification
  - Nonspecific amplification is another concern in PCR. Some technologies have been invented to prevent this phenomenon
    - Hot start PCR
      - Hot start PCR can improve PCR specificity by controlling mispriming events
      - In hot start PCR, reactions are designed such that the polymerase

only becomes active at a high temperature, thus ruling out the possibility of nonspecific amplification at lower temperature
  - Hot start PCR technique can be realized by manually adding the key components at higher temperature, which is cumbersome and time-consuming
  - Various commercially available reagents can simplify the hot start procedure. For example, AmpliTaq Gold® DNA polymerase (Applied Biosystems) is a chemically modified form of AmpliTaq DNA polymerase which is only active at high temperature. Binding to an aptamer blocks polymerase activity at low temperatures
- UNG (described above) not only can prevent postamplification amplicon contamination but also can destroy any nonspecific, misprimed products prior to specific amplification, thus improving specificity and sensitivity

8.2.2.2 Variations of PCR

Reverse Transcription PCR (RT-PCR)

- General information
  - Reverse transcription coupled with the polymerase chain reaction (RT-PCR) has proven extremely useful in the study of gene expression or detection of RNA virus
  - In RT-PCR, an RNA template is reverse transcribed to generate a DNA transcript (cDNA) using a retroviral reverse transcriptase
  - The cDNA sequence of interest is then amplified exponentially using PCR. Detection of the PCR product is typically performed by fluorescence dye for real-time PCR or agarose gel electrophoresis followed ethidium bromide staining
- Principle
  - RT-PCR is reverse transcription and PCR amplification combined together
  - Two basic reactions are involved in RT-PCR
    - cDNA is synthesized from an mRNA template usually by avian myeloblastosis virus (AMV) or Moloney murine leukemia virus (MMLV or MuLV) reverse transcriptases
    - The second cDNA strand is synthesized, and subsequent PCR amplification is performed with *Taq* DNA polymerase as with conventional PCR
- Procedure (Fig. 8.10)
  - Synthesis of the first DNA strand in the presence of primers, dNTPs, and an RNA-dependent DNA polymerase, that is, reverse transcriptase
  - RNaseH (an RNA digestion enzyme) is added to digest the RNA away from the RNA-cDNA hybrid
  - Second strand synthesized with DNA polymerase
  - Standard PCR is conducted using DNA oligo primers specific for the sequence of interest
- Applications
  - RT-PCR is the most sensitive technique for mRNA detection and quantitation
  - The procedure can be either qualitative or quantitative
  - It is also used to create cDNA libraries
- Advantages
  - Sensitivity: RT-PCR is the most sensitive technique for mRNA detection and quantitation currently available. Theoretically, a single copy of mRNA can be detected by this technique. In practice, tens to hundreds of copies are required for reliable quantitation
  - Sample integrity requirements: Since most RT-PCR methods amplify only a few hundred bases rather than the complete mRNA sequence, the sample RNA can be slightly degraded
  - Quantitation
    - Like other methods of mRNA analysis, RT-PCR can be used for relative or absolute quantitation

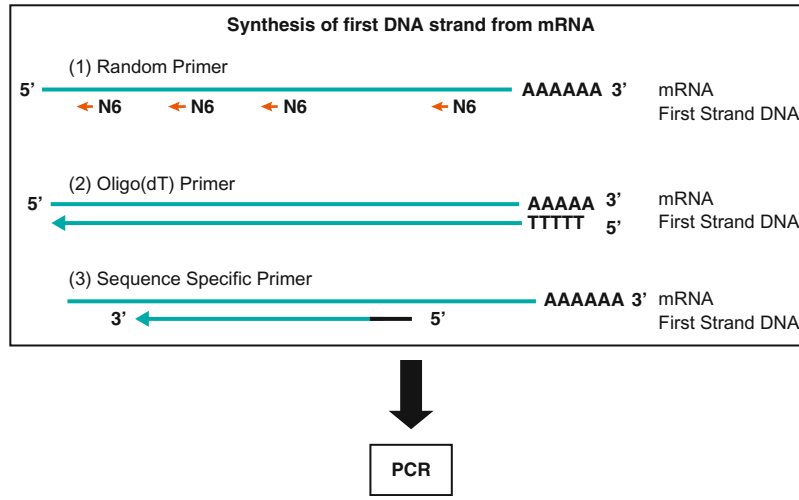


Fig. 8.10 Principle and procedure of RT-PCR (see text for description)

- Relative quantitation compares transcript abundance across multiple samples, using a coamplified internal control, which ideally has invariant expression within those samples, for sample normalization
- Absolute quantitation using competitive RT-PCR measures the absolute amount of a specific mRNA sequence in a sample. Dilutions of a synthetic RNA molecule (identical sequence in PCR primer binding site but slightly shorter than or different from the endogenous target) are added to sample RNA replicates and are coamplified with the endogenous target. The PCR product from the endogenous transcript is then compared to the concentration curve created by the synthetic "competitor RNA"
- It is also possible to do real-time RT-PCR quantitation by measuring an internal control in replicate samples

- Limitations
  - Sample purity requirements
    - Because of its sensitivity, the technique of RT-PCR requires that samples be free of genomic DNA or other DNA contaminants
    - Special care must be taken during RNA isolation to ensure that the sample RNA is DNA-free
  - Optimization of RT-PCR assays can be technically challenging. The design of suitable primers and controls often requires substantial preexperimental planning
  - In relative RT-PCR, the choice of internal standard is critical. An ideal internal standard is one with invariant expression during the cell cycle, between cell types, and in response to the experimental treatment under analysis
  - Competitive RT-PCR makes use of an exogenous RNA transcript (competitor) that must be accurately quantitated and added to replicate samples in amounts that span the range of the target mRNA levels

Real-Time PCR

- General information
  - Real-time PCR systems are based on the detection and quantitation of a fluorescence signal (Fig. 8.11a)
  - The fluorescence signal increases in direct proportion to the amount of PCR product formed. Fluorescence is measured repeatedly with each PCR cycle. Quantitation of product is therefore based on measurements performed during the amplification process, as opposed to the endpoint detection of conventional PCR
  - A fixed fluorescence threshold is set significantly above the baseline fluorescence level and can be altered by the operator

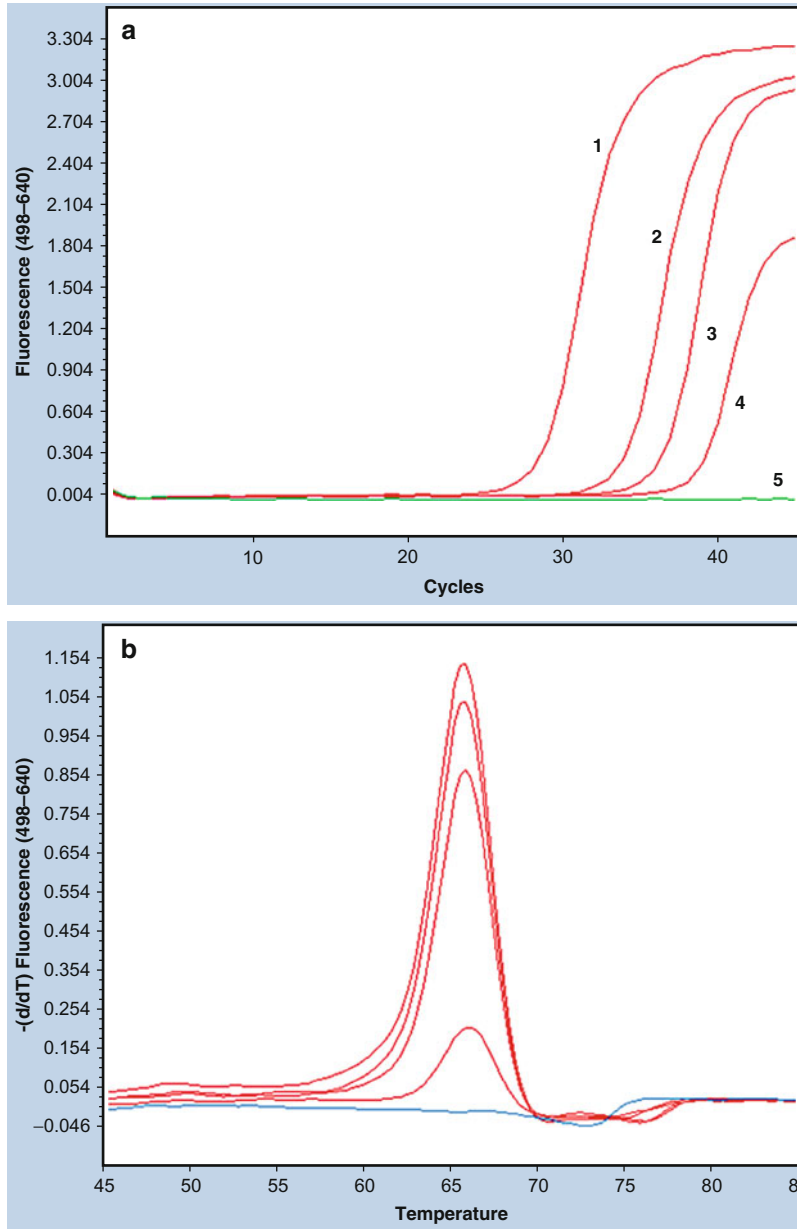


Fig. 8.11 (**a**) Detection of HPV 18 using a real-time PCR assay on the LC 480 instrument. (*1*) 10,000 copies/reaction, (*2*) 1,000 copies/reaction, (*3*) 100 copies/reaction, (*4*) 10 copies/reaction, and (*5*) 0 copies/reaction. (**b**) Fluorescence melting curve analysis

- The parameter C_T (threshold cycle) is defined as the cycle number at which the fluorescence emission exceeds the fixed threshold
- The first cycle at which the level of fluorescence exceeds the threshold correlates to the initial amount of target template. The higher the starting copy number of the nucleic acid target, the sooner a significant increase in fluorescence is observed
- Real-time PCR quantitation eliminates post-PCR processing of PCR products. This helps to increase throughput and reduces the chances of carryover contamination

  - Thus, accurate quantitation can be achieved over a very large range of initial target concentration
  - Real-time PCR can also be used to detect mutations and SNPs, with the help of melting curve analysis (Fig. 8.11b)
- Principle and procedure
  - Real-time PCR is based on the same fundamental principles as conventional PCR; however, the incorporation of fluorescent markers in the reaction mixture permits real-time monitoring of amplification
  - There are two types of probes used in real-time PCR: (1) DNA-binding dyes, such as SYBR® Green (molecular probes), and (2) sequence-specific probes, including TaqMan® (Applied Biosystems), hybridization probes, molecular beacons, and Scorpions® (DxS Ltd.) (see below for detailed information)
- Applications
  - Real-time PCR has a broad application and is replacing conventional PCR due to its advantages
    - Quantitation of gene expression
      - Viral quantitation and pathogen detection
      - Methylation detection
      - Detection of inactivation of X chromosome
      - Monitoring minimal residual disease after hematopoietic stem cell transplantation
      - Genotyping by fluorescence melting curve analysis (FMCA) or high-resolution melting analysis (HRMA)
      - Haplotyping
      - Quantitative microsatellite analysis
      - Prenatal diagnosis/sex determination using single cell isolated from maternal blood or fetal DNA in maternal circulation
      - Intraoperative cancer diagnostics
- Advantages
  - Homogeneous assay: That is, amplification and detection steps occur simultaneously and in the same reaction vessel. Since the PCR products do not need to be manipulated after amplification in order to perform a separate detection step, this limits the potential for contamination
  - Elimination of post-PCR processing saves time and labor costs, facilitating high throughput
  - Measurement of PCR product during exponential phase of amplification (rather than the plateau phase) permits precise quantitation, as during the exponential phase, none of the reaction components are limiting
  - Large dynamic range (7–8 log) allows for very high sensitivity
  - Detection is possible down to a twofold change
  - High degree of reproducibility
  - Multiplexing capability (internal control, single or multiple targets may be amplified and detected simultaneously in one reaction tube using multiple fluorescent probes of different color)
  - Can use integrated quantification standard (obviates need for external standard curve)
- Limitations
  - Special instruments are required to perform real-time PCR
  - Real-time PCR is susceptible to PCR inhibition by compounds present in certain biological samples. To circumvent this problem, alternative DNA polymerases (e.g., Tfl, Pwo, and Tth) that are resistant to particular inhibitors can be used
  - Increased risk of false-negative values without use of an internal standard
  - Increase risk of false-positive values if DNA-binding dye is used (due to nonspecific amplification)

Nested PCR

- General information
  - Nested polymerase chain reaction (nested PCR) is a modification of conventional PCR
  - Two internal primers are used to amplify the PCR products of two external primers
  - Nested PCR is intended to increase fold of amplification and amplification specificity

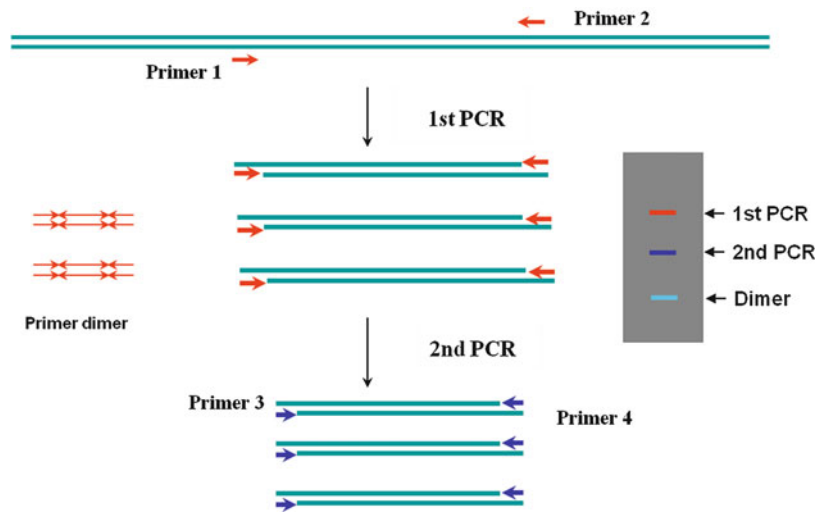


Fig. 8.12 Principle and procedure of nested PCR (see text for description)

- Principle
  - Nested PCR involves two sets of primers and two successive PCR procedures
  - The second primer set functions in the amplification of a secondary target within the first run product
- Procedure (Fig. 8.12)
  - The target DNA undergoes the first round of PCR with the first set of primers
  - The amplicons from the first reaction undergo a second round amplification with the second set of primers
- Advantages
  - Increased sensitivity over conventional PCR
  - Increased specificity due to enrichment for the target sequence with the first round of amplification
- Limitations
  - Because the first-round PCR products need to be transferred prior to the second round, nested PCR is highly susceptible to contamination
  - Cannot be used for quantitative analysis of target

Amplification Refractory Mutation System (ARMS)

- General information
  - ARMS is a well-established, PCR-based method for the detection of SNPs and other genetic variations
  - It is one of the most sensitive and specific method for detection SNP (0.01% mutant allele in background of wild-type allele)
- Principle
  - Multiplex reaction is carried out in the same tube using four primers (Fig. 8.13)
    - Two outer primers to amplify control template
    - One primer specific for wild type which is able to amplify the template DNA with one of the outer primer
    - One primer specific for mutant allele (the terminal 3′-nucleotide of which overlaps the mutation site) which is able to amplify the template DNA with the other outer primer
    - Therefore, three PCR products are expected in heterozygous condition which can be detected by gel electrophoresis
  - Since the primer sets are allele specific, the reaction containing the wild-type primer set is refractory to PCR amplification of mutant template DNA. Conversely, the reaction containing the mutant primer set is refractory to PCR amplification wild-type DNA. By adjusting the primer concentration, mutant allele can be preferably amplified, resulting in increased sensitivity
  - To increase specificity (by preventing nonspecific priming), an additional deliberate mismatch near the 3′ end (in most

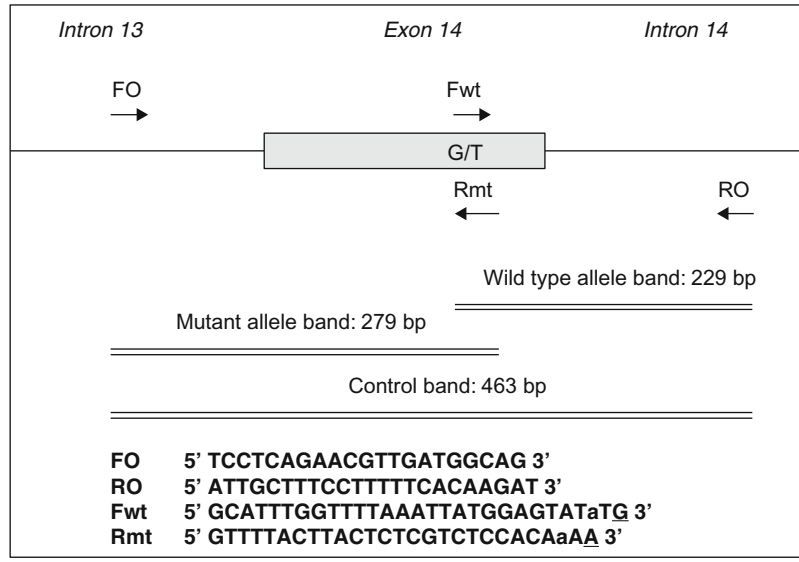


Fig. 8.13 Schematic representation of primer design for an amplification refractory mutation system (ARMS) PCR assay used to detect the Jak2 mutation (1,849 G > T)

situations, two bases way from the 3′ terminus) of the mutation-specific and wild-type-specific primers is introduced
- Applications
  - Genotyping: excellent for SNP detection in the background of high concentration of wide-type allele
- Limitations
  - Optimization is not easy to perform and validate

Coamplification at Lower Denaturation Temperature PCR (COLD-PCR)

- General information
  - COLD-PCR is a modified PCR method that enriches minority variant alleles, such as mutant DNA from a mixture of wild type
  - The DNA can be used for downstream analysis including sequencing, to improve detection sensitivity
- Principle
  - Mismatch dsDNA has slightly altered the melting temperature (Tm) (about 0.2–1.5 °C in ~200 bp)
  - Each dsDNA has a "critical temperature" (Tc) lower than its melting temperature Tm. The PCR amplification efficiency drops measurably below the Tc
  - Two template DNA fragments differing by only one or two nucleotide mismatches will have different amplification efficiencies if the denaturation step of PCR is set to the Tc
- Procedure
  - Denaturation: DNA is denatured at a high temperature – usually 94 °C (Fig. 8.14)
  - Intermediate annealing: Set an intermediate annealing temperature that allows hybridization of mutant and wild-type allele DNA to form mismatch heteroduplex DNA
  - Melting: The heteroduplexes will more readily melt at lower temperatures. Hence, they are selectively denatured at the Tc
  - Primer annealing: The homoduplex DNA will preferentially remain double stranded and not be available for primer annealing, but heteroduplex DNA is readily bound by primers
  - Extension: The DNA polymerase will extend complementary to the template DNA. Since the heteroduplex DNA is used as template, a larger proportion of minor variant DNA will be amplified and be available for subsequent rounds of PCR
  - There are two forms of COLD-PCR
    - Full COLD-PCR utilizes five steps are used for each round of amplification

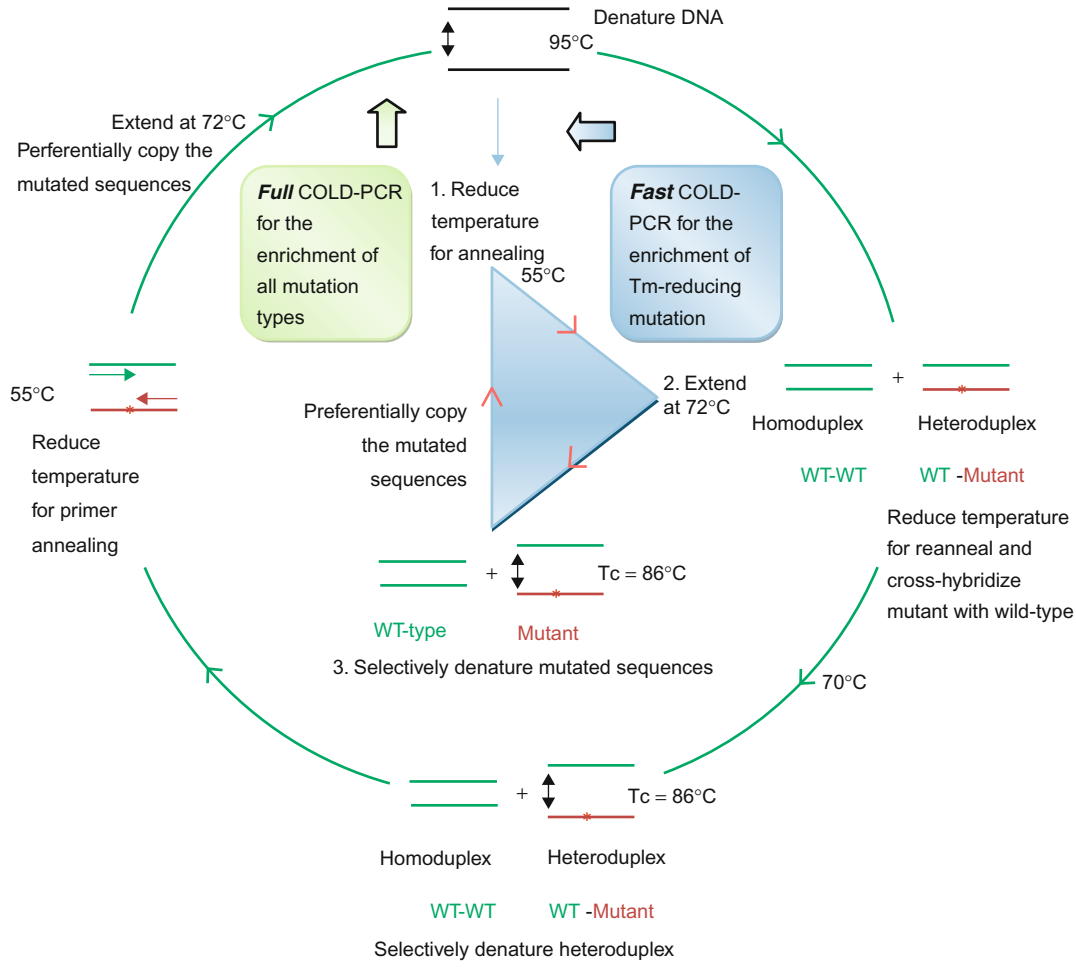


Fig. 8.14 Principle and procedure of coamplification at lower denaturation temperature PCR (COLD-PCR) (see text or description)

  - Fast COLD-PCR differs from full COLD-PCR in that the denaturation and intermediate annealing steps are skipped
- Applications
  - The COLD-PCR products can be used for
    - Restriction fragment length polymorphism results
    - Sequencing
- Advantages
  - Single-step method capable of enriching both known and unknown minority alleles irrespective of mutation type and position
  - Does not require any extra reagents or specialized machinery
  - Better than conventional PCR for the detection of low-level mutations in a mixed sample
  - Compatible with many downstream analysis methods
- Limitations
  - Optimal Tc must be determined for each amplicon
  - Requirement for precise denaturation temperature control during PCR to within ±0.3 °C (0.54 °F)
  - A suitable Tc may not be available that differentiates between mutant and wild-type DNA sequences
  - Restricted to analyzing sequences smaller than approximately 200 bp

- Variable overall mutation enrichment which depends on DNA position and nucleotide substitution
- No guarantee that all low-level mutations will be preferentially enriched
- Adding extra steps increases assay time

8.2.2.3 Strand Displacement Amplification (SDA)

- General information
  - SDA is an isothermal nucleic acid amplification method
  - 10^9 copies of target DNA sequence can be generated in one reaction
  - Requires use of special polymerases (e.g., Bst DNA polymerase or Phi 29 polymerase) and special thiolated dCTP
- Principle and procedure (Fig. 8.15)
  - SDA is a two-step procedure: target creation and exponential target amplification
    - Target creation: Two sets of primers are applied in this step – B1 and S1; B2 and S2
      - S1 and S2 are special primers that contain a restriction enzyme recognition sequence (enzyme BsoBI) 5′ to the target-binding region
      - S1 primed product is displaced and serves as a template for another set of primers, B2 and S2
      - BsoBI cleaves between the first and second nucleotide at the 5′ end of S1 and S2 but cannot cleave between nucleotides joined by a phosphorothioate linkage
    - Exponential target amplification: *Bso*BI enzyme is used in this step to nick the S1- and S2-primed amplicons
      - After nicking, the DNA polymerase binds to this nick and begins the synthesis of a new strand while simultaneously displacing the downstream strand
      - This cycle of alternate nicking and displacement repeats. The displaced strands are capable of binding to opposite strand primers, which produces exponential amplification at 52.5 °C
- Applications
  - The broadest application of SDA is in the diagnosis of infectious diseases
  - The BD ProbeTec ET system is widely used with screening assays for *Chlamydia trachomatis* (CT) and *Neisseria gonorrhoeae* (GC) in urogenital specimens
  - SDA is readily coupled with fluorescence energy transfer (FRET) probes to permit homogenous real-time detection of amplified products
- Advantages
  - A rapid DNA amplification method that can produce target in excess of 10^9 copies in less than 15 min
- Limitations
  - Not completely isothermal; an initial heat denaturation step is required to separate dsDNA
  - Complicated reaction mechanisms and experimental designs

8.2.2.4 Transcription-Mediated Amplification (TMA)

- General information
  - TMA is an isothermal nucleic acid amplification method
  - TMA uses RNA transcription (RNA polymerase) and DNA synthesis (reverse transcriptase) to produce RNA amplicons from a target nucleic acid
  - It can be used to target both RNA and DNA
  - TMA produces 100–1,000 copies per cycle in contrast to PCR that produces only two copies per cycle
  - TMA has been combined with the GEN-PROBE® hybridization protection assay (HPA) detection technique in a single tube format
- Principle
  - TMA is an RNA transcription amplification system using two enzymes, RNA polymerase and reverse transcriptase, and two primers, a promoter primer and a regular primer
    - RNA polymerase mediates transcription, resulting in 100–1,000 copies of

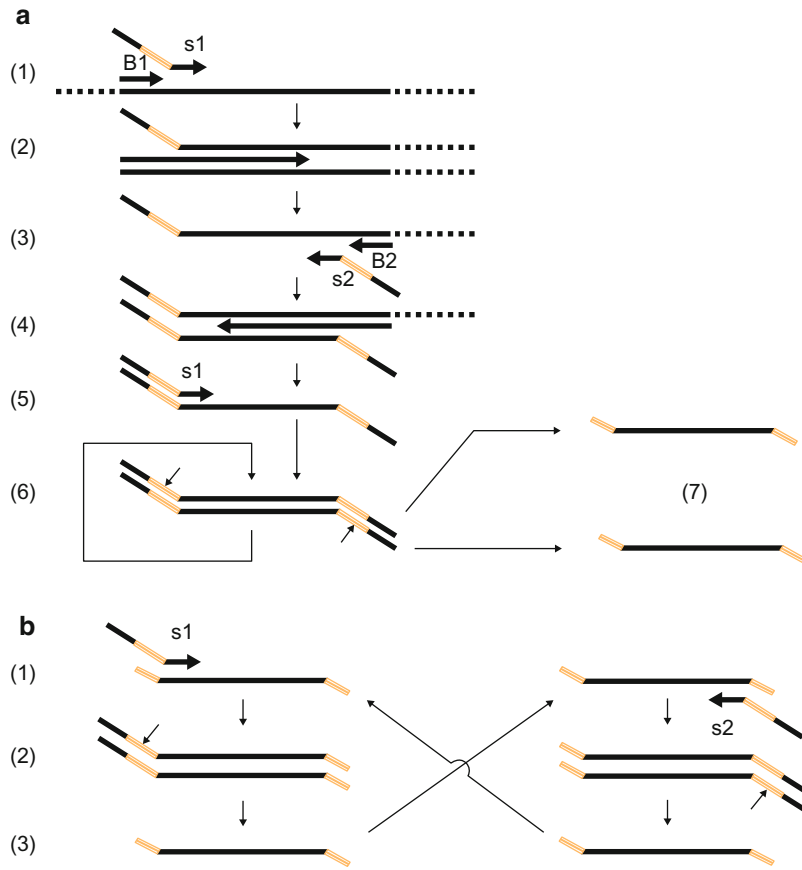


Fig. 8.15 Principle of strand displacement amplification (SDA). (**a**) Target generation: The S1 primer (which contains a restriction enzyme recognition site, indicated by double-line) binds to the target DNA sequence. The B1 "bumper" primer binds upstream to the S1 primer (*1*). Extension from the B1 primer displaces the extension product of the S1 primer (*2*). The S2 primer (which also contains a restriction enzyme recognition site) then binds to the S1 extension product. The B2 "bumper" primer then binds upstream of the S2 primer (*3*). The S2 extension product is in turn displaced by extension from the B2 primer (*4*). Binding of an S1 primer to the S2 extension product (*5*), with subsequent extension, produces a double-stranded molecule with a restriction site at each end (*6*). Since extension incorporates phosphorothioate-modified nucleotides, newly synthesized portions of the strands are resistant to cleavage, and restriction enzymes can only cleave restriction sites that were present in the original S1 or S2 primers. Therefore, the action of the restriction enzyme produces single-strand nicks (designated by *arrows*) in the double-stranded molecule. Extension from the nick sites displaces strands into solution which contain partial restriction sites (*7*); it is these strands that enter the exponential amplification stage of the reaction (see **b**). Meanwhile, the double-stranded molecule can regenerate itself, in order to undergo additional cycles of strand displacement. (**b**) Exponential amplification of the target sequence: single strands produced by the target-generation reaction (in **a**) hybridize with S1 and S2 primers (*1*). Both the primer and its corresponding target strand then undergo extension from their 3′ ends. The resulting double-stranded molecule (*2*) is nicked by the restriction enzyme. 3′ extension from the nicked site releases a single strand into solution (*3*). Binding of an S1 or S2 primer to the displaced strand continues the cycle

RNA amplicon per DNA template. Reverse transcriptase creates a DNA copy of the target RNA

- The promoter primer contains a promoter sequence recognized by RNA polymerase. The other primer works for reverse transcriptase
  - First, the promoter primer binds to the target rRNA and reverse transcriptase creates a DNA copy of the target rRNA. The RNA

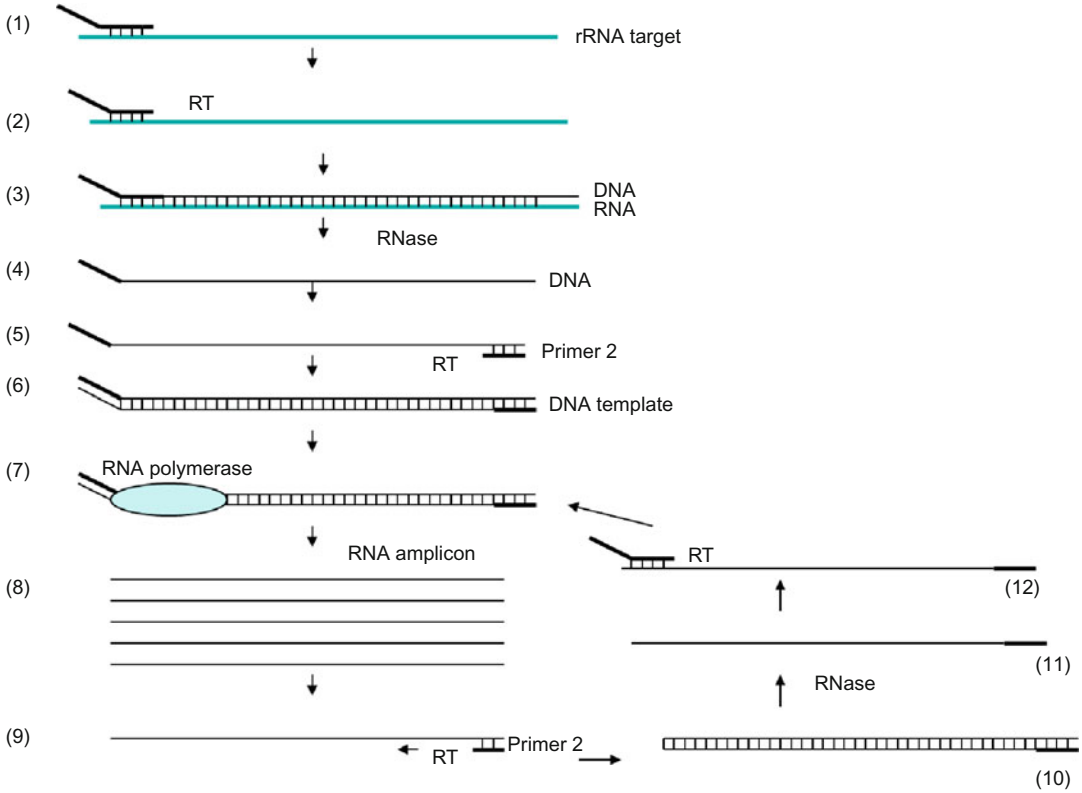


Fig. 8.16 Mechanism of transcription-mediated amplification (TMA)

in the resulting RNA:DNA hybrid is cleaved by the reverse transcriptase
  - The regular primer hybridizes to the DNA copy, and reverse transcriptase produces the second DNA strand. RNA polymerase recognizes the promoter sequence in the DNA template and initiates transcription
  - An exponential amplification is achieved by repeated cycles of reverse transcription and DNA synthesis
- Procedure (Fig. 8.16)
  - Promoter primer binds to rRNA target
  - Reverse Transcriptase (RT) creates a DNA copy of the rRNA target
  - RNA:DNA duplex is formed
  - RNAse H activity of RT degrades the RNA strand
  - Primer 2 binds to the DNA, and RT creates a new DNA copy
  - Double-stranded DNA template with a promoter sequence is formed
  - RNA polymerase (RNA Pol) initiates transcription of RNA from DNA template
  - 100–1,000 copies of RNA amplicon are produced
  - Primer 2 binds to each RNA amplicon, and RT creates a DNA copy
  - RNA:DNA duplex is formed
  - RNAse H activity of RT degrades the RNA strand
  - Promoter primer binds to the newly synthesized DNA. RT creates a double-stranded DNA, and the autocatalytic cycle continues, resulting in exponential amplification
- Applications
  - The amplified *Mycobacterium tuberculosis* direct test (Gen-Probe) for detection of *Mycobacterium tuberculosis* in clinical samples is routinely used in hundreds of laboratories worldwide and provides accurate same day test results

  - Other assays are direct assays for *Chlamydia trachomatis*, Gonococcus, Trichomonas, and HPV
- Advantages
  - Improved reliability by targeting abundant ribosomal RNA. Since rRNA is present in thousands of copies per cell, the likelihood of initiating amplification is greater than when single-copy DNA targets are used. This advantage is very important when organisms are present in low numbers, which is when target amplification methods are most useful
  - Single temperature exponential amplification. The procedure is simple to perform, does not require costly thermocycler equipment, and provides rapid amplification of target sequence present in the sample
  - Primary RNA amplicon. The RNA product of the amplification system is more labile outside the reaction tube than DNA product made by other amplification systems. The risk of laboratory contamination and false-positive results is thus substantially reduced
  - Single tube solution format with no-wash steps. Reagents are only added to the amplification tube and never removed or transferred. This again minimizes the chance of cross contamination and false-positive results. The single tube, no-wash format also allows for the development of relatively simple instrumentation to automate the amplification and detection steps
  - Simplicity. With few reagent additions and HPA detection, the format is user-friendly and familiar to laboratories already using the Gen-Probe DNA probe assays
- Limitations
  - Complicated reaction mechanisms and experimental designs
  - Inefficient in amplification of long target sequences

8.2.2.5 Loop-Mediated Isothermal Amplification (LAMP)

- General information
  - Loop-mediated isothermal amplification (LAMP) is an isothermal DNA amplification method (Fig. 8.17)
  - The LAMP method employs a DNA polymerase and a set of four specially constructed primers
  - The amplification product can be detected by photometry or fluorometry for increasing quantity of magnesium pyrophosphate in solution or with addition of DNA-binding dyes such as SYBR green
- Principle
  - The LAMP method employs a set of four specially constructed primers that recognize six distinct sequences on the target DNA
  - An inner primer with sequences of sense and antisense strands of the target initiates LAMP
  - A pair of outer primers then displaces the amplified strand to release a single-stranded DNA
  - The released ssDNA then forms a hairpin to initiate the starting loop for cyclic amplification
  - *Bst* DNA polymerase is used in this assay which has a high displacement activity
  - Amplification proceeds in cyclical order, each strand being displaced during elongation with the addition of new loops with every cycle
  - The final product is a mixture of stem-loop DNA with various stem length and cauliflower-like structures with multiple loops formed by annealing between alternately inverted repeats of the target sequence in the same strand
- Advantages
  - It is an isothermal amplification and therefore does not need expensive thermocycler
- Limitation
  - Must know the sequence to design the primers
  - Since multiple probes and primers are involved, designing and validation can be a challenge

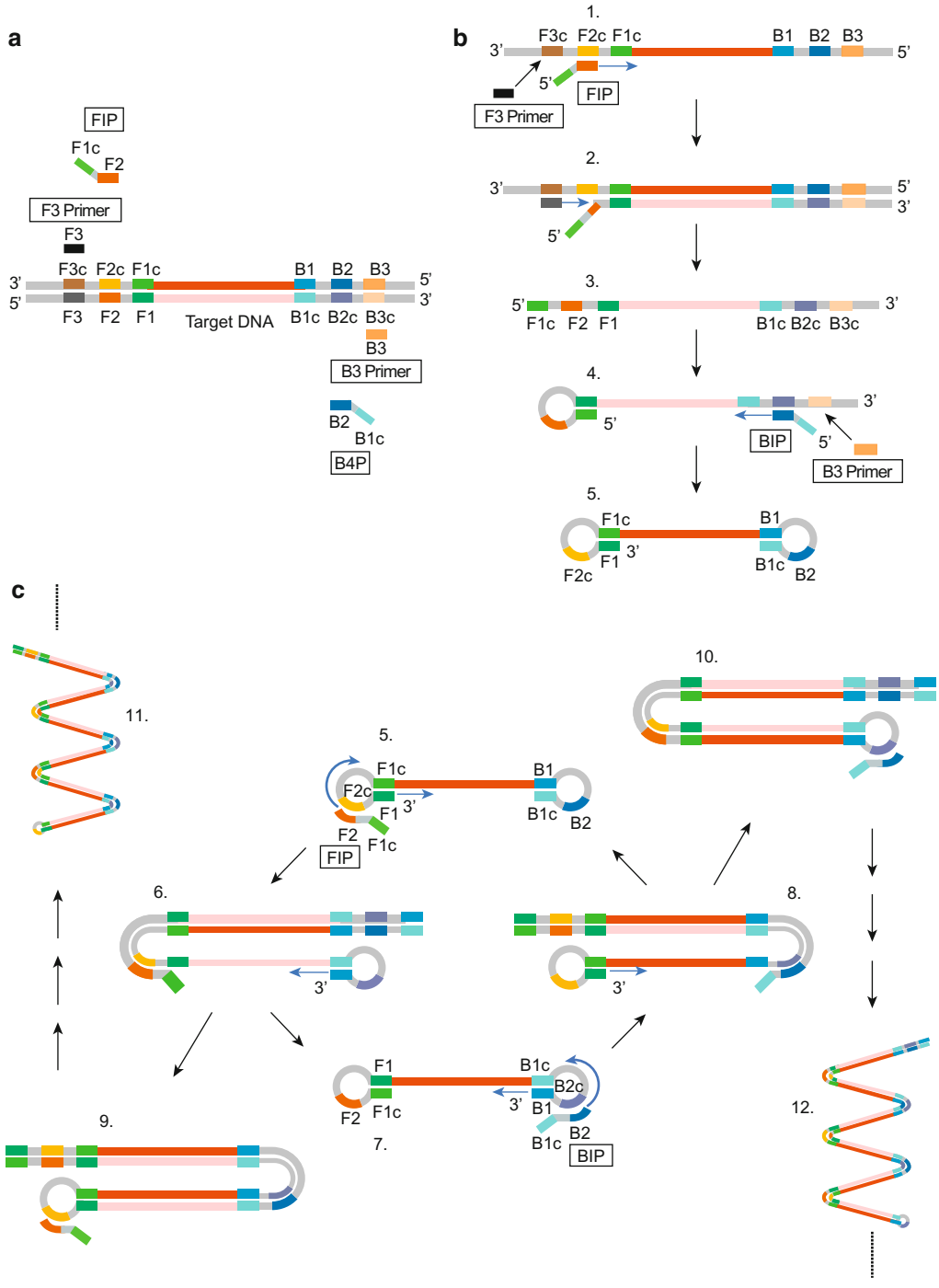


Fig. 8.17 Loop-mediated isothermal amplification (Courtesy of EIKEN CHEMICAL CO., LTD)

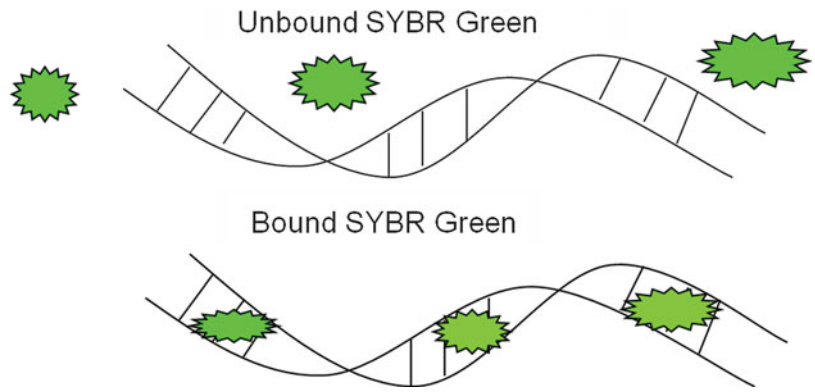


Fig. 8.18 SYBR® Green binding to DNA

- Requires a special DNA polymerase with high strand displacement activity
- May have high background due to nonspecific amplification

8.3 Signal Detection Methods

- Before the application of fluorescence detection methods, gel electrophoresis was the standard method for detection of amplified nucleic acid products (see below). Currently, most real-time amplification detection is via fluorescence. Real-time fluorescence detection of amplification products is rapid, reducing assay time. An additional benefit is that the product can be detected within the reaction tube, markedly reducing the potential for carryover contamination

8.3.1 DNA-Binding Dyes

- DNA-binding dyes such as SYBR® green and LC480 HRM dye are cost-effective and easy to use, especially for researchers who are new to using real-time PCR

8.3.1.1 SYBR® Green

- General information
  - SYBR® green is a double-stranded DNA-binding dye. When free in solution, SYBR® Green I displays relatively low fluorescence, but when bound to double-stranded DNA, its fluorescence increases by over 1,000-fold
  - The greater the quantity of double-stranded DNA, present, the more binding sites are available for the dye; thus, the intensity of the fluorescent signal is proportional to the DNA concentration (Fig. 8.18)
- Application
  - SYBR® Green is often used for initial expression validation screening of microarray samples as well as for other gene expression applications not requiring exceptional sensitivity and specificity
  - Real-time quantification of target DNA
  - Scanning SNPs in target DNA using melting curve analysis
- Advantages
  - Several kits are commercially available for optimization of PCR reactions using SYBR green dye
  - Optimization of primers to use with SYBR® green is straightforward and provides a high level of experimental design success
- Limitations
  - At high concentrations, SYBR® Green inhibits enzymes, including reverse transcriptase and Taq DNA polymerase. Therefore, optimal concentration of SYBR® Green must be experimentally determined in order to balance amplification efficiency and sensitivity with fluorescence intensity

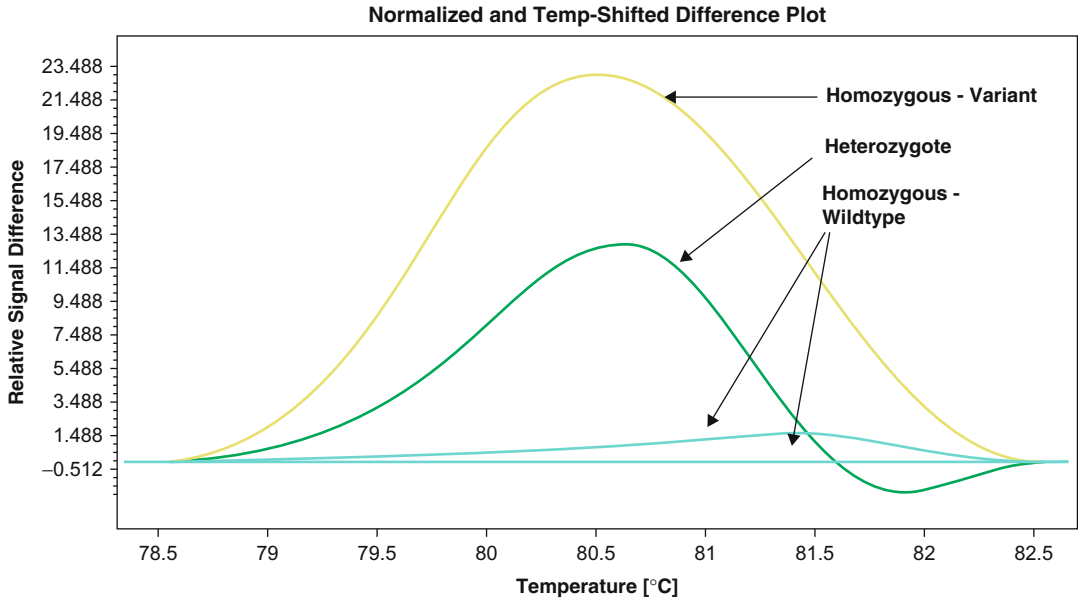


Fig. 8.19 High-resolution melting curve analysis using a LC 480 instrument to detect Kras mutation

- The dye can bind to any double-stranded DNA in the reaction mixture, including primer dimers and other nonspecific reaction products. Therefore, the dye cannot distinguish between specific and nonspecific products accumulated during PCR. However, though nonspecific signal cannot always be prevented, its presence can be easily and reliably detected by performing melting curve analysis on the PCR products

8.3.1.2 High-Resolution Dye (Roche, LC480 HRM Dye)

- General information
  - High-resolution melting (HRM) analysis which is based on detecting small differences in PCR melting (dissociation) curves is a relatively noble, post-PCR analysis method used to identify variations in nucleic acid sequences (Fig. 8.19)
  - It is enabled by improved dsDNA-binding dyes (Roche, LC480 HRM dye, and so on) used in conjunction with real-time PCR instrumentation that has precise temperature ramp control and advanced data capture capabilities
  - The dye enables detection of sequence variations by different melting curve shapes. This feature is not shared with other dyes used in real-time PCR (e.g., SYBR Green I)
  - Samples with variations in DNA sequence are distinguished by discrepancies in melting curve shape. Particularly, heterozygous DNA variants forming mismatched homozygous and wild type because of their different melting behavior
- Application
  - Cost effectively canning mutations including Indel, SNP in target sequence
- Advantages
  - High-resolution dye (Roche, LC480 HRM dye) does not inhibit PCR and thus can be used in high concentrations
  - The characteristics of this dye enable it to demonstrate subtle, but reproducible, changes in the melting curve shape, even with a single-nucleotide change
  - Have the unique ability to detect heteroduplexes during melting analysis after PCR

- Limitations
  - Different heterozygotes may produce melting curves so similar to each other that, although they clearly vary from homozygous variants, they are not differentiated from each other
  - Unable to differential real target from nonspecific amplification if the melting curves are similar

8.3.2 Probe-Based Chemistries

- As compared with nonspecific chemistries such as SYBR® Green, a higher level of detection specificity is provided by use of fluorescent sequence-specific probes to detect PCR products. For the most part, these detection mechanisms rely on the principle of fluorescence resonance energy transfer (FRET), in which a fluorescent dye, in the excited state, can transfer energy to either another (acceptor) fluorophore, with subsequent emission of a fluorescent signal from the receptor molecule, or to a quencher dye which dissipates the energy without emission of a detectable fluorescence signal. In order for the energy transfer to take place, the donor and acceptor molecules must be situated in close physical proximity
- Probes are designed such that in the absence of a specific target sequence in the reaction mixture, the desired fluorescence signal is not produced. However, when the probe (or probes) hybridizes to the target sequence, the level of fluorescence detected is directly related to the amount of amplified target in each PCR cycle. In addition to enhanced specificity, another significant advantage of using sequence-specific probes is that multiple probes can be labeled with different reporter dyes and combined to allow detection of more than one target in a single reaction
- There are two general types of sequence-specific fluorescent probes: linear probes, which include TaqMan probes (hydrolysis probes) and hybridization probes, and structured probes, which include molecular beacons and scorpions
- MGB (minor groove binder) probe is similar to the standard TaqMan probes, but they include the addition of a minor groove-binding moiety on the 3′ end that acts to stabilize annealing to the template. The stabilizing effect that the MGB group has on the Tm of the probe allows for the use a much shorter probe (down to around 13 bp). The shorter probe sequence is more susceptible to the destabilizing effects of single bp mismatches, which makes these probes better than standard TaqMan® probes for applications that require discrimination of targets with high sequence homology

8.3.3 Linear Probes

- The linear probes (TaqMan® probes and hybridization probes) are the most widely used detection methods in real-time PCR

8.3.3.1 TaqMan® Probes

- General information
  - The TaqMan® probe is a linear oligonucleotide with one end labeled with a fluorescent reporter dye and the other end bonded to a quencher. More specifically, a fluorescent dye, typically FAM (reporter), is attached to the 5′ end of the probe, and a quencher, historically TAMRA, is attached at the 3′ end. Increasingly, more effective quenchers such as the Black Hole Quenchers (BHQ) are replacing the use of TAMRA because they provide lower background fluorescence
  - In the absence of gene targets, the reporter and quencher dyes are maintained in close proximity, such that FRET takes place and no fluorescence is detected at the reporter dye's emission wavelength. TaqMan probes use a FRET quenching mechanism where quenching can occur over a fairly long distance (100 Å or more, depending upon the fluorophore and quencher used). Thus, as long as the quencher is on the same oligonucleotide as the fluorophore, quenching will occur

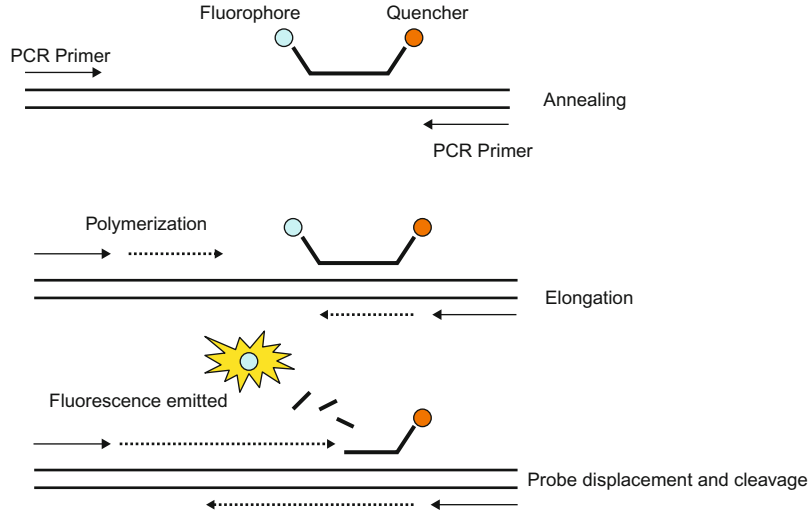


Fig. 8.20 TaqMan probe chemistry mechanism

  - The probe is designed to anneal to one strand of the target sequence just slightly downstream of one of the primers. During PCR, when the polymerase replicates a template on which a TaqMan probe is bound, the 5′-exonuclease activity of the polymerase cleaves the probe into separate single bases, releases free reporter dye into solution, and the fluorescence is detected (Fig. 8.20)
  - It is very important to adjust the thermal profile to facilitate both the hybridization of probe and primers and the cleavage of the probe. To meet these requirements, probes will generally have a two-step thermal profile with a denaturing step (usually at 95 °C) and a combination annealing/extension step at 60 °C, 7–10 °C below the Tm of the probe. If the temperature in the reaction is too high when Taq DNA polymerase extends through the primer (such as at a standard extension temperature of 72 °C), the probe will be strand-displaced rather than cleaved and no increase in fluorescence will be seen
- Application
  - DNA or RNA quantification in biological samples
  - Single-nucleotide polymorphism (SNP) detection or mutation analysis (two probes can be designed, each for different alleles)
- Advantages and limitations
  - It is easy to incorporate into PCR assay
  - Challenge to optimize conditions as to prevent the probes from cross-reacting with the wrong allele
  - Unable to perform melting curve analysis

8.3.3.2 Hybridization Probes

- As opposed to hydrolysis probe methodologies, where only one probe is applied, hybridization probe methodologies utilize two oligonucleotide probes
- One probe is labeled with a donor fluorophore at the 3′ end, and the other is labeled at the 5′ end with an acceptor fluorophore
- The probes have sequence specificities such that when both probes bind to the target sequence, the donor and acceptor fluorophores are brought into close proximity, allowing FRET to occur
- When the donor fluorophore absorbs energy from an excitatory signal, the energy is then transferred to the acceptor fluorophore, which in turn emits light at a specific wavelength that can be measured
- One of the advantages of hybridization probes can be used to analyze the melting curve in addition to monitor PCR reaction

8.3.4 Structured Probes

- Structured probes contain stem-loop structure regions that confer enhanced target specificity when compared to traditional linear probes. This characteristic enables a high level of discrimination between similar sequences. Therefore, structured probes are well suited for SNP and allele discrimination applications. Molecular beacons and scorpions are common types of structured probes

8.3.4.1 Molecular Beacons

- General information
  - Molecular beacons form a stem-loop structure, where the central loop sequence is complementary to the target of interest and the stem arms are complementary to each other. One end (typically 5′) of the stem is labeled with a fluorescence reporter, and the other end is bound to a quencher
  - Instead of using FRET, molecular beacons use ground-state or static quenching, which requires the fluorophore and quencher to be in very close proximity for quenching to occur. Historically DABCYL, or methyl red, has been used for this application
  - In the absence of target, the close proximity of the reporter and quencher prevents the probe from fluorescing. However, when a molecular beacon hybridizes to a target, the probe becomes linear, the fluorescent dye and quencher are separated, and a fluorescence signal is omitted. The probe remains intact during the amplification reaction (Fig. 8.21)
  - Careful design of the molecular beacon stem is critical. If the stem structure is too stable, target hybridization can be inhibited. Conversely, if the molecular beacon probe does not fold in the expected stem-loop conformation, it will not quench properly
- Application
  - Well suited for applications such as SNP detection and allele discrimination
- Advantages and limitations
  - Increased mismatch discrimination due to stable stem structure of the probe
  - Melting curve analysis can be performed

8.3.4.2 Scorpions

- General information
  - Scorpions are bifunctional molecules containing a PCR primer covalently linked to a probe (Fig. 8.22). The fluorophore in the probe interacts with a quencher which reduces fluorescence. During a PCR reaction, the fluorophore and quencher are separated which leads to an increase in light output from the reaction tube
  - There are two formats of scorpions. One is the bimolecular linear scorpion format (Fig. 8.23). The alternative is the unimolecular stem-loop format. The fluorescence of scorpion primer probes is normally quenched. Upon primer-mediated DNA synthesis of the gene targets, the scorpion probes hybridize to the newly formed complementary sequences, separating the fluorescent reporters from the quenchers thus producing fluorescence
  - Sensitivity – the limit of detection is a few molecules even in the presence of very high levels of background DNA
  - Specificity – both the scorpion primer region and probe region can be made sequence specific. This gives unparalleled discrimination allowing the detection of single-nucleotide changes even in admixtures where the alternative sequence is in vast excess
  - Speed – signal generation is exceptionally fast. This means that the technology can support very rapid PCR, allowing detection in less than 10 min
  - Difficult sequences – the intramolecular signal generation mechanism means that scorpions are particularly good at detecting targets in regions of high G + C content or secondary structure

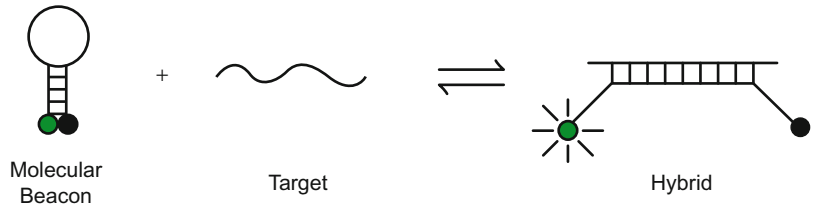


Fig. 8.21 Molecular beacon chemistry mechanism (http://www.stratagene.com/tradeshows/Introduction_to_Quantitative_PCR_web.pdf)

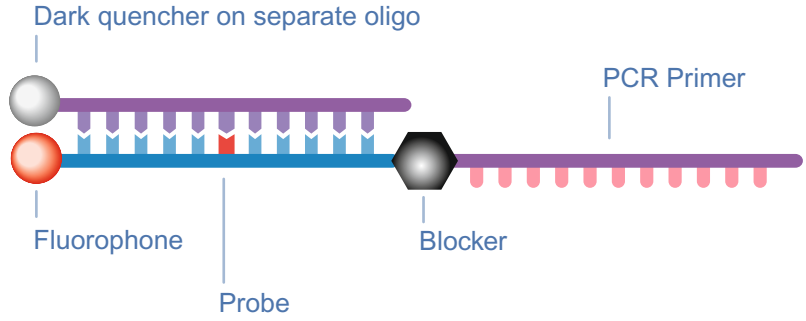


Fig. 8.22 Elements of a scorpion primer (Courtesy of DxS, used with permission)

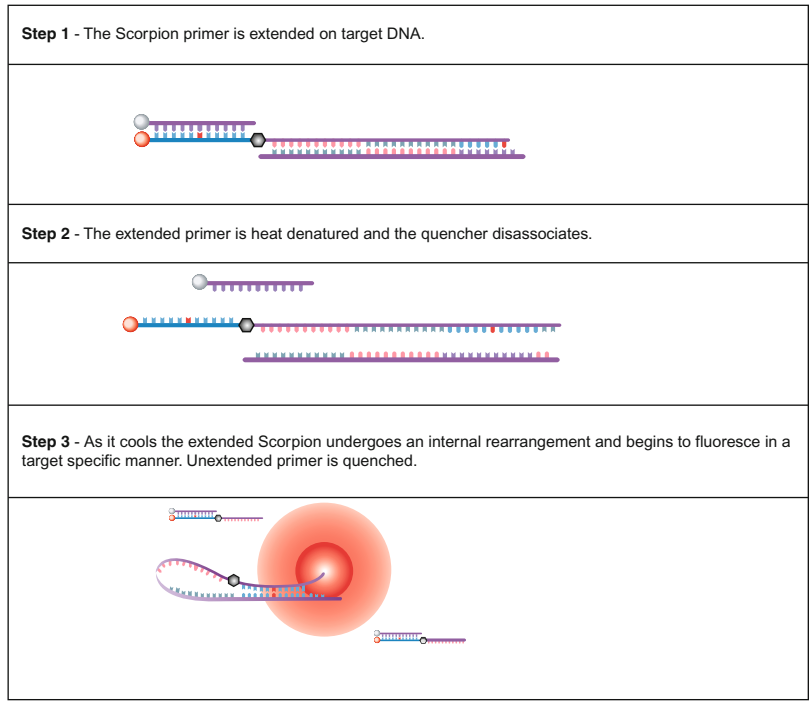


Fig. 8.23 Scorpion primer fluorescence mechanism (Courtesy of DxS, used with permission)

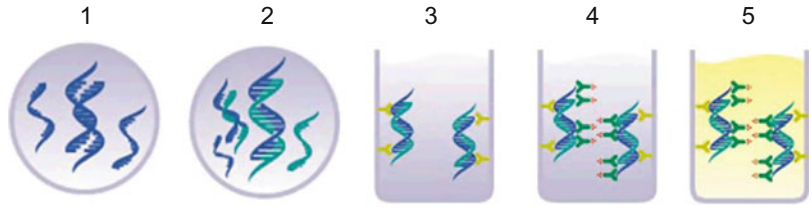


Fig. 8.24 Principle and procedure of hybrid capture hybridization (see text for description) (Adapted from http://www.qiagen.com/hpv/hc2technology.aspx)

8.4 Nonamplification Methods for the Detection of Nucleic Acid Targets

8.4.1 Hybrid Capture (HC)

- General information
  - An in vitro, solution hybridization, signal amplification test for detecting DNA or RNA targets, mainly used in molecular microbiology
  - Probe: target-specific, single-stranded RNA probes (riboprobes)
  - Target: usually the entire genome
- Principle and procedure (Fig. 8.24)
  - Target DNA is isolated
  - Double-stranded DNA is denatured (producing single-stranded DNA)
  - Sample DNA in solution is hybridized with an RNA probe or probe cocktail to form specific DNA–RNA hybrids
  - Hybrids are immobilized by antibodies bound to the wells of a microtiter plate that specifically recognize RNA–DNA hybrids
  - Multiple detector antibodies bind with each immobilized target probe hybrid (first signal amplifier). The detector antibody is a second RNA–DNA-specific antibody, with each antibody molecule conjugated to multiple molecules of alkaline phosphatase (second signal amplifier)
  - After removal of excess antibodies and unhybridized probes, each immobilized alkaline phosphatase enzyme reacts with numerous molecules of dioxetane substrate to produce a chemiluminescent product (third signal amplifier). The resulting light signal is detected by the photomultiplier tube of a luminometer
  - The intensity of emitted light, expressed as relative light units, is proportional to the amount of target DNA present in the specimen, providing a semiquantitative measurement
- Applications
  - The major clinical use is to detect human papillomavirus (HPV) in cervical specimens to determine appropriate followup, primarily in older patients and in patients with ASCUS
    - First introduced by Digene in 1995 to detect 14 HPV types including high-risk types (HPV 16, 18, 31, 33, 35, 45, 51, 52, and 56) and low-risk types (HPV 6, 11, 42, 43, and 44)
      - Can detect viral loads as low as 1 pg of viral DNA per milliliter
      - Has a higher diagnostic sensitivity (exceeding 90%) than cervical cytology
- Advantages
  - No target amplification, no cross contamination like PCR
  - RNA–DNA hybrids more stable than DNA–DNA hybrids
  - Semiquantitative measurement
  - Easy to perform in clinical settings and suitable for automation (hybrid capture II)
- Limitations
  - Lower sensitivity than PCR or other target amplification methods
  - The procedure is laborious, but automation improves throughput
  - For HPV test, high level of retest (RLU in gray zone)

8.4.2 In Situ Hybridization (ISH)

- General information
  - A hybridization technique first introduced independently by Gall and Pardue, Buongiorno-Nardelli and Amaldi, and John et al. in 1969
  - It uses a DNA or RNA probe to detect the target DNA or RNA in the cell cytoplasm or nucleus (i.e., in situ)
  - A signal amplification method, no sequence amplification involved
  - Probes
    - A short DNA or RNA sequence complimentary to the target DNA or mRNA sequence
    - Usually between 20 and 50 bp long (oligonucleotide) or hundreds of nucleotides (DNA probes) but can be as long as thousands of nucleotides
  - Specimens suitable for in situ hybridization test
    - Cells in suspension or fixed on glass slides
    - Frozen-sectioned tissue
    - Paraffin-embedded tissue sections
- Principle
  - Probe (DNA or RNA) undergoes hydrogen bonding (annealing) to complimentary target sequences in fixed or fresh tissue, thus revealing the location and quantity of DNA or RNA
  - Factors that affect hybridization
    - Temperature. The binding of probe to the target DNA or RNA depends upon the Tm or melting point, the temperature at which 50% of double strands are denatured. The optimal temperature for hybridization is just below the Tm
    - Others include pH, monovalent cation concentration, and presence of organic solvents in the hybridization solution
  - Hybridization conditions
    - High-stringency conditions allow highly specific hybridization of the probe with the identical or very similar homology to the target sequence to increase the specificity
    - Low-stringency conditions allow less specific binding of probe with lower homology to the target sequence to increase the sensitivity
- Procedure (Fig. 8.25)
  - Prepare probe. There are four common types of probes. The advantages and disadvantages of those probes are listed in Table 8.2
    - Oligonucleotide probe
      - Synthesized by an automated DNA synthesizer
      - Size: around 40–50 bp
    - Single-stranded DNA probe
      - Produced by PCR or reverse transcription of RNA
      - Size: 200–500 bp
    - Double-stranded DNA probe
      - Produced by bacteria or PCR
      - Size: in the range of hundreds of base pairs
    - RNA probe (cRNA probe or riboprobe)
      - Prepared by an RNA polymerase-catalyzed transcription of mRNA in the 3′ to 5′ prime direction or in vitro transcription of lineralized plasmid DNA with RNA polymerase (T3, T7, or SP6)
      - Size: in the range of hundreds of nucleotides
  - Probe labeling
    - Radiolabeling with ^{3}H, ^{32}P, ^{35}S, or less commonly ^{14}C, ^{125}I: quick, easy, and sensitive but needs long exposure. Due to radio material, it is no longer used routinely in laboratories
    - Fluorescence labeling with FITC, Texas red, or rhodamine among others. The fluorescence is detected with fluorescence microscopy
    - Nonradioactive, nonfluorescence labeling using biotin, digoxin, digoxigenin (DIG), etc. This type of probe labeling needs an intermediate step, such as avidin or antibodies
  - Prehybridization (RNase-free if RNA–DNA or RNA/RNA hybridization).

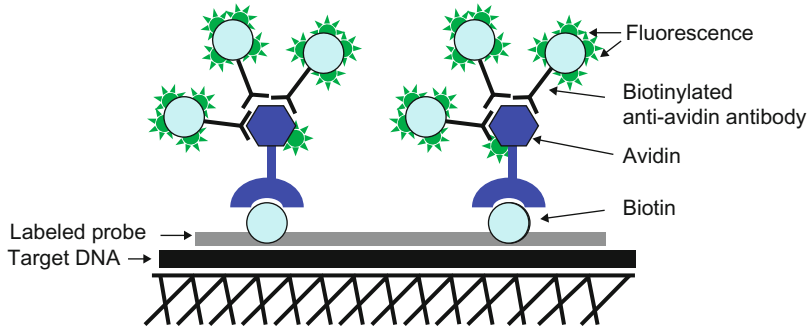


Fig. 8.25 Schematic presentation of in situ hybridization (FISH as an example). Biotin-labeled probe first hybridizes with the target DNA sequence, and fluorescence-labeled avidin subsequently binds to biotin on the probes. A fluorescence-labeled and biotinylated anti-avidin antibody can be used to further amplify signals

Table 8.2 Comparison of probes used for in situ hybridization

Probe type	Advantages	Disadvantages
DNA (double strand)	Easy to use Subcloning unnecessary Choice of labeling methods High specific activity Possibility of signal amplification (networking)	Reannealing during hybridization (decreased probe availability) Probe denaturation required, increasing probe length and decreasing tissue penetration Hybrids less stable than RNA probes
DNA (single strand)	No probe denaturation needed No reannealing during hybridization (single strand)	Technically complex Subcloning required Hybrids less stable than RNA probes
RNA	Stable hybrids (RNA–RNA) High specific activity No probe denaturation needed No reannealing Unhybridized probe enzymatically destroyed, sparing hybrid	Subcloning needed Less tissue penetration
Oligonucleotide	No cloning or molecular biology expertise required Stable Good tissue penetration (small size) Constructed according to recipe from amino acid data No self-hybridization	Limited labeling methods Lower specific activity, so less sensitive Dependent on published sequences Less stable hybrids Access to DNA synthesizer needed

Reproduced from Feldman et al. (1997)

The tissue or cells are fixed with paraformaldehyde, treated with acid (HCl), proteinase K or detergents to increase signal, and H_2O_2 or acetylated to chemically modify proteins and reduce nonspecific binding

– Hybridization (RNase-free if RNA/DNA or RNA/RNA hybridization): Labeled probe binds to the target DNA/RNA sequence in a hybridization solution under a cover slip
– Posthybridization: Rinse with washing solution (SSC) to remove excess or nonspecific bound probes or eliminate single-stranded RNA by adding RNase
– Detection
  • Radioactive detection: Bound radiolabeled probes are directly detected using either photographic film or photographic emulsion

- FISH: Bound fluorescent labels are detected either directly (directly labeled probe) or indirectly (biotin-avidin labeled probe or DIG–anti-DIG antibody labeled probe) by using a fluorescent microscope or plate reader. More than one different probe can be visualized at one time in the same location
- CISH. Bound probes are detected by chromogenic reaction. This usually involves three steps: (1) probes labeled with either biotin or DIG bind to the target sequences, (2) avidin or anti-DIG antibody labeled with peroxidase or alkaline phosphates binds to the probe-target complex, and (3) enzymatic reaction catalyzes the substrates to display the color in situ

– Setting up controls
  - Control for DNA/RNA quality in tissue section or cells to avoid false-negative results
    – Poly(dT) probe to check the quality of mRNA by detecting mRNA poly A tails
    – Probes against constitutively expressed housekeeping genes such as actin or beta-tubulin
  - Positive control (test efficacy)
    – Use of a probe to hybridize with the target sequence in tissue section or cells known to have the target sequence
    – Northern analysis of tissue or cells to show that the labeled probe binds to an mRNA target of the correct molecular size or immunohistochemical detection of the gene product to correlate the coexpression of the protein and mRNA. Not commonly used
  - Negative control (test specificity)
    – Digest the target mRNA with RNase prior to hybridization with the oligonucleotide/DNA probe. RNase treatment will abolish the mRNA in the tissue or cells, and no hybridization shall occur. It will also reduce the background
    – Hybridize labeled sense probe in parallel with labeled antisense probe with the target sequence. Detection of signal only by the antisense probe indicates specific binding
    – Pretreat tissue sections or cells with excess nonlabeled probe (usually 10 times more) before hybridization with labeled probe (competition)
– Interpreting the results should follow general guidelines
  - A positive result should be reported only if the negative control is negative
  - A negative result should be reported only if the positive control is positive
  - If there is discrepancy between the control and test results, the result is "inconclusive" or should be repeated
  - If there are mixed populations of cells (normal and abnormal cells), hybridization results should be correlated with morphology
  - It is always advised to read multiple foci of tissue sections after ISH. Heterogeneity of gene expression in different parts of a lesion may lead to misinterpretation

- Applications
  – To locate DNA sequences or genes on chromosomes (normal, mutations, translocations, deletion, etc.). Useful in genetic testing, diagnosis, monitoring early relapse, and assessing efficacy of therapeutic regimens
  – To study the macroscopic distribution and cellular localization of DNA and RNA sequences in a heterogeneous cell population
  – To quantify the DNA or RNA in situ
  – To study gene expression (mRNA)
  – To detect RNA or viral DNA/RNA
- Advantages
  – Relatively rapid and precise localization of DNA or RNA of interest in anatomic location or specific types of cells
  – Great sensitivity to detect a target present in only a small fraction of the cells in the tissue or mixture

  - Higher specificity than immunohistochemistry
  - Suitability to study the gene expression by detection of mRNA in situ
  - Both qualitative and quantitative
  - Automation and high-throughput analysis possible
- Limitations
  - The sequence of DNA or RNA of interest must be known
  - Probes must be designed for specific applications
  - The sensitivity is relatively lower than that of PCR. A low level of gene expression in a small number of cells can go undetected
  - Significant experience is required to perform these tests and interpret the results
  - Special precautions are required for handling radioactive probes
  - Relatively expensive equipment for FISH

8.4.3 Southern Blot

- General information
  - A hybridization technique developed by Edwin Southern in 1975
  - Used to locate a particular sequence of DNA within a complex mixture or genome and measure its relative quantity
  - Probe: a short segment of single-stranded DNA or oligonucleotide
  - The minimal detectable amount of DNA by Southern blot in general is approximate 0.1 pg under the optimal conditions. However, it is dependent on the size and specificity of the probe
- Principle
  - DNA transferred onto a membrane after agarose gel electrophoresis is hybridized with a DNA probe usually labeled with ^{32}P, or a biotin/streptavidin-enzyme complex
  - The hybridization is based upon the complementarity between the target DNA fragments and probe
  - The degree of hybridization of a complimentary probe is proportional to the amount of the specific target sequence in the sample
  - The signal of autoradiography or amount of enzymatic reaction measured by color or chemiluminescence is used to measure the degree of hybridization
  - The presence of a specific pattern by the above methods is used to detect or rule out a mutation in the target gene or sequence
- Procedure (Fig. 8.26)
  - Isolate and digest genomic DNA or large DNA fragments with restriction enzymes
  - Separate denatured DNA fragments by gel electrophoresis, usually agarose gel, based upon the sizes of DNA fragments
  - Transfer separated DNA fragments onto a charged membrane, usually nitrocellulose or nylon
  - Hybridize labeled probe with the target DNA on the membrane
  - Visualize hybridization using autoradiography (radioactive probe), colorimetric or bioluminescent methods (enzyme or biotin-avidin enzyme complex labeled probes)
- Applications
  - To detect a particular gene and the number of gene copies in the genome (genetic testing, oncogene/tumor suppressor gene detection, etc.)
  - To analyze restriction fragment length polymorphisms and VNTR (forensic DNA testing, transplantation, genetic epidemiology, etc.)
  - To identify the degree of similarity between a gene and probe sequence (molecular detection of intraspecies variations, gene mutation, etc.)
  - To detect gene rearrangements (T cell receptor gene rearrangement, etc.)
  - Used in gene cloning, chromosome walking, genome mapping, etc.
  - Rarely used in the laboratory since high-speed, next-generation sequencing is available

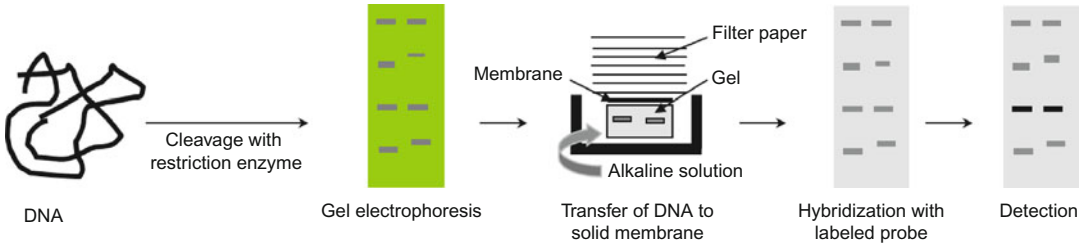


Fig. 8.26 Schematic presentation of Southern blot using a radiolabeled probe (see text for description of steps)

- Advantages
  - Qualitative and quantitative method
  - High analytic sensitivity and specificity
  - Detection of major deletions and rearrangements in large DNA fragments
  - Analysis of multiple samples at the same time
- Limitations
  - Time-consuming and laborious
  - Requires a relatively large amount (microgram scale) of high-quality DNA, usually from fresh or frozen tissue
  - May not detect a clonal population if it represents $<10\%$ of the total cells in the specimen
  - Cannot detect a single or a small number of nucleotide differences in the total genome or in large fragments of DNA

8.4.4 Northern Blot

- General information
  - A hybridization technique developed by Alwine and coworkers in 1977
  - Used to identify RNA (mRNA) and its relative quantity on a membrane using a DNA or RNA probe
- Principle and procedure
  - Similar to those of Southern blotting except the following
  - The target sequence is RNA, not DNA
  - Northern blot uses formaldehyde in the electrophoresis gel as a denaturant. Sodium hydroxide, as used in the Southern blot procedure, would degrade RNA
  - The procedure requires an RNase-free environment. Maintaining the purity and integrity of RNA is essential to performing efficient Northern blots
  - Molecular weight is measured in nucleotides (nt) or kilonucleotides (knt) and occasionally in bases (b) or kilobases (kb)
  - More than one mRNA may be probed after stripping off the bound probes
- Applications
  - Rarely used in research laboratory, not used in clinical diagnosis
  - Detection and quantification of gene expression (mRNAs)
  - Comparison of mRNA abundance among different samples based on the intensity of the signal on a single membrane
  - Determination of transcript size and detection of alternatively spliced transcripts
- Advantages
  - Qualitative and quantitative method
  - Relatively sensitive
  - Exceptionally versatile. Virtually any type of probe can be used
  - Multiple samples can be analyzed at the same time
  - Relatively fast
- Limitations
  - RNase-free environment required
  - Easy degradation of RNA, often fresh or frozen tissue is necessary
  - Less sensitive than nuclease protection assays and RT-PCR
  - Difficult to use multiple probes to detect multiple mRNA transcripts

8.4.5 Allele-Specific Oligonucleotide (ASO) and Sequence-Specific Oligonucleotide (SSO) Hybridization

- General information
  - Allele-specific oligonucleotide (ASO) hybridization was introduced to identify specific alleles by Saiki et al. in 1986. ASO is sometimes mistakenly referred to as sequence-specific oligonucleotide (SSO) hybridization, which is similar to ASO, but used to detect specific sequences instead of alleles
  - ASO and SSO probes are short and specific for particular DNA sequences, typically 15–20 nucleotides long
  - ASO and SSO probes are usually designed such that a single-nucleotide difference between alleles occurs in a central segment of the oligonucleotide sequence, so that the single-nucleotide mismatch maximally enhances thermodynamic instability
- Principle
  - Hybridization between the probe and the target sequence is stable only if there is perfect base complementarity and unstable even if there is a single mismatch between the probe and the target sequence (Fig. 8.27)
- Procedure
  - It is often used in combination with nanobeads (Luminex), dot blot assay (dot blot-ASO assay), Southern blot (see Southern blot), or PCR. Below is the procedure of ASO/SSO-nanobead assay (Luminex)
    - Capture of denatured target DNA onto color-coded beads which labeled with specific capture oligo probes (ASO or SSO probe)
    - The target DNA sequence is labeled with biotin via PCR reaction
    - Addition of fluorescent labeled streptavidin which can bind to biotin
    - Detection of captured DNA-streptavidin complex by flow cytometry
- Applications
  - Genetic testing to discriminate alleles with a single-nucleotide difference (can identify diseased patients and carriers)
  - Determine the allelic difference between individual patients (HLA typing)
  - Pharmacogenetic testing, such as detection of drug resistance alleles due to single-nucleotide polymorphisms (SNPs)
  - Forensic DNA testing based upon multiple single-nucleotide polymorphisms
  - Genetic epidemiology (population genetics and evolution)
- Advantages
  - Suitable for both screening and diagnostic purposes
  - Highly specific
  - Easy, rapid (nanobead-ASO), efficient, and relatively inexpensive
  - Adaptable to automation and high-throughput detection (SNP scanning, BeadArray, etc.)
- Limitations
  - The sequences of the target alleles or loci must be known

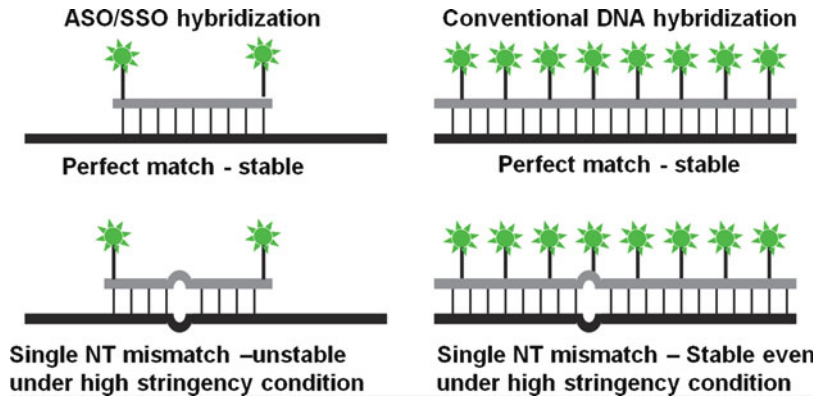


Fig. 8.27 Comparison of hybridization using a conventional DNA probe and an allele-specific/sequence-specific oligonucleotide probe

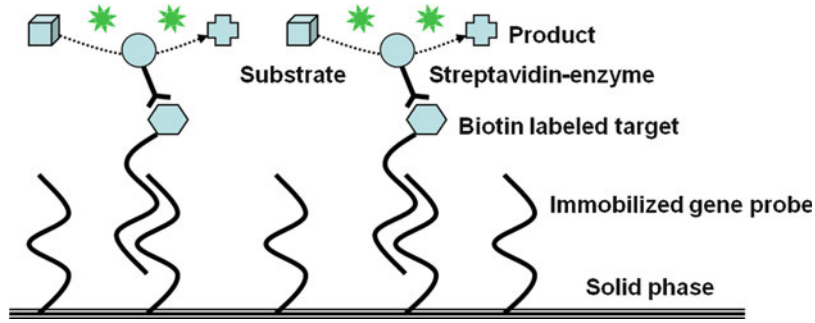


Fig. 8.28 Schematic presentation of reverse hybridization assay using a biotin-avidin-enzyme system

- Only a specific nucleotide difference or mutation is detected. However, many diseases are caused by multiple mutations
- Its sensitivity is relatively lower than that of PCR

8.4.6 Reverse Hybridization

- General information
  - A hybridization technique using the opposite order of the conventional method, that is, the target DNA/RNA binds probes which are fixed on a membrane or solid phase matrix
  - Probes usually are oligonucleotides or single-stranded DNA (or denatured double-stranded DNA)
  - Common types of reverse hybridization techniques
    - Reverse dot blot
    - Reverse Northern blot
    - DNA/oligonucleotide microarray
    - Line probe assays (Lipa)
- Principle and procedure (Fig. 8.28)
  - Similar to other hybridization methods, based upon the complementarity between probe sequence and target DNA or RNA sequence. The major differences from conventional hybridization include
  - Labeling of the target (DNA or RNA) rather than probes. The target sequences may be directly labeled or amplified by PCR before labeling
  - Dotting, blotting, or linking unlabeled probes onto the charged membrane or solid phase matrix
  - Hybridizing labeled target sequences to immobilized probes. Target DNA sequences must be denatured before hybridization
  - Using multiple probes to examine the target sequences simultaneously, thus allowing efficient high-throughput detection and automation
- Applications
  - Large-scale genetic testing, genotyping of viruses, etc.
  - Other uses similar to corresponding hybridization methods
  - Advantages
  - Quick and suitable for automation and high-throughput analysis
- Limitations
  - Similar to corresponding hybridization methods
  - Difficult to correctly identify and distinguish between two closely spaced mutations
  - More costly than the conventional hybridization methods

8.5 DNA Separation Methods

8.5.1 Gel Electrophoresis

8.5.1.1 Agarose Gel Electrophoresis

- General information
  - The commonest and easiest electrophoretic method to separate DNA or RNA by size
  - Agarose acts as a molecular sieve, through which nucleic acids are driven by an electric field

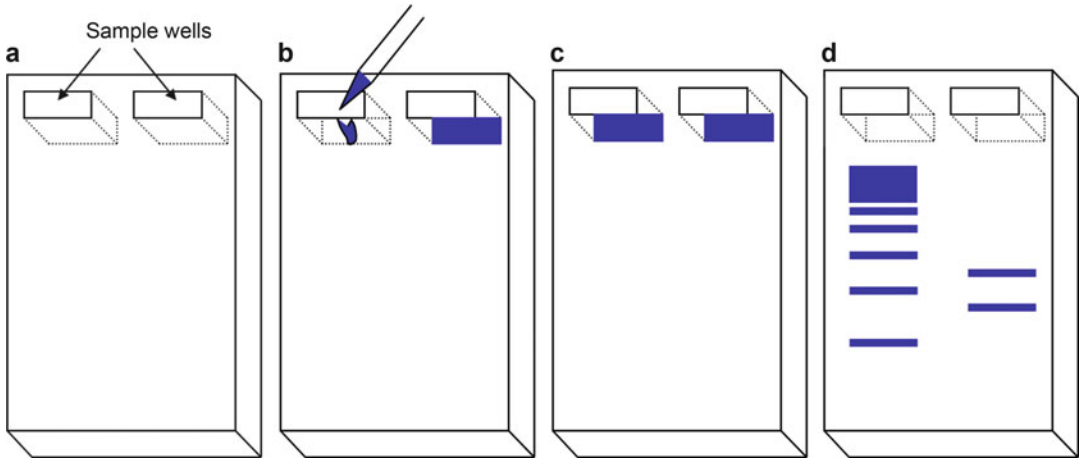


Fig. 8.29 Schematic presentation of agarose gel electrophoresis. Agarose gel is prepared (**a**). DNA markers and samples are loaded into wells (**b**, **c**). Fragments of DNA (or RNA) are separated by electrophoresis (**d**)

  - The pore size is determined by the agarose concentration. The pore size decreases as the agarose concentration increases
  - Most agarose gels are made between 0.7% and 2%. The resolution ranges from 0.2 to 10 kb depending upon the gel concentration
- Principle
  - DNA and RNA molecules are negatively charged at neutral pH due to the presence of phosphate groups. Consequently, these molecules will migrate toward the positive pole when an electrical potential is applied
  - The migration rate of DNA in agarose gel electrophoresis depends upon four main factors
    - The agarose concentration
      - Inversely related to the logarithm of the electrophoretic mobility
      - Lower concentration gels provide better resolution for larger DNA fragments, and vice versa
    - The molecular sizes of the DNA fragments
      - The migration rate of linear duplex DNA is inversely proportional to the logarithm of the DNA fragment size
      - Smaller fragments migrate faster than larger fragments
    - The conformation of the DNA
      - The migration rate of individual DNA fragments differs with their conformations in the agarose gel
      - Among three conformations (closed circular or typically supercoiled DNA, nicked circular DNA, and linear DNA), the supercoiled form migrates the fastest, and the linear form migrates the slowest
    - The applied voltage
      - For a given DNA fragment, the higher the voltage, the faster the migration
      - In general, the voltage applied on an agarose gel is 5 V per cm (length of gel)
  - The DNA is visualized by staining with PicoGreen or ethidium bromide (EB). Both dyes bind strongly to DNA by intercalating between the bases and emit visible fluorescent or orange light when they absorb invisible UV light
  - The molecular size of an unknown band of linear DNA is estimated by comparison of its traveling distance with that of molecular weight standards. Supercoiled DNA must be linearized before it can be evaluated for molecular weight
- Procedure (Fig. 8.29)
  - Prepare the agarose gel by mixing agarose powder in buffer solution, then heating, and

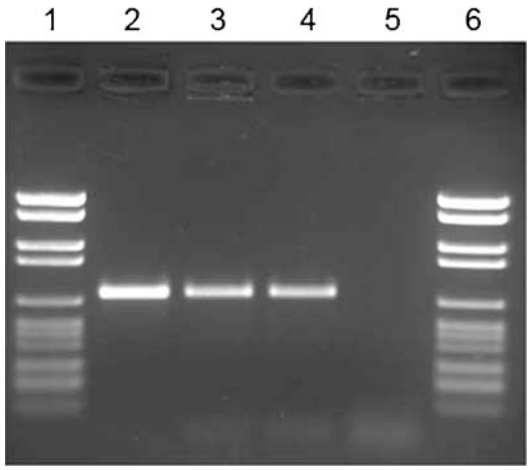


Fig. 8.30 An example of agarose gel electrophoresis of DNA fragments. Lanes 1 and 6: molecular weight markers; lanes 2 and 3: two patient samples; lane 4: positive control; lane 5: negative control (Adapted from Richter et al. 2002)

pouring the liquefied gel into the electrophoresis apparatus and allowing it to solidify
  - Load DNA ladder into the gel. The DNA ladder (molecular weight markers) is pretreated with loading buffer containing a marker dye
  - Mix DNA samples with loading buffer and load gel
  - Start electrophoresis by applying electric field
  - Nucleic acid fragments will separate in the gel on the basis of size
  - Stop electrophoresis when the marker dye approaches the end of the gel
  - Stain DNA bands in the gel with PicoGreen or EB. Optionally, EB can also be added into electrophoresis buffer to stain the DNA during electrophoresis
  - Visualize the stained DNA under fluorescent or ultraviolet light (Fig. 8.30)
  - Cut the band(s) of interest in the gel and retrieve or purify the DNA if needed
- Applications
  - The commonest analytic method to determine the presence, the size, and the amount of DNA fragments
  - Preparative method for the isolation or purification of a particular DNA species
  - The method that can also be used for RNA separation
- Advantages
  - Versatile (wide range of separation)
  - Easy to use, fast separation, relatively inexpensive, and nontoxic
- Limitations
  - Only useful in separation of relatively large fragments of DNA/RNA due to large pore size of gels
  - Lower resolving power than PAGE (see below)
  - Low sensitivity

8.5.1.2 Polyacrylamide Gel Electrophoresis (PAGE)

- General information
  - An electrophoretic technique introduced by Raymond and Weintraub in 1959
  - Uses polyacrylamide gel as a molecular sieve to separate DNA when an electric field is applied
- Principle
  - Similar to agarose gel electrophoresis. Unique features include
    - Polyacrylamide gels are formed from the polymerization of two compounds, acrylamide and a cross-linking agent, N, N_1-methylene-bis-acrylamide (or bis)
    - The size of the pore of polyacrylamide is determined by the gel concentration and usually ranges from 3% to 30% of acrylamide and bis. The separation of molecules within a gel is determined by the relative sizes of the pores formed within the gel
    - The resolution of PAGE ranges from 10 to 2,000 bp and is dependent upon the gel concentration
- Procedure
  - Prepare polyacrylamide gel. Acrylamide and bis solution is mixed in the proper proportion. Ammonium persulfate along with either β-dimethyl amino-propionitrile (DMAP) or N,N,N_1,N_1,-tetramethylethylenediamine

(TEMED) is added to the mixture for polymerization
  - Pour the gel into the plate
  - Load the sample and DNA ladder with marker dye (bromophenol blue). If necessary, DNA can be denatured by heating to allow analysis of confirmation change of ssDNA or heteroduplex
  - Separate DNA fragments by electrophoresis
  - Stop electrophoresis when marker dye approaches the bottom of the gel
  - Stain the gel with ethidium bromide
  - Illuminate the gel with ultraviolet light
  - Alternately, if ^{35}S or ^{33}P nuclides were incorporated into the DNA during synthesis, autoradiography can be performed after the gel is dried
- Applications
  - Primarily used for DNA sequencing
  - Analysis of small fragments of DNA (quantification, size determination)
  - Small-scale preparation
  - Analysis of DNA confirmation change for SNP (ssDNA conformational change) or clonality (T cell receptor gene rearrangement)
- Advantages
  - Ability to separate smaller DNA fragments than agarose gel, attributed to finer porosity
  - Very high resolution; separation of strands with single base pair difference, under optimal conditions
  - Stronger than agarose gel
  - Dried gel can be used for permanent archival record
- Limitations
  - Small range of separation
  - More difficult to use and more expensive than agarose gels
  - Acrylamide is neurotoxic; appropriate precautions must be taken

8.5.1.3 Capillary Electrophoresis

- General information
  - Capillary electrophoresis (CE), introduced by Hjerten in 1983, is a family of separation technologies with wide-ranging applications in chemistry and the biological sciences. Separation is achieved by the flow of analytes through a narrow capillary, propelled directly or indirectly by a strong electric field
- Principles
  - Capillaries are very narrow (normally ranging from 25 to 100 μm in internal diameter and 0.5–1.0 m in length) and are usually comprised of fused silica
  - Capillaries may be filled with either buffer solution or replaceable liquid polymers
  - Depending on the precise technique employed, separation occurs on the basis of charge-to-mass ratio, size, or other physical properties of the analyte
  - In fused silica capillaries, cations from the buffer solution are attracted to negatively charged silanoate groups that line the inner capillary wall, producing a layer of positively charged mobile ions. These ions are drawn toward the negative pole (cathode) along with associated water molecules, producing a bulk flow of buffer which carries all analytes (regardless of charge) toward the cathode, a phenomenon known as electroosmotic flow or EOF
  - Capillary gel electrophoresis (CGE) is a variant of CE which has broad applicability in the analysis of nucleic acids and is particularly useful in DNA sequencing
  - In CGE, the capillary is filled with solution-phase polymers which act as a molecular sieve, resolving DNA fragments on the basis of their sizes
  - The gel matrix used in CGE includes cross-linked polyacrylamide and noncross-linked matrix linear polymers such as polyacrylamide, polyvinyl alcohol, dextran, and agarose
  - In CGE, it is necessary to minimize EOF so that the polymer is not drawn out of the capillary during the run; this is achieved by coating the capillary walls

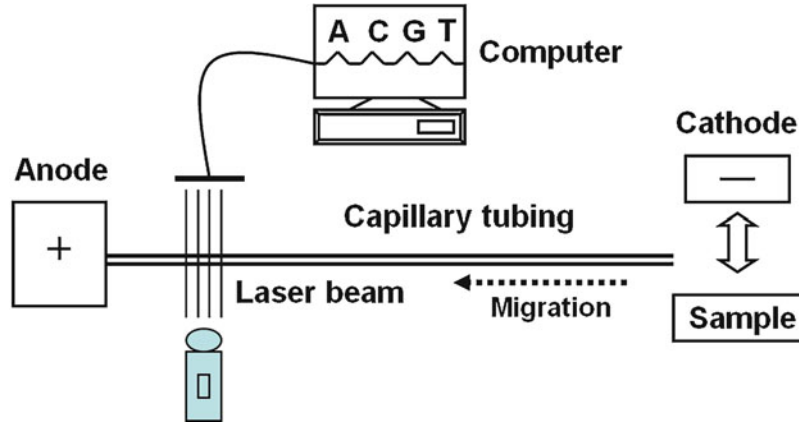


Fig. 8.31 Schematic presentation of CE as used in DNA sequencing (see text for description of steps). A light (usually laser) is used to elicit a fluorescence signal that is captured by a detector

  - In the absence of EOF, nucleic acids, which are negatively charged, migrate toward the anode. Smaller fragments, which are less encumbered by the gel matrix, move more quickly than larger fragments
  - Detection is usually fluorescence-based (e.g., by incorporation of dye-labeled terminators in sequencing reactions)
  - The liquid gel matrix is replaced after each run, minimizing the potential for carryover contamination
- Procedure (CGE) (Fig. 8.31)
  - DNA fragments are prepared (isolation, synthesis, purification, and/or enzyme digestion)
  - The DNA sample is introduced into the capillary at the cathodic (negative) end either by application of electrical current or mechanical pressure
  - DNA fragments are resolved as they migrate through the matrix
  - Resolved DNA fragments are detected near the anode (positive pole) of the capillary. Detection methods include fluorescence, absorbance, electrochemical, and mass spectrometry
  - Data is analyzed and electropherogram is produced
- Applications
  - Separation of dsDNA: the most frequent application of CE. It can separate dsDNA up to 40 kbp in a homogenous electric field using a mixture of polyethylene oxide (PEO)
  - Separation of DNA sequencing fragments: It can separate DNA fragments up to 1,000 bp under certain conditions
  - Genotyping, allele analysis, or restriction fragment length polymorphism detection
  - Combination with other DNA separation methods
- Advantages
  - Highly efficient and quick separation of a large number of DNA fragments at the same time
  - High resolution
  - Small capillary diameter permits rapid dissipation of heat
  - A small sample volume (pico to nanoliter) sufficient
  - Automated and high-throughput applications are possible
  - Quantitative and qualitative detection
  - Replaceable gel or matrix
  - Minimal quantities of reagents required
  - Low cost of individual runs, short analysis time, and simplicity
- Limitations
  - Cannot perform separations at preparative scales
  - High concentration of DNA in a small volume required
  - Relative low sensitivity compared to HPLC
  - Initial high cost of instrumentation

- Laborious and low throughput
- Interpretation of electropherogram sometimes can be challenging

8.5.1.4 Gradient Gel Electrophoresis (GGE) and Denaturing Gradient Gel Electrophoresis (DGGE)

- General information
  - A variant of gel electrophoretic technique first described by Fischer and Lerman in 1983
  - It uses a gradient gel with pore size decreasing from the top to the bottom to increase the resolution. The gradient gel is almost invariably polyacrylamide gel
  - The concentration of gel typically varies from 5% at the top to 25% at the bottom
  - Denaturation is accomplished by heat, or heat in conjunction with a chemical or pH gradient. As the DNA migrates, fragments with minor, even single-nucleotide, changes can be separated. This technique is called denaturing gradient gel electrophoresis or DGGE (see Fig. 8.32 for a schematic presentation)
  - In practice, the denaturants used are heat (a constant temperature of 60 °C) and a fixed ratio of formamide (ranging from 0% to 40%) and urea (ranging from 0 to 7 M)
  - PCR coupled with DGGE permits the separation of almost all DNA variants. A GC-rich sequence is incorporated into one of the primers used for enzymatic amplification, producing PCR products with altered melting points
- Principle (Fig. 8.32)
  - The migration rate of the small DNA fragments (100–700 bp) through a low to high conventional or denaturant gradient acrylamide gel initially is determined by the molecular size
  - With heightening denaturing conditions, individual DNA fragments reach a point where the duplex DNA fragment is melted
  - The partial melting severely retards the migration of the DNA fragments in the gel, and a mobility shift is observed

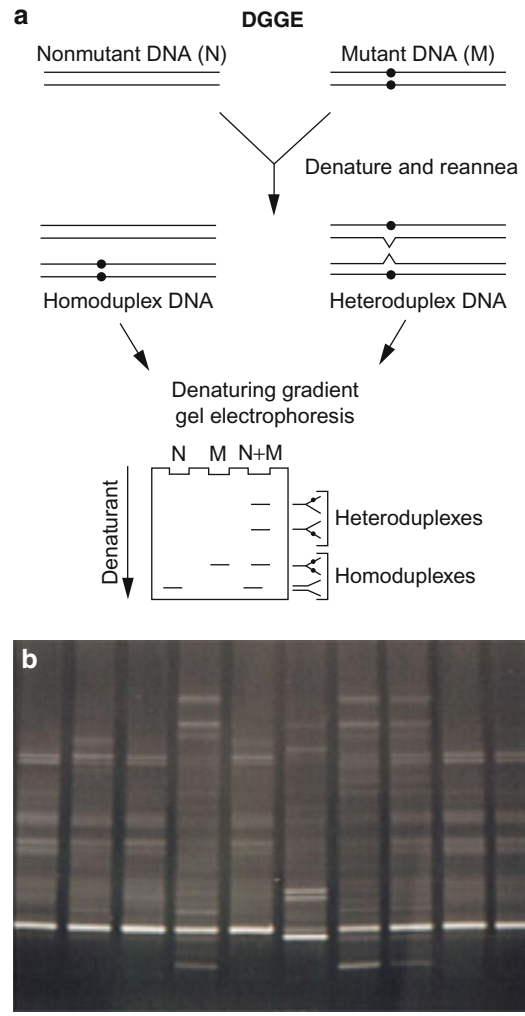


Fig. 8.32 Mechanism of denaturing gradient gel electrophoresis (DGGE) and an example of its application (analysis of sequence variation in the second internal transcribed spacer (ITS-2) of ribosomal DNA of strongyloid nematodes). Schematic representation of heteroduplex formation and principle of DGGE (**a**). Mutant homoduplexes (lane M) melt at a lower denaturant concentration than nonmutant homoduplexes (lane N) and are consequently retarded at a higher position in the gel. Heteroduplexes (lane N + M) melt at even lower denaturant concentrations. DGGE analysis of sequence variation within ITS-2 PCR products (350 bp in size including the GC clamp) amplified from single adults of the nematode Haemonchus contortus originating from a natural population. (**b**) Each bright (homoduplex) band in DGGE represents an enriched clonal sequence type of ITS-2 as determined by direct sequencing of excised bands. Bands in the upper third of the gel represent heteroduplex molecules produced during PCR (Reprinted from Gasser et al., Parasitology Today 1999;15:462–465 with permission from Elsevier)

  - As the denaturing conditions become more extreme, the partially melted fragment completely dissociates into single strands
  - The denatured DNA is essentially stopped (arrested) at the melting point in the gel and forms a sharp band
  - Minimal differences between DNA fragments (as little as a single base pair change) can result in a significant mobility shift
  - The final mobility of DNA in DGGE depends upon the molecular size and composition of each fragment
- Procedure
  - Prepare gradient gel with the aid of a gradient maker at room temperature
  - Place the solidified gel into a tank containing gel buffer at a temperature of 60 °C, with recirculation pump turned on
  - Load the samples
  - Run electrophoresis at an appropriate voltage (usually 65–75 V)
  - Stop electrophoresis and stain the gel with ethidium bromide
  - Examine the gel by UV illumination
  - Transfer the DNA onto nylon blots if necessary
- Applications
  - To type alleles of a polymorphism (e.g., alleles of the ABO blood group system)
  - To detect non-RFLP polymorphisms
  - To screen exons of a mutated gene in human genetics and oncology
  - To be used in molecular microbiology such as identification of microorganisms and epidemiological studies
- Advantages
  - High sensitivity DGGE can detect virtually all mutations in a given piece of DNA. It displays the highest detection rate among mutation scanning methods
  - High reproducibility in experienced hands
  - Rapid and relatively inexpensive
  - Possibility of optimizing the analysis by computer simulation
  - A nonradioactive protocol
- Limitations
  - The denaturant gradient slope and running times vary for every DNA region to be analyzed, greatly affecting the routine application of the method
  - It requires great personal experience in the preparation of denaturants and pouring of acrylamide gels to yield reproducible results
  - The long running time is needed for resolving heteroduplexes and homoduplexes, and often produce curtains and smears instead of sharp zones
  - Without DNA sequencing, heteroduplexes and homoduplexes that comigrate in the gel can confound DGGE interpretations

8.5.1.5 Pulsed-Field Gel Electrophoresis (PFGE)

- General information
  - An electrophoresis technique introduced by Schwartz and Cantor in 1984
  - It uses alternating electrical fields to separate especially long strands of DNA by their sizes in a low-density agarose gel matrix
  - It raises the upper size limit of DNA separation in agarose from 20 kb to above 10 Mb (10,000 kb)
- Principle
  - DNA molecules greater than 30–50 kb migrate with the same mobility regardless of size during continuous field electrophoresis; thus, they are seen in the gel as a single large diffuse band. Multiple fragments of DNA in such a band can be separated from each other if an electric field is applied to change direction of electrophoresis. With each reorientation of the electric field relative to the gel, smaller-sized DNA will begin to move in a new direction more quickly than the larger DNA, thereby leaving the larger DNA fragments behind. This provides a separation of long strands of DNA based upon their sizes
  - PFGE instrumentation generally falls into two categories (Fig. 8.33)
    - Field inversion gel electrophoresis (FIGE) works by periodically inverting the polarity of the electrodes during electrophoresis

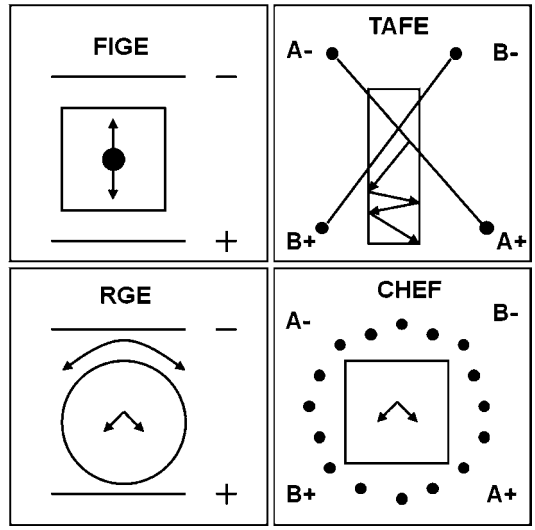


Fig. 8.33 Electrode configurations of commonly used pulsed-field gel electrophoresis units

  - The other type of instrumentation functions to reorient the DNA at smaller oblique angle, to move the DNA fragments forward in a zigzag pattern down the gel. This type of instrumentation includes transverse alternating field gel electrophoresis (TAFE), contour-clamped homogeneous electric field (CHEF), and rotating gel electrophoresis (RGE)
- PFGE special equipment consists of a gel box, a chiller and pump, a switch unit, a programmable high-voltage power supply, and computer software
- Procedure
  - Prepare intact or unsheared DNA by embedding intact cells in agarose plugs and digesting away the proteins in the plugs using enzymes to avoid shearing of large DNA fragments
  - Digest intact DNA in the plugs with a rare-cutting restriction endonuclease
  - Separate DNA fragments by gel electrophoresis using PFGE special equipment
  - Detect and interpret the banding patterns
- Applications (Fig. 8.34)
  - Identifying restriction fragment length polymorphisms (RFLPs) and fingerprinting
  - Determining the number and size of chromosomes (electrophoretic karyotype) from fungi and parasites
  - Detecting in vivo chromosome breakage and degradation
  - Epidemiological studies, for example, to establish the degree of relatedness among different strains of the same species
  - The use of this method has been largely replaced by next-generation sequencing
- Advantages
  - Ability to separate and characterize large DNA fragments, for example, chromosomes of microorganisms
  - Ability to examine the elongated and oriented configuration of large DNA molecules in agarose gels at finite field strengths
  - High sensitivity and reproducibility. It is one of the most reliable techniques for determining strain genetic similarity in many bacteria
- Limitations
  - Intact DNA is required. Special care must be taken not to shear or damage the DNA
  - Faulty results may arise due to (1) easy contamination of the agarose plugs with DNases by accidental introduction of nonspecific DNA-degrading enzymes and (2) incomplete (partial) digestion of DNA

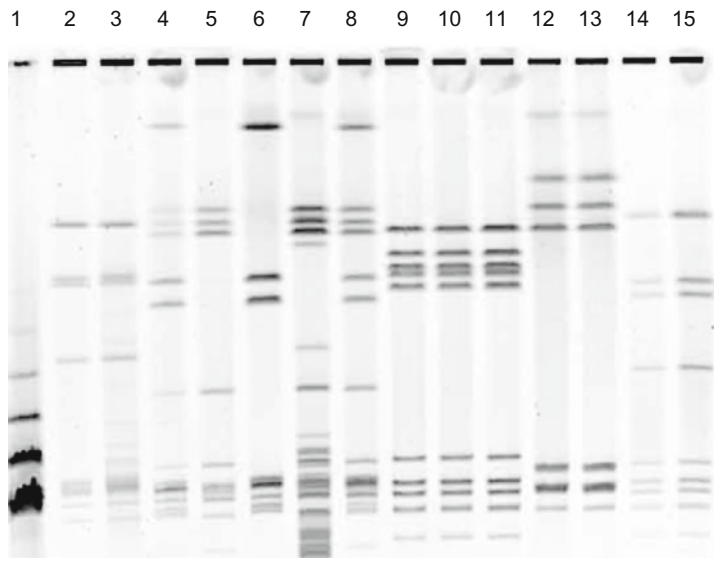


Fig. 8.34 Pulsed-field gel electrophoresis (PFGE) of group B Streptococcus (GBS) strains isolated from college women with urinary tract infections and their most recent sex partner. Lane 1 is a DNA ladder. Lanes 2–3 and 14–15 represent rectal and vaginal GBS isolates from two different women. Lanes 4–6 are two urine isolates and one rectal isolate from one woman, and lanes 7 and 8 are the urine and rectal isolate from her sex partner, respectively. Lanes 9–11 represent a female vaginal isolate and her sex partner's urine and rectal isolate (Adapted from Manning 2003)

by the restriction enzyme that generates unusually large fragments
  - Methodology is complicated, time-consuming
  - High cost

8.5.1.6 Single-Strand Conformation Polymorphism (SSCP)

- General information
  - An electrophoretic technique introduced by Sunnucks et al. in 1989
  - Separates single-stranded DNA or RNA based on mutation-related conformational differences
- Principle (Fig. 8.35)
  - Single-stranded DNA molecules assume unique conformations that depend on their nucleotide sequences under nondenaturing conditions and reduced temperature. Subtle differences in sequence, often a single base pair variation, may cause a different secondary structure by altering intrastrand base pairing, resulting in a measurable difference in electrophoretic mobility. Small or tightly packed molecules migrate more quickly through the gel than large or loosely packed molecules
  - Most SSCP protocols are designed to analyze the polymorphism at a single locus using a specific pair of PCR primers bracketing the target region. The target single-stranded DNA is amplified by asymmetric PCR, in which one primer is present in excess over the other. After the low-concentration primer supply is exhausted, continued PCR only allows amplification of the target single-stranded DNA
- Procedure
  - Prepare DNA fragments by digestion of genomic DNA with restriction endonucleases and/or PCR amplification of the target single-stranded DNA fragment or PCR amplification of the dsDNA followed by denaturation into single strands
  - Denature DNA samples in an alkaline (basic) solution

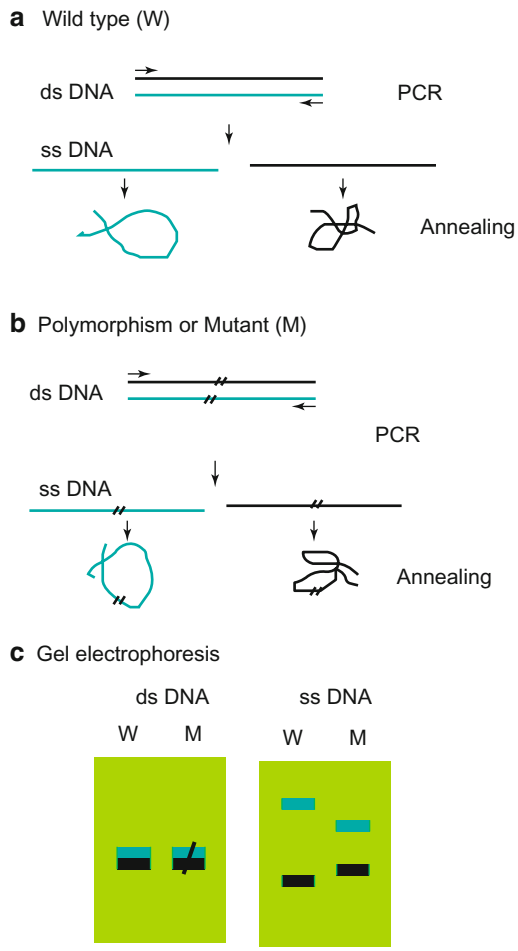


Fig. 8.35 SSCP principle. Wild type (**a**) and polymorphism or mutant (**b**) double-stranded DNAs have equal length, and their corresponding single-stranded fragments show different conformations. The banding result after gel electrophoresis of both double-stranded and single-stranded DNA fragments is shown (**c**). The migration patterns of the double-stranded fragments are indistinguishable. However, three single-stranded fragments display different migration patterns, or mobilities, according to their conformations

  - Separate DNA fragments by neutral polyacrylamide gel electrophoresis
  - Stain the gel or transfer DNA onto a nylon membrane followed by hybridization with probes
  - Compare the mobility of control with unknown DNA sample fragments
  - Confirm identity of mutations by DNA sequencing

- Advantages
  - Simple procedure
  - Inexpensive equipment
  - No need for precise knowledge of the sequence polymorphism
- Limitations
  - Limited size range of DNA fragments (150–300 bp) for optimal separation results. Adding certain reagents such as glycerol to the gel may increase the size limit. RNA–SSCP electrophoresis may allow for separation of larger-sized fragments
  - Limited sensitivity. Under optimal conditions, approximately 80–90% of the potential base exchanges are detectable by SSCP. Changing the pH and adding glycerol may increase sensitivity
  - No information provided for the position of the change
  - A constant temperature is required during the electrophoresis for best results, because single-stranded DNA mobilities are temperature dependent
- Applications
  - Detection of DNA polymorphisms and mutations at multiple sites in DNA fragments
  - Serving as genetic markers because they are allelic variants similar to restriction fragment length polymorphisms (RFLPs)
  - Use in molecular genetics with modifications, such as heteroduplex analysis, ribonuclease protection assay, and SNP (single-nucleotide polymorphism) techniques

8.5.1.7 Restriction Fragment Length Polymorphism (RFLP) Analysis

- General information
  - RFLP analysis is a methodology that combines restriction enzymatic digestion and conventional gel electrophoresis to separate and analyze the DNA fragments. The resolved DNA fragments can either be visualized with ethidium bromide or by Southern blot

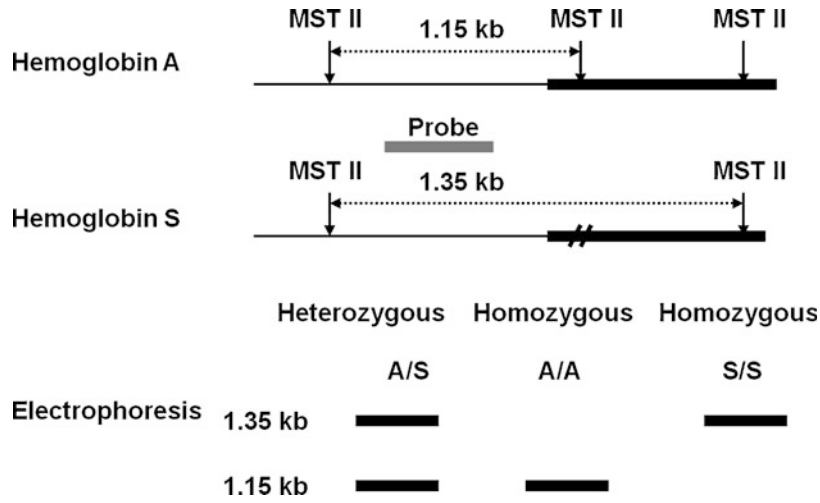


Fig. 8.36 Schematic presentation of RFLP (an example of detecting a missing site (for Mst II) in hemoglobin S). Hemoglobin S produces a longer DNA fragment (1.35 kb) than hemoglobin A (1.15 kb) after Mst II restriction enzyme digestion due to mutation at an Mst II cleavage site (*upper panels*). Gel electrophoresis can separate two types of hemoglobin genes based upon their sizes (*lower panel*)

  - Differences in the size and/or number of restriction fragments result from sequence changes (base substitutions, additions, deletions, etc.) that involve restriction enzyme recognition site(s)
  - The term "RFLP" also refers to these sequence changes (polymorphisms) themselves
  - Special types of RFLPs include (a) minisatellites, detecting hypervariable RFLP loci (variable number tandem repeat loci) and (b) restriction landmark genomic scanning, two-dimensional RFLP analysis
- Principle
  - Each restriction enzyme recognizes a specific sequence (4–6 bp or occasionally 8 bp in length). Any change of this sequence (e.g., by mutation, insertion, or deletion) results in loss of the splice site due to lack of recognition by the enzyme
  - Digestion of a DNA sample by restriction enzymes produces a collection of DNA fragments of precisely defined length (restriction fragments) which can be resolved by gel electrophoresis. RFLPs can represent normal variation or be associated with disease
- Procedure (Fig. 8.36)
  - Isolate and purify genomic DNA with or without PCR amplification
  - Digest DNA with restriction enzyme(s)
  - Separate DNA fragments according to their sizes by gel electrophoresis
  - Detect DNA bands by staining the gel or blotting DNA onto a membrane followed by hybridization using a labeled probe (Southern blot)
- Applications
  - Genetic testing or screening of human DNA for the presence of potentially deleterious mutations
  - Forensic DNA testing or fingerprinting
  - Genetic epidemiology or population studies
  - Molecular microbiology to identify related species
  - This method is largely replaced by direct sequencing or high-resolution melting analysis
- Advantages
  - No prior sequence information required
  - Inexpensive
- Limitations
  - RFLP sites recognized by enzymes must be present in association with known mutations
  - The informativeness is limited. Polymorphisms are only detected if they affect restriction sites
  - It requires a large amount of DNA with high integrity

Table 8.3 Comparison of selected sequencing technologies

Name	Technology principle	Read length (bp)	Throughput (Mb/h)
454 Pyrosequencing	Pyrophosphate fluorescence/emulsion PCR	200–300	13
Solid ABI	Ligation sequencing/emulsion PCR	35	21–28
Illumina	Bridge amplification/bridge PCR	32–40	25

  - Sensitivity is variable
  - Difficult to standardize
  - Time-consuming and laborious

8.6 Sequencing of Nucleic Acids

- DNA sequencing is the process of determining the nucleotide order of a given DNA fragment
  - Currently, almost all DNA sequencing is performed using the chain-termination method developed by Frederick Sanger. This technique uses sequence-specific termination of DNA synthesis reactions using modified nucleotide substrates
- Another method was developed by Maxam and Gilbert, hence called Maxam and Gilbert sequencing
  - Base-specific chemical modification followed by cleavage of modified DNA
  - This method is no longer used for routine DNA sequencing because the reagents are toxic and the procedure is not amenable to automation
  - However, it is still valuable in special applications such as assaying protein and DNA interactions known as "footprinting"
- Other technologies, such as pyrosequencing, SNaPshot, and Invader, are also being used for sequencing and SNP (single-nucleotide polymorphism) detection
- More recently, newer techniques, named "next-generation sequencing," have been developed which allow for an automated means of generating thousands of reactions to produce a much more elaborate amount of data. A selected number of techniques are described below
- A comparison of selected sequencing technologies can be seen in Table 8.3

8.6.1 Sanger Sequencing

- General information
  - "Classic" method of sequencing using chain termination and four types of deoxynucleotides (dATP, dCTP, dGTP, and dTTP)
  - Utilizes four separate DNA polymerase reactions for each of the deoxynucleotides
- Principle
  - Extension is initiated at a specific site on the template DNA by using a short oligonucleotide primer complementary to the template at that region
  - Primer is extended using a DNA polymerase and the four aforementioned deoxynucleotides along with a low concentration of a chain-terminating nucleotide (most commonly a dideoxynucleotide, ddNTP)
  - Limited incorporation of the chain-terminating nucleotide by the DNA polymerase results in a series of related DNA fragments that are terminated only at positions where that particular nucleotide is used
  - Four separate DNA polymerase reactions, each using a different one of the four dideoxynucleotides, are carried out
  - The products of these reactions are then size separated by electrophoresis in separate lanes of a slab polyacrylamide gel, or more commonly now, by capillary electrophoresis
- Procedures
  - The DNA sequencing reactions are somewhat analogous to PCR (see Sect. 8.2)
    - The reaction mix includes the template DNA, free nucleotides, an enzyme (usually a variant of *Taq* polymerase) and a "primer" – a short piece of

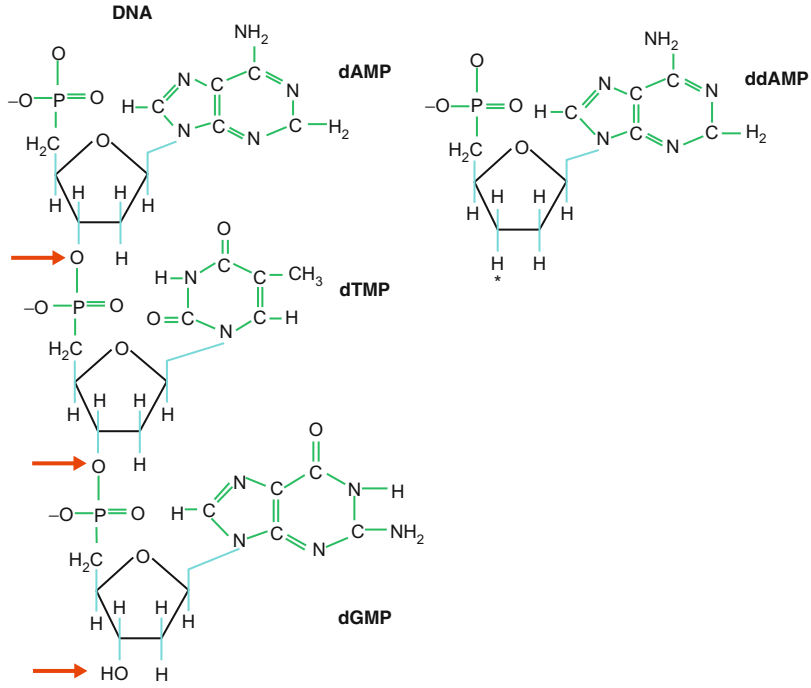


Fig. 8.37 Comparison of normal nucleotide structure with the structure of a chain-terminating dideoxynucleotide. *Asterisk* indicates the absence of a hydroxyl group at the 3′ position of the dideoxynucleotide that prevents formation of DNA. *Red arrow* highlights hydroxyl group position

single-stranded DNA about 20–30 nt long that can hybridize to one strand of the template DNA
- For detection purposes, the primers can be labeled with radioactive isotope S32
- DNA needs to be purified to remove unincorporated nucleotides and primers which can interfere with the sequencing reaction and results (see below)

– The Sanger technique utilizes 2′, 3′-dideoxynucleotide triphospates (ddNTPs), molecules that differ from deoxynucleotides by having a hydrogen atom attached to the 3′ carbon rather than an OH group (Fig. 8.37)
- These molecules terminate DNA chain elongation because they cannot form a phosphodiester bond with the next deoxynucleotide

– The sequencing reaction is conventionally performed in four separate tubes using a different ddNTP for each reaction
- A mixture of a particular ddNTP with its normal counterpart, and the other three dNTPs
  – The concentration of ddCTP should be 1% of the concentration of dCTP
  – DNA polymerization will take place and will terminate whenever a ddCTP is incorporated into the growing strand
- If the ddNTP is only 1% of the total concentration of dNTP, a whole series of labeled strands will result (Fig. 8.38)
- Note that the lengths of these strands are dependent on the location of the base relative to the 5′ end
  – These reactions can be done as PCR to improve the sensitivity and specificity

– When these reactions are completed, polyacrylamide gel electrophoresis (PAGE) is performed. The products of each reaction are loaded into separate lanes for a total of four lanes (Fig. 8.39)
- The DNA is transferred to a nitrocellulose filter. Autoradiography is performed so that only bands containing DNA with radioactive label will appear

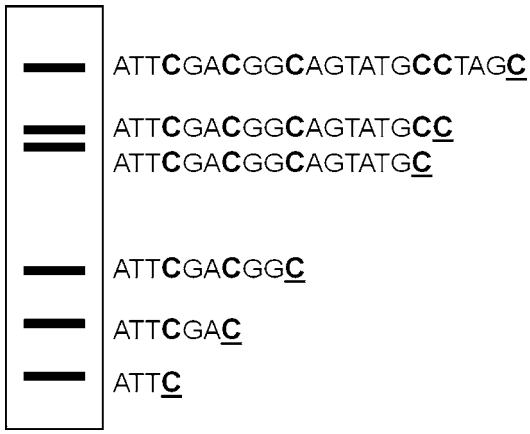


Fig. 8.38 Schematic diagram of one lane of a sequencing gel containing products of a sequencing reaction containing ddCTP. The position of the chain-terminating dideoxynucleotide in each strand is underlined

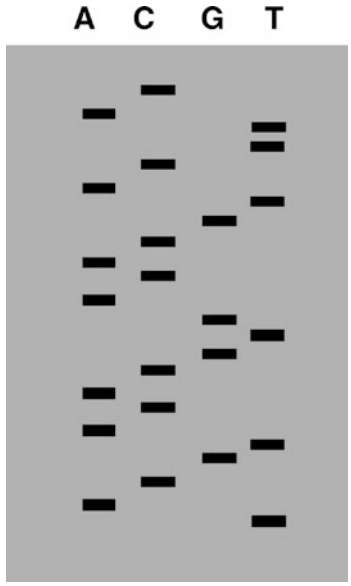


Fig. 8.39 Conventional sequencing gel. Sequence output: 5′ TACGTACACGTGACACGTACTTAC 3′

- In PAGE, the shortest fragments will migrate the farthest
- The bottommost band indicates the particular dideoxynucleotide that was added first to the labeled primer
- In Fig. 8.33, for example, the product that migrated the farthest was from the ddTTP reaction mixture
  - Thus, ddTTP must have been added first to the primer, and its complementary base, thymine, must have been the base present on the 3′ end of the sequenced strand
  - The sequenced strand can be read 5′ to 3′ by reading top to bottom the bases complementary to those on the gel
- Applications
  - Sanger sequencing is used a number of applications, but not in high-throughput testing
  - This method is largely replaced by dye terminator method
- Advantages
  - Sanger sequencing allows for analysis of very small changes to the target in question
- Limitations
  - More cumbersome, cannot do high-throughput testing
  - More expensive and time-consuming

8.6.2 Dye Terminator Sequencing

- General information
  - An alternative method which using labeled terminators instead, commonly called “dye terminator sequencing”
  - Principle
  - Similar to above-described conventional Sanger sequencing, however, fluorescence-labeled nucleotides are used instead of radioisotope
  - The four fluorescence-labeled ddNTP are used, therefore, the reaction can be carried out in a single tube
  - To increase efficiency and signal, cyclic extension of primer is performed as in PCR reaction using thermostable polymerase
- Procedure
  - Remove free PCR primers using SAP (shrimp alkaline phosphatase) and Exo (exonuclease)
  - Set up big dye terminator reaction using terminator ready reaction mix (Applied

Biosystems) which contains fluorescence-labeled ddNTP
  - Cycle primer extension (usually 25 cycles) using a PCR instrument
  - Remove ddNTP using Agencourt CleanSEQ
  - Separate the products using capillary or slab gel electrophoresis
  - Laser detection is used to identify the bases at each position
  - The sequence is "read" from the bottom up, using a key for the established colors for each probe
  - Using software provided by the manufacturers of sequencing machines, the signal/noise ratios of the dyes are determined for each position so that the proper base can be "called."
  - The order of the bases is displayed in a format known as a "chromatogram," "electropherogram," or "trace" file (Fig. 8.40)
- Applications
  - This method is now used for the vast majority of sequencing reactions as it is both simpler and cheaper
  - Advantages
  - Quicker and less technically demanding than Sanger sequencing
  - The complete sequencing set can be performed in a single reaction, rather than the four needed with the labeled primer approach. This is accomplished by labeling each of the dideoxynucleotide chain terminators with a differently colored fluorescent dye
- Limitations
  - This method is easier and quicker than the dye primer approach but may produce more uneven data peaks (different heights), due to a template dependent difference in the incorporation of the large dye chain terminators. However, this problem significantly reduced with the introduction of new enzymes and dyes that minimize incorporation variability
  - The context of a given nucleotide has a significant influence on peak height (Fig. 8.41). This feature may be useful in the interpretation of a trace, since artifactual peaks have no effect on the heights of adjacent peaks
  - Common artifacts in dye terminator sequencing traces include "dye blobs" (Fig. 8.42) and large peaks that occur if bubbles are present in the capillary (Fig. 8.43)
  - Usually can only read out 500–700 nucleotides accurately. Therefore, one needs to break the DNA into several fragments for sequencing if the sequence is too long
  - Limited sensitivity for SNP detection if the SNP is present in less than 25% of the sample DNA
  - Misreading by DNA polymerase may occur. Therefore, reading from both sides is necessary

8.6.3 The SNaPshot Method (Applied Biosystems, ABI)

- General information
  - This method is a modified form of Sanger sequencing used to detect single-nucleotide polymorphisms (SNPs)
- Principle
  - The SNaPshot primer targets a sequence immediately upstream of the SNP site and is extended by a single base in the presence of all four fluorescently labeled dideoxynucleotides (ddNTPs)
  - Each fluorescent ddNTP emits a different wavelength, which is translated into a specific color for each base. The size of the product is the size of the initial probe plus one fluorescent base
  - The reactions are run on an ABI 3700, and genotypes are determined by the color and location of the peak that is generated from the emitted fluorescence
  - Data are then analyzed with the ABI Gene Scan software package using size standards for verification of the peaks. Primer design and DNA template purification can significantly affect genotyping accuracy

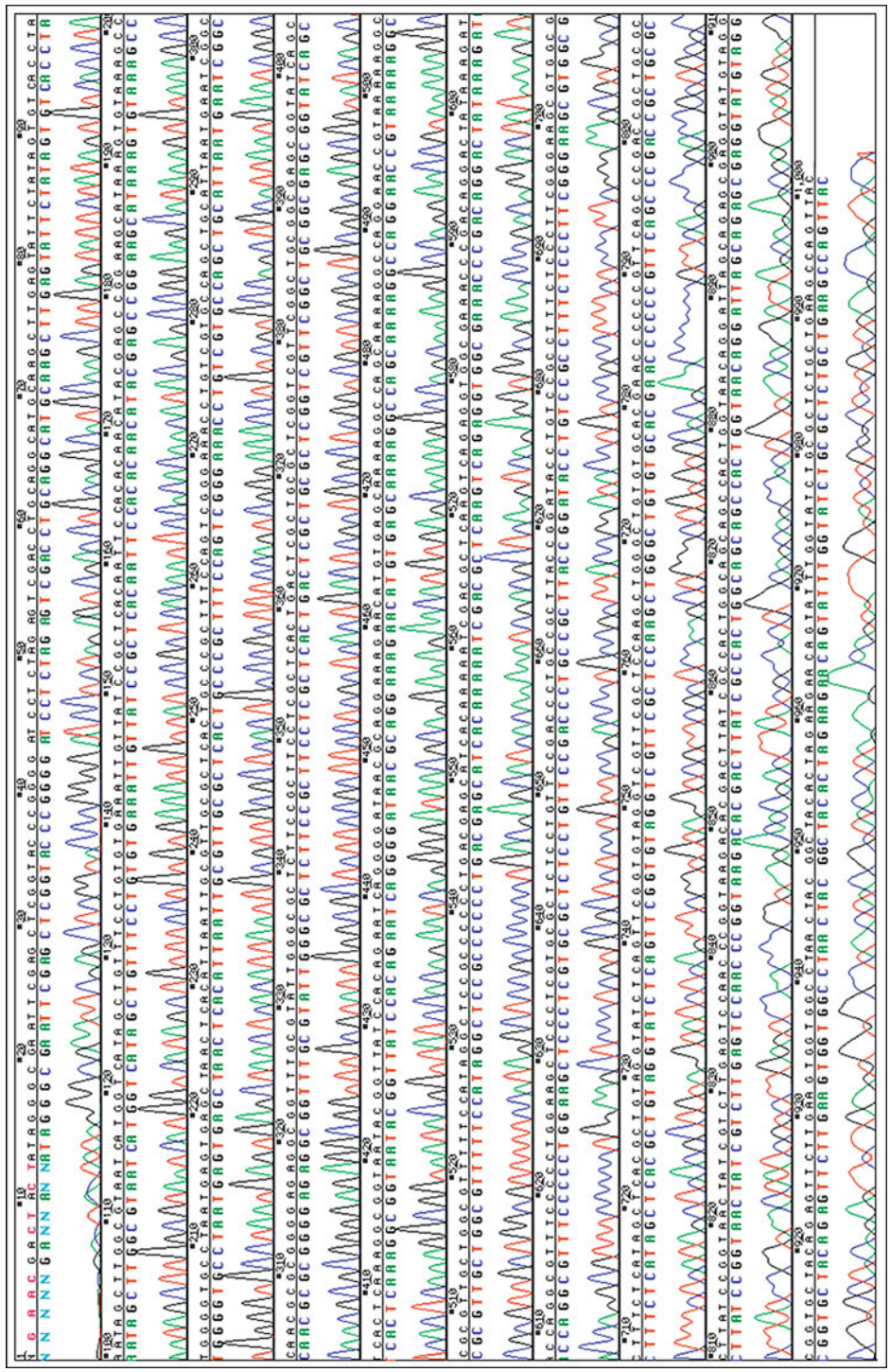


Fig. 8.40 An example of a dye terminator read. This 3.2 kb pGEM3zf-plasmid template was sequenced using the m13 forward −21 primer on an ABI 3730xl sequencer. The sequence could be read unambiguously starting about 20 nt from the primer with 100% accuracy out to around 1,000 bp. The dye terminators were completely removed by absorption to silica-coated paramagnetic particles (Copyrighted 2007© by David F. Bishop, used with permission)

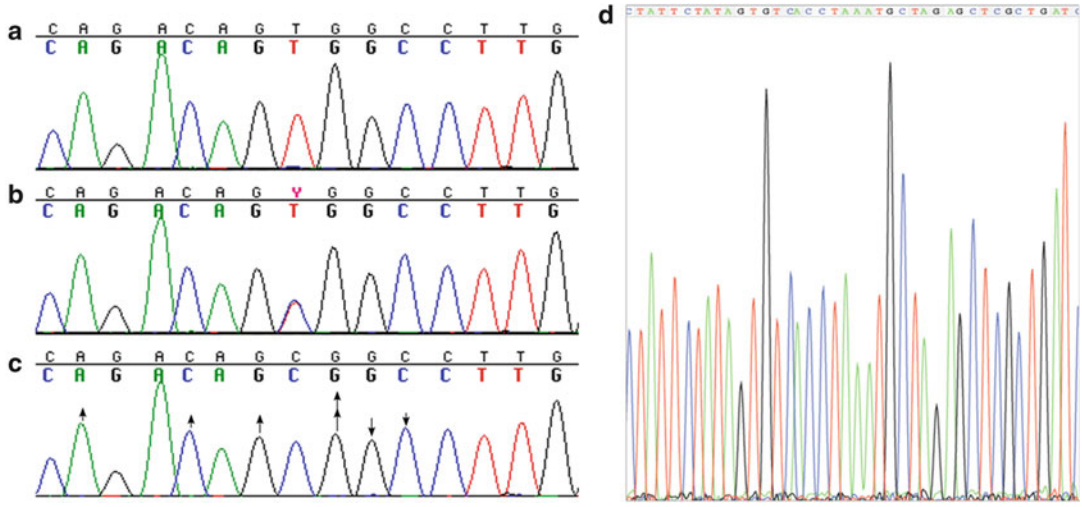


Fig. 8.41 Due to the different kinetics of fluorescent dye terminator incorporation, electropherogram peak heights are not uniform and vary by context due to the enzymatic properties of DNA polymerases. While in general, the peak height variability is not too great, the most significantly attenuated signals are found for Gs following after As. In electropherograms (**a**–**c**), the G after the first A is weak, as are the two G's after A's in electropherogram (**d**). If there are significant background problems, the weak G after A peaks may be ambiguous, especially if the sequence is unknown. Sequencing in both orientations obviates this problem, as the corresponding Cs after Ts are strong. Electropherograms (**a**) to (**c**) also demonstrate the specific effects of sequence context on peak height. Sequence B is that for an individual heterozygous for a T to C polymorphism. Note when comparing the homozygous C sequence in (**c**) to the homozygous T in (**a**), the peak heights for at least six nucleotides surrounding the altered base change in height with the arrows approximating the magnitude of change. Such changes can help one discriminate artifactual peaks from real ones as the artifacts do not result in other peak changes (Copyrighted 2007© by David F. Bishop, used with permission)

- Limitations
  - Failure to remove unincorporated ddNTPs can yield extraneous fluorescence
  - This can prevent a sample from being genotyped or cause it to be genotyped incorrectly. GeneMapper software is available to automatically determine sample genotypes

8.6.4 Pyrosequencing

- General information
  - The technique was developed by Pål Nyrén and Mostafa Ronaghi at the Royal Institute of Technology in Stockholm in 1996
  - Pyrosequencing is a method of DNA sequencing based on the "sequencing by synthesis" principle
  - It differs from Sanger sequencing, in that it relies on the detection of pyrophosphate release on nucleotide incorporation, rather than chain termination with dideoxynucleotides
- Principle
  - Pyrosequencing is a recently developed DNA sequencing method based on detecting the formation of pyrophosphate, the byproduct of DNA polymerization
  - In a number of enzymatic steps, pyrophosphate is converted to ATP, which fuels a luciferase reaction and converts luciferin to oxyluciferin (Fig. 8.44)
  - Light is generated with each addition of a nucleotide in the growing DNA chain
    - The intensity of light generated is proportional to the amount of nucleotide incorporated
    - Scores are determined by computer-automated comparison of predicted SNP patterns with raw data

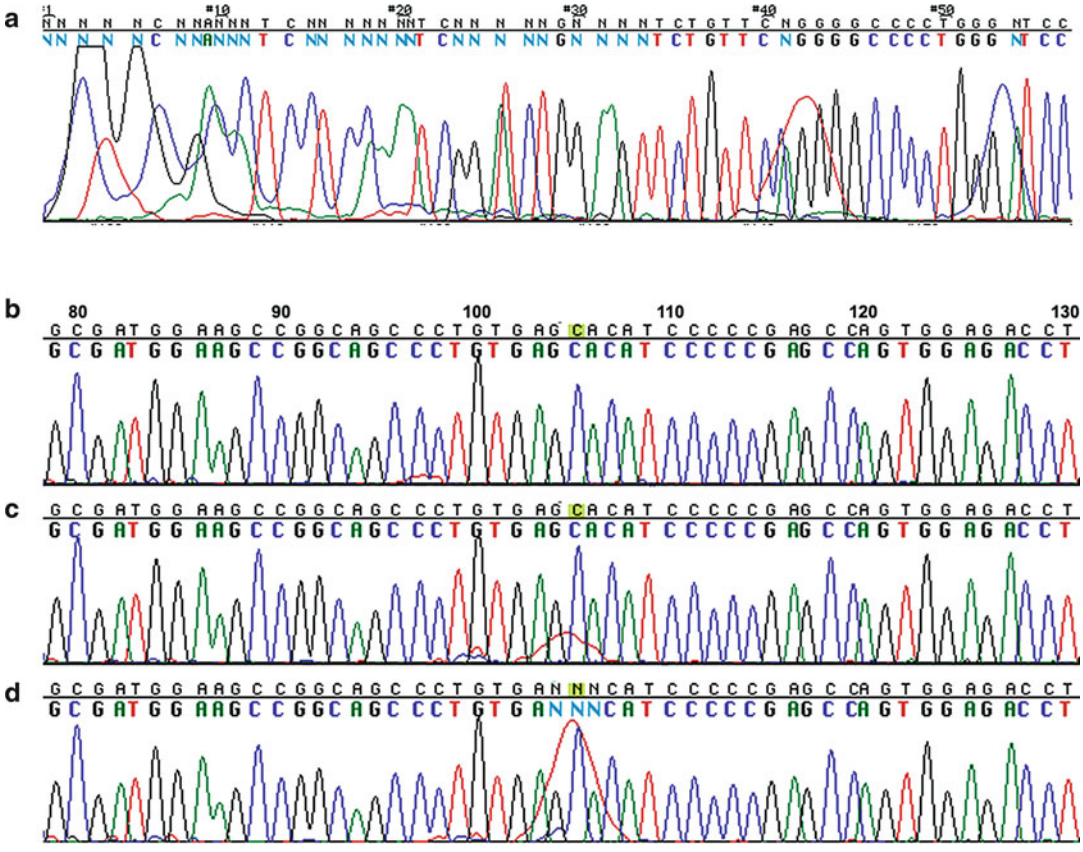


Fig. 8.42 If the fluorescent dye terminators used in the cycle sequencing reaction are not completely removed, they will show up as multiple broad peaks in the first 50 bases of the sequence, as seen in (**a**). The dye blob peaks are recognized by their broad elution profile relative to the sharper peaks of each fluorescent sequence fragment. For example, there is a broad T blob centered at 42 nt in electropherogram (**a**) that spans from around nt 40–46. Even though the base-calling software is confused, it is possible to read the correct sequence by ignoring the dye blob. Even after dye terminators are completely removed, there is a slow hydrolysis of dye moieties with storage of the reaction products. This frequently shows up as a broad T peak around nt 100. Electropherograms (**b**) through (**d**) show varying degrees of this degradation product, but the sequence can be easily read manually (Copyrighted 2007© by David F. Bishop, used with permission)

  - Samples generally do not require manual interpretation, which provides reliability and accuracy in scoring
- Procedure
  - Step 1: A sequencing primer is hybridized to a single-stranded DNA template and incubated with DNA polymerase, ATP sulfurylase, luciferase, and apyrase enzymes, as well as the substrates adenosine 5′ phosphosulfate (APS) and luciferin
  - Step 2: One of four deoxyribonucleotide triphosphates (dNTPs) is added to the reaction. DNA polymerase catalyzes the incorporation of the deoxyribonucleotide triphosphate into the DNA strand only if it is complementary to the base in the template strand
    - Each incorporation event is accompanied by release of pyrophosphate (PPi) in a quantity equimolar to the amount of incorporated nucleotide

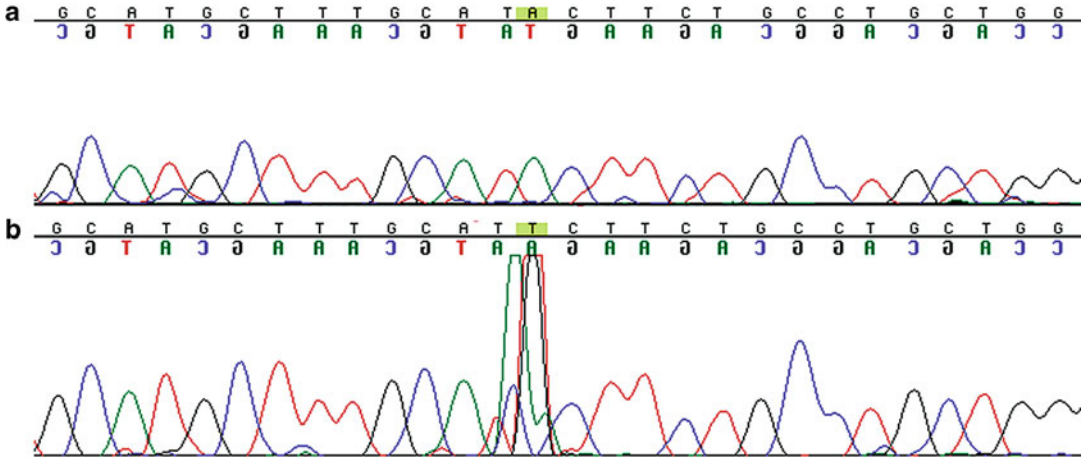


Fig. 8.43 Large peaks corresponding to all four nucleotides in the same base position are most likely caused by a bubble migrating though the capillary (**a**, **b**). The optical refraction caused by the bubble disperses the wavelengths and focuses them causing an intensity increase. This is most frequently observed after installation of a new capillary (Copyrighted 2007© by David F. Bishop, used with permission)

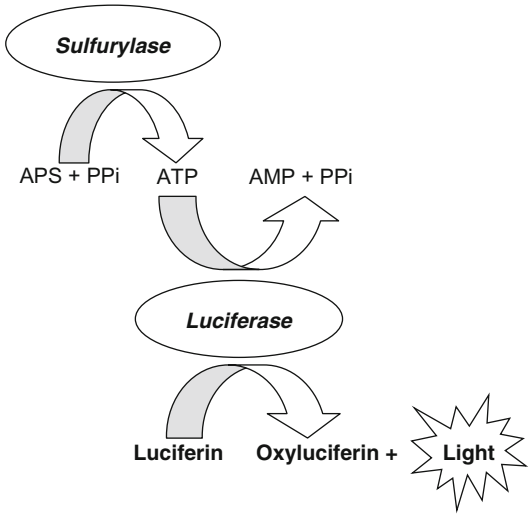


Fig. 8.44 Mechanism of light generation in pyrosequencing reactions (see text for explanation)

  - Step 3: ATP sulfurylase quantitatively converts PPi to ATP in the presence of APS. This ATP drives the luciferase-mediated conversion of luciferin to oxyluciferin that generates visible light in amounts that are proportional to the amount of ATP
    - The light produced in the luciferase-catalyzed reaction is detected by a charge coupled device (CCD) camera and seen as a peak in a Pyrogram™
    - The height of each peak (light signal) is proportional to the number of nucleotides incorporated
  - Step 4: Apyrase, a nucleotide-degrading enzyme, continuously degrades ATP and unincorporated dNTPs. This switches off light production and regenerates the reaction solution. The next dNTP is then added
  - Step 5: Addition of dNTPs is performed one at a time. It should be noted that deoxyadenosine alfa-thio triphosphate (dATPaS) is used as a substitute for the natural deoxyadenosine triphosphate (dATP) since it is efficiently used by the DNA polymerase, but not recognized by the luciferase
  - As the process continues, the complementary DNA strand is built up and the nucleotide sequence is determined from the signal peaks in the Pyrogram
- Advantages
  - Pyrosequencing not only generates sequence information but also it is

quantitative, ideal for measuring the relative amounts of alleles
  - This property allows the quantification of DNA methylation, heterozygosity, ploidy levels, multicopy genes, pooled DNA samples, hematopoietic chimerism, and mixed genotypes in heterogeneous samples (e.g., tumor and normal cells)

8.6.5 RNA Sequencing

- As RNA is generated by transcription from DNA, the information is already present in the cell's DNA. Therefore, RNA sequence can be deduced from the coding DNA sequence
- However, it is sometimes desirable to sequence RNA molecules. In particular, in eukaryotes RNA molecules are not necessarily colinear with their DNA template, as introns are excised
- To sequence RNA, the usual method is first to reverse transcribe the sample to generate DNA (cDNA) fragments. The DNA can then be sequenced as described above

8.6.6 Next-Generation Sequencing

- A new method of automated sequencing which uses various technological applications to sequence a larger area of interest, producing a vast amount of data that is analyzed using data analysis programs

8.6.6.1 Massively Parallel Signature Sequencing (Lynx Therapeutics, Hayward, CA)

- General information
  - An open-ended platform that analyzes genetic expression by examining individual mRNA molecules using thousands of PCR reactions
- Principle
  - Tagged PCR products (produced from extracted cDNA) are amplified, producing thousands of PCR products from each mRNA
  - Tags are used to attach PCR products to microbeads. After several rounds of ligation-based sequence determination, an average sequence of 17 bp is identified from each bead
  - The analysis of each bead is performed in parallel, producing hundreds of thousands of signature sequences
  - Each signature sequence in a dataset is analyzed and compared with all of the other signatures, allowing a level of speed and sensitivity not normally available
  - This process has sensitivity of a few molecules of mRNA per cell and produces results in a digital format
- Procedure
  - cDNA fragments are cloned onto microbeads using Lynx Megaclone technology. The Megaclone produces one million beads, each containing 100,000 cloned copies of cDNA
  - All molecules are covalently attached to the microbeads at the polyA ends with DpnII end available for sequencing reactions
  - Sequencing is then performed by first hybridizing the beads and performing fluorescence-activated cell sorting (FACS) separation
  - cDNA templates are initially cleaved at the DpnII end and converted to 3-base overhangs. A set of encoded adaptors are added, and complement adaptors are ligated
  - Identity and ordering of nucleotides are read by 16 decoder probes that are sequentially hybridized, providing a fluorescent signal that is recorded
  - BbvI digest exposes the next 4 nucleotides and shortens the template. The round is then repeated until the entire desired sequence is produced
- Applications
  - Useful for sequencing large areas of interest by examining individual mRNA molecules

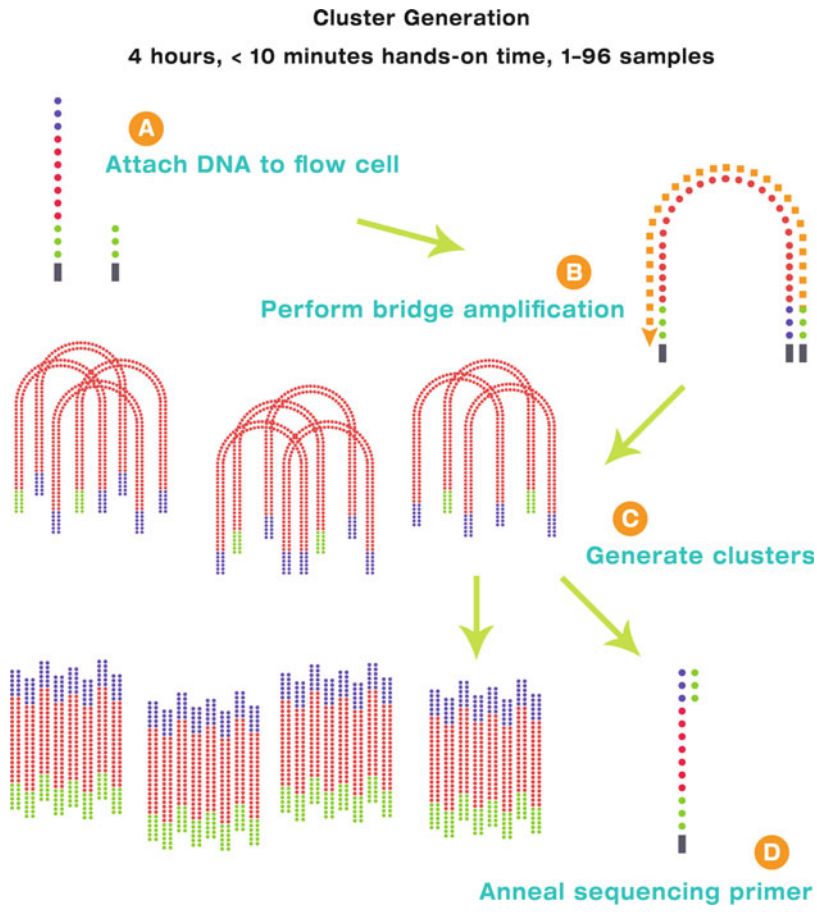


Fig. 8.45 Bridge amplification used in Illumina Genome Analyzer

- Advantages
  - Fast sequencing producing thousands of PCR reactions
- Limitations
  - Data must still be analyzed. Nonsensical or superfluous information must be extracted from the large amount of data produced

8.6.6.2 Genome Analyzer (Illumina, San Diego, CA)

- Principle
  - Utilizes bridge amplification (Fig. 8.45)
    - A single strand flips over to hybridize to adjacent primers forming a "bridge"
    - The hybridized primer is then extended by polymerase
  - "TruSeq technology": DNA library samples are bound to complementary adapter oligos grafted on the surface of the Illumina Genome Analyzer flow cell
  - The templates are copied from the hybridized primer by 3′ extension using a high-fidelity DNA polymerase to prevent misincorporation errors. These are isothermally amplified to create clonal clusters of ~1,000 copies each, ready for sequencing on the Genome Analyzer
  - The paired-end flow cell contains proprietary oligos that enable selective cleavage of the forward DNA strand after resynthesis of the reverse strand
  - The reverse strand is regenerated by bridge amplification with the paired-end module (attached to the Genome Analyzer)

  - After resynthesis of the reverse strand, the original forward strand is cleaved and the reverse strand is sequenced for the second read
- Procedure
  - Prepare library
    - Fragment DNA
    - Repair ends and add A overhang
    - Ligate adapters
    - Select ligated DNA
  - Automated cluster generation
    - Hybridize single molecule to flow cell
    - Extend hybridized oligos
    - Perform bridge amplification to generate >1,000 copies
  - Sequencing using sequencing by synthesis chemistry
    - Perform sequencing on forward strand using fluorescence-labeled terminator
    - Regenerate reverse strand
    - Perform sequencing on reverse strand
- Applications
  - Whole genome sequencing
    - Resequencing
    - De novo sequencing
    - Targeted sequencing
    - Metagenomic sequencing
  - Transcriptome
    - ENA-seq
    - DGE
    - Small RNA
    - miRNA
  - Methylation
  - Chip-seq
- Advantages
  - Allows for significantly larger runs (Gb vs Mb available in Pyrosequencing)
  - More expensive per run but less expensive per Mb
  - Does not use emulsion PCR, so allows for self-contained system that does not need a clean room
  - More accurate sequence data
- Limitations
  - Relatively short length (about 100 bp), therefore, difficult to align large indel
  - Coverage is about 99.8%

8.6.6.3 SOLiD System (Applied Biosystems, Carlsbad, California)

- General information
  - Sequencing produced using ligation to beads fixed to a glass slide
- Principle
  - ABI SOLiD technology uses ligation sequencing, producing labeled beads fixed to a surface, which then produce an extension product over multiple rounds (Fig. 8.46)
  - Beads labeled with the library of interest are deposited on to a glass slide
  - Primers hybridize to the adapter sequence on the templated beads
  - A set of four fluorescently labeled di-base probes compete for ligation to the sequencing primer. Specificity of the di-base probe is achieved by interrogating every first and second base in each ligation reaction
  - Multiple cycles of ligation, detection, and cleavage are performed with the number of cycles determining the eventual read length
  - Following a series of ligation cycles, the extension product is removed and the template is reset with a primer complementary to the n-1 position for a second round of ligation cycles
- Procedure
  - Library preparation: producing sequencing fragments or mate-paired
  - Emulsion PCR and bead enrichment: generating clonal bead population
  - Bead deposition: depositing 3′ modified beads onto a glass slide
  - Performing sequencing by ligation
- Advantages
  - Not as expensive as pyrosequencing but more expensive than Illumina genomic analyzer
  - Approximately 3,000 Mb per run
  - 99.99% call accuracy

8.6.6.4 HeliScope Genetic Analysis System (Helicos BioSciences Corporation, Cambridge, MA)

- General information
  - Single molecule technology, no amplification
  - Sequencing by synthesis with virtual terminators

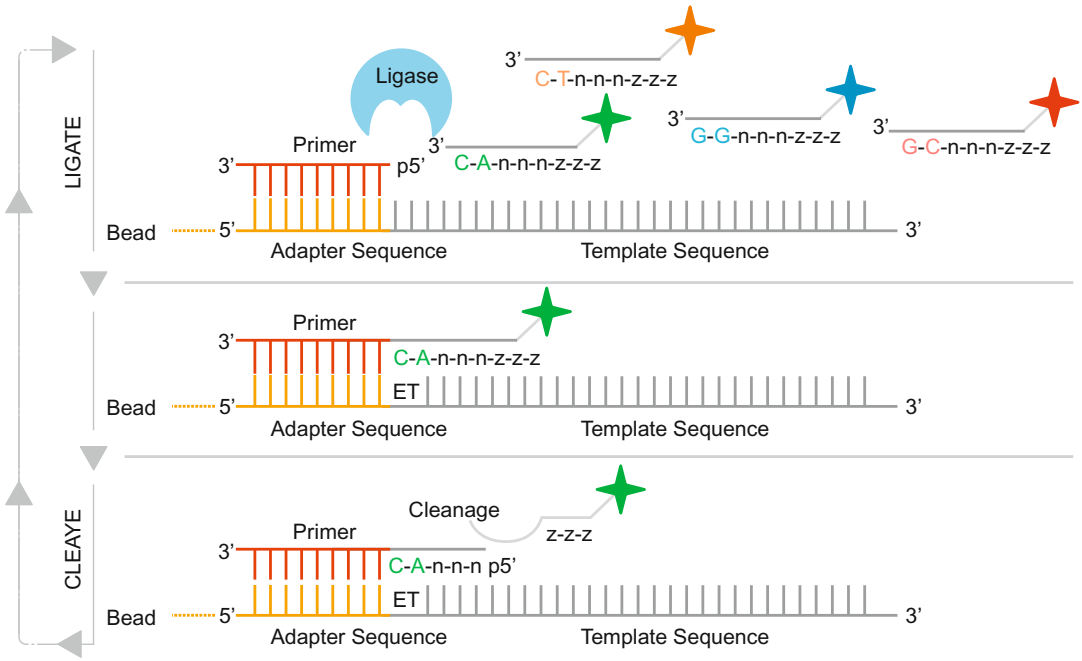


Fig. 8.46 Ligation sequencing used in Applied Biosystems SOLiD System (Courtesy of Applied Biosystems, used with permission. See text for description)

- Principle
  - Fluorescently labeled nucleotides are added sequentially to DNA fragments captured onto a flow cell
  - The image produced is then analyzed, and another round of imaging is done, eventually sequencing the entire DNA of interest
- Procedure
  - DNA samples are fragmented, and then the DNA is melted into single strands
  - A poly A tail is added
  - These DNA molecules are captured on the surface of a flow cell
  - Fluorescently labeled nucleotides are then added sequentially and incorporated into the strand by DNA polymerase
  - Excess nucleotides are washed away, and the surface is then illuminated with a laser
  - The light emission is detected by the sequencer, which then analyzes the images produced
  - The label is removed, and then the next nucleotide is added to continue the cycle
  - Samples are then loaded onto the flow cells, whose product is then analyzed
- Advantages
  - Much larger sequencing output (83 Mb/h) compared to other next-generation technology
  - Does not need ligation or amplification, also requires less material
  - Degraded or modified specimens can be directly used as templates

8.6.7 Target Capture

- General information
  - Target capture is used to select genomic regions of interest from full complexity of human genomic DNA
  - Target capture eliminates the necessity of setting up thousands of PCR reactions, instead allowing for parallel enrichment of target regions in a single experiment
  - Targeted sequence capture can reduce the cost, improve throughput, and save time for next-generation sequencing

- Principle
  - A set of oligonucleotide probes is used to capture the desired sequences from total human genomic DNA
  - These captured sequences are then amplified in a single PCR reaction using common linkers or adaptors, originally attached either to the probes or to the genomic DNA, as primers (The unwanted sequences are discarded)
- Procedure
  - Genomic DNA is hybridized to either the oligo pool or the capture array specific regions of genomic DNA
  - If solution capture is used, then streptavidin beads are used to pull down the complex of captured oligo and genomic DNA
  - After washing, the target fragments are eluted and recovered from the solution. Amplification by PCR is then performed, and the success of the enrichment is measured by qPCR
  - The end product is a sequencing library for the target regions
  - The array version varies in that capture probes are synthesized on an array. Unbound fragments are removed by washing, and then target fragments are recovered from the array
- Application
  - Sequencing specific DNA region of interest
- Advantages
  - Much faster than standard sequencing, as it uses parallel sequencing against a targeted DNA region
- Limitations
  - Targeted area must be identified and made into a library before analysis can begin

8.7 Protein Detection Methods

8.7.1 Enzyme Immunoassay (EIA)

- General information
  - Enzyme immunoassay (EIA) is the newer name of enzyme-linked immunosorbent assay (ELISA), which was developed in 1971 as a way of detecting proteins in serum
  - EIA is an immunoassay technique combining an in vitro antigen–antibody reaction with a subsequent enzymatic reaction
  - EIA is a highly sensitive and specific assay and compares favorably with other methods used to detect substances in the body, such as radio immunoassay (RIA)
  - EIA tests are usually performed in microwell plates, and use an enzyme linked to an antibody or antigen as a marker for detection of a specific protein, especially an antigen or antibody
- Principle (Fig. 8.47)
  - EIA techniques have many variations. Three most commonly used formats are

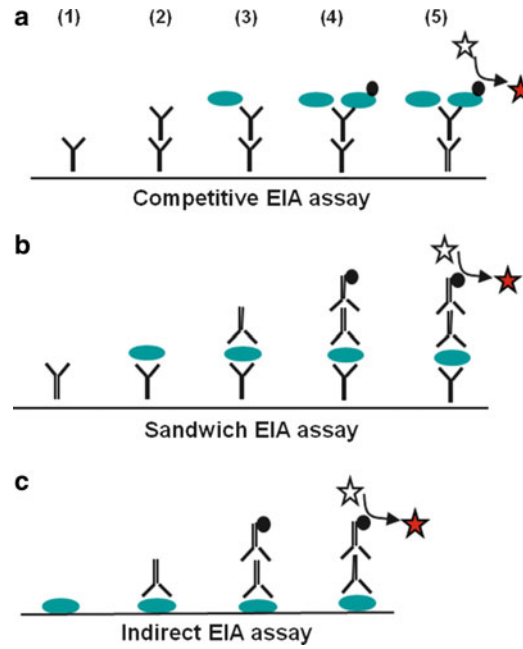


Fig. 8.47 Three types of EIA assay. (**a**) Competitive EIA assay: (*1*) coat plate with secondary antibody; (*2*) add primary antibody; (*3*) add antigen (sample); (*4*) add antigen conjugate; (*5*) add substrate to produce color. (**b**) Sandwich EIA assay: (*1*) coat plate with primary antibody #1; (*2*) add antigen (sample); (*3*) add primary antibody #2; (*4*) add secondary antibody conjugate; (*5*) add substrate to produce color. (**c**) Indirect EIA assay: (*1*) coat plate with antigen; (*2*) add primary antibody; (*3*) add secondary antibody conjugate; (*4*) add substrate to produce color

- Competitive immunoassay relies on the principle of competition between antigen in a test specimen and antigen–enzyme conjugate for binding with a constant amount of antibody. Usually, secondary antibody (such as goat anti-rabbit IgG or goat anti-mouse IgG) is precoated on microtiter plates and is used to bind with an antigen-specific mono- or poly-antibody. In the meantime, a fixed amount of enzyme-labeled antigen competes with antigen in the test specimen for a fixed number of binding sites of the antigen-specific mono- or poly-antibody. Thus, the amount of enzyme-labeled antigen immunologically bound to the well progressively decreases as the concentration of antigen in the specimen increases
- Sandwich immunoassay is based on the capture of antigen by one antibody and the detection by another antibody. In general, a monoclonal antibody, which is precoated on microtiter plates, is used to capture antigen in a test specimen. Antigen-specific polyclonal antibody, which is conjugated to an enzyme, then binds with the immobilized antigen
- Indirect immunoassay is used to detect specific antibody and is somewhat different from the first two immunoassays, which are for detection of antigen. Antibody in a test specimen is captured by antigen precoated on microtiter plates. Detection occurs by means of a secondary antibody conjugated to an enzyme

– The enzyme label: Most of the immunoassays employ horseradish peroxidase (HRP), alkaline phosphatase (AP), or β-galactosidase. Substrates used with HRP include 2,2′-azo-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), o-phenylenediamine (OPD), and 3,3′5,5′-tetramethylbenzidine base (TMB), which yield green, orange, and blue colors, respectively. TMB is gradually replacing mutagenic substrates such as OPD, leading to increased sensitivity and safety
– Color development: With the addition of antibody or antigen conjugated to enzyme (such as HRP) and followed by the addition of substrate (such as TMB), the amount of antigen or antibody is detected through measurement of the color intensity with a spectrophotometer. This results in a direct or inverse relationship between optical density (OD) and concentration: The higher the OD, the more antigen or antibody (for sandwich or indirect EIA type), or the higher the OD, the less antigen (for competitive EIA)

- Procedure
  – An EIA is a five-step procedure: (1) coat the microtiter plate wells with antigen or antibody; (2) block all unbound sites to prevent false-positive results; (3) add antibody or antigen to the wells; (4) add primary or secondary antibody conjugated to an enzyme; and (5) introduce substrate
    - General procedure for the competitive EIA method
      – To coat the plate with secondary antibody, add diluted secondary antibody to each well. The appropriate dilution should be determined using a checkerboard titration prior to testing samples. A microtiter plate will bind approximately 100 ng/well (300 ng/cm^2). The amount of antibody used will depend on the individual assay, but if maximal binding is required, use at least 1 μg/well. Allow to incubate for 4 h at room temperature or 4 °C overnight
      – Wash the coated plate by filling wells with PBS. Flick the plate over a suitable container, and rinse with PBS two more times
      – To block residual binding capacity of the plate, fill each well to the top with

blocking buffer (3% BSA/PBS with 0.02% sodium azide) and incubate for at least 2 h at room temperature. Rinse plate three times with PBS as in step b. After the last rinse, remove residual liquid by wrapping each plate in a large paper tissue and gently flicking it face down onto several paper towels laid out on a benchtop
  - Add antigen-specific antibody to the coated wells. The antibody should be diluted in blocking buffer (3% BSA/PBS with 0.05% Tween-20)
  - Add standard or sample (antigen) to plate. Standard or antigen sample should be diluted in blocking buffer
  - Add antigen-conjugate solution to the coated wells. All dilutions should be done in the blocking buffer. Incubate for at least 2 h at room temperature in a humid atmosphere
  - Wash the plate four times with PBS as in step b
  - Add substrate (such as TMB) and measure optical density (OD): TMB is added to each well and incubated for 30 min at room temperature, resulting in the development of blue color. The color development is stopped with the addition of 1N H_2SO_4 or 3N HCl, and the OD is measured spectrophotometrically at 450 nm
  - A standard curve is obtained by plotting the concentration of standards versus OD. Competitive EIA yields an inverse curve, where higher values of standards or antigen in the samples yield a lower amount of color change
  - The concentration of the antigen specimen can be calculated from the standard curve
- General procedure for the sandwich EIA method
  - Coat plate with antigen-specific first monoclonal antibody
  - Wash the wells three times with PBS
  - Block residual binding capacity of plate
  - Add standard or sample (antigen) to plate: Incubate for at least 2 h at room temperature in a humid atmosphere
  - Wash the plate four times with PBS
  - Add antigen-specific second polyclonal antibody to plate: The amount to be added can be determined in preliminary experiments. For accurate quantitation, the second antibody should be used in excess. Incubate for at least 2 h at room temperature in a humid atmosphere
  - Wash with several changes of PBS
  - Add HRP-conjugated secondary antibody (such as goat anti-rabbit antibody) to plate: Incubate for 1 h at room temperature in a humid atmosphere
  - Wash with several changes of PBS
  - Add substrate (such as TMB) and measure optical density (OD)
  - Make a standard curve
  - Perform calculation
- General procedure for the indirect EIA method
  - Coat plate with antigen
  - Wash the wells three times with PBS
  - Block residual binding capacity of plate
  - Add standard or sample (antibody) to plate
  - Wash the plate four times with PBS
  - Add HRP-conjugated secondary antibody to plate
  - Wash with several changes of PBS
  - Add substrate (such as TMB) and measure optical density (OD)
  - Make a standard curve
  - Perform calculation
- Applications
  - EIA can be performed to evaluate the presence of antibody in a sample, which is recognized by an antigen. Therefore, it is a useful tool for determining serum antibody concentrations, such as with the

human immunodeficiency virus (HIV) test or West Nile virus
  - EIA also can be used to detect and quantify the concentrations of antigens that are recognized by antibodies, such as disease-related substances
  - The most commonly used EIA assay format is the sandwich assay. It can be used to measure antigens that are bound between two antibodies. Competitive assay is often used when the antigen is small and has only one epitope, or antibody-binding site
- Advantages
  - EIA is a sensitive and specific assay for the detection and quantitation of antigens or antibodies. It combines the specificity of antibodies with the sensitivity of simple enzyme assays
  - EIA is also a rapid and relatively easy assay when compared to conventional GC/MS and HPLC. Whereas a conventional method may require 1 h to analyze one sample, EIA can analyze about 30 samples per h. The immunoassay instrumentation kit is portable and can be used for testing right at the sampling site
  - EIA can be performed without the use radioactive materials and is also considerably less expensive than radioimmunoassay (RIA)
  - ELISA may be run in a qualitative or quantitative format. Qualitative results provide a simple positive or negative result for a sample. In quantitative ELISA, the optical density of the sample is interpolated into a standard curve which is typically based on a serial dilution of the target
- Limitations
  - EIA requires skilled laboratory technicians and specialized laboratory equipment
  - Cross-reactivity may occur with the secondary antibody, resulting in nonspecific signal
  - Because of the design of the immunoassay, sample contaminants that might interfere with the antigen–antibody reaction can produce positive readings when samples are indeed negative (false-positive results)
  - Lower molecular weight molecules often lack specific antigenic sites, and sometimes there are cross-linking problems

8.7.2 Protein Electrophoresis (SDS-PAGE)

- General information
  - SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) is a very common method for separating proteins using a polyacrylamide gel as the support medium and sodium dodecyl sulfate (SDS) as a denaturing agent
  - SDS-PAGE is used to separate proteins based on molecular weight
  - Gel density can be controlled by varying the monomer concentration. Gels can be of constant density or they can be variable (gradient gels)
- Principle
  - SDS has two important features
    - It is an anionic detergent that binds quantitatively to proteins, giving them uniform negative charge which means they will all migrate toward the positive pole when placed in an electric field. The number of SDS molecules that bind to a protein is proportional to the number of amino acids that make up the protein. Each SDS molecule contributes two negative charges, overwhelming any charge the protein may have
    - SDS also disrupts the forces that contribute to protein folding (tertiary structure), ensuring that all proteins are denatured to the same linear configuration (Fig. 8.48)
  - Since all proteins become linearized and uniformly negatively charged, separation occurs solely on the basis of size. Smaller molecules are able to navigate the gel faster than larger ones; thus, they migrate more rapidly

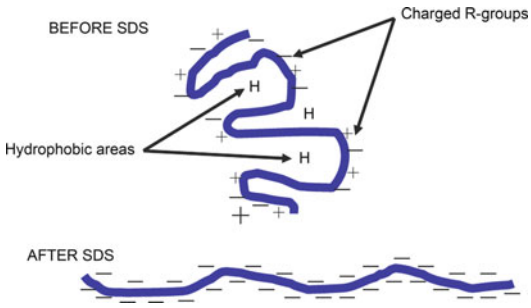


Fig. 8.48 Cartoon depicting changes in a protein (*blue line*) when incubated with the denaturing detergent SDS. The *top* portion of the figure shows a protein with negative and positive charges due to charged R groups. The *lower* diagram shows how SDS coats protein molecules with a negatively charged layer which overwhelms any positive charges intrinsic to the protein. SDS also disrupts hydrophobic interactions. As a result, proteins are denatured (reduced to their primary structures) and thus linearized

  - The polyacrylamide gel is a cross-linked matrix that functions as a sieve that differentially retards the motion of molecules as they move through the electric field
  - The preparation of the gel requires casting two different layers of acrylamide between glass plates. The lower layer (separating, or resolving, gel) is responsible for actually separating polypeptides by size. The upper layer (stacking gel) includes the sample wells; it is designed to sweep up proteins in a sample between two moving boundaries so that they are compressed (stacked) into micrometer thin layers when they reach the separating gel
  - Staining of proteins in gels may be done using the standard Coomassie Brilliant Blue, Amido black, or silver stain reagents of various kinds. Coomassie Brilliant Blue G-250 is probably the most widely used due to its convenience. It binds nonspecifically to virtually all proteins and can visualize bands containing as little as 0.3 μg protein
  - Protein molecular weight standards are used to measure the relative sizes of the unknown proteins

- Procedures
  - A sample of proteins is first denatured. An appropriate amount of electrophoresis sample buffer (1X = 125 mM Tris-HCl pH 6.8, 2% SDS, 5% glycerol, 0.003% bromophenol blue, and 1% beta-mercaptoethanol) is then added to all proteins. The mixture is then heated to 95 °C for 3–5 min
  - The SDS-PAGE gel is prepared in two steps. First the resolving gel is prepared and then the stacking gel, which sits atop the resolving gel
  - The gel is placed into the gel apparatus, in an appropriate buffer, and protein samples are loaded onto the stacking gel. Load 5–100 ug total protein in a volume that is appropriate for the well size. Be sure to use protein markers which produce bands of known size
  - When power is supplied, the proteins migrate from the cathode (upper chamber of gel apparatus) to the anode (lower chamber)
  - Electrophoresis is stopped when the bromophenol blue dye front reaches the bottom of the gel. The protein bands are then stained (e.g., with Coomassie Brilliant Blue G-250)
  - The size of proteins corresponding to each band is determined based on comparison with molecular weight markers which are run along with the sample proteins on the gel
- Applications
  - Estimating relative molecular mass of proteins
  - Determining protein purity in a sample
  - Identifying the composition of protein complexes
  - Preparation for blotting
- Advantages
  - One-dimensional gel electrophoresis is a relatively easy technique and is very reliable
  - Can be performed using large gels or smaller “micro” gels, giving a choice as to how much sample and reagent(s) will be utilized

- Limitations
  - Occurrence of false-positives and false-negatives due to comigrating contaminants
  - Abnormal migration of proteins due to presence of large numbers of charged amino acids

8.7.3 Western Blotting (WB)

- General information
  - Western blotting was first introduced in the late 1970s. Since then, it has become rapidly accepted and widely applied
  - Western blotting is used to detect a specific target protein in a sample containing a complex mixture of proteins by using a polyclonal or monoclonal antibody specific to that protein
  - Western blotting allows investigators to determine the molecular weight of a protein and to measure relative amounts of the protein present in different samples
- Principle
  - SDS, a reducing agent such as dithiothreitol (DTT) or 2-mercaptoethanol, and heat are responsible for the actual denaturation of proteins
    - SDS breaks up the two- and three-dimensional structure of proteins by adding negative charge to the amino acids
    - Disulfide bonding is covalent and is not disrupted by SDS. DTT or 2-mercaptoethanol is a strong reducing agent. Its specific role in protein denaturation is to remove the last remnants of tertiary and quaternary structure by reducing disulfide bonds
    - Many proteins have significant hydrophobic properties and may be tightly associated with other molecules. Heating the proteins to at least 60 °C separates the molecules, allowing SDS to bind in the hydrophobic regions to complete the process of denaturation
  - The denatured proteins are separated based on weight and electrical properties by gel electrophoresis, usually SDS-polyacrylamide (SDS-PAGE) (see above)
  - The proteins are then electrophoretically transferred to a membrane of charged nylon, nitrocellulose, or polyvinylidene fluoride (PVDF)
  - The membrane must be incubated with generic proteins (such as milk proteins) to block remaining hydrophobic binding sites on the membrane. This reduces background and prevents binding of the primary antibody to the membrane itself
  - A primary antibody with specificity for the protein of interest is introduced, and antibody–protein complexes are formed. The appropriate working concentration of the primary antibody is dependent upon its binding characteristics
  - The method of detection is dependent upon the label that has been conjugated to the primary (or secondary) antibody. The most common antibody label used in Western blotting is an enzyme such as alkaline phosphatase or horseradish peroxidase (HRP), which can be detected visually through the conversion of a colorimetric substrate (chromagen) to a colored precipitate at the site of antibody binding. Alternatively, chemiluminescent substrates may be employed which emit light upon conversion by the enzyme. The light emitted at the site of substrate conversion can be captured on x-ray film. Chemiluminescent substrates are much more sensitive than colorimetric substrates
- Procedure (Fig. 8.49)
  - Preparation of cell lysates
    - Process cells by trypsinization and spin or tissue by mincing and digestion
    - Lyse the pellet or tissue with lysis buffer on ice for 10 min
    - Sonication for 10–30 s
    - Spin at 14,000 rpm in an Eppendorf® microfuge for 20 min at 4 °C
    - Transfer the supernatant to a new tube and discard the pellet

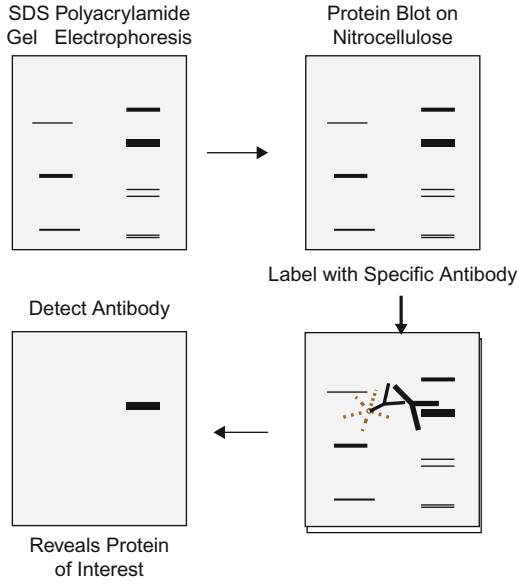


Fig. 8.49 Schematic presentation of how western blotting is performed (see text for description of steps)

- Determine the protein concentration using Bradford assay or BCA method

– SDS-PAGE is performed (as above)
– Membrane transfer
  - Prechill transfer buffer at −20 °C
  - Cut a piece of a suitable membrane (e.g., nitrocellulose and PVDF)
  - Prewet the membrane, sponges and filter papers in transfer buffer
  - Assemble a "sandwich" following the sequences: sponge–filter paper–gelmembrane–filter paper–sponge
  - Transfer proteins from the gel to the membrane using electrophoresis for 1 h at 100 V at 4 °C. Bigger proteins might take longer to transfer. The negative pole must be on the side of the gel, and the positive pole must be on the side of the membrane to drive the negatively charged proteins over to the positively charged membrane. One must ensure that there are no air bubbles between the membrane and the gel or the proteins will not transfer
  - When finished, the membrane is imprinted with the same protein bands as the gel
– Blocking
  - Immerse the membrane in blocking buffer
  - Block for 30 min at 37 °C, 1 h at room temperature, or overnight at 4 °C
– Incubation with primary antibody
  - Decant the blocking buffer, and add a primary antibody, diluted in blocking buffer as suggested in its product description sheet
  - Incubate with gentle shaking for 30 min at 37 °C, 1 h at room temperature, or overnight at 4 °C
  - Decant the antibody, and wash for 30 min with agitation in wash buffer (TBS or PBS with 0.1% Tween 20), changing the wash buffer every 5 min
– Incubation with secondary antibody
  - Decant the wash buffer, and add secondary antibody conjugated to HRP, diluted in blocking buffer, for 1 h at room temperature
  - Decant the antibody conjugate, and wash for 30 min with agitation in wash buffer (TBS or PBS with 0.1% Tween 20), changing the wash buffer every 5 min
– Enzymatic chemiluminescence (ECL) substrate incubation and visualization
  - Decant the wash buffer, and place the membrane in a plastic bag or clean tray containing the development working solution (0.125 ml/cm^2) for 1–5 min following the manufacturer's instructions for specific ECL reagents and procedures. Agitate the bag or tray to cover the surface of the membrane
  - Remove the membrane from the bag or tray, and wrap it using plastic paper
  - Expose to x-ray film or any sensitive screen for the appropriate time period. For best results, use a range of exposures (10 s, 1 min, 5 min, and 20 min) to visualize the chemiluminescence signal

corresponding to the specific antibody–antigen complex
    - Bands appear wherever there is a protein-primary antibody-secondary antibody–enzyme complex. These bands correspond to the location of the target protein
- Applications
  - Western blotting provides information about presence and concentration of an antigen in a sample
  - Western blotting also reveals data about the nature of the antigen detected, such as its molecular weight, tertiary structure, and, in some cases, its biological activity
  - In clinical settings, Western blotting is routinely used to confirm serious diagnoses suggested by ELISA such as HIV seroconversion
- Advantages
  - Western blotting is a sensitive and specific technique
  - Western blotting is especially helpful when dealing with antigens that are insoluble, difficult to label, or easily degraded and thus not amenable to procedures such as immunoprecipitation
  - Western blotting, unlike EIA, can be used to detect multiple protein antigens
- Limitations
  - Western blotting is a relatively imperfect quantitative technique. For example, variations in epitopes can affect the intensity of staining or derivatization, and chemiluminescence exposure times are known to vary from blot to blot
  - Western blotting sometimes produces serious errors such as background staining or extra bands in the blot

8.7.4 Technologies Used In Proteomics

- Proteomics is the large-scale study of proteins, particularly their structures and functions. This term was coined by Marc Wilkins and colleagues in the early 1990s and made an analogy with "genomics," which describes the entire collection of genes in an organism. Proteomics is much more complicated than genomics. Most importantly, while the genome is a rather constant entity, the proteome differs from cell to cell and is constantly changing through its biochemical interactions with the genome and the environment. One organism has radically different protein expression patterns in different parts of its body, in different stages of its life cycle, and in different environmental conditions
- Since proteins play a central role in the life of an organism, proteomics is instrumental in the discovery of biomarkers, such as markers that indicate the presence of a particular disease. Proteomics studies usually require three stages of sample preparation (Fig. 8.50). Several key technologies are involved, including two-dimensional electrophoresis and mass spectrometry. Proteins are first separated using two-dimensional electrophoresis, and then individual protein spots of interest are cut from the gel and digested into smaller polypeptide fragments (5–10 amino acids in length) by enzymes. The polypeptide fragments are then analyzed by mass spectrometry. Matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry is used primarily to measure the masses of peptides. Electrospray ionization (ESI) tandem mass spectrometry is used to obtain peptide structure and sequence data

8.7.4.1 Two-Dimensional Electrophoresis

- General information
  - Two-dimensional electrophoresis is a method of protein separation, by which proteins in a mixture are separated according to their isoelectric point (pI) in the horizontal direction (isoelectric focusing, or IEF) and molecular weight in the vertical direction (SDS-PAGE)
  - Two-dimensional electrophoresis is the most effective means of resolving complex protein mixtures and was first introduced in the early 1970s

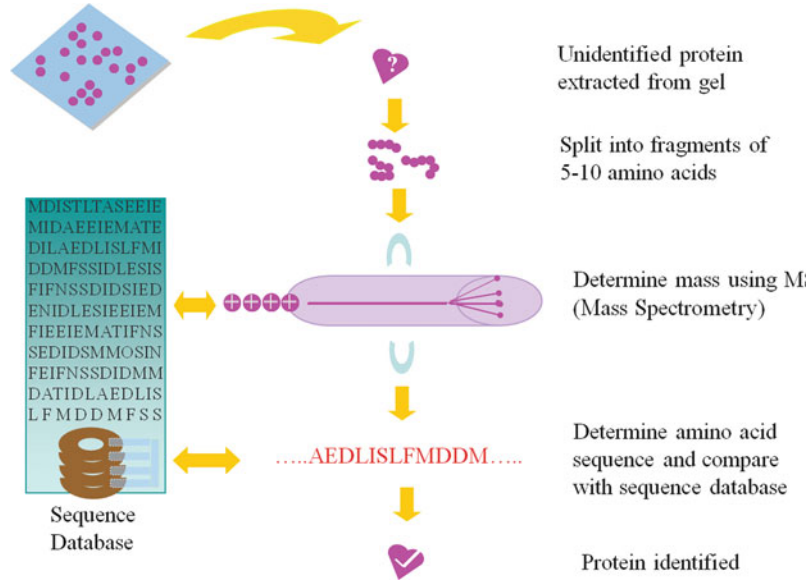


Fig. 8.50 General flow scheme for proteomic analysis

  - Two-dimensional electrophoresis is used for the isolation/separation/purification of proteins for further characterization with mass spectrometry and identification of specific proteins. Thus, this separation method has become synonymous with proteomics
  - Two-dimensional electrophoresis can effectively separate multiple isoforms of a protein
- Principle
  - Sample preparation and solubilization are crucial factors for the overall performance of two-dimensional electrophoresis. Protein complexes and aggregates should be completely disrupted in order to prevent the appearance of artifactual spots due to incomplete protein solubilization. Proteins can be completely solubilized, typically by a reagent such as urea
  - Two-dimensional electrophoresis separates proteins based on size, as in regular electrophoresis, but also based on charge, or isoelectric point (pI)
  - The first step is isoelectric focusing (IEF). The mixed protein sample is run on an immobilized pH gradient; the range of the gradient used depends on the expected proteins in the sample. The sample is added to the gradient and an electric current is applied. Proteins will be positively charged at pHs below their pI and negatively charged at pHs above their pI. When a given protein is at the point in the gradient where the surrounding pH is equal to its pI, the protein will have a net charge of zero and it will stop moving
  - The second step is SDS-PAGE. Once enough time has passed for the proteins to settle in the gradient, the current is removed and the gradient is laid horizontally along an SDS-PAGE gel. An electric current is then applied, and the proteins move horizontally out of the IEF gradient and into the SDS-PAGE gel where they are separated based on molecular weight
  - Once the proteins have been separated, they can be visualized by conventional staining techniques, including the standard Coomassie Brilliant Blue, Amido black, and silver stains. Silver staining is one of the most sensitive protein detection methods, with sensitivity greater than 100-fold that of Coomassie Brilliant Blue staining. A stained two-dimensional gel is shown in Fig. 8.51

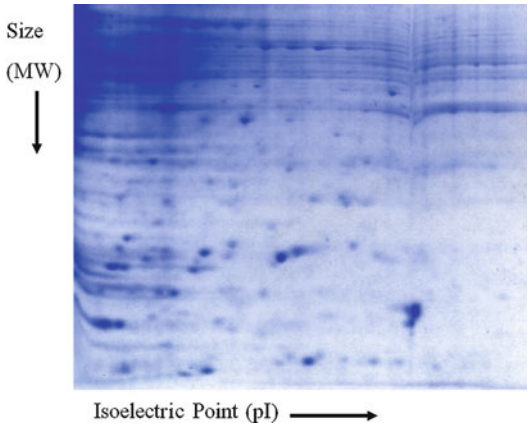


Fig. 8.51 Two-dimensional protein electrophoresis (polyacrylamide gel). Proteins are separated based on two different physical properties: isoelectric focusing is followed by standard separation based on size

- Procedures
  - Sample preparation: Proteins for two-dimensional electrophoresis are first extracted from tissue or cells using an appropriate amount of urea lysis buffer (8 M urea, 4% CHAPS, 65 mM DTE, 40 mM Tris, and a trace of bromophenol blue) depending on the size of tissue or the number of cells and strip size
  - Immobilized pH gradient (IPG) as first dimension
    - IPG gel strip rehydration: 100 ug to 15 mg of protein sample is pipetted into rehydration tray; IPG gel strips are positioned such that the gel of the strips is in contact with the protein samples (upside down). The gels and samples are covered with several ml of low-viscosity paraffin oil to prevent evaporation. The strips are then left at room temperature for rehydration. A minimum of 10 h is required for rehydration, and overnight is recommended
    - Sample application: The rehydrated IPG gels carrying the proteins are removed from the rehydration tray with forceps, rinsed with water, and positioned on a focusing tray as specified by manufacturer (e.g., Bio-rad, Amersham, Pharmacia Biotech)
    - Running conditions: Depending on different IEF equipment, follow manufacture's protocol for IEF run. For example, based on Pharmacia-Hoeffer Biotechnology AB, the voltage is linearly increased from 300 to 3,500 V during 3 h, followed by three additional hours at 3,500 V, whereupon the voltage is increased to 5,000 V. Focusing is carried out for a total of 100 kVh in an overnight run. After running, strips can be frozen (−20 °C) for several weeks (remove oil) or used immediately for the second dimension
    - IPG gel strips equilibration: After the first dimension of electrophoresis is carried out, the strips are equilibrated in order to resolubilize the proteins and to reduce disulfide bonds. The strips are transferred into an equilibration tray and equilibrated in equilibration buffer I (50 mM Tris-HCl pH 6.8, 6 M urea, 30% glycerol, 2% SDS, and 2% DTE) for 10 min and then equilibrated in equilibration buffer II (50 mM Tris-HCl pH 6.8, 6 M urea, 30% glycerol, 2% SDS, 2.5% iodoacetamide, and a trace of bromophenol blue) for another 10 min
  - SDS-PAGE as second dimension
    - Prepare resolving gels (SDS-PAGE) as described in Sect. 8.7.2
    - IPG strip transfer: After equilibration, the IPG strips are picked up using two pairs of clean forceps and are carefully placed atop the SDS-PAGE (resolving) gels facing the front of the gel cassettes, by carefully sliding them down the gel cassette via the plastic laminated side of the IPG strips. For consistency, the IPG strips are positioned with the basic side closer to the anode (Red/+) and the acidic side closer to the cathode (Black/–)
    - The tops of the gel cassettes are sealed with 0.5–1% agarose in electrophoresis buffer. Ensure that there is no formation of bubbles between the IPG strips and the SDS-PAGE gel. The combination of

the IPG strips and agarose avoids the need for a stacking gel
    - Running conditions: Run the gels at 40 mA/gel at 8–12 °C until the tracking dye reaches the bottom of the gels
- Protein detection
    - At the end of the second dimension run, the proteins can be detected by conventional staining techniques, such as Coomassie Brilliant Blue, Amido black, and silver stains. Silver staining is popular owing to its high sensitivity. The procedure for silver staining is described below (the gel should be gently agitated throughout the various steps)
    - After the SDS-PAGE is carried out, the gels are removed from the glass plates and washed in deionized water for 5 min
    - Soak in ethanol:acetic acid:water (40:10:50) for 1 h
    - Soak in ethanol:acetic acid:water (5:5:90) for 2 h or overnight
    - Wash in deionized water for 5 min
    - Soak in a solution containing glutaraldehyde (1%) and sodium acetate (0.5 M) for 30 min
    - Wash three times in deionized water for 10 min each
    - Soak twice in a 2,7 naphthalene-disulfonic acid solution (0.05%) for 30 min each
    - Rinse four times in deionized water for 15 min each
    - Stain in a freshly made ammoniacal silver nitrate solution for 30 min
    - After staining, the gels are washed four times in deionized water for 5 min each
    - The images are developed in a solution containing citric acid (0.01%) and formaldehyde (0.1%) for 5–10 min
    - When a slight background stain appears, development is stopped with a solution containing Tris (5%) and acetic acid (2%)
- Image analysis
    - Stained gels are scanned with a scanning device, and the proteins of interest are marked for excision
- Applications
    - Resolution and analysis of highly complex protein mixtures
    - Separation of the isoforms of a protein
    - Proteomic analysis
    - If sufficient sample is present on the gel (>300 μg of total protein), then proteins can be excised from the gel, subjected to in-gel proteolysis, and analyzed by mass spectrometry
- Advantages
    - The main advantage of using two-dimensional electrophoresis is the large mass range and the sheer amount of proteins that can be analyzed at any one time. Two-dimensional electrophoresis is particularly effective in the analysis of proteins within the mass range of 20–250 kDa and pI of 3–8 and can separate 2,000–3,000 proteins in one gel
    - Two-dimensional electrophoresis is the single best method for resolving highly complex protein mixtures
- Limitations
    - It is difficult to isolate proteins with isoelectric points outside of the range of 3.0–8.0. This is due to problems associated with creating stable pH gradients outside that range
    - Two-dimensional electrophoresis is a low-throughput, time-consuming process (3–4 days per run) that involves many steps and requires a high level of laboratory skill to obtain good results
    - Two-dimensional electrophoresis has diminished utility in the analysis of extremely acidic, basic, or hydrophobic proteins such as membrane-bound proteins and also in the analysis of smaller proteins and peptides (<15 kDa)

8.7.4.2 Protein Array

- General information
    - Protein microarray technology is a powerful emerging analytic strategy for interrogating the proteomes of tissues and cells

  - Measures the alterations in protein expression levels and detects posttranslational modification and mRNA processing events
  - There are several different types of protein arrays, including reverse-phase protein array, antibody array, and immunoblot array
  - The throughput can be low (<100 samples/run) or high (>100 samples/run), and the density of protein/antibody spots can be low (<100 elements/array) or high (>1,000 elements/array)
  - To detect proteins that are bound to the array, the samples must be labeled directly with a fluorophore or a hapten
  - The major limitation of antibody based array is the availability of high affinity and specific antibodies. Since over 6,000 antibodies are available and more will be available in future, this limitation become less important
- Protein array technologies
  - Reverse phase protein array
    - Made by arraying many protein samples or serum sample on a supporting material such as a glass slide. The proteins presented in sample can be probed with a small number of antibodies
  - Antibody arrays and microspot ELISA
    - Low density (9–100 elements/array). In these arrays, known antibodies are arrayed and used to capture antigens from unknown samples. To detect antigen that is bound to the array, the antigen either needs to be labeled directly with a fluorophore or a second binder/antibody can be used. The latter option creates a sandwich assay similar to a traditional ELISA only in a microspot format (hence, microspot ELISA)
  - Bead-based array (i.e., Luminex bead array)
    - Use multiple, different fluorescent beads, each spectrally resolvable from the other and each coated with a different capture antibody
      - The beads are incubated with a sample to allow protein binding to the capture antibodies, and the mixture is incubated with a mixture of detection antibodies, each corresponding to one of the capture antibodies
      - The detection antibodies are tagged to allow fluorescent detection
      - The beads are passed through a flow cytometer, and each bead is probed by two lasers: one to read the color, or identity, of the bead and another to read the amount of detection antibody on the bead
  - Immunoblot array (i.e., protein pathway array)
    - A powerful tool to analyze the expressed proteins with excellent flexibility and simplicity. It allows global screening of changes in protein expression and posttranslational modification (i.e., phosphorylation)
      - The focus of the pathway array is to determine the signaling pathways that control cancer development (initiation, promotion, progression, and metastasis)
      - The proteins selected for study in the array are highly expressed in cancer cells and are functionally linked to angiogenesis, apoptosis, cell cycle regulation, DNA repair, migration, proliferation, signaling, stem cell association, and transcription activity
      - The pathway array system has three integrated components (Fig. 8.52)
        - One-dimensional gel electrophoresis/multiplex protein immunoblot
        - Image acquisition and data analysis
        - Computational analyses that integrate the results with known protein–protein, cell signaling, and gene regulation cancer biology pathways
- Procedures
  - The total proteins are extracted from each fresh tissue sample or cell line

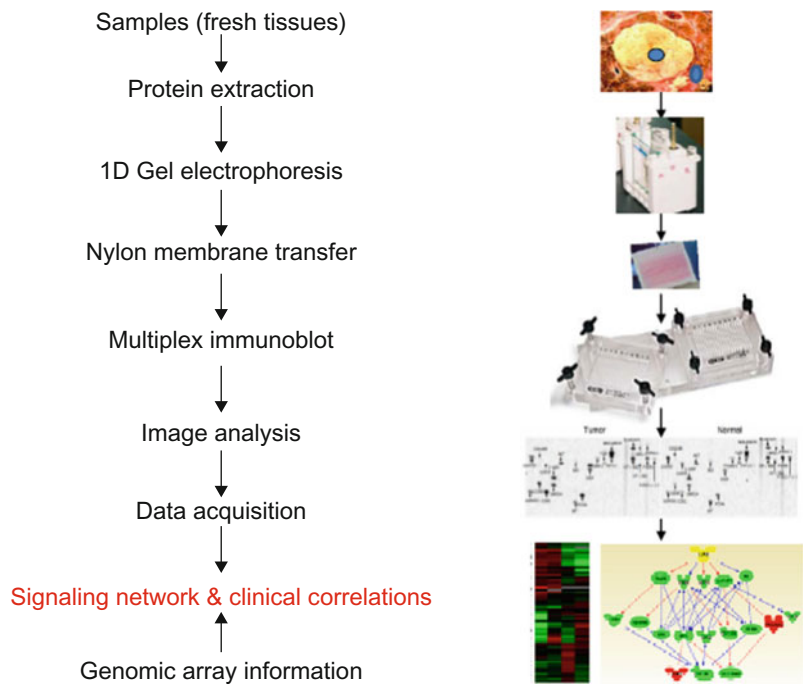


Fig. 8.52 Flowchart and schematic representation of protein pathway array for protein and phosphoprotein analysis

  - Proteins (200–300 ug) are loaded to single well across a SDS-PAGE and separated via electrophoresis
  - The proteins are then transferred to a nitrocellulose membrane and blotted using a Western blotting manifold that isolates 20 channels across the membrane
  - Each channel includes 2–4 antibodies (a total of 40–80 antibodies) for immunoblot
  - The proteins specific to the antibody can be detected using a chemiluminescent method
  - The images can be acquired using the ChemiDoc XRS System (Bio-Rad), and the correct band for each protein/phosphoprotein can be determined by molecular weight
  - The volume of each band can be recorded
  - The bound antibodies on the membrane can then be stripped off and blotted with another set of antibodies. This process can be repeated several times so that up 300–400 antibodies can be blotted on the same sample
- Advantages
  - A large percentage of recovered differentially expressed proteins and phosphoproteins (20–40% of the total proteins tested) compared to other gene expression- or proteomics-based approaches (2–6% of the genes or proteins)
  - A wide detection dynamic range (10–10^4 relative density unit) and ability to detect a 1.5-fold change between two samples
  - An improved sensitivity to detect low-abundance proteins (lower limit of detection = 0.1–1 ng and lower detection concentration = 0.05%)
  - High accuracy and reproducibility (the interrun and intrarun variations: CV = 25% and 35%, respectively, and an R square = 0.933 between runs)
  - The major limitation is the availability of high affinity and specific antibodies

8.7.4.3 Mass Spectrometry (MS)

- Mass spectrometry is an analytical tool which can be used to determine chemical structure on the basis of the mass/charge (m/z) ratio of molecular ions derived from a fragmented parent molecule. Mass

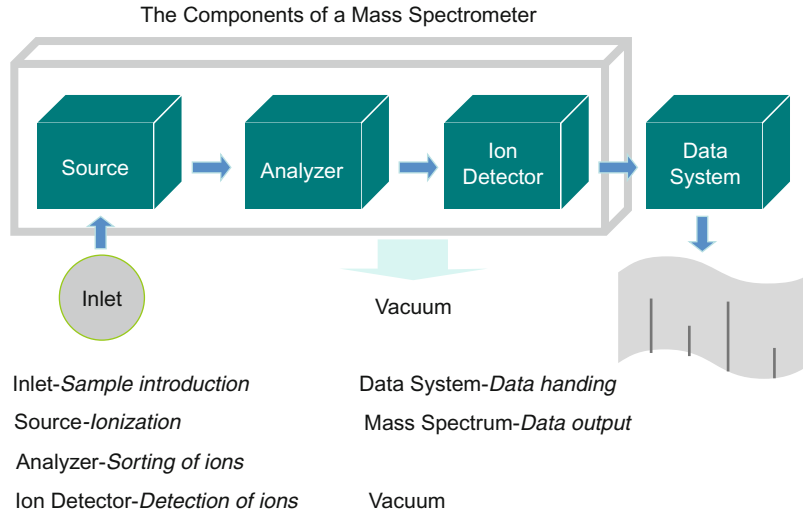


Fig. 8.53 The components of a mass spectrometer (MS)

spectrometer instruments have three fundamental components (Fig. 8.53). The first component is the ion source, which converts sample molecules into fragmented molecular ions. The second component is the mass analyzer, which resolves these ions based on their m/z ratios. The third component is the detector, which detects the ions resolved by the mass analyzer. In brief, the sample has to be introduced into the ionization source of the instrument. Once inside the ion source, the sample molecules are ionized. The resulting ions are extracted into the analyzer region of the mass spectrometer where they are separated according to m/z ratio. The separated ions are detected, and information pertaining to the m/z ratio and relative abundance of each ion is stored for presentation as a mass spectrum. Some instruments are capable of analyzing intact molecules with little fragmentation

- Two different types of instrument are used for most proteomics MS work: MALDI-TOF instruments, which measure peptide masses, and ESI-tandem MS instruments, which are used to obtain structure and sequence data. The two types operate in entirely different ways and generate different but complementary information. Indeed, the best-equipped proteomics laboratories have both types of instruments available

MALDI-TOF (Matrix-Assisted Laser Desorption Ionization-Time of Flight) Mass Spectrometry

- General information
  - The first term "MALDI" (matrix-assisted laser desorption ionization-time of flight) refers to the ion source and describes a method of ionization, whereas the term "TOF" (time of flight) refers to the mass analyzer
  - MALDI-TOF mass spectrometry is a method used for measuring the mass of a sample. For large samples such as biomolecules (proteins, peptides, oligosaccharides, and oligonucleotides), molecular masses can be measured to within an accuracy of 0.01% of the total molecular mass of the sample. For small organic molecules, the molecular masses can be measured to within an accuracy of 5 ppm or less, which is often sufficient to confirm the molecular formula of a compound
- Principle
  - In-gel digestion of proteins separated by two-dimensional electrophoresis
    - Before peptide masses can be obtained using MALDI-TOF, the proteins must be cleaved into peptides. This is because there are errors in the measurement of intact proteins. The greater the

mass of the protein, the greater the absolute magnitude of the error
- Proteins can be digested into smaller polypeptide fragments (5–10 amino acids) with a suitable enzyme. Trypsin is useful for MS studies because each proteolytic fragment contains a basic arginine (R) or lysine (K) amino acid residue and thus is eminently suitable for positive ionization MS analysis

– MALDI principle (Fig. 8.54)
  - The sample to be analyzed is dispersed in a large excess of a chemical matrix, which typically is composed of a small molecular weight organic molecule which functions as a chromophore that strongly absorbs applied laser energy. Common matrix compounds include sinapinnic acid (SA) for protein samples and alpha-Cyano-4-hydroxycinnamic acid (ACH) for peptide samples
  - The admixture of the sample and matrix is then spotted onto a small plate or slide and left to evaporate in air. The evaporation of residual water or other solvent from the sample allows the formation of a crystal lattice into which the peptide sample is integrated
  - The target (plate or slide) is then placed into the MALDI source. The source is equipped with a laser, which fires a beam of light at the target. The matrix chemicals absorb photons from the beam and become electronically excited. This excess energy is then transferred to the peptides or proteins in the sample, thereby ionizing them. The matrix also functions to help overcome molecular photodissociation of the sample ions induced by direct laser irradiation, thereby preventing unwanted fragmentation
  - The ionization process can produce either positive or negative ions, depending on the nature of the sample. Positive ionization is used in general for protein and peptide analyses. In positive ionization mode, the protonated molecular ions $(M+H)^+$ are usually the dominant species. Negative ionization can be used for the analysis of oligosaccharides and oligonucleotides. In negative ionization mode, the deprotonated molecular ions $(M–H)^-$ are usually the most abundant species

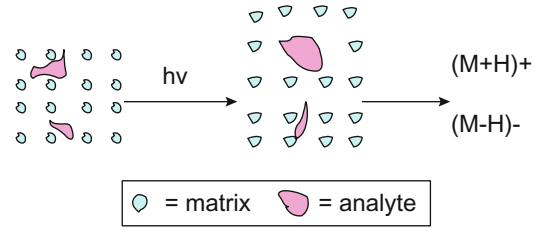


Fig. 8.54 Schematic presentation of MALDI

  - The ions formed in the MALDI source are then extracted and directed into the TOF mass analyzer

– TOF principle (Fig. 8.55)
  - The TOF analyzer separates ions according to their m/z ratios by measuring the time it takes for ions to travel through a field-free region known as the flight or drift tube, before striking a detector. The heavier ions are slower than the lighter ones. In this way, each molecule yields a distinct signal
  - TOF analyzers can operate in either of two ways. In "linear mode," formed ions are extracted from the MALDI source and then directly sent down the flight tube (TOF analyzer) to the detector (Fig. 8.55a). The resolution of instruments running in linear mode with continuous extraction of ions is, however, relatively poor ("resolution" refers to the extent to which the instrument can distinguish between ions of slightly different m/z values). In "reflecting mode," a reflectron (a device which uses static electric fields to alter the paths of ions) located at the end of the flight tube is used to compensate for differences in flight times of ions having the same m/z value (Fig. 8.55b). This results in ions of the same m/z value reaching the detector at

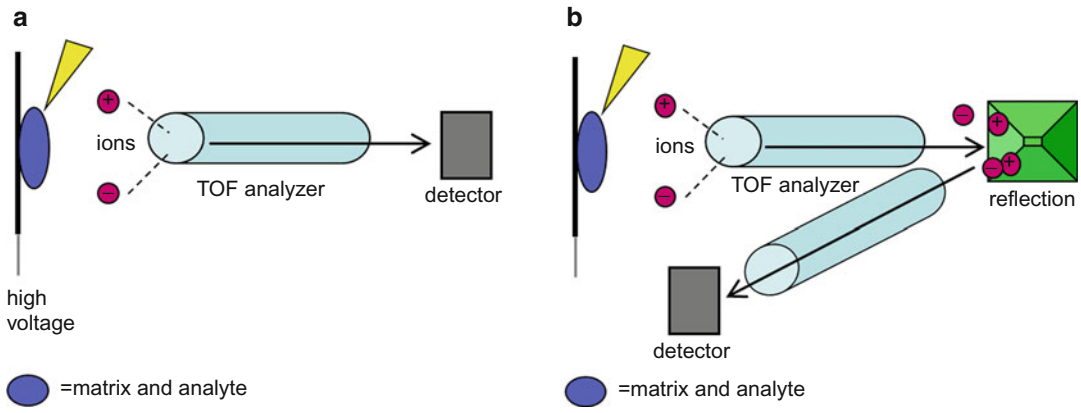


Fig. 8.55 Simplified schematic of MALDI-TOF mass spectrometry. (**a**) Linear mode and (**b**) reflecting mode

the same time. The reflectron dramatically improves resolution in TOF analyzers

- Procedures
  - Spot excision
    - Excise spots of interest from two-dimensional gel by using a clean pipette tip, cutting as close to the edge of the spot as possible, or by using a spot cutter
    - Transfer the gel spots into 0.5-ml siliconized tubes or a 96-well plate, one spot for one tube or one well
    - Rinse the spots using water; then remove water from tubes or wells
  - In-gel digestion of proteins
    - Add 50 µl of 50% acetonitrile/100 mM ammonium bicarbonate (ABC) buffer to each tube or well, and incubate for 5 min
    - Remove supernatant using pipette
    - Repeat steps 2(i) and 2(ii)
    - Add 50 µl of acetonitrile to each tube or well
    - Remove the acetonitrile
    - Dry in speed vacuum
    - Swell gel pieces in 10–20 µl 50 mM ABC for 4 min at 37 °C
    - Add equal volume of trypsin solution to soaked gel pieces
    - Digest at least 1 h at 37 °C
  - Extraction of tryptic peptides from gel
    - Add 50 µl of 60% acetonitrile/5% trifluoroacetic acid (TFA) to the gel pieces
    - Sonicate for 15 min in ice bath
    - Centrifuge to bring down liquid for 30 s
    - Transfer supernatant to 0.5-ml siliconized tube
    - Repeat steps 3(i)–3(iv), pool the supernatants
    - Dry pooled supernatants in speed vacuum
  - Sample preparation for MALDI analysis
    - The above supernatants may be concentrated by speed vacuum to dryness and then brought up in no more than 5–10 µl of 0.1% TFA/60% acetonitrile. A sample concentration of 1 mg/ml is ideal. Usually from 1 to 10 picomoles of sample is required for analysis
    - Take an aliquot (1–2 µl) of this peptide solution and mix with an equal volume of a solution containing a vast excess of a matrix. The two most commonly used matrices are alpha-Cyano-4-hydroxycinnamic acid (ACH) and sinapinic acid (SA). ACH is usually dissolved at 30 mg/ml in 0.1% TFA/60% acetonitrile. SA can be prepared at 50 mg/ml in 0.1% TFA/60% acetonitrile
  - MALDI-TOF analysis
    - The mixture is applied to a stainless steel sample target and allowed to dry. It is essential that all spots on the target be completely dry before the target is inserted into the sample chamber.

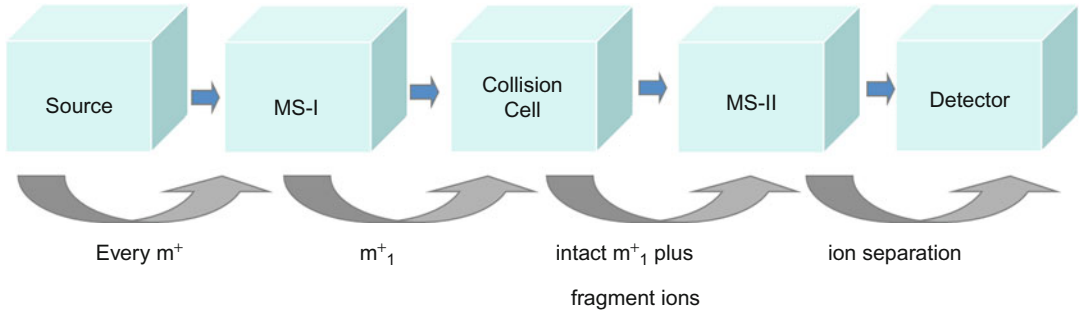


Fig. 8.56 Simplified schematic of tandem mass spectrometry

Moisture on the target leads to rapid degradation of the electrodes
- Insert target into MALDI-TOF MS instrument
- With the target under high vacuum, the laser is fired. The energy arriving at the sample/matrix surface is optimized
- Data is accumulated until an m/z spectrum of reasonable intensity has been amassed
- The m/z scale of the MS is calibrated with a known sample that can either be analyzed independently (external calibration) or premixed with the sample and matrix (internal calibration)

- Applications
  - MALDI-TOF MS is used for molecular mass measurements of various analytes such as peptides, proteins, oligosaccharides, and oligonucleotides
  - It is an important analytical tool in proteomics
  - It is also used in the analysis of the products of peptide synthesis
  - It is used as a method of N-terminal and C-terminal protein/peptide sequencing
- Advantages
  - MALDI-TOF MS is a very sensitive technique, allowing detection at the low fmole level
  - Produces highly accurate data with high resolution
  - MALDI-TOF instruments are among the easiest of MS instruments to operate
  - MALDI-TOF MS is well suited to high-throughput proteomics work
- Limitations
  - MALDI-TOF instrument is best suited to measuring peptide masses. This type of information, although useful for protein identification, is nevertheless limited
  - The success of MALDI-TOF analyses is highly dependent on the quality of the sample. Contamination of the peptide digest sample with significant levels of detergents, buffer salts, metals, or organic modifiers may greatly inhibit peptide ionization in the MALDI source

Tandem Mass Spectrometry (MS/MS)

- General information
  - The tandem mass spectrometer is an instrument consisting of two mass spectrometers in series connected by a chamber known as a collision cell. The sample to be examined is essentially sorted and weighed in the first mass spectrometer, then broken into pieces in the collision cell, and the piece or pieces subsequently sorted and weighed in the second mass spectrometer (Fig. 8.56)
  - The term "tandem mass spectrometry" is often abbreviated as "tandem MS" or "MS/MS"
  - Tandem MS can be used for structural and sequencing studies
  - Tandem mass spectrometry was first introduced in the 1970s and was quickly accepted in the analytical community

- Principles
  - The underlying principle of the MS/MS method is the use of filiation relations: The first mass spectrometer (MS1) is used to select, from the primary ions, those of a particular m/z value which then pass into the fragmentation region. The ion selected by the MS1 is the parent ion or a precursor ion and can be a molecular ion or an ion resulting from primary fragmentation. Dissociation occurs in the fragmentation region. The daughter ions are analyzed in the second mass spectrometer (MS2). In fact, MS1 can be viewed as an ion source for MS2
    - In MS1, ionization is produced by electron impact. One ion is then selected (by a quadrupole, magnetic sector, or ion cyclotron resonance mechanism). The selected ion proceeds through the outlet to the fragmentation region
    - In the fragmentation region, there is a neutral gas (e.g., He, Ar, and N2) under high pressure. The ions selected interact with these molecules of gas. During the collision, the kinetic energy of ions is transformed into internal energy which leads to fragmentation into daughter ions
    - In MS2, the daughter ions are detected and the final spectrum shows the peaks of the selected ion and all its daughters
  - Instrumentation for MS/MS: A tandem mass spectrometer can be thought of as two mass spectrometers in series connected by a chamber that can break an ion into pieces
    - Types of sources: The abbreviations ESI (electrospray), FAB (fast atom bombardment), or MALDI (matrix-assisted laser desorption ionization) before the term "tandem MS" indicate the manner in which the analyte is introduced into the tandem mass spectrometer
      - Dissolution–nebulization sources (ESI, "electrospray"): This type of source is used to produce ions from molecules in solution. The solution is placed into a metallic capillary. An electric field is applied between the capillary's point and an electrode; multicharged droplets are produced and accelerated toward the electrode
      - Ionization desorption sources: based on secondary emission. The bombardment of a solid or liquid sample with a primary beam (ion, atom, or photon) induces the secondary emission of particles: electrons, neutral particles, and ions. Only these particles are analyzed by mass spectroscopy. The name of the source depends on the nature of the incident beam: ions in SIMS ("secondary ion mass spectrometry"), atoms in FAB ("fast atom bombardment"), photons in LD ("laser desorption") and MALDI ("matrix-assisted laser desorption ionization")
    - Analyzers: A tandem mass spectrometer is a mass spectrometer that has more than one analyzer, in practice usually two. The two analyzers are separated by a collision cell into which an inert gas (e.g., argon or xenon) is admitted to collide with the selected sample ions bringing about their fragmentation. The analyzers can be of the same or different types, the most common combinations being
      - Quadrupole – quadrupole
      - Magnetic sector – quadrupole
      - Magnetic sector – magnetic sector
      - Quadrupole – time of flight
    - Collision cell: Collisional activation can be divided into two categories, involving high or low energy, to which different types of collision cells are appropriate. In sector instruments, where high-energy collisions are most common, the cell is usually a tight chamber of 1–3 cm length with entrance and exit slits which transmit the ion beam. Good pumping is essential to maintain a low pressure outside the cell. In some instruments, the collision cell is electrically insulated from the

mass spectrometer and can be held at a high potential to retard the ion beam and reaccelerated it on exit. This allows control of the collision energy and also reduces the kinetic energy spread of daughter ions formed

- Types of detectors: The detector's purpose is to translate ion arrival into electrical signals measured by an electronic system. There are two different classical types of detector
  - Electron multiplying: The principal sorts of electron multiplying detectors are the "channeltron" and the "microchannel plate." In these devices, ions impact upon a surface composed of half conductors. The impact releases electrons which are accelerated to another surface where additional electrons are released, thereby amplifying the signal. The resulting current pulse is proportional to the original signal intensity
  - Photodetectors: Electrons are created in the same manner as above but then interact with a phosphorescent surface which generates photons. The photons are recovered; their number is proportional to the signal intensity

- Procedure
  - Spot excision: same as "d-2-1 procedures 1"
  - In-gel digestion of proteins: same as "d-2-1 procedures 2"
  - Extraction of tryptic peptides from gel: same as "d-2-1 procedures 3"
  - Sample preparation for tandem MS analysis: For maximum sensitivity, thoroughly desalt the tryptic digest before analysis using HPLC, which gives considerable advantage in sensitivity
    - HPLC separations are performed with a CapLC (Waters) using trapping guard and analytical columns in series
    - With the eluent of the guard column directed to waste, sample (20 ml) is loaded at a flow rate of 30 ml/min with pump C (aqueous 0.1% formic acid), washed for 5 min at 30 ml/min, and then the 10 port valve is actuated so the acetonitrile-water-formic acid mixture from pumps B and A is directed to the trapping column to elute the peptides in-line with the analytical column, which is preequilibrated with 5% solvent B from the last run and the flow reduced to 200 nl/min. The flow (1 ml/min) from pump A and B is reduced to 200 nl/min by splitting the flow with a Valco Tee and 5 m of fused silica (75-mm inner diameter) on the waste arm
  - Tandem MS analysis
    - The eluent from the analytical column is ionized by electrospray ionization (ESI) with tandem MS analysis
    - Carry out the analysis using automated data analysis. The SEQUEST computer program is used to match the sequence information in the spectra to a database of known protein sequences. SEQUEST is a powerful suite of programs that can take the information in the peptide MS/MS spectra and correlate it to a database of known DNA or protein sequences
- Applications
  - An important application of tandem mass spectrometry is protein identification by peptide sequencing
  - Oligonucleotide sequencing can also be achieved by tandem mass spectrometry
  - Tandem mass spectrometry has been used for structure elucidation of unknowns and for analysis of complex mixtures
- Advantages
  - Tandem mass spectrometry does not require extensive sample purification
  - It can produce protein sequence data that cannot be obtained using other mass spectrometry methods
  - Produces data much faster than other mass spectrometry methods due to its use of nonsequential sample analysis

  - Tandem mass spectrometry is a very sensitive technique
- Limitations
  - With this methodology it is difficult to elucidate the identity of N-terminal and C-terminal residues in a peptide
  - The instrument is expensive and complex. Performing analyses demands a very high level of expertise

Further Reading

Andras SC, Power JB, Cocking EC, et al. Strategies for signal amplification in nucleic acid detection. Mol Biotechnol. 2001;19:29–44.

Arshad MF, Dunn FJ, Vega R, et al. Progress in developing improved programs for pulsed field agarose gel electrophoresis of DNA. Electrophoresis. 1993;14:344–8.

Barany F. Genetic disease detection and DNA amplification using cloned thermostable ligase. Proc Natl Acad Sci USA. 1991;88:189–93.

Bogner PN, Killeen, AA. An overview of nucleic acid chemistry, structure, and function: The foundation of molecular biology. In: Coleman WB, Tsongalis GJ, editors. Molecular diagnostics for the clinical laboratorian. 2nd ed. Totowa, NJ: Humana Press; 2006. p. 25–29.

Cariello NF, Skopek TR. Mutational analysis using denaturing gradient gel electrophoresis and PCR. Mutat Res. 1993;288:103–12.

Collection, transport, preparation, and storage of specimens for molecular methods; approved guideline. Wayne: Clinical and Laboratory Standards Institute, MM13-A. 2006;25(31).

Darby IA, Bisucci T, Desmouliere A, et al. In situ hybridization using cRNA probes: isotopic and nonisotopic detection methods. Methods Mol Biol. 2006;326:17–31.

Dovichi NJ. DNA sequencing by capillary electrophoresis. Electrophoresis. 1997;18:2393–9.

Dunbar B. Troubleshooting and artifacts in two-dimensional polyacrylamide gel electrophoresis. In: Dunbar BS, editor. Two-dimensional electrophoresis and immunological techniques. New York: Plenum; 1987. p. 173–95.

Eldering E, Spek CA, Aberson HL, et al. Expression profiling via novel multiplex assay allows rapid assessment of gene regulation in defined signaling pathways. Nucleic Acids Res. 2003;31:e153.

Feldman RS, et al., editors. Principles of neuropsychopharmacology. Sunderland: Sinauer Associates; 1997. p. 31–5.

Fenyo D, Qin J, Chait BT. Protein identification using mass spectrometric information. Electrophoresis. 1998;19:998–1005.

Fodde R, Losekoot M. Mutation detection by denaturing gradient gel electrophoresis (DGGE). Hum Mutat. 1994;3:83–94.

Gillespie PG, Hudspeth AJ. Chemiluminescence detection of proteins from single cells. Proc Natl Acad Sci USA. 1991;88:2563–7.

Godde R, Akkad DA, Arning L, et al. Electrophoresis of DNA in human genetic diagnostics – state-of-the-art, alternatives and future prospects. Electrophoresis. 2006;27:939–46.

Goldsby RA, Kindt TJ, Osborne BA, Kuby J. Enzyme-linked immunosorbent assay. In: Marcus WH, editor. Immunology. 5th ed. New York: Freeman; 2003. p. 148–50.

Hames BD, Rickwood D, editors. Gel electrophoresis of proteins: a practical approach. 2nd ed. New York: Oxford University Press; 1990.

Harlow E, Lane D. Immunoblotting in antibodies: a laboratory manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1988. p. 471–510.

Hedges DJ, Guettouche T, Yang S, et al. Comparison of three targeted enrichment strategies on the SOLID sequencing platform. PLoS One. 2011;6: e18595.

Hicks DG, Longoria G, Pettay J, et al. In situ hybridization in the pathology laboratory: general principles, automation, and emerging research applications for tissue-based studies of gene expression. J Mol Histol. 2004;35:595–601.

Hochstrasser DF, Harrington MC, Hochstrasser AC, et al. Methods for increasing the resolution of two-dimensional protein electrophoresis. Anal Biochem. 1998;173:424–35.

Hubbard RA. Human papillomavirus testing methods. Arch Pathol Lab Med. 2003;127:940–5.

Iftner T, Villa LL. Chapter 12: Human Papillomavirus Technologies. J Natl Cancer Inst Monogr. 2003; (31):80–88.

Immunoglobulin and T-cell receptor gene rearrangement assays; approved guideline. NCCLS, MM2-A. 1995;15(18).

Jalali GR, Vorstman JAS, Errami AB, et al. Detailed analysis of 22q11.2 with a high density MLPA probe set. Hum Mutat. 2008;29:433–40.

Kemp G. Capillary electrophoresis: a versatile family of analytical techniques. Biotechnol Appl Biochem. 1998;27:9–17.

Li J, Milbury CA, Li C, et al. Two-round coamplification at lower denaturation temperature-PCR (COLD-PCR)-based sanger sequencing identifies a novel spectrum of low-level mutations in lung adenocarcinoma. Hum Mutat. 2009;30:1583–90.

Maniatis T, Jeffrey A, van deSande H. Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. Biochemistry. 1975;14:3787–94.

Manning SD. Molecular epidemiology of Streptococcus agalactiae (group B Streptococcus). Front Biosci (Review). 2003;8:s1–s18.

Margulies M, Egholm M, Altman W, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005;441:376–80.

Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res. 2008;18:1509–17.

Milbury CA, Correll M, Quackenbush J, et al. COLD-PCR enrichment of rare cancer mutations prior to targeted amplicon resequencing. Clin Chem. 2012;58:580–90.

Mullis KB, Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol. 1987;155:335–50.

Nataraj AJ, Olivos-Glander I, Kusukawa N, et al. Single-strand conformation polymorphism and heteroduplex analysis for gel-based mutation detection. Electrophoresis. 1999;20:11771185.

Notomi T, Okayama H, Masubuchi H, et al. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. 2000;28:E63.

Nygren AO, Ameziane N, Duarte HM, Vijzelaar RN, Waisfisz Q, Hess CJ, et al. Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. Nucleic Acids Res. 2005;33:e128.

Olson MV. Separation of large DNA molecules by pulsed-field gel electrophoresis. A review of the basic phenomenology. J Chromatogr. 1989;470:377–83.

Patterson SD. Matrix-assisted laser-desorption/ionization mass spectrometric approaches for the identification of gel-separated proteins in the 5–50 pmol range. Electrophoresis. 1995;16:1104–14.

Porchet N, Aubert JP. Northern blot analysis of large mRNAs. Methods Mol Biol. 2000;125:305–12.

Porstmann T, Kiessig ST. Enzyme immunoassay techniques, an overview. J Immunol Methods. 1992;150: 5–21.

Pourzand C, Cerutti P. Genotypic mutation analysis by RFLP/PCR. Mutat Res. 1993;288:113–21.

Princivalle AP, Parker RM, Dover TJ, et al. mRNA: detection by in situ and northern hybridization. Methods Mol Biol. 2005;306:51–91.

Ragoussis J, Elvidge G. Affymetrix GeneChip system: moving from research to the clinic. Expert Rev Mol Diagn. 2006;6:145–52.

Reinartz J, Bruyns E, Lin JZ, Burcham T, Brenner S, Bowen B, et al. Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms. Brief Funct Genomic Proteomic. 2002;1:95–104.

Richter E, et al. Mycobacterium avium subsp. paratuberculosis infection in a patient with HIV. Emerg Infect Dis. 2002;8(7):729–31

Rogers Y, Venter JC. Genomics massively parallel sequencing. Nature. 2005;437:326–7.

Schweitzer B, Predki P, Snyder M. Microarrays to characterize protein interactions on a whole-proteome scale. Proteomics. 2003;3:2190–9.

Serwer P. Sieving of double-stranded DNA during agarose gel electrophoresis. Electrophoresis. 1989;10:327–31.

Smith JA. Electroelution of proteins from stained gels. In: Ausubel FA, Brent R, Kingston RE, et al., editors. Current protocols in molecular biology. New York: Wiley; 1993. p. 10.5.1–10.5.5.

Southern EM. Blotting at 25. Trends Biochem Sci. 2000;25:585–8.

Stoneking M, Hedgecock D, Higuchi RG, et al. Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. Am J Hum Genet. 1991;48:370–82.

Stults JT. Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS). Curr Opin Struct Biol. 1995;5:691–8.

Summerer D. Enabling Technologies of genomic-scale sequence enrichment for targeted high-throughput sequencing. Genomics. 2009;94(6):363–8.

Tang YW, Stratton CW, editors. Advanced techniques in diagnostic microbiology. New York: Springer; 2010. p. 222.

Thompson JF, Steinmann KE. Single molecule sequencing with a HeliScope genetic analysis system. Curr Protoc Mol Biol. 2010;Chapter 7:Unit7.10.

Tomita N, Mori Y, Kanda H, et al. Loop-mediated isothermal amplification (LAMP) of Gene sequences and simple visual detection of products. Nat Protoc. 2008;3:877–82.

Urdea MS. Branched DNA signal amplification. Nat Biotechnol. 1994;12:926–8.

Vandenbroucke I, Van Marck H, Verhasselt P, et al. Minor variant detection in amplicons using 454 massive parallel pyrosequencing: experiences and considerations for successful applications. Biotechniques. 2001;51:167–77.

Wulfkuhle JD, Aquino JD, Calverts VD, et al. Signal pathway profiling of ovarian cancer from human tissue specimens using reverse-phase protein microarrays. Proteomics. 2003;3:2085–90.

Clinical Flow Cytometry in Molecular Genetic Pathology

9

Magdalena Czader

Contents

9.1 Definition and Applications of Flow Cytometry 302

9.2 Technical Aspects of Flow Cytometry 302
9.2.1 Principle and Instrumentation 302
9.2.2 Sample Processing 304
9.2.3 Selection of Antibody Panel 304

9.3 Analysis and Interpretation of Flow Cytometric Data 305

9.4 Basic Cell Populations Identified by Flow Cytometry 307
9.4.1 Granulocytic Lineage 308
9.4.2 Monocytic Lineage 308
9.4.3 Erythroid Lineage 309
9.4.4 Megakaryocytic Lineage 309
9.4.5 Lymphoid Lineage 309

9.5 Flow Cytometric Analysis of Myeloid Disorders 311
9.5.1 Acute Myeloid Leukemias 312
9.5.2 Chronic Myeloproliferative Neoplasms and Myelodysplastic Syndrome (MDS) 317

9.6 Flow Cytometric Analysis of Lymphoid Neoplasms (Lymphoblastic Leukemia/Lymphoma and Mature Lymphoid Neoplasms) 319
9.6.1 B Lymphoblastic Leukemia/Lymphoma (B-LL) 319
9.6.2 T Lymphoblastic Leukemia/Lymphoma (T-LL) 321
9.6.3 Mature Lymphoid Neoplasms 321

9.7 Other Clinical Applications of Flow Cytometry 330
9.7.1 Primary and Secondary Immunodeficiencies 330
9.7.2 Paroxysmal Nocturnal Hemoglobinuria (PNH) 331
9.7.3 Stem Cell Transplantation 332
9.7.4 Novel Applications of Flow Cytometry 333

Further Reading 333

M. Czader, MD, PhD
Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_10,

9.1 Definition and Applications of Flow Cytometry

- Flow cytometry is the technique that measures the physical and antigenic properties of particles
- Any particle that can be suspended in a fluid, that is, cells, chromosomes, and individual molecules, can be detected and characterized by flow cytometry
- The most significant discovery that led to the advancement of flow cytometry and its subsequent widespread application to clinical practice was the development of monoclonal antibodies, for which Georges J.F. Köhler and César Milstein received a Nobel Prize in 1984
- Currently, immunophenotyping of hematolymphoid neoplasms using fluorochrome-conjugated monoclonal antibodies is the most common clinical application of flow cytometry
- In contrast to other applications of monoclonal antibodies, such as immunohistochemistry or Western blotting, flow cytometry examines antigens in their native (nonfixed) state
- Multicolor immunophenotyping is the current standard in clinical flow cytometry. Most commonly, 4–10 antibodies are analyzed at the same time. However, technical advances allow for the simultaneous detection of up to 17 antigens on an individual cell
- The most common applications of clinical flow cytometry include
  - Diagnosis and subclassification of malignant hematologic disorders such as leukemias and lymphomas
  - Detection of minimal residual disease in acute leukemia
  - Enumeration of T cell subsets for followup of HIV-positive patients
  - Determination of immunophenotypic/functional abnormalities in congenital immunodeficiencies
  - Enumeration of hematopoietic stem cells for bone marrow transplantation
  - Diagnosis of platelet disorders
  - Detection of fetal hemoglobin in fetomaternal hemorrhage
- The technique of flow cytometry can also be applied to
  - Cell sorting
  - Detection of chromosomal abnormalities based on in situ hybridization or polymerase chain reaction (PCR)
  - Functional assays
    - Proliferation
    - Apoptosis
    - Calcium efflux
    - Phosphorylation (cell signaling)

9.2 Technical Aspects of Flow Cytometry

9.2.1 Principle and Instrumentation

- Flow cytometry measures light scattering and fluorescence of individual particles as they are illuminated by a light (laser) source
  - In flow cytometer, individual particles are suspended in a fluid and pass one by one in front of a light source (Fig. 9.1)
  - As particles are illuminated, they scatter light and emit fluorescent signals
    - Light scattering
      - Forward scatter signal (FSC) measures cross-sectional area and is roughly proportional to size of a particle
      - Side scatter signal (SSC) reflects the internal complexity of a cell (cytoplasmic granules, vacuoles, and organelles). Cell size and refractive index may contribute to SSC characteristics
    - Fluorescence
      - The principle of fluorescence is illustrated in the simplified Jablonski diagram (Fig. 9.2): as an electron in its ground state absorbs light, it is raised to the excited state. The excess energy is emitted as nonradiative transition in a process of internal conversion and vibrational relaxation.

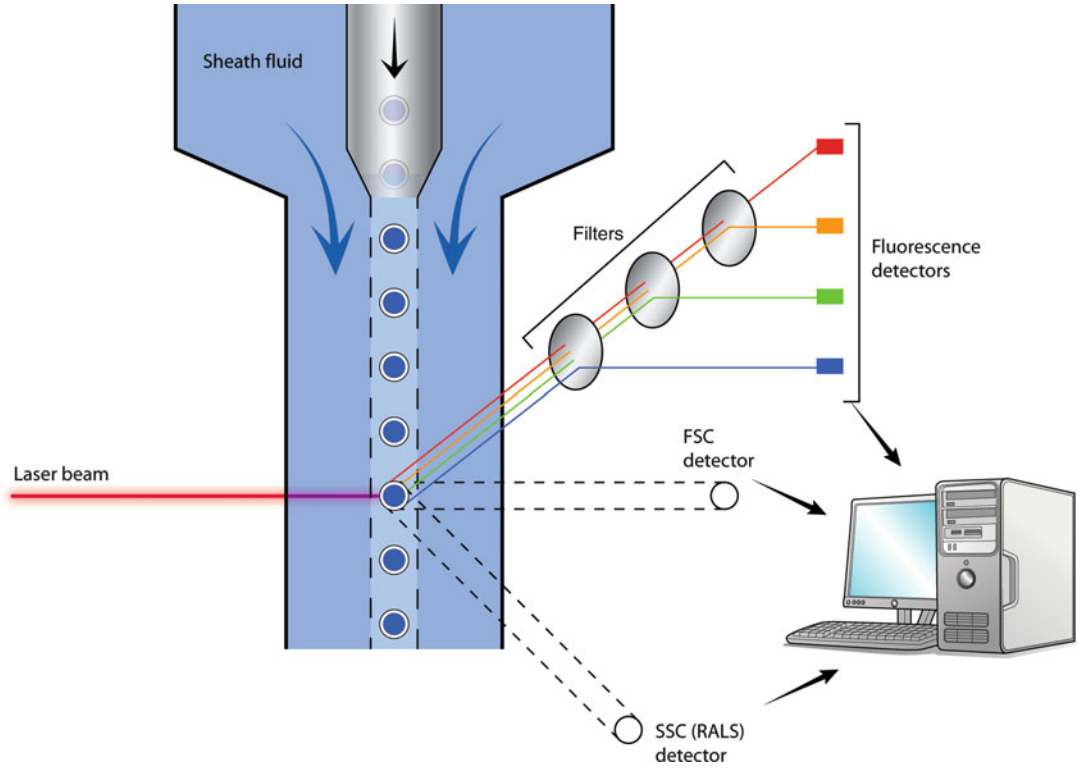


Fig. 9.1 Diagram of flow cytometer. Cell suspension is injected into sheath fluid under pressure, which positions the cells in a single file in the center of the stream for interrogation by the laser. FSC and SSC signals and fluorescent signals of specific wavelength are recorded by separate detectors

Upon subsequent return to the ground state, the fluorescence is emitted. Since some energy is lost during nonradiative transitions, the energy content of emitted fluorescence is lower than the energy absorbed, resulting in the emission at a longer wavelength than the absorption (Stokes shift)

– Each fluorochrome is characterized by a distinct spectral pattern of absorption and emission (fluorescence). The fluorochromes must be specifically selected to absorb a certain wavelength of light emitted by the laser available in the instrument (some flow cytometers are equipped with more than one laser). Currently, a wide variety of monoclonal antibodies conjugated to various fluorochromes are available, which allows simultaneous detection of multiple antibodies bound to a single cell. However, great care should be taken to select fluorochromes with minimal overlap in the emission spectra to optimally resolve individual antibodies

- Flow cytometer consists of fluidics, a light source (laser), a detection system, and a computer (Fig. 9.1)
- Steps in the flow cytometric analysis of the sample:
  – Aspiration of the stained cell suspension into a stream of sheath fluid
  – Alignment of a single cell file centrally in the sheath fluid through the hydrodynamic focusing
  – Illumination of cells passing individually in front of the laser source

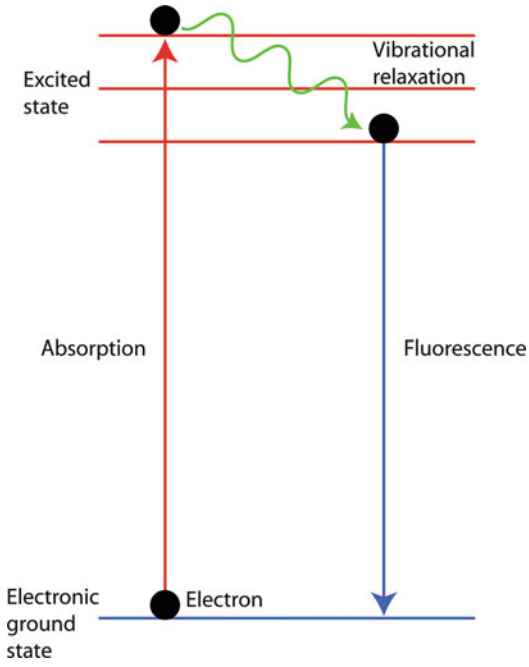


Fig. 9.2 Jablonski diagram. As a result of energy absorption, electrons are raised to the excited state. The energy is emitted in a process of internal conversion or vibrational relaxation and subsequently as fluorescence when electrons return to their ground state

  - Registration of light scatter and fluorescence signals from individual cells by dedicated photodetectors (separate detectors for light scatter and each fluorochrome; partitioning into different wavelengths is achieved by a series of dichroic mirrors)
  - As the sample is run, the data is digitized and simultaneously displayed and stored for subsequent analysis

9.2.2 Sample Processing

- Any specimen in a form of single cell suspension is suitable for flow cytometric analysis
- Most common clinical samples analyzed by flow cytometry include
  - Bone marrow and peripheral blood (collected with an anticoagulant, i.e., sodium heparin, ethylenediamine tetra acetic acid, or acid citrate dextrose)
  - Solid tissues (lymph nodes and extranodal samples suspected to harbor hematologic malignancy should be submitted in culture media such as RPMI 1640 to maintain viability and subsequently mechanically dissociated usually by mincing with a scalpel and filtering to yield a single cell suspension)
  - Body cavity fluids
- Quality of the sample is critical for accurate analysis
  - Prolonged transport or transport in inappropriate conditions may render a sample unsuitable for analysis
  - Peripheral blood and bone marrow specimens should be processed within 24–48 h from the time of collection and, if transported, should be kept at room temperature
  - Certain samples, such as body cavity fluids or specimens with a high proliferation rate, may require even shorter time intervals between collection and processing
- Steps in sample processing:
  - Hypotonic lysis for specimens with an admixture of red blood cells
  - Determination of the cellularity and viability of all submitted samples
    - Cell count can be obtained using flow cytometry with standardized beads or using automated cell counters
    - Flow cytometry of a sample stained with DNA dyes (e.g., propidium iodide, 7-AAD) or a manual method utilizing trypan blue exclusion can be used to test viability
  - Preparation of a cytospin for the morphologic inspection of the cell suspension
  - Staining with a cocktail of fluorochrome-conjugated monoclonal antibodies (both surface, i.e., membrane bound, and intracellular antigens can be analyzed)

9.2.3 Selection of Antibody Panel

- Comprehensive antibody panels with multiple markers for myeloid and lymphoid lineage are recommended by 2006 Bethesda International

Consensus on the immunophenotypic analysis of hematolymphoid neoplasia

- Antibody panels are designed to identify multiple cell subpopulations expected to be present in the sample. Both terminally differentiated cell populations and successive developmental stages should be covered
- Numerous hematopoietic cell antigens and the corresponding antibodies have been cataloged by Workshops on Human Leukocyte Differentiation Antigens (HLDA) held regularly since 1982
- These workshops provide a forum for reporting new antigens/antibodies and defining a cluster of antibodies, which recognize the same antigen (cluster of differentiation [CD], Tables 9.1 and 9.2). Consecutive numbers are assigned to each new reported antigen. The recent, HLDA9 workshop, currently known as HCDM for Human Cell Differentiation Molecules, brought the number of characterized antigens to 350
- The selection of an antibody panel is based on the properties of the antibodies and fluorochromes
  - The selection of an antibody clone is often critical because antibodies can recognize different epitopes on the same antigen with different distributions in hematopoietic cells (e.g., antibodies against CD34 antigen recognize three different epitopes of this molecule). Additionally, clones may differ in binding capacity
  - The choice of the fluorochrome should be related to the density of a given antigen. For example, when only a few molecules of the antigen are expected, one should select the antibody coupled with a strong fluorochrome to enhance the detection
- Selected markers most commonly analyzed by flow cytometry in hematopathology are presented in Tables 9.1 and 9.2
- Dependent on the laboratory, a comprehensive antibody panel can be analyzed upfront. Alternatively, a limited screening panel is utilized initially with the subsequent addition of selected markers

Table 9.1 Lineage-associated markers commonly analyzed in clinical flow cytometry

Immature	Granulocytic/monocytic	Erythroid
CD34	CD33	CD71
CD117	CD13	CD235a
TdT	CD15	
	CD16	
	CD14	
Megakaryocytes	B cell	T cell
CD41	CD19	CD2
CD42b	CD20	CD3
CD61	CD22	CD4
	κ-light chain	CD5
	λ-light chain	CD7
		CD8

9.3 Analysis and Interpretation of Flow Cytometric Data

- The evaluation of flow cytometric data is based upon analysis of the patterns of antigen expression presented graphically in the form of scattergrams and histograms and their comparison to normal expression (maturation and lineage specific)
- The detailed knowledge of immunophenotypic characteristics of normal hematopoietic differentiation, as well as normal variations (e.g., age-related), is critical for optimal interpretation of flow cytometric data
- Both qualitative (positive/negative, homogeneous vs heterogeneous expression in a given population) and semiquantitative information on antigen expression (low/moderate/high intensity) should be recorded since patterns of antigen expression are diagnostically significant
- The concept of gating:
  - Cells with similar physical properties (size, complexity/granularity, and the presence/absence of a specific antigen) form clusters on the displays of flow cytometric data
  - Gate is a borderline that identifies these clusters of cells
  - In the gating process, a population of interest is selected (outlined with a cursor) for further analysis (i.e., to determine antigen

Table 9.2 Select hematolymphoid antigens commonly used in clinical flow cytometry

Cluster of differentiation	Cellular expression
CD1a	Precursor T cells, Langerhans cells
CD2	Precursor and mature T cells, NK cells, thymic B cells
CD3	Precursor and mature T cells
CD4	Precursor T cells, helper T cells, monocytes
CD5	Precursor and mature T cells, subset of B cells (B1a cells)
CD7	Precursor and mature T cells, NK cells
CD8	Precursor T cells, suppressor/cytotoxic T cells, subset of NK cells
CD10	Precursor B cells, germinal center B cells, granulocytes
CD11b	Granulocytic and monocytic lineage, NK cells, subset of T- and B cells
CD13	Granulocytic and monocytic lineage
CD14	Mature monocytes, neutrophils (dim expression)
CD15	Granulocytic and monocytic lineage
CD16	Granulocytic and monocytic lineage, NK cells
CD19	Precursor and mature B cells
CD20	Precursor and mature B cells
CD22	Precursor and mature B cells
CD31	Megakaryocytes, platelets, leukocytes
CD33	Granulocytic and monocytic lineage
CD34	Hematopoietic stem cells
CD36	Megakaryocytes, platelets, erythroid precursors, monocytes
CD38	Hematopoietic cells including activated lymphocytes and plasma cells
CD41	Megakaryocytes, platelets
CD42b	Megakaryocytes, platelets
CD45	Hematopoietic cells
CD56	NK cells, subset of T cells
CD61	Megakaryocytes, platelets
CD62P	Megakaryocytes, platelets
CD63	Megakaryocytes, platelets
CD64	Granulocytic and monocytic lineage
CD71	High density on erythroid precursors, low density on other proliferating cells
CD79a	Precursor and mature B cells
CD117	Hematopoietic stem cells and mast cells

expression for the selected – "gated" population)

– Gating can also be applied at the time of data acquisition (so-called live-gating) to selectively collect high number of cells from a specific subpopulation, for example, CD19 positive cells, to facilitate the detection of a small number of monoclonal B cells
– For diagnostic purposes, the data is most commonly collected ungated, that is, all events detected by a flow cytometer are recorded, to comprehensively analyze the entire sample and retain internal positive and negative controls. A separate portion of the sample can be live-gated to better visualize specific cell populations (e.g., lymphocytes, blasts, or plasma cells)

- Steps in the analysis of flow cytometric data
  – Inspection of dot plots presenting cell size (FSC), internal complexity (SSC), and the expression of panhematopoietic antigen CD45
    - Specific cell populations can be identified based on their size and cytoplasmic complexity (granules/vacuoles) (Fig. 9.3a)
    - The identification is confirmed and further resolved on the display of CD45 antigen and SSC (Fig. 9.3b). This scattergram provides information on the relative proportion of specific cell populations in the flow cytometric sample and is of particular value when analyzing bone marrow/peripheral blood specimens
      – CD45 is a surface protein tyrosine phosphatase expressed at different levels on all hematopoietic cells
      – Lymphocytes show the highest density of CD45 expression with approximately 10% of the cell membrane occupied by this antigen
      – Granulocytic series including myeloid blasts, B cell precursors, and proerythroblasts show intermediate CD45 density
      – Late erythroid precursors along with megakaryocytes are negative for the CD45 antigen

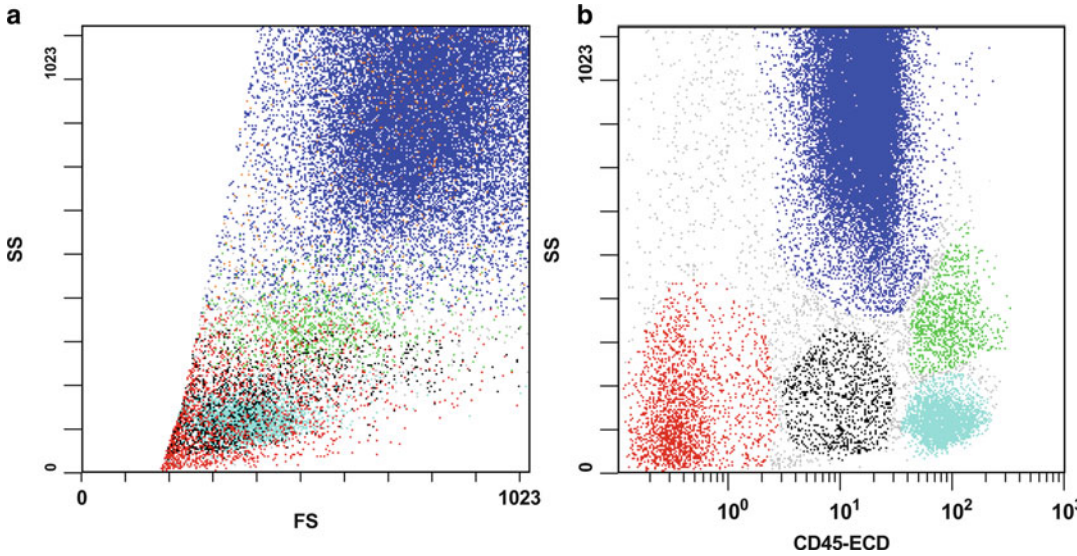


Fig. 9.3 Main hematopoietic populations of normal bone marrow. (**a**) Scattergram of FSC (cell size) vs SSC (internal complexity) reflects the heterogeneity of bone marrow. Lymphocytes as smallest with negligible amount of cytoplasm are located closest to the origins of the axes (*shown in aqua*). Monocytes are slightly larger with occasional granules and vacuoles (*green*). Granulocytic series shows prominent granularity (*navy*). (**b**) Differential density of panhematopoietic marker CD45 on marrow leukocytes. Lymphocytes (*aqua*) and monocytes (*green*) show highest density of CD45 antigen. Intermediate expression of CD45 is seen in granulocytic population (*navy*) and blasts (*black*). Late erythroid precursors (*red*) are negative for CD45 antigen

- Focused analysis of the patterns of antigen expression including both qualitative data (antigen present/absent) and fluorescence intensity (on the logarithmic scale) as a relative measure of the antigen density on the cell surface
  - Residual normal cells present in a sample can be used as an internal negative and positive control and to gauge the intensity of staining
  - The intensity of staining is dependent on technical variables including antibody clone and type of fluorochrome. Thus, levels defining bright, moderate, and dim expression should be established by individual laboratories taking into account specific antibodies and fluorochromes used and previous experience
  - The autofluorescence and nonspecific background staining due to Fc receptors should be taken into account when evaluating antigen expression

9.4 Basic Cell Populations Identified by Flow Cytometry

- Genetically controlled differentiation program and bone marrow environment govern the expression of surface and cytoplasmic molecules that define hematopoietic cell populations
- Specific morphologic stages of development are accompanied by distinct changes in immunophenotypes. However, even though approximate morphologic–immunophenotypic correlates exist, transitions between immunophenotypes of various developmental phases are best described as a continuum
- All hematopoietic progeny are derived from pluripotent stem cells
  - These cells are morphologically unrecognizable and are defined by their functional and antigenic characteristics
  - They usually express a combination of intermediate density CD45, CD34,

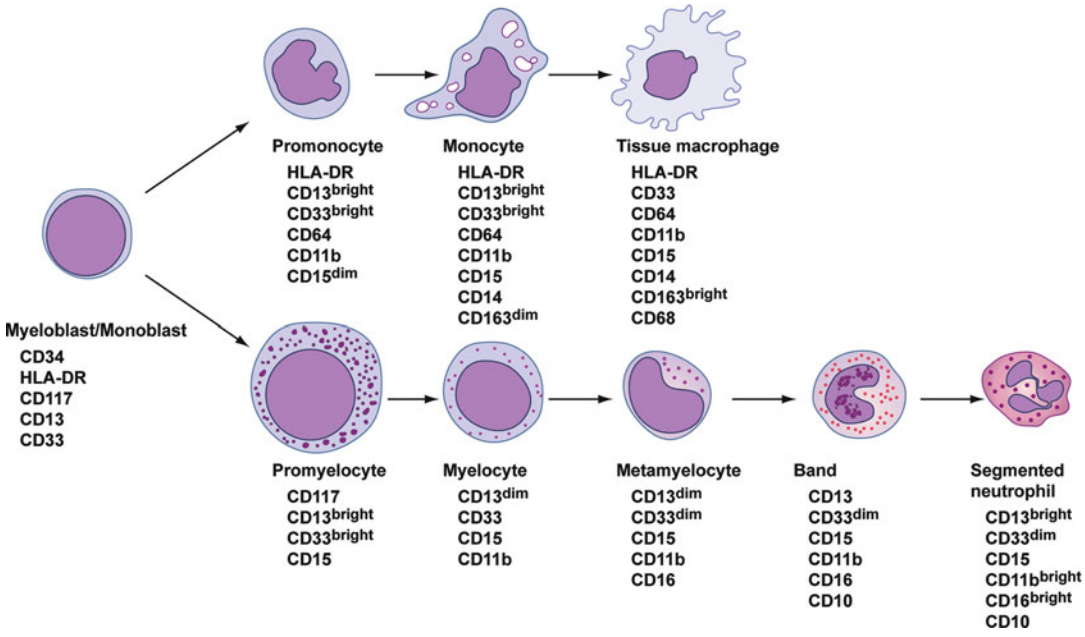


Fig. 9.4 The maturation of myeloid series is a genetically driven developmental program characterized by the continuum of phenotypic and functional changes. Discrete morphologic stages correspond to specific immunophenotypes

CD133, CD117 (c-kit), variable density of CD38 and CD90, CD123, and HLA-DR

- As hematopoietic cells mature, they lose stem cell markers and acquire lineage-specific antigens
- Neoplastic hematopoietic cells to a certain extent mimic normal maturation stages; however, they frequently display aberrant antigen expression patterns

9.4.1 Granulocytic Lineage

- The differentiation of granulocytic lineage, as defined by the expression of specific antigens, corresponds closely to the morphologic maturation stages as depicted in Fig. 9.4
- Myeloblast is the first morphologically recognizable cell committed to the myeloid lineage and typically expresses immature cell markers CD34, CD38, HLA-DR, and stem cell factor receptor CD117, and panmyeloid markers, CD13 and CD33
- As the myeloblast matures to a promyelocyte, it loses CD34 and HLA-DR and gradually acquires the CD15 antigen
- Further maturation to myelocytes results in a complete loss of CD117 and downregulation of CD13 antigen. The decrease in the density of CD33 is also seen at this stage. CD15 and CD11b are positive
- Finally, as myeloid cells near the band stage, CD16 and CD10 are acquired and the density of CD13 increases. Further decrease of CD33 intensity is also noted
- The segmented neutrophil is characterized by high-density CD13, CD11b, and CD16 and dim CD33

9.4.2 Monocytic Lineage

- The immunophenotype of the earliest stage of monocytic development, a monoblast, overlaps with that of myeloblast and includes the expression of CD34, HLA-DR, CD117, and panmyeloid markers CD33 and CD13 (Fig. 9.4)

- Further monocytic maturation is marked with the increase in density of CD33 and CD13, appearance of CD64 and CD11b, and subsequently low density CD15 (promonocyte)
- Subsequent acquisition of CD14 and further increase in density of CD45 define the transition point to a mature monocyte
- The expression of CD163 and CD68 antigens is strongest on tissue macrophages

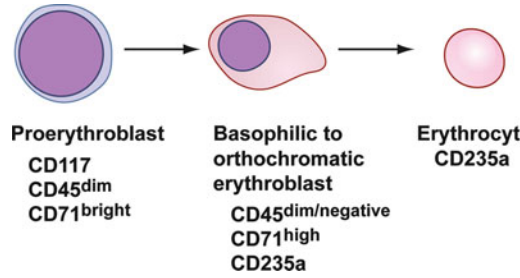


Fig. 9.5 The development of erythroid series is defined by stepwise loss of CD45 along with acquisition of erythroid markers

9.4.3 Erythroid Lineage

- Erythroid precursors are characterized by a gradual decrease in the density of CD45 antigen to the undetectable level in reticulocytes. Thus, late erythroid precursors are one of the few cells in the bone marrow that express a negligible number of CD45 molecules
- The earliest marker of erythroid differentiation is the transferrin receptor, CD71 (Fig. 9.5). This marker increases in density starting from the proerythroblast stage and is rapidly downregulated in reticulocytes. Mature erythrocytes are negative for the CD71 antigen
- The decrease in CD45 intensity seen in basophilic erythroblasts is accompanied by the emergence of CD235a (glycophorin A). The latter marker persists through erythroid maturation and is also present in erythrocytes

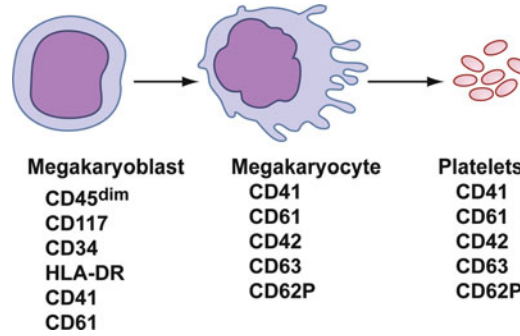


Fig. 9.6 The sequence of immunophenotypic changes of megakaryocytic lineage is characterized by early appearance of CD41 and CD61, that is, gpIIb/IIIa complex

9.4.4 Megakaryocytic Lineage

- The identification of megakaryocyte population by flow cytometry is not done routinely
- CD41 and CD61 (gpIIb/IIIa complex) appear as the first markers of megakaryocytic differentiation and are present on a small subset of CD34 and CD117 positive cells believed to represent early megakaryoblasts
- CD31 and CD36, although not entirely specific for megakaryocytic lineage, are also present on megakaryoblasts
- As megakaryoblasts mature to megakaryocytes and platelets, additional antigens appear including CD42b, CD62P, and CD63
- CD41 and CD61 persist through the megakaryocyte differentiation (Fig. 9.6)
- The precise sequence of expression of megakaryocyte-associated antigens has not been well studied

9.4.5 Lymphoid Lineage

- The B and T lymphocytes are derived from lymphoid progenitors expressing CD34, terminal deoxynucleotidyl transferase (TdT), and HLA-DR

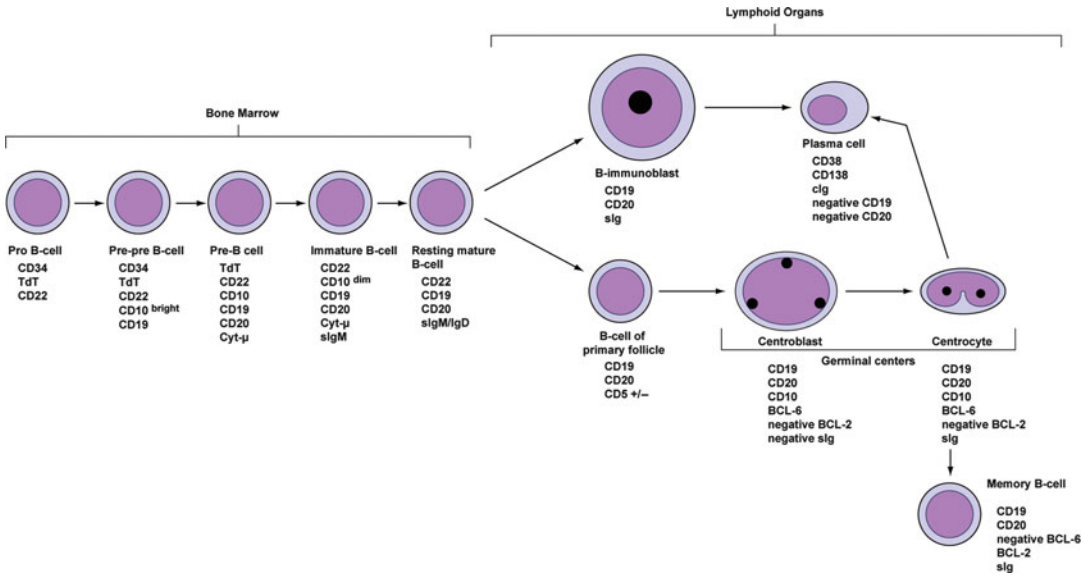


Fig. 9.7 The early stages of B cell maturation are completed in the bone marrow. Subsequent maturation of B cell lineage occurs in lymph nodes and extranodal lymphoid tissues and results in the production of plasma cells and memory B cells

- The number of CD45 molecules steadily increases with B cell maturation and reaches characteristic high-density expression at the level of mature B cells. Early B cell precursors show low-density CD45
- The lymphoid differentiation represents a continuum of changes in the expression of surface and cytoplasmic antigens

9.4.5.1 B Cell Lineage

- The simplified schema of B cell differentiation is presented in Fig. 9.7
- The earliest B cell markers include cytoplasmic CD22, CD19, and cytoplasmic CD79a
- As B cell precursors proceed in their maturation, they acquire the CD10 antigen, which is initially expressed at high levels
- Subsequent appearance of the CD20 antigen is accompanied by the decrease in CD10 intensity and its subsequent loss
- The μ-heavy chain, a portion of immunoglobulin (Ig) molecule, is initially expressed in the cytoplasm and eventually transported to the cell surface where it forms a B cell receptor (BCR)
- The mature naïve B cell population expresses heterogeneous (polyclonal) surface light chains and, in this respect, differs from neoplastic B cells, which are restricted to only a single κ- or λ-light chain
- The mature B cells circulate to the secondary lymphoid organs including lymph nodes, spleen, and lymphoid tissues of extranodal sites, where they settle in the follicles and mantle zones
- Further differentiation of mature naïve B cells occurs upon antigen exposure and includes passage through germinal centers marked by the signature coexpression of CD10 and Bcl-6 antigens
- Plasma cells, the terminal stage of B cell differentiation, lose CD20 and surface Ig chains and can be identified by the high-density expression of CD38 and CD138 (syndecan-1)

9.4.5.2 T Cell Lineage

- Early stages of T cell development take place in the bone marrow (Fig. 9.8)
- The first committed T cell precursor (prothymocyte) expresses immature markers

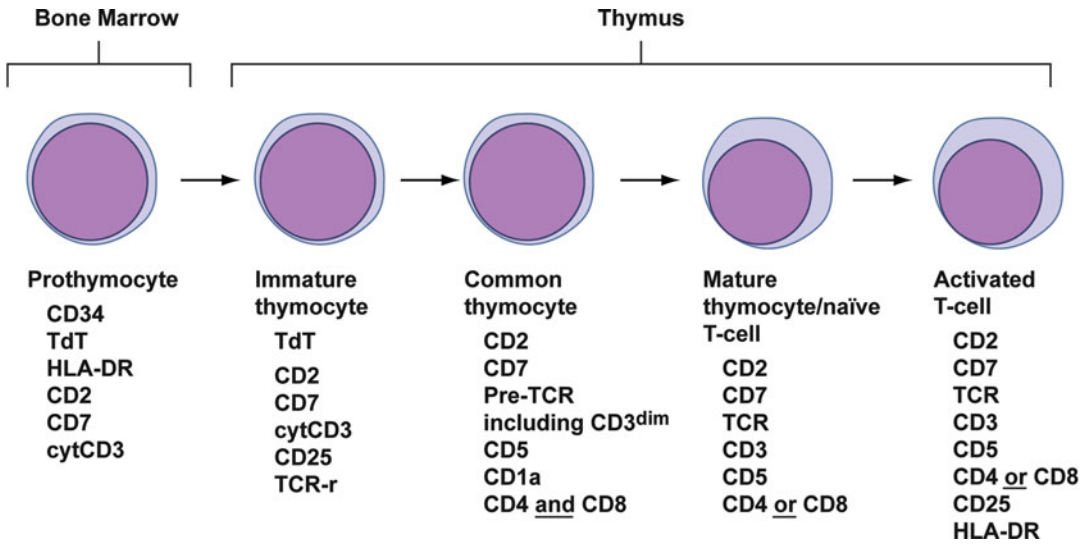


Fig. 9.8 The early T cell precursors are generated in the bone marrow and migrate to thymus to complete their maturation

(CD34, TdT, and HLA-DR) and T cell-associated antigens including CD2, CD7, and cytoplasmic CD3
- Prothymocytes migrate to the thymus to complete T cell development
- Successive steps of T cell receptor (TCR) gene rearrangement with the production of TCR complex, expression of CD1a and CD5, and coexpression of CD4 and CD8 antigens define immature and common thymocyte stages
- As the double-positive (CD4+, CD8+) common thymocyte matures, the density of CD3 antigen increases and CD4 or CD8 is lost, giving rise to mature helper (CD4+) and suppressor (CD8+) T cells

9.4.5.3 Natural Killer (NK) Cells

- NK cells are positive for CD2, CD7, CD56, and CD16
- Different densities of surface antigens allow for the separation of two functionally distinct NK subsets
  - Cytotoxic NK cells show expression of $CD56^{dim}$, $CD16^{bright}$, KIR^{bright}, and CD94/$NKG2A^{dim}$ (90% of NK population)
  - Immunoregulatory NK cells show expression of $CD56^{bright}$, $CD16^{dim/-}$, CD117, KIR^{dim}, and CD94/$NKG2A^{bright}$ (10% of overall NK population, however, represent the majority of the NK cells in lymph nodes)
- Published data suggests that the development of NK cells starts in the bone marrow from CD34+ progenitor cell
- It has been suggested that the later stages of development of immunoregulatory NK subset occur in the lymph nodes from $CD34^{dim}$ progenitor cells
- The site and exact sequence of the development of cytotoxic NK cells are unknown

9.5 Flow Cytometric Analysis of Myeloid Disorders

- In myeloid malignancies, flow cytometry can be used for the initial diagnosis, followup, and prognostication (specific immunophenotypes are associated with prognostically significant cytogenetic abnormalities) and less commonly for a determination of treatment targets (e.g., CD33 antigen for gemtuzumab ozogamicin therapy)

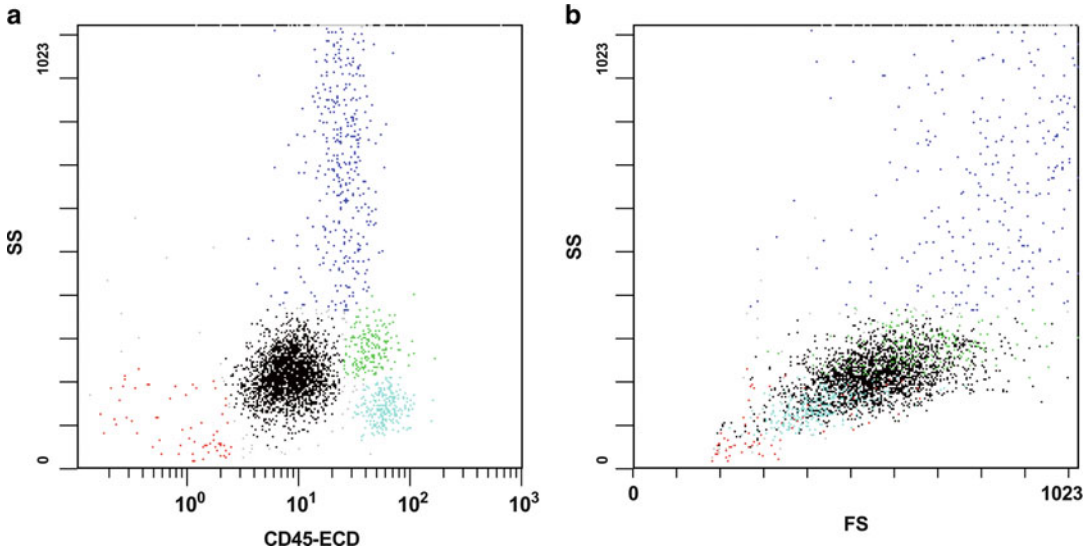


Fig. 9.9 (**a**) The best resolution of blast population (shown in *black*) is achieved using scattergram of SSC vs CD45. The intermediate density of CD45 antigen together with low SSC results in clear delineation of blast population in the majority of cases of acute leukemia. The lymphoid cells (*aqua*), monocytes (*green*), and maturing granulocytic series (*navy*) localize outside the blast gate. (**b**) On the scattergram demonstrating cell size (FSC) and granularity (SSC), blasts usually overlap with other cell populations

- The majority of myeloid neoplasms are regarded as disorders of hematopoietic stem and progenitor cells
- In acute myeloid leukemias (AML), the maturation arrest leads to the accumulation of a homogeneous population of cells demonstrating an immature myeloid immunophenotype; thus, the blast region, best demonstrated on a CD45/SSC scattergram, is densely populated in most cases (Fig. 9.9a, compare with Fig. 9.3b). The increased number of blasts and the paucity of maturing marrow elements can also be visualized on SSC/FSC display (Fig. 9.9b)
- In myelodysplastic syndrome (MDS) and chronic myeloproliferative disorders, a myeloid maturation is at least partially preserved; thus, both evaluation of immature and maturing cells is considered essential for the diagnosis
- In this chapter, the immunophenotypic features of AML and chronic myeloid disorders are briefly discussed in the context of the 2008 World Health Organization (WHO) classification
- The WHO classification introduced new categories of AML defined by recurrent cytogenetic abnormalities. These leukemias often show specific immunophenotypes and will be presented separately in the following outline

9.5.1 Acute Myeloid Leukemias

9.5.1.1 AML with Recurrent Cytogenetic Abnormalities

AML with t(8;21)(q22;q22); *RUNX1-RUNX1T1*

- Majority of cases show an immature myeloid immunophenotype with a high density of CD34 and coexpression of low-density CD19 (Fig. 9.10)
- Numerous myeloid antigens, including CD33, CD13, and myeloperoxidase, are expressed
- Frequently, there is asynchronous coexpression of CD34 and CD15
- The presence of TdT is common
- The coexpression of CD56 has been reported to be associated with worse prognosis

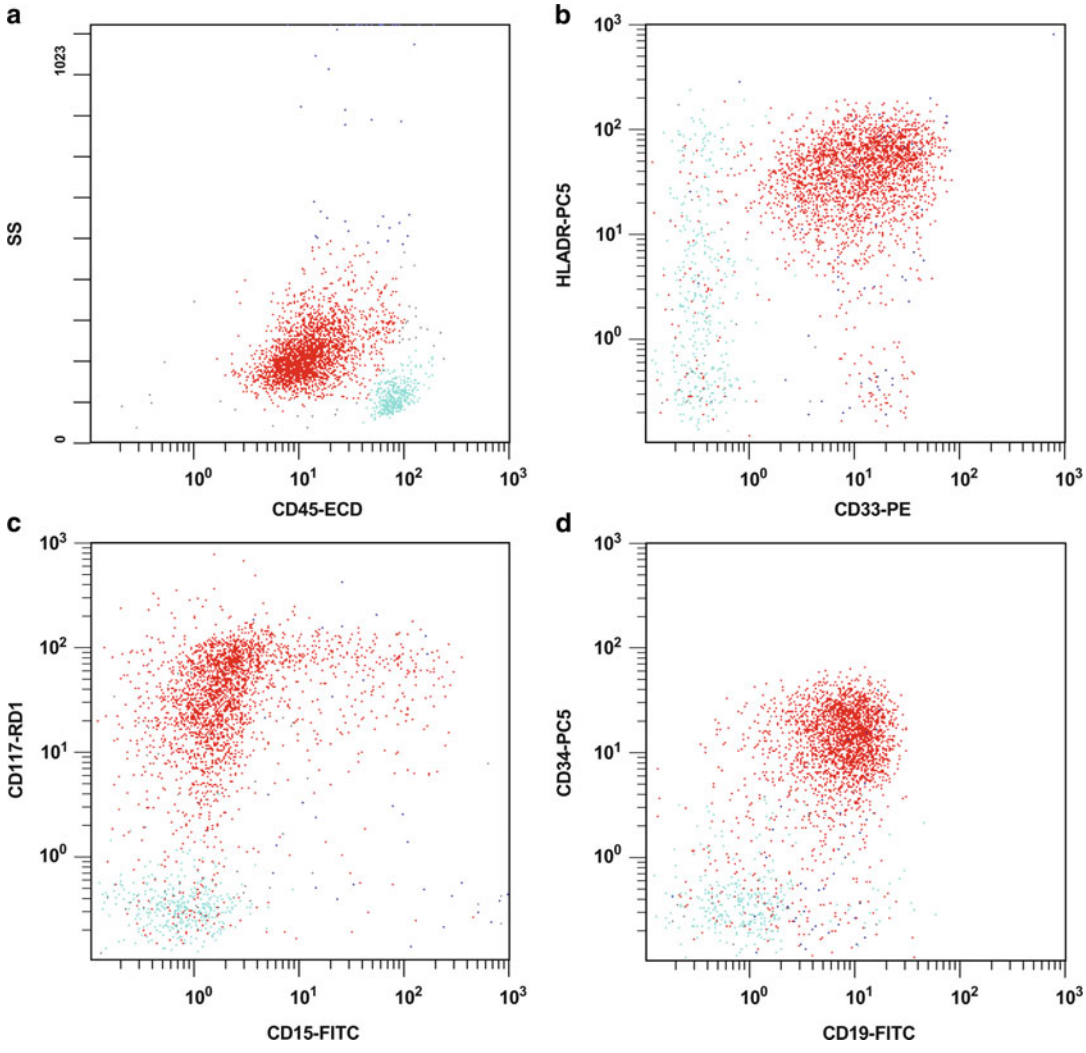


Fig. 9.10 AML with t(8;21)(q22;q22); (*AML1*/*ETO*). Blasts are shown in *red* and residual lymphocytes in *aqua*. (**a**) CD45 vs SSC demonstrates a distinct blast population with marked decrease in other hematopoietic cells. The residual lymphocytes are present. (**b**–**d**) Blasts express immature markers frequently present in progenitor cells (CD34, HLA-DR, and CD117) and show characteristic coexpression of CD19

AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB-MYH11*

- Immature cells show expression of CD34, CD117, and TdT
- Subpopulation of maturing cells expresses monocytic (CD14, CD11b, $CD4^{dim}$) and granulocytic (CD15) markers
- The coexpression of CD2, antigen normally seen on T and NK cells, is common (Fig. 9.11)

Acute Promyelocytic Leukemia with t(15;17)(q22;q12); *PML-RARA* (APL)

- In contrast to most less differentiated myeloid leukemias, APL presents with high SSC reflecting the granular cytoplasm of leukemic cells (Fig. 9.12)
- A constellation of immunophenotypic features used to diagnose APL includes lack of CD34 and HLA-DR antigens

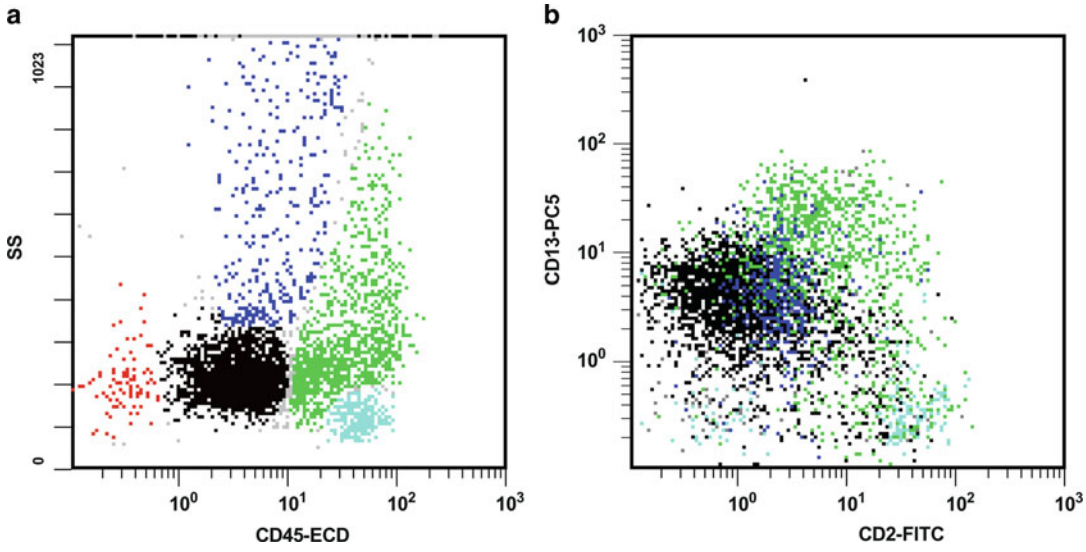


Fig. 9.11 The AML with inv(16)(p13q22) or t(16;16)(p13;q22)/(*CBF-β/MYH11*) most commonly presents a myelomonocytic leukemia. (**a**) The CD45/SSC scattergram shows two merging populations of myeloid blasts (in *black*) and monocytic cells (in *green*). (**b**) The monocytic component with typical high-density CD13 coexpresses CD2 antigen

- Expression of homogeneous bright CD33 along with myeloperoxidase and variable expression of CD13 and CD15 are present
- CD2 and higher incidence of CD34 expression have been reported in APL with microgranular morphology

AML with t(9;11)(p22;q23); *MLLT3-MLL*

- Constitute a heterogeneous group most commonly presenting with monocytic differentiation
- The immunophenotypic features are not specific and can be seen in any acute myelomonocytic or monocytic leukemias
- Most commonly, these leukemias show variable expression of immature markers such as CD34 and CD117, and positive for CD33, CD13, CD14, $CD4^{dim}$, CD11b, and CD64

9.5.1.2 AML Not Otherwise Specified

AML with Minimal Differentiation and AML Without Maturation

- Blasts show low-density CD45 antigen expression and display low SSC reflecting their relatively agranular cytoplasm (Fig. 9.9a)
- The majority of even least differentiated AMLs express myeloid markers such as CD13, CD33, and/or CD117
- Primitive hematopoietic antigens, CD34 and HLA-DR, are often seen
- Myeloperoxidase is negative or only expressed in a minority of cells

AML with Maturation

- In addition to primitive hematopoietic and early myeloid antigens, more mature myeloid markers such as CD15 and myeloperoxidase are often expressed
- Occasionally, there is an asynchronous coexpression of antigens, which in normal hematopoietic cells are not expressed simultaneously (e.g., exclusive early and late myeloid markers such as CD34 and CD15 are expressed at the same time)
- Coexpression of markers associated with other lineages, for example, lymphoid, may be seen on the myeloid blasts. The most common example is the CD7 antigen

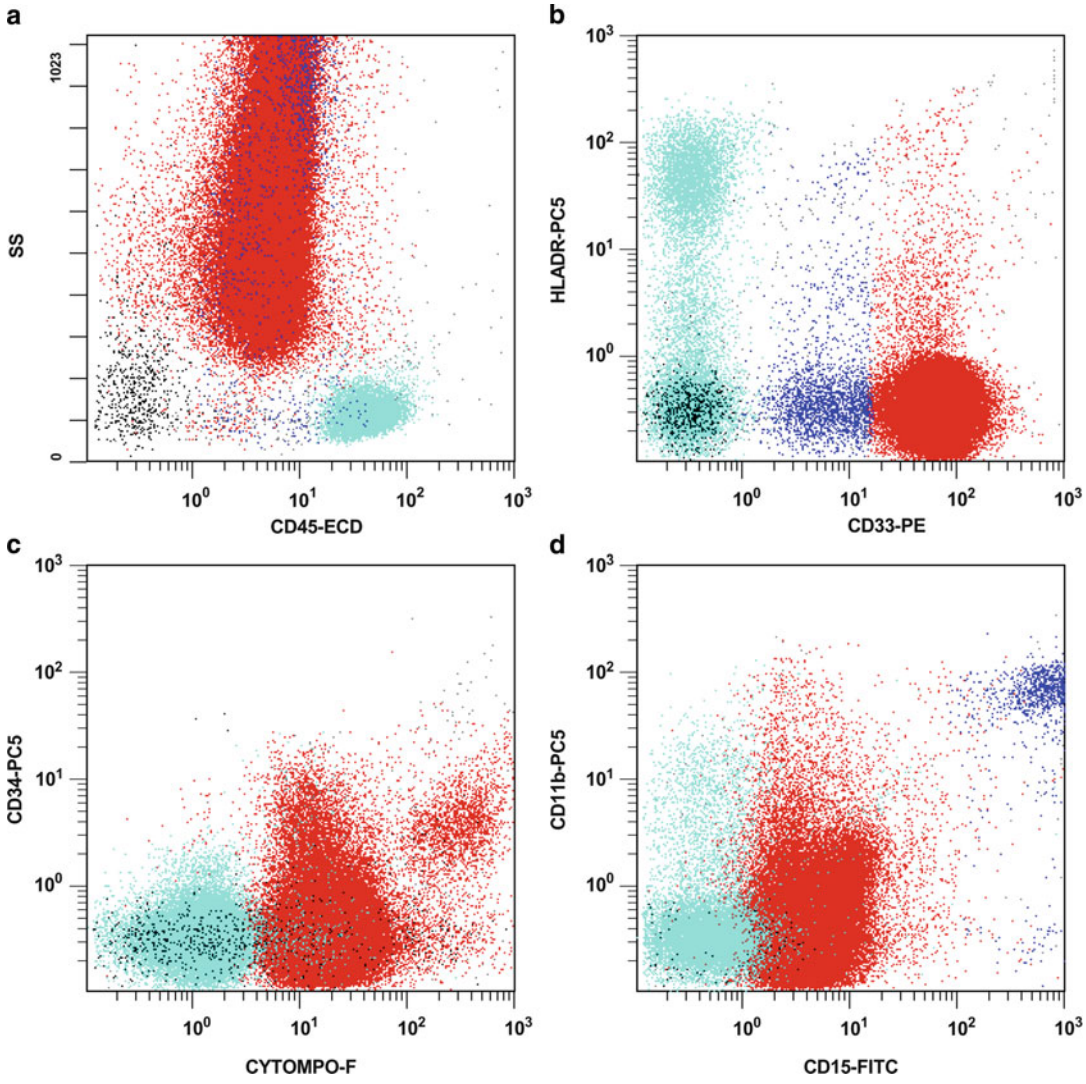


Fig. 9.12 Acute promyelocytic leukemia. (**a**) High SSC of APL corresponds to prominent granularity of leukemic cells (in *red*). Residual lymphocytes are presented in *aqua*. (**b**) Leukemic promyelocytes are positive for CD33 (bright as compared with low-density CD33 of residual normal neutrophils, in *navy*) and negative for HLA-DR. (**c**) CD34 is negative and myeloperoxidase is expressed by the majority of cells. (**d**) Low-density CD15 antigen can be present

Acute Leukemias with Monocytic Differentiation (Acute Myelomonocytic Leukemia and Acute Monoblastic/ Monocytic Leukemia)

- The SSC and CD45 expression in acute leukemias with monocytic differentiation is variable and dependent on the relative proportion of primitive myeloid blasts and the degree of differentiation of neoplastic monocytes. Patterns with a distinct and separate population of blasts and monocytes or a large merging cluster of cells, starting in the blast region and extending upward to monocyte region, can be seen (Fig. 9.13)
- In acute myelomonocytic leukemia, a population of primitive myeloid blasts is often distinct
- The expression of myeloid markers and antigens associated with monocytic lineage such as

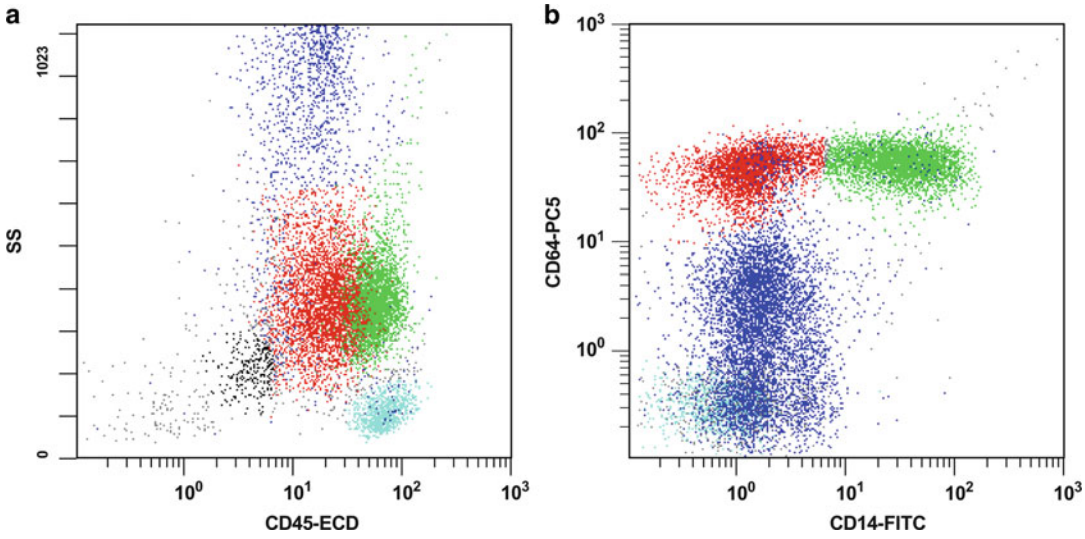


Fig. 9.13 Acute leukemias with monocytic component frequently show a spectrum of monocytic maturation. (**a**) Both uncommitted myeloid blasts and different stages of monocytic maturation (in *red* and *green*) are present. (**b**) Immature monocytic cells, presented in *red*, show high-density CD64. More mature monocytic component expresses CD14 antigen (in *green*)

CD14, CD4, CD11b, and CD64 is commonly seen
- Despite the CD14 antigen being present on all mature monocytes, it can be negative in monocytic leukemias. Frequently, a heterogeneous pattern of CD14 expression is seen reflecting a maturation spectrum of neoplastic monocytes
- Markers ordinarily present on both immature and mature monocytes, such as high-density CD64, are more consistently expressed

9.5.1.3 Acute Erythroid Leukemias

- Acute erythroid leukemias are categorized into two subtypes: pure erythroid leukemia and erythroid/myeloid leukemia (erythroleukemia)
- In erythroleukemia, both primitive myeloid blasts and erythroid precursors are present
- Erythroid markers CD71, glycophorin A (CD235a), and hemoglobin can be present
- When glycophorin A and hemoglobin are absent, the diagnosis is based on the absence of myeloid markers, the presence of bright CD71, and the scatter characteristics of leukemic cells

9.5.1.4 Acute Megakaryoblastic Leukemia

- Usually shows low SSC and dim to absent CD45
- Early megakaryocytic markers, CD41 and CD61, are frequently expressed (Fig. 9.14)
- Occasionally, the late megakaryocytic marker, CD42b, is present
- There is variable expression of stem cell markers, CD34 and HLA-DR, on the population of leukemic megakaryoblasts

9.5.1.5 Blastic Plasmacytoid Dendritic Cell Neoplasm

- Initially, this neoplasm was thought to have originated from NK cells and had been previously referred to as blastic NK cell lymphoma (agranular CD4+, CD56+ hematodermic neoplasm)
- The most current evidence suggests its origin from the plasmacytoid dendritic cells
- The immunophenotype reflects the cell of origin: CD56, CD4, HLA-DR, and CD123 antigens are positive. The latter, along with the expression of blood dendritic cell

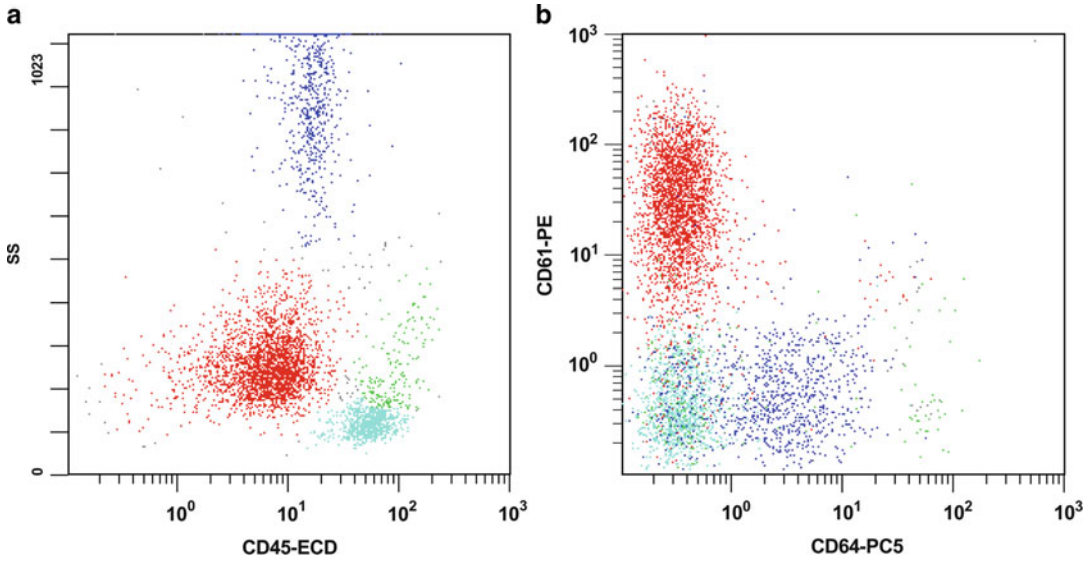


Fig. 9.14 Acute megakaryocytic leukemias can be diagnostically challenging. (**a**) Leukemic cells usually show low SSC. (**b**) The expression of megakaryocytic markers, such as CD61, in otherwise undifferentiated acute leukemia proves the megakaryocytic lineage of leukemic cells (in *red*)

antigens 2 and 4, is highly specific for dendritic cells and their precursors
- Rare cases showing positivity for TdT and CD34 were reported
- The presence of T cell, B cell, and myelomonocytic lineage-associated markers should be excluded

9.5.2 Chronic Myeloproliferative Neoplasms and Myelodysplastic Syndrome (MDS)

- Multiparameter flow cytometry is helpful in diagnosis, prognostication, and prediction of a response to therapy in patients with chronic myeloid neoplasms, particularly in MDS. The contribution of flow cytometry in chronic myeloproliferative and myelodysplastic/myeloproliferative neoplasms is less well established
- The abnormalities detected by flow cytometry reflect abnormal morphologic features and abnormal maturation
- Flow cytometric abnormalities fall in two categories: quantitative (enumeration of marrow components, light scatter changes) and qualitative (altered antigen expression in maturation and lineage-specific hematopoietic compartments such as abnormal antigen density, asynchronous maturation, lineage "infidelity")
- A detailed review is beyond the scope of this text; however, a few examples are presented next to illustrate the contribution of flow cytometry to diagnosis of chronic myeloid neoplasms

9.5.2.1 Myelodysplastic Syndromes

- MDS is characterized by ineffective hematopoiesis with abnormalities in maturation and, often, decreased survival of hematopoietic progeny
- Diagnosis of MDS is based on the morphologic, immunophenotypic, and genetic features as well as clinical manifestations
- Flow cytometry contributes to the diagnosis of MDS through the identification of aberrant maturation and other immunophenotypic abnormalities, some of which are typical for MDS
- Aberrant immunophenotypes demonstrated by flow cytometry are seen in up to 98% of

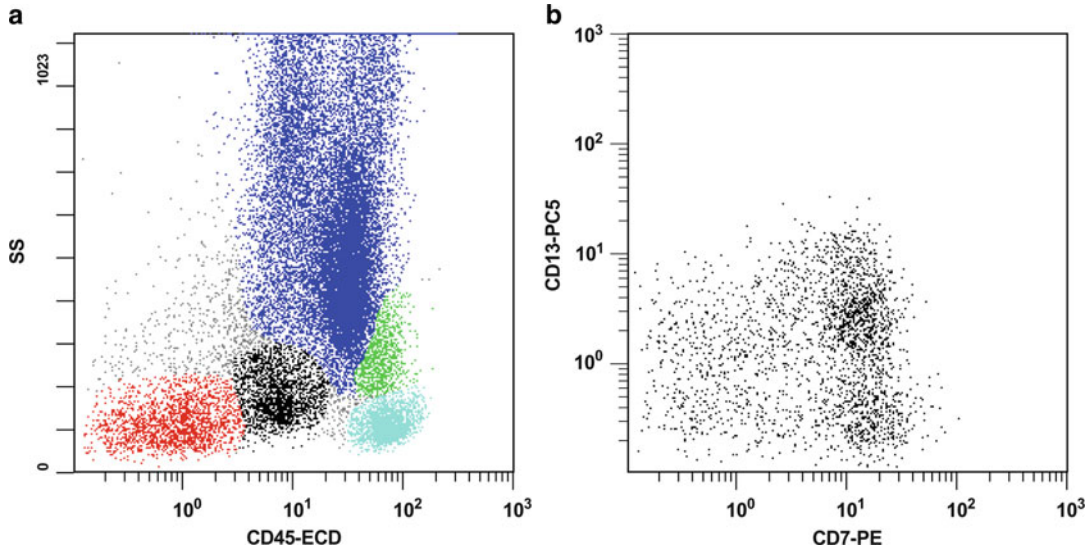


Fig. 9.15 Flow cytometry demonstrates phenotypic abnormalities in the majority of MDS cases. (**a**) Hypogranulated neutrophils (in *navy*) can be visualized by their abnormally low SSC. Slight increase in the number of blasts is also seen in this case. (**b**) Coexpression of CD7 on myeloid blasts is frequent in MDS

MDS cases and can be routinely analyzed in granulocytic, monocytic, and erythroid lineages

- It has been previously reported that even in cases with minimal or no morphologic features indicative of MDS, flow cytometry was predictive of future cytogenetic abnormalities and diagnosis of myelodysplasia
- Recent studies underscored the prognostic significance of specific immunophenotypic abnormalities for the natural course of the disease or the outcome after bone marrow transplantation
- The abnormalities detected by flow cytometry in MDS include
  - Aberrant SSC reflecting morphologically identified dysplasia: hypogranulated neutrophils can be visualized by their abnormally low SSC in up to 70% of cases (Fig. 9.15a)
  - Changes in the relative proportion of cells at specific stages of myeloid maturation: the high-grade MDS usually demonstrates increased number of immature cells. A significant left shift in granulocytic maturation and an increase in blast percentage can be demonstrated by flow cytometry (Fig. 9.15a)
  - The disruption of normal maturation patterns as reflected by the asynchronous expression of myeloid markers
    - Appearance of late myeloid markers inappropriately early in the differentiation (i.e., CD15 on myeloblasts)
    - Persistent expression of immature markers in late granulocytic stages (e.g., retention of CD34 and HLA-DR on mature granulocytes)
    - Uncoupling of the normal sequence of CD13 and CD16 expression in granulocytic differentiation
  - In addition, blasts in MDS (and other myeloid neoplasms) can demonstrate
    - Detectable early progenitor population positive for CD34 and negative for CD38 antigen
    - Decreased or increased antigen density such as CD45
    - Aberrant expression of immature markers (CD34, CD117, HLA-DR)
    - Absence of myeloid markers CD13 or CD33

- Expression of markers associated with lymphoid lineage (CD7, CD56, CD4, CD2, CD5; Fig. 9.15b)

– The lymphoid populations in bone marrow of patients with MDS show the following changes
  - The paucity of B-lymphoid precursors, which can help to differentiate MDS from nonclonal cytopenias
  - The significant increase in CD8 positive T cells, expansion of CD4 positive cells, and reduced Th2 and regulatory T cells seen in low-risk MDS
  - V beta-restricted CD8 positive T cells in patients who respond well to immunosuppressive therapy
  - Increased number of CD3 + CD4 + IL-17 producing T cells (Th17) and regulatory T cells in high-risk MDS

9.5.2.2 Chronic Myeloproliferative Neoplasms

- The utility of flow cytometry in chronic myeloproliferative neoplasms is less well established

Chronic Myelogenous Leukemia, *BCR-ABL1* Positive

- Application of flow cytometry as a diagnostic tool in chronic myelogenous leukemia is limited to the accelerated phase or blast crisis, in which the lineage of an expanding blast population needs to be determined
- In the chronic phase, the presence of the Philadelphia chromosome (as seen on conventional karyotyping or molecular analysis) remains the defining feature of this disorder

***BCR-ABL1* Negative Chronic Myeloproliferative Neoplasms**

- In general, flow cytometric abnormalities are seen in the majority of cases with cytogenetic abnormalities
- However, no consistent set of abnormalities to routinely subclassify neoplastic myeloproliferative states was described

9.6 Flow Cytometric Analysis of Lymphoid Neoplasms (Lymphoblastic Leukemia/Lymphoma and Mature Lymphoid Neoplasms)

- The diagnosis of lymphoid malignancies relies on the presence of lineage-associated markers corresponding to specific stages of lymphoid development
- No single marker can be used for a definite diagnosis; thus, the presence of several B cell or T cell-associated antigens is used for lineage assignment
- The sentinel feature of mature B cells and T cells is the presence of surface receptor complexes
- The immune system has to respond to a wide array of antigens. In consequence, in healthy individuals, B cells and T cells express a great diversity of surface receptor complexes (Ig and TCRs). This diversity defines reactive polyclonal lymphoid populations
- On the contrary, the neoplastic lymphoid cells are characterized by monoclonal B cell and T cell receptors. In the majority of cases, the presence of clonality is a definite confirmation of the malignant nature of lymphoid proliferation
- Lymphoid precursors often show an absence of surface receptor complexes. Thus, in precursor-derived neoplasms, the homogeneous expression of specific markers on lymphoblast population, rather than the presence of clonal surface receptors, is considered diagnostic of malignancy. The following paragraph presents the key immunophenotypic features of lymphoblastic and mature lymphoid malignancies

9.6.1 B Lymphoblastic Leukemia/Lymphoma (B-LL)

- B-LL shows expression of markers seen in normal B cell differentiation such as CD19, CD22 (cytoplasmic or membranous), CD79a, HLA-DR, CD34, and TdT (Fig. 9.16)

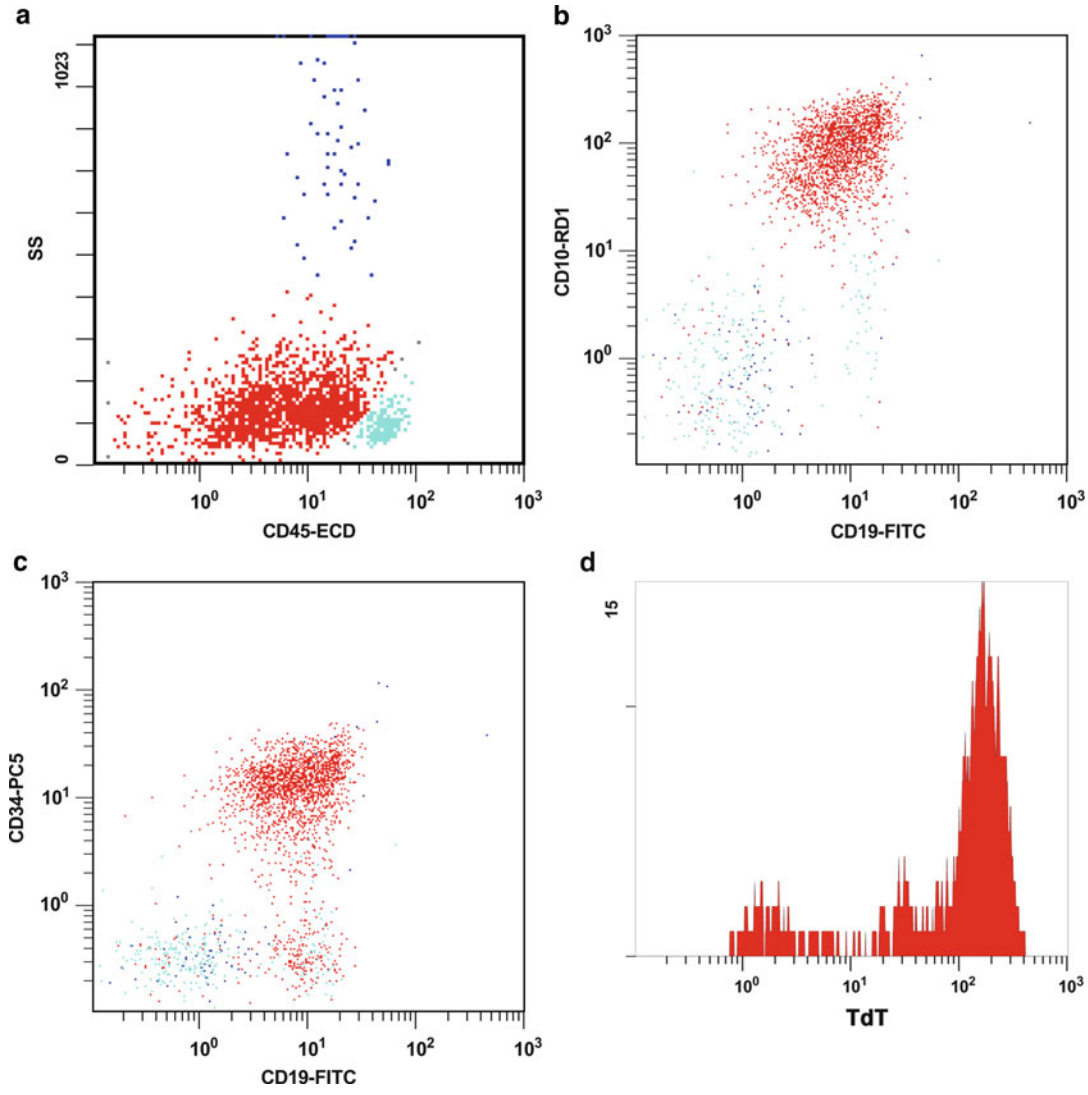


Fig. 9.16 Pre-B ALL. (**a**) A distinct blast population is characterized by low SSC and low-density CD45 antigen expression. (**b**) The CD10 antigen tends to be expressed at higher levels than in normal B cell precursors. (**c**) The majority of leukemic cells are positive for CD34. (**d**) TdT is brightly positive

- High-density CD10 antigen is frequently seen (Fig. 9.16b)
- Surface Ig light chains are not present; however, cytoplasmic μ-chain or IgM may be detected
- Immunophenotypes of leukemic lymphoblasts resemble stages of normal B cell differentiation. However, even though maturation sequence is roughly reproduced, the majority of cases show aberrant expression of select markers. In addition, the categorization according to maturation stage is of limited practical value as clinical behavior is influenced mostly by clinical features and genetic abnormalities
- Specific immunophenotypes frequently correlate with cytogenetic and clinical features. However, in routine practice, the confirmation of cytogenetic abnormality with either conventional karyotyping and/or molecular techniques is necessary. Select examples of immunophenotypic-genotypic associations are presented below

9.6.1.1 B-LL with t(v;11q23); *MLL* Rearranged

- This rearrangement commonly occurs in infant B-LL, while the frequency of *MLL* involvement in older children and adults is much lower (<5%)
- The most frequent fusion partner for *MLL* gene in ALL is *AF4* gene on chromosome 4q21 (t[4;11])
- Rarely other genes are involved including *ENL* (19p13.3) and *AF9* (9q21–22)
- CD19, CD34, and TdT are positive. The more mature B cell marker, CD20, is negative
- Contrary to the majority of B-LL cases, blasts in this leukemia are negative for the CD10 antigen, which indicates the early stage of B cell maturation
- Myeloid markers, CD15 and CD65, are frequently positive

9.6.1.2 B-LL with t(9;22)(q34;q11.2); *BCR-ABL1*

- Philadelphia chromosome [t(9;22(q34;q11.2); *BCR-ABL1*] is a hallmark of chronic myelogenous leukemia. However, a *BCR-ABL1* translocation with a breakpoint in the minor breakpoint region (m-*BCR*) occurs in both pediatric and adult B-LL
- Cases of LL with *BCR-ABL1* translocation carry a particularly dismal prognosis; thus, their prompt identification is essential for effective treatment decision making
- Most *BCR-ABL1* cases have a classic intermediate (common) LL immunophenotype with the expression of CD19 and TdT
- The homogeneous CD10 and CD34 and dim/heterogeneous CD38 expressions are seen
- The expression of myeloid markers, especially CD13, is common
- The expression patterns of CD34 and CD38 allow differentiation between *BCR-ABL1-positive* and negative LL in multivariate analysis

9.6.1.3 B-LL with t(12;21)(p13;q22); *TEL-AML1 (ETV6-RUNX1)*

- *TEL-AML1* translocation occurs in 25% of childhood B-LL cases and is associated with a favorable prognosis
- The detection of t(12;21) by conventional karyotyping can be challenging; thus, the identification of this subset of B-LL by flow cytometry adds to a diagnostic accuracy
- The immunophenotype is that of a precursor B cell with expression of CD19, CD34, CD10, and TdT. CD20 antigen is negative
- The CD45 is more commonly positive and the aberrant coexpression of CD13 is frequent
- The most specific immunophenotypic features predictive of *TEL-AML1* fusion are negative or partially positive CD9 and negative CD20

9.6.2 T Lymphoblastic Leukemia/Lymphoma (T-LL)

- Both T lymphoblastic leukemia and T lymphoblastic lymphoma are derived from immature cells committed to T cell lineage
- Similar to B-LL, dependent on the primary site of involvement, that is, bone marrow or lymph node, the designation of leukemia or lymphoma is used
- The most specific marker of T cell differentiation is the CD3 antigen. Similar to normal T cells, in pre-T ALL, this antigen is initially seen in the cytoplasm before the transfer to the cell surface as a portion of the TCR complex
- Other T cell antigens include CD2, CD7, CD5, CD1a, CD4, and CD8
- T-LL mimics the normal maturation of T cells; however, aberrant antigen densities and/or antigen loss are frequent (Fig. 9.17)
- CD34 and CD10 may also be present; however, HLA-DR is typically absent
- In T-LL, the correlation of the immunophenotype with specific genetic lesions is not clear

9.6.3 Mature Lymphoid Neoplasms

- The flow cytometric diagnosis of lymphomas is based on the presence of clonal lymphoid population bearing numerous lymphoid markers
- Gating based on SSC/FSC characteristics is most commonly used

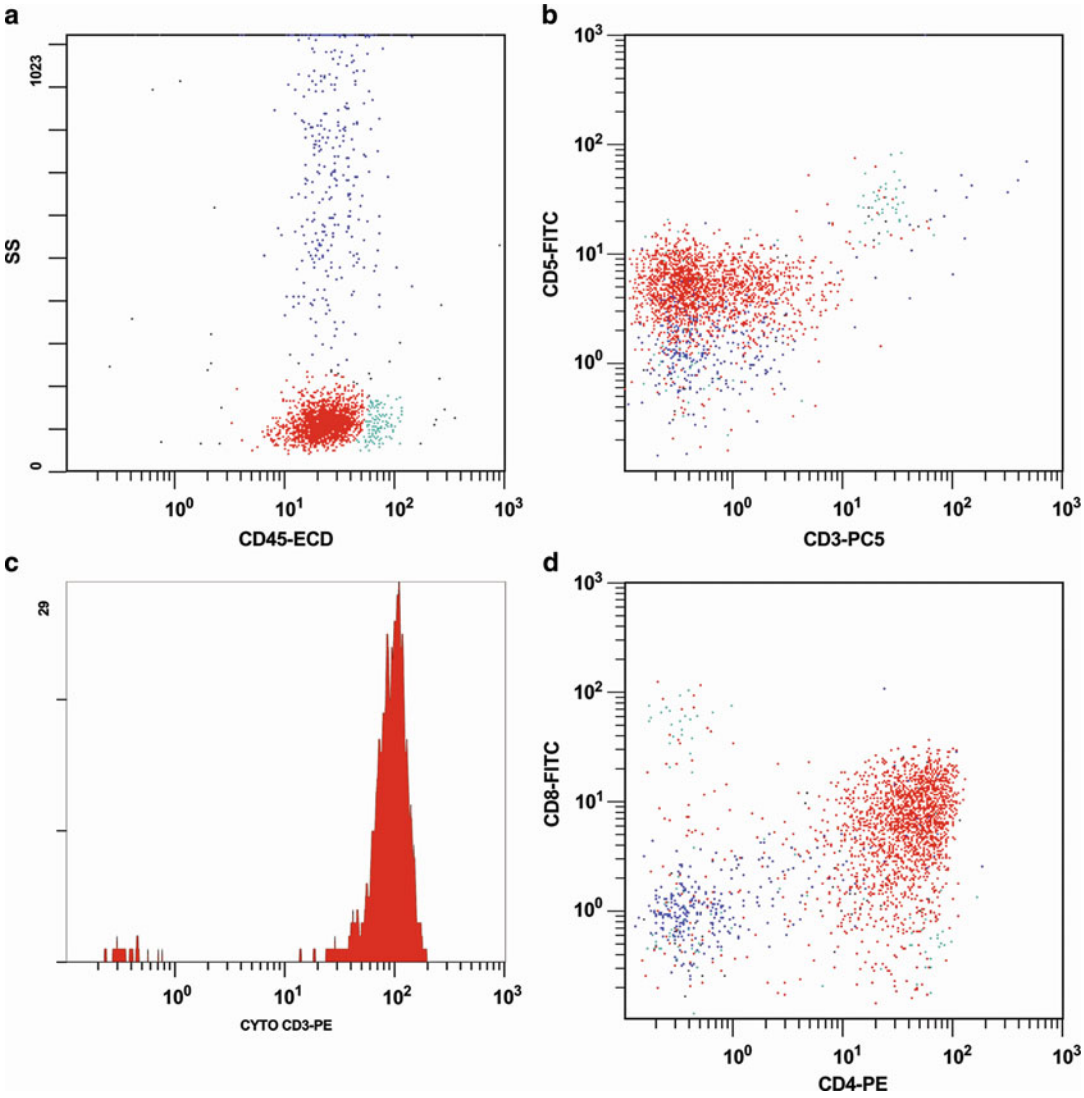


Fig. 9.17 T-LL in the bone marrow. (**a**) Predominant cell population is seen in the blast gate. (**b**) CD3 antigen can be negative on the surface of leukemic cells. Note internal control, residual normal T cells (in *aqua*), positive for surface CD3 and CD5. (**c**) In surface CD3 negative cases, the CD3 antigen can be demonstrated in cytoplasm of leukemic cells. (**d**) The coexpression of CD4 and CD8 can be seen

- On FSC/SSC dot plots, neoplastic lymphoid cells are seen in the area of small or large lymphocytes. In large cell lymphomas or hairy cell leukemia, the neoplastic population can overlap with the monocyte region
- The display of SSC vs CD45 is not typically used for gating of lymphomas since most mature lymphoid malignancies display high-density CD45 antigen. Only rare cases of lymphoma show slightly dimmer CD45 or, even more infrequently, are negative for the CD45 antigen as in lymphomas with plasmablastic differentiation
- The designation "clonal" implicates that the entire lymphoma population is derived from a single lymphoid cell that underwent malignant transformation. Thus, all neoplastic cells

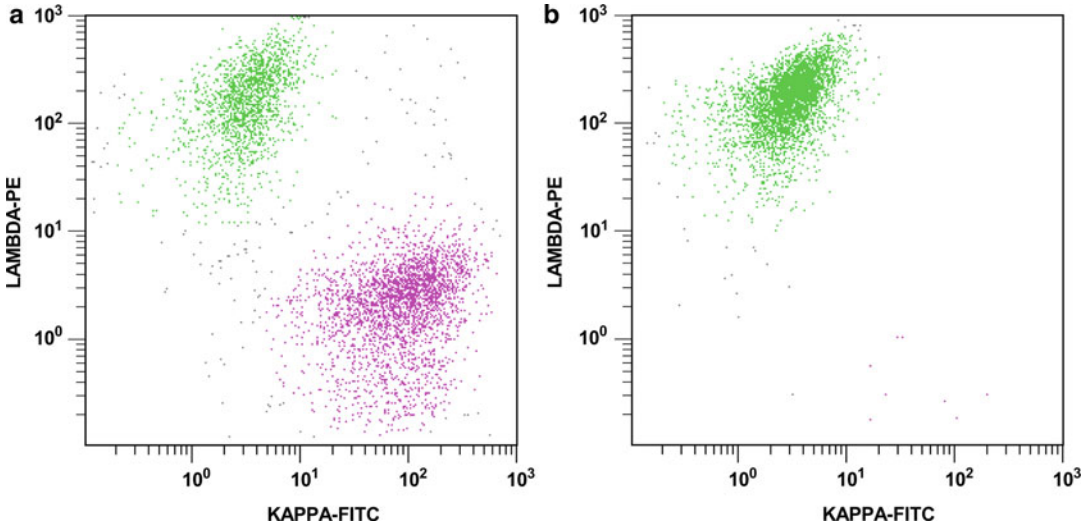


Fig. 9.18 Surface Ig light chain expression in reactive and malignant B cells. (**a**) Reactive B cells show heterogeneous (polyclonal) expression of κ and λ. (**b**) Mature B cell neoplasms are monoclonal with the entire lymphoma population expressing only one type of Ig light chain

should demonstrate similar genetic and immunophenotypic features. The clonality is best represented as an expression of uniform (monoclonal) surface light chain or TCR. This stands in stark contrast to the highly variable, polyclonal immunophenotype of normal lymphocytes, which reflects a random receptor gene rearrangement as a response to a variety of antigenic stimuli

9.6.3.1 Mature B Cell Neoplasms

- Normal precursor B cells randomly rearrange Ig heavy and light chains. As a result, mature B cells show a polyclonal pattern of Ig heavy and light chains (Fig. 9.18a). In contrast, a monoclonal surface light chain expression (exclusively κ or λ) is seen in the majority of B cell lymphomas (Fig. 9.18b)
- The light chain monoclonality along with the expression of pan-B cell markers is in most instances diagnostic of B cell lymphoma. The lymphoma subclassification is based on the presence and density of specific lymphoid markers
- Rarely, mature lymphoid neoplasms lose their surface Igs, a feature not commonly seen in normal mature B cells
- On the contrary, neoplastic plasma cells typically lack surface Igs and show only cytoplasmic expression of κ or λ

Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma (CLL/SLL)

- CLL and SLL are derived from recirculating CD5+, IgM+, and IgD ± B cells normally present in the peripheral blood (Fig. 9.19)
- The WHO classification considers CLL and SLL as one entity with different presentations. The diagnosis is based on the predominant site of involvement (bone marrow/peripheral blood vs lymphoid organs)
- CLL/SLL is positive for pan-B cell antigens including CD19 and CD20 (dim CD20 expression, Fig. 9.19b and c)
- In addition, the expression of CD5, CD23, and weak surface monoclonal (κ or λ) light chain is seen
- The absence of FMC7 and cyclin D1 and presence of CD23 are features distinguishing CLL/SLL from mantle cell lymphoma (MCL)
- Currently, two groups of CLL/SLL are recognized

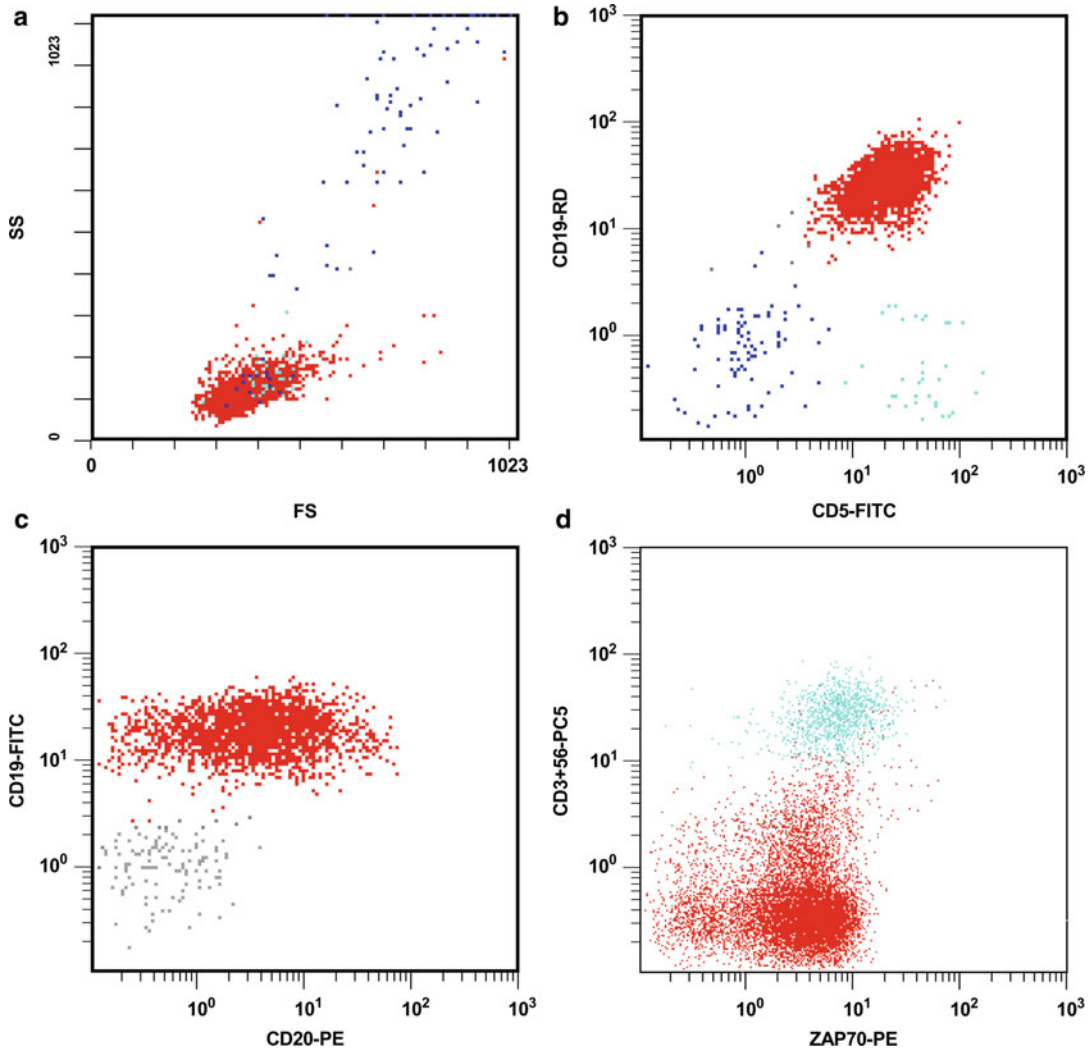


Fig. 9.19 (**a**) CLL/SLL are composed of small lymphoid cells (in *red*). (**b**) Neoplastic cells are positive for CD19 and coexpress CD5. (**c**) The expression of CD20 antigen is dim. (**d**) ZAP-70, the expression which has been linked to the unmutated variable region of Ig heavy chain, is expressed in a proportion of cases. Note internal positive control in aqua (T and NK cells)

– One, corresponding to the pre-germinal center phenotype (naïve, showing no mutations in the variable region of Ig heavy chain [V_H] gene)
– The second type is derived from memory B cells (post-germinal center, mutated V_H gene)
– The subclassification roughly corresponds to the expression of ZAP-70 (Fig. 9.19d) and CD38 molecules, which can be quantified by flow cytometry

Monoclonal B Cell Lymphocytosis

- Occasionally, low numbers of monoclonal B cells are seen in peripheral blood of asymptomatic elderly individuals without any other site of involvement by lymphoma. This condition, named monoclonal B cell lymphocytosis, has to be distinguished from circulating lymphoma cells
- Monoclonal B cell lymphocytosis (MBL) is defined as monoclonal B cell population in

peripheral blood of a patient without signs and symptoms of B cell lymphoproliferative neoplasm, autoimmune disease, or infectious process
- These cells can demonstrate an immunophenotype similar to that seen in CLL/SLL and are present at the number below the threshold for the diagnosis of CLL/SLL (less than 5,000/ul)
- The incidence is variable; overall 3.5% of healthy individuals over 40 years of age are diagnosed with MBL
- The prevalence is highly dependent on the sensitivity of flow cytometric methodology
- In a proportion of patients, MBL can precede the overt CLL/SLL or other B cell lymphoproliferative disorder
- The level of MBL is linked to the risk of development of CLL/SLL

Mantle Cell Lymphoma (MCL)
- MCL demonstrates expression of pan-B cell markers (CD19, CD20) and clonal surface light chains
- There is coexpression of CD5 antigen; however, CD23 is negative
- In contrast to CLL/SLL, the high-density CD20 and high-density light chains are seen, and there is coexpression of FMC7, the antigen that is invariably negative in typical cases of CLL/SLL
- The defining feature of MCL is the t(11;14), in which the *cyclin D1 (CCND1)* gene is translocated into the proximity of the Ig heavy chain gene promoter, resulting in the constitutive expression of this protein. The presence of cyclin D1 can be demonstrated by flow cytometry

Follicular Lymphoma (FL)
- The immunophenotype reflects the follicle center cell origin of FL
- Pan-B cell markers (CD19, CD20) are present along with the coexpression of CD10 and clonal surface Ig
- The coexpression of CD10, similar in density to that seen in reactive follicular hyperplasia, and relative low-density CD19 are characteristic features of FL
- In contrast to reactive germinal centers, neoplastic follicular cells express BCL2, which results in decreased sensitivity to apoptosis and allows for the accumulation of neoplastic lymphocytes. The expression of BCL2 by FL cells is due to the t(14;18)(q32;q21), which places the *BCL2* gene under a promoter of the Ig heavy chain gene

Marginal Zone Lymphomas
- Three subtypes of marginal zone lymphomas are recognized: nodal, extranodal (mucosa-associated lymphoid tissue lymphoma), and splenic. They all share similar immunophenotype
- Marginal zone lymphomas express the B cell markers (CD19, CD20, and CD22) and clonal surface light chains
- CD5 and CD10 are absent in the majority of cases
- No specific surface markers, routinely analyzed by flow cytometry, allow for subclassification of this lymphoma based on immunophenotype; therefore, the diagnosis is based on the absence of immunophenotypic features specific for other subtypes of lymphoma

Lymphoplasmacytic Lymphoma/Waldenstrom Macroglobulinemia
- Lymphoplasmacytic lymphoma is a B cell lymphoproliferative disorder composed of a heterogeneous proliferation of small B cells, lymphoplasmacytoid lymphocytes, and plasma cells
- The defining feature is the demonstration of monoclonal IgM protein in the serum
- B cell-associated antigens, including CD19, CD20, and CD22, are consistently expressed
- In the majority of the cases, the B cell immunophenotype is nondescript
- Frequency of coexpression of other antigens seen in B cell lymphomas such as CD5, CD23, and FMC7 is variable. CD10 is most frequently negative

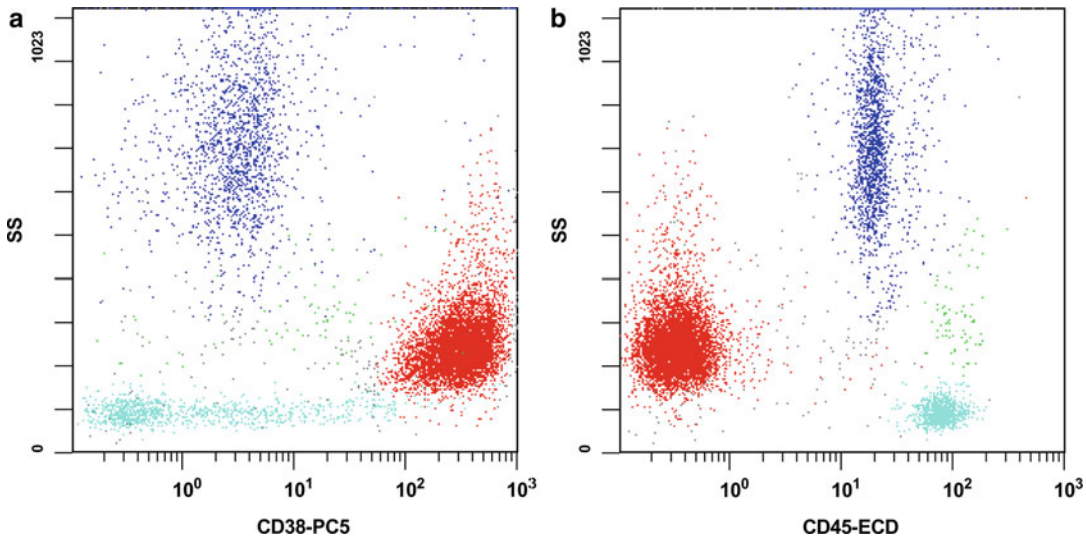


Fig. 9.20 Plasma cells, both neoplastic and reactive, are best visualized using CD38 antigen. (**a**) The reactive plasma cells and the majority of their neoplastic counterparts express characteristic high-density CD38 (in *red*). (**b**) The neoplastic plasma cells are frequently negative or dim positive for CD45 antigen

- The surface IgM expression can be demonstrated in all cases
- The majority of cases show dim positivity for CD25 antigen
- In a proportion of cases a second subset of monoclonal B cells is identified, corresponding to lymphoplasmacytoid lymphocytes (intermediate FSC and SSC, positive for CD19, CD20, FMC7, and bright CD38 and negative for CD138)
- In addition, as reported previously, a minute monoclonal plasma cell component can be identified in the majority of cases

Plasma Cell Neoplasms

- Plasma cell neoplasms are characterized by a monoclonal proliferation of terminally differentiated B cells, that is, plasma cells
- These disorders can present as a localized or disseminated process most commonly involving bone marrow, bone, and, more infrequently, extramedullary sites and peripheral blood
- Neoplastic plasma cells demonstrate a highly heterogeneous immunophenotype, different from that of their normal counterpart
- Due to overlap with various hematopoietic populations on SSC/FSC and SSC/CD45, the identification of plasma cells is best accomplished through their expression of CD138 and uniquely bright CD38 (Fig. 9.20a)
- The majority of neoplastic plasma cells, unlike their normal counterpart, show a decreased density of CD45 antigen (negative to weakly positive, Fig. 9.20b). Only about 20% of plasma cell myelomas demonstrate homogeneous bright to heterogeneous expression of CD45, similar to that of normal plasma cells
- Neoplastic plasma cells are most often negative for the pan-B cell markers, CD19 and CD20. The expression of CD20 is retained in approximately 20% of cases. CD19 is seen in $<5\%$ of neoplastic plasma cell proliferations
- Neoplastic plasma cells present with monoclonal cytoplasmic and occasionally surface Igs
- The CD56 antigen is seen in 70% of myeloma cases, and its absence has been associated with adverse clinical outcome
- The presence of myeloid markers including early antigens, such as CD117, is frequently reported

Diffuse Large BCL (DLBCL)

- The defining morphologic feature of DLBCL is a large cell size that can be appreciated on the displays of FSC vs SSC
- As in other B cell lymphomas, pan-B cell antigens such as CD19, CD20, and CD22 are positive
- DLBCL can originate from different stages in B cell development; hence, the coexpression of other markers is heterogeneous. The CD5, CD10, BCL-6, CD30, and CD138 can all be present

Burkitt Lymphoma (BL)

- This lymphoma is composed of medium-sized, highly proliferating lymphoid cells with basophilic vacuolated cytoplasm. The WHO classification distinguishes three variants of this lymphoma: endemic (occurring predominantly in Africa), sporadic, and immunodeficiency associated. All variants show similar immunophenotypic features
- The immunophenotype of BL reflects germinal center origin
  - CD19, CD20, CD10, and BCL-6 antigens are positive
  - BCL-2 is negative
- As in the majority of mature B cell neoplasms, there is surface expression of monoclonal Ig light chains
- Even though the immunophenotype of BL alone is not specific enough for the definitive subclassification of this lymphoma, the confirmation of high-proliferative activity associated with the expression of CD71 can be used to support the diagnosis. This feature is linked to the constitutive expression of *MYC* gene (cell cycle gate-keeping gene) due to its translocation under the promoter of Ig heavy or light chain genes (t(8;14), t(2;8), t(8;22)). The above translocations are pathognomonic of BL

9.6.3.2 Mature T and NK Cell Lymphomas

- Lymphomas derived from mature T and NK cells are much less common than the previously discussed mature B cell neoplasms and show greater geographic and ethnic variability
- The immunophenotypic features of T and NK cell malignancies are overlapping and frequently less specific than those seen in B cell lymphomas
- Expansion of a specific T cell subset (e.g., CD4 or CD8) with loss or altered intensity of T cell-associated markers (most commonly CD7, CD3, and CD5) is seen in the majority of T cell lymphomas (TCL)
- Aberrant immunophenotype is a reliable diagnostic feature only when the neoplastic population is significant since small numbers of T cells with unusual antigen makeup can also be seen in inflammatory conditions, autoimmune disorders, or viral infections. Thus, the aberrant immunophenotype alone cannot be considered pathognomonic of a T cell malignancy
- Considering these factors, an integration of morphologic, immunophenotypic, cytogenetic, molecular, and clinical information, as stressed by the WHO classification, is of particular importance in diagnosing T and NK cell malignancies
- Until recently, the demonstration of clonality in T cell proliferations was limited to molecular methods detecting *TCR* gene rearrangements (PCR or Southern blot analysis). The development of multiple Vβ-family antibodies directed against the variable region of the TCR β-chain allows for the determination of T cell clonality by flow cytometry. The determination of clonality is based on the preferential usage of a single Vβ-family and has close to 90% sensitivity and specificity in diagnosing T cell lymphoproliferative disorder. The results of this assay must be correlated with additional immunophenotypic and clinical data since rare cases with Vβ-family expansion have been reported in patients without a diagnosis of malignant lymphoma

T Cell Prolymphocytic Leukemia

- This is a mature TCL derived from helper T cells

- The majority of cases retain the mature T cell immunophenotype with expression of CD4 antigen
- Rare cases (~10%) can be double-positive or double-negative for CD4 and CD8
- CD25 can be coexpressed with the density similar to that of activated mature T cells

T Cell Large Granular Lymphocytic Leukemia (T-LGL)

- T-LGL is an indolent lymphoproliferative disorder derived from cytotoxic T cells
- The demonstration of a monoclonal aberrant T cell population in bone marrow and peripheral blood in a patient with cytopenias is a hallmark of the disease
- Considering the variable morphology of T-LGL, flow cytometric immunophenotyping plays a key role in diagnosing this disease
- The immunophenotype is similar to CD8+ T cells. Varying degrees of loss or decreased density of CD7, CD2, and CD3 were reported
- CD57 is positive in the majority of cases. A small number of cases express CD56 and/or CD16
- NK receptors for class I major histocompatibility molecules (both of killer cell Ig-like receptor type, CD158 antigens; and C-type lectin type, CD94, and NKG2 molecules) showed aberrant expression in the majority of reported cases supporting the diagnosis of T-LGL

Aggressive NK Cell Leukemia

- This is a rare neoplasm of NK cells typically presenting with systemic involvement and an aggressive clinical course
- As in normal NK cells, surface CD3 and CD5 are absent
- The CD2, CD56, and CD16 are present in the majority of cases. CD57 can be absent
- There is a varying degree of CD7 loss

Extranodal NK/T Cell Lymphoma, Nasal Type

- Nasal type NK/T cell lymphomas are EBV-positive extranodal proliferations derived from NK or cytotoxic T cells. Both upper respiratory tract and other extranodal sites can be involved
- The immunophenotypic features include positivity for CD56, CD2, frequently CD7, and the cytoplasmic ϵ-chain of CD3 antigen
- CD8 and CD4, as well as other T cell-associated markers, are negative
- Rare cases demonstrate cytotoxic T cell immunophenotype

Adult T Cell Leukemia/Lymphoma

- Adult T cell leukemia/lymphoma is a mature T cell neoplasm caused by human lymphotropic virus type 1 (HTLV-1)
- Even though classically it has been considered to be a neoplasm of helper T cells, subsequent studies demonstrated its immunophenotypic and functional similarity to regulatory T cells
- In addition to pan-T cell markers, CD3, CD5, and CD2, the expression of CD4, bright CD25, and regulatory T cell marker FoxP3 can be demonstrated by flow cytometry
- CD7 antigen and cytotoxic/NK cell markers are absent

Hepatosplenic TCL

- Hepatosplenic TCL is a disseminated TCL originating from γδ T cells
- The neoplastic cells are positive for surface CD3 and associated γδ TCR. Only rare cases of αβ TCR type were also reported
- Typically, CD5, CD4, and CD8 antigens are negative
- CD56 and CD16 are expressed in some cases. CD57 is negative

Angioimmunoblastic T Cell Lymphoma (AILT)

- AILT is most commonly diagnosed by the combination of morphologic, immunophenotypic, and clinical features. The origin of this lymphoma is a CD4+ T cell from the germinal center
- The pathologic diagnosis can be challenging due to morphologic heterogeneity and immunophenotypically mixed T cell

proliferation with a significant admixture of reactive component
- Even though the number of malignant T cells may be low, flow cytometry demonstrates the immunophenotypically aberrant T cell population in >90% of cases
- The immunophenotype is that of CD4+ mature T cells with a varying loss of the CD3 and CD7 antigens
- CD8 and CD56 antigens are absent
- Characteristic coexpression of CD10 is seen in 80% of cases

Mycosis Fungoides (MF) and Sezary Syndrome (SS)

- MF is the most common cutaneous lymphoma. SS presents as a disseminated disease with widespread skin involvement, lymphadenopathy, and circulating lymphoma cells
- Flow cytometry is most commonly utilized to demonstrate circulating MF/SS cells with a specific T cell immunophenotype that includes the expression of pan-T cell markers CD3, CD5, and CD2 along with CD4 antigen
- An important feature is the absence of the CD7 and CD26 antigens. Other T cell-associated antigens such as CD2, CD3, and CD5 can also be negative (Fig. 9.21)

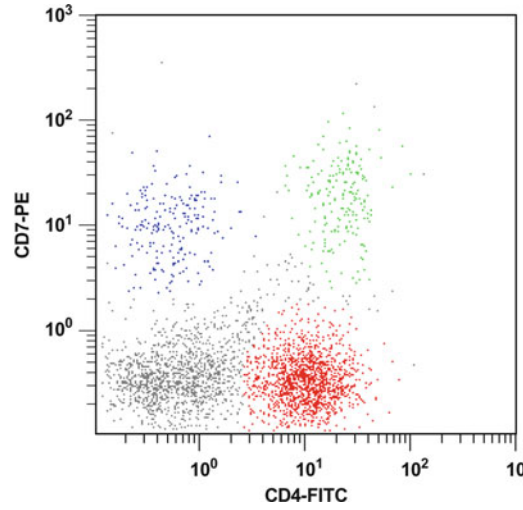


Fig. 9.21 MF and SS are positive for CD4 and commonly lose CD7 antigen (in *red*; residual helper [T cell] cells are presented in *green* and cytotoxic T cells are in *navy*)

Enteropathy-Associated TCL

- Enteropathy-associated TCL is derived from intraepithelial T cells and involves predominantly small bowel
- Flow cytometric immunophenotype was reported in rare cases and demonstrated neoplastic lymphoid cells positive for CD3, CD2, CD7, CD11c, and CD103 with variable loss of pan-T cell antigens, most commonly CD5
- Reported cases were positive for CD8 or double-negative for CD8 and CD4

Peripheral T Cell Lymphoma, Not Otherwise Specified (PTCL)

- This is a morphologically heterogeneous group of lymphomas with mature T cell phenotype
- The majority of the cases are derived from CD4+ T cells and retain this immunophenotype. pproximately 10% of cases show CD8 expression. Double-negative cases have also been reported
- Variable loss of pan-T cell antigens is seen. Most frequently CD7 (in 50% of cases) and surface CD3 are lost

Anaplastic Large Cell Lymphoma (ALCL)

- ALCL is composed of large pleomorphic cells, which, as other lymphomas with this morphology, often fall in the gate overlapping with large lymphoid cells and/or monocytes (Fig. 9.22)
- The presence of CD30 antigen and, in up to 70% of cases, ALK-1 protein is the defining immunophenotypic features of this lymphoma and can be demonstrated by flow cytometry. The overexpression of ALK-1 is most often due to t(2;5)(p23;35), between *ALK* gene and *nucleophosmin (NPM1)* gene. Alternative fusion partners for the *ALK* translocation have also been identified
- Various combinations of the CD30 antigen and T cell markers are seen including CD2, CD4, CD3, CD7, CD5, and CD8
- CD7 antigen is lost most frequently followed by CD5, CD3, and CD2
- CD25 antigen is positive in up to 90% of cases

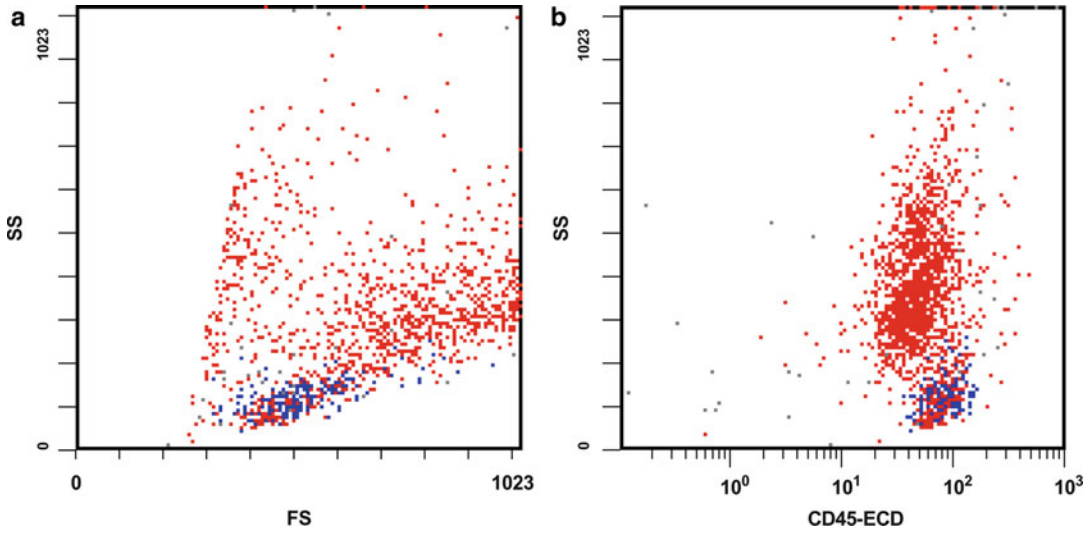


Fig. 9.22 (**a**) Cells of ALCL show high FSC and frequently increased SSC. (**b**) SSC/CD45 shows an overlap with monocyte gate

- In ALK positive cases, ALK protein can be demonstrated by flow cytometry
- Interestingly, the expression of myeloid markers (CD13, CD33, and CD15) has also been reported, which, especially in the rare cases with leukemic involvement, can bring a myeloid neoplasm into the differential diagnosis

9.7 Other Clinical Applications of Flow Cytometry

9.7.1 Primary and Secondary Immunodeficiencies

- Flow cytometry is commonly used to diagnose and immunophenotype primary and secondary immunodeficiencies. The detailed discussion of this topic is beyond the scope of this text. The following examples are provided to illustrate the most common applications
  - Immunophenotyping: the loss of specific antigens, such as β_2 integrins, is easily demonstrated by flow cytometry and is used to diagnose leukocyte adhesion deficiencies
  - Functional defects: an absence or low levels of NADPH oxidase, an enzyme involved in oxidative burst, occurs in chronic granulomatous disease. The level of enzymatic activity can be assayed using flow cytometry and correlate with specific genetic lesions (Fig. 9.23)
    - In the presence of adequate levels of NADPH oxidase, the nonfluorescent compound, dihydrorhodamine 123, converts to fluorescent rhodamine. The phorbol 12-myristate-13-acetate-stimulated granulocytes from healthy volunteers serve as a positive control (Fig. 9.23b)
    - The X-linked recessive form of the disease most frequently results in the complete absence of the enzymatic activity (Fig. 9.23c), whereas the autosomal recessive-type such as defect of p47-phox enzyme subunit presents with markedly decreased level of fluorescence (Fig. 9.23d)
  - Enumeration of CD4+ helper T cells in human immunodeficiency virus infection is performed using flow cytometry and serves as an indicator of disease progression and response to treatment
    - The absolute number of helper T cells in peripheral blood correlates with the stage of the disease and patient prognosis

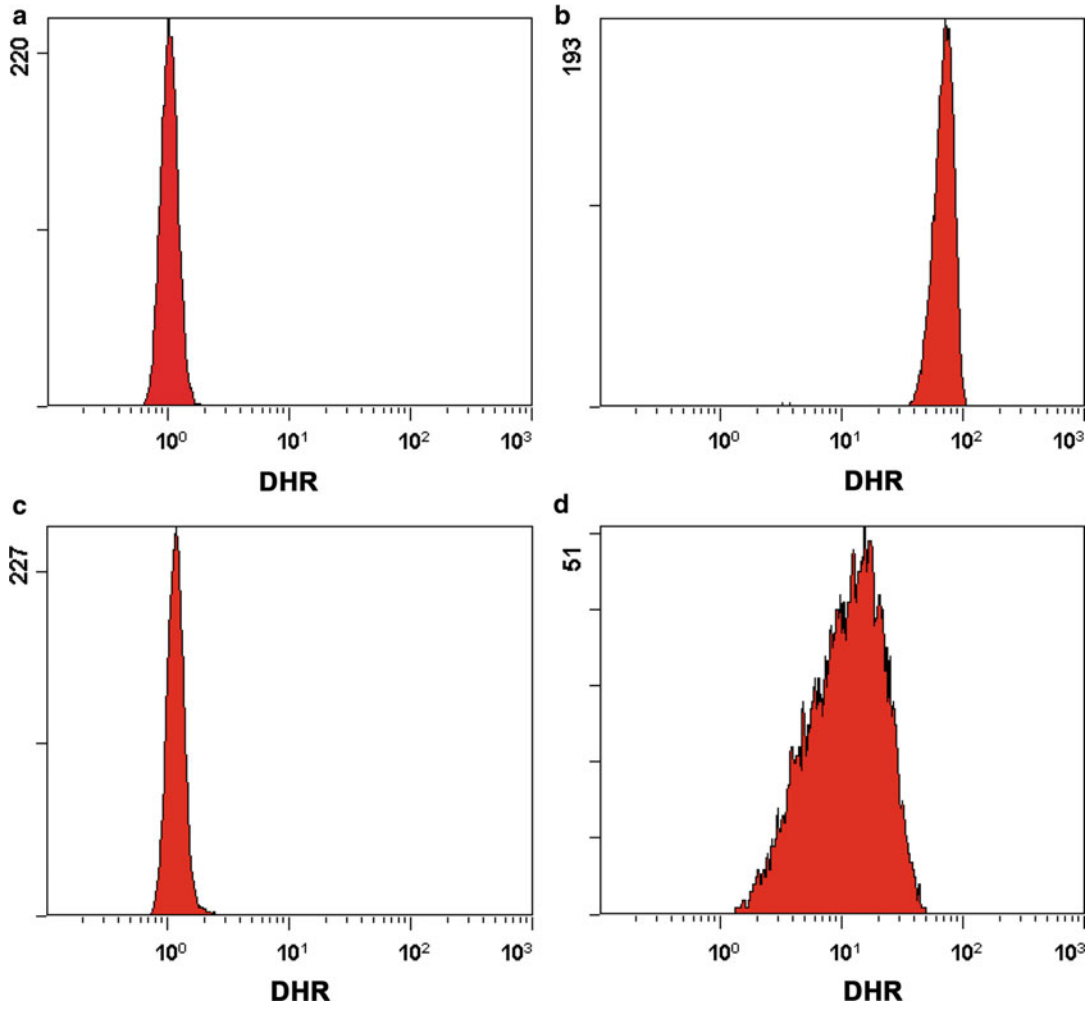


Fig. 9.23 Chronic granulomatous disease. (**a**) Unstimulated neutrophils from a healthy volunteer show baseline level of dihydrorhodamine 123 fluorescence. (**b**) Normal granulocytes with adequate activity of NADPH oxidase show a distinct shift in fluorescence intensity upon phorbol 12-myristate-13-acetate stimulation. (**c**) X-linked recessive chronic granulomatous disease shows complete loss of enzyme activity (no fluorescence shift). (**d**) Low-level enzyme activity in a patient with p47-phox deficiency corresponds to decreased fluorescence intensity as compared to a normal control

- The enumeration of T cells and their subsets is easily accomplished by flow cytometry using a simple combination of antibodies against CD3, CD4, and CD8 antigens
- The absolute numbers are derived either from a routine white blood cell count of the concurrent peripheral blood specimen (dual platform) or from calibrating beads run simultaneously with the patient sample (single-platform method)

9.7.2 Paroxysmal Nocturnal Hemoglobinuria (PNH)

- PNH is caused by an absence or decreased numbers of membranous glycosylphosphatidylinositol (GPI) anchor, which results in a loss of GPI-linked proteins
- Diagnosis of PNH by flow cytometry relies on the demonstration of decreased expression of two GPI-anchored proteins on two cell populations. Various combinations of

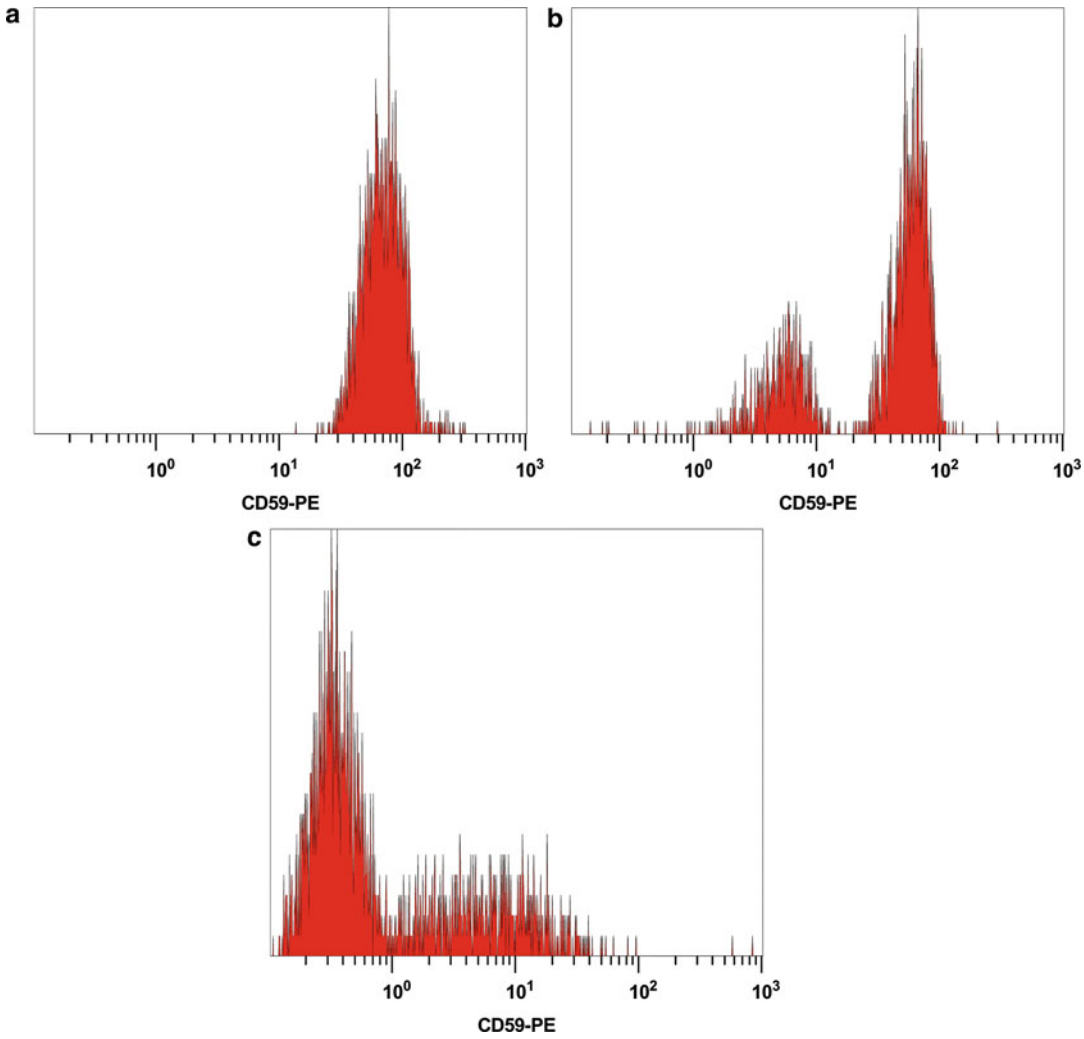


Fig. 9.24 The diagnosis of PNH is based on the absence of GPI-anchored molecules. Type I (normal), II (partial deficiency), and III (completed loss) cells are best recognized among red blood cells using antibody against CD59 antigen. (**a**) Red blood cells from healthy volunteer, used as a positive control, show high-intensity CD59 (correspond to type I cells). (**b**) Type I and II cells (small peak with decreased expression) in a patient with PNH. (**c**) Complete loss of CD59 (type III cells) in a patient with PNH supported by red blood cell transfusion

antibodies are applied, most commonly against the following antigens: CD55, CD59, CD24, CD14, CD66b, CD16, CD48, and CD157
  - The analysis of CD59 expression on red blood cells provides the best discrimination between type I, II, and III cells (Fig. 9.24)
  - Granulocytes or monocytes are frequently analyzed in addition to red blood cells
- FLAER reagent, which binds directly to GPI-anchors, is also frequently used

9.7.3 Stem Cell Transplantation

- Flow cytometry is utilized for enumeration of CD34+ stem cells and cell sorting
  - CD34 population can be enriched using flow cytometric sorting. Using this approach, heterogeneous populations can be physically separated into cell subsets with different physical or immunophenotypic properties

– High-speed flow cytometric sorting is achieved through charging of droplets containing individual cells with a specific polarity. As a charged droplet passes through the electrostatic field, it is isolated from the remainder of the sample and collected into a separate container

9.7.4 Novel Applications of Flow Cytometry

- Flow cytometry-based molecular testing: detection of PCR target amplicons using liquid bead array systems
- Tissue typing
- Assaying response to medications, for example, monitoring platelet activation after antiplatelet therapy

Further Reading

Adriaansen H, Boekhorst PAW, Hagemeijer AM, et al. Acute myeloid leukemia M4 with bone marrow eosinophilia (M4Eo) and inv(16)(p13q22) exhibits a specific immunophenotype with CD2 expression. Blood. 1993;81:3043–51.

Andrieu V, Radford-Weiss I, Troussard X, et al. Molecular detection of t(8;21)/AML1-ETO in AML M1/M2: correlation with cytogenetics, morphology and immunophenotype. Br J Haematol. 1996;92:855–65.

Beck RC, Stahl S, O'Keefe CL, et al. Detection of mature T cell leukemias by flow cytometry using anti T cell receptor V beta antibodies. Am J Clin Pathol. 2003;120:785–94.

Borowitz MJ, Rubnitz J, Nash M, et al. Surface antigen phenotype can predict TEL-AML1 rearrangement in childhood B-precursor ALL: a Pediatric Oncology Group study. Leukemia. 1998;12:1764–70.

Borowitz MJ, Craig FE, Digiuseppe JA, et al. Guidelines for the diagnosis and monitoring of paroxysmal nocturnal hemoglobinuria and related disorders by flow cytometry. Cytometry B Clin Cytom. 2010;78:211–30.

Chen I-M, Chakerian A, Combs D, et al. Post-PCR multiplex fluorescent ligation detection assay and flow cytometry for rapid detection of gene-specific translocations in leukemia. Am J Clin Pathol. 2004;122:783–93.

Chen W, Kesler MV, Karandikar NJ, et al. Flow cytometric features of angioimmunoblastic T cell lymphoma. Cytometry Part B (Clin Cyto). 2006;70B:142–8.

Ciudad J, Orfao A, Vidriales B, et al. Immunophenotypic analysis of CD19+ precursors in normal human adult bone marrow: implications for minimal residual disease detection. Haematologica. 1998;83:1069–75.

Domingo-Claros A, Larriba I, Rozmann M, et al. Acute erythroid neoplastic proliferations. A biological study based on 62 patients. Haematologica. 2002;87:148–53.

Edwards BS, Oprea T, Prossnitz ER, et al. Flow cytometry for high-throughput, high-content screening. Curr Opin Chem Biol. 2004;8:392–8.

Gorczyca W. Differential diagnosis of T cell lymphoproliferative disorders by flow cytometry multicolor immunophenotyping. Correlation with Morphology. Methods Cell Biol. 2004;75:595–621.

Harbott J, Mancini M, Verellen-Dumoulin C, et al. Hematological malignancies with a deletion of 11q23: cytogenetic and clinical aspects. Leukemia. 1998;12:823–7.

Helleberg C, Knudsen H, Hansen PB, et al. CD34+ megakaryoblastic leukaemic cells are CD38−, but CD61+ and glycophorin A+: improved criteria for diagnosis of AML-M7? Leukemia. 1997;11:830–4.

Human cell differentiation molecules. http://www.hcdm.org.

Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature. 1975;256:495–7.

Kohno T, Yamada Y, Akamatsu N, et al. Possible origin of adult T cell leukemia/lymphoma cells from human T lymphotropic virus type-1-infected regulatory T cells. Cancer Sci. 2005;96:527–33.

Krasinskas AM, Wasik MA, Kamoun M, et al. The usefulness of CD64, other monocyte-associated antigens, and CD45 gating in the subclassification of acute myeloid leukemias with monocytic differentiation. Am J Clin Pathol. 1998;110:797–805.

Kussick SJ, Wood BL. Four-color flow cytometry identifies virtually all cytogenetically abnormal bone marrow samples in the workup of non-CML myeloproliferative disorders. Am J Clin Pathol. 2003;120:854–65.

Li S, Eshleman JR, Borowitz MJ. Lack of surface immunoglobulin light chain expression by flow cytometric immunophenotyping can help diagnose peripheral B cell lymphoma. Am J Clin Pathol. 2002;118:229–34.

Lo Coco F, Avvisati G, Diverio D, et al. Rearrangements of the RAR-alpha gene in acute promyelocytic leukaemia: correlations with morphology and immunophenotype. Br J Haematol. 1991;78:494–9.

Loken MR, Shah VO, Dattilio KL, et al. Flow cytometric analysis of human bone marrow: I. normal erythroid development. Blood. 1987;69:255–63.

Macedo A, Orfao A, Ciudad J, et al. Phenotypic analysis of CD34 subpopulations in normal human bone marrow and its application for the detection of minimal residual disease. Leukemia. 1995;9:1896–901.

Matreo G, Castellanos M, Rasillo A, et al. Genetic abnormalities and patterns of antigenic expression in multiple myeloma. Clin Cancer Res. 2005;11:3661–7.

Ogata K, Nakamura K, Yokose N, et al. Clinical significance of phenotypic features of blasts in patients with myelodysplastic syndrome. Blood. 2002;100: 3887–96.

Perfetto SP, Chattopadhyay PK, Roederer M. Seventeen-colour flow cytometry: unraveling the immune system. Nat Rev Immunol. 2004;4:648–55.

Petrella T, Bagot M, Willemze R, et al. Blastic NK cell lymphomas (Agranular CD4 + CD56 + hematodermic neoplasms). Am J Clin Pathol. 2005;123:662–75.

San Miguel JF, Vidriales MB, Ocio E, et al. Immunophenotypic analysis of Waldenstrom's macroglobulinemia. Semin Oncol. 2003;2:187–95.

Stelzer GT, Shults KE, Loken MR. CD45 gating for routine flow cytometric analysis of human bone marrow specimens. Ann N Y Aca Sci.. 1993;677:265–81.

Stetler-Stevenson M, Arthur DC, Jabbour N, et al. Diagnostic utility of flow cytometric immunophenotyping in myelodysplastic syndrome. Blood. 2001;98: 979–87.

Stewart CC, Behm FG, Carey JL, et al. U.S.-Canadian consensus recommendations on immunophenotypic analysis of hematologic neoplasia by flow cytometry: selection of antibody combinations. Cytometry. 1997;30:231–5.

Swerdlow SH, Campo E, Harris NL, et al., editors. World Health Organization classification of tumours of haematopoietic and lymphoid tissues. Lyon: IARC Press; 2008.

Tabernero MD, Bortoluci AM, Alaejos I, et al. Adult precursor B-ALL with BCR/ABL gene rearrangements displays a unique immunophenotype based on the pattern of CD10, CD34, CD13 and CD38 expression. Leukemia. 2001;15:406–14.

Terstappen LWMM, Safford M, Loken MR. Flow cytometric analysis of human bone marrow III. Neutrophil maturation. Leukemia. 1990;4:657–63.

Terstappen LWMM, Huang S, Picker LJ. Flow cytometric assessment of human T cell differentiation in thymus and bone marrow. Blood. 1992;79:666–77.

van Lochem EG, van der Velden VHJ, Wind HK, et al. Immunophenotypic differentiation patterns of normal hematopoiesis in human bone marrow. Reference patterns for age-related changes and disease-induced shifts. Cytometry. 2004;60B:1–13.

Wells DA, Benesch M, Loken MR, et al. Myeloid and monocytic dyspoiesis as determined by flow cytometric scoring in myelodysplastic syndrome correlates with the IPSS and with outcome after hematopoietic stem cell transplantation. Blood. 2003;102:394–403.

Wood BL, Arroz M, Barnett D, et al. 2006 Bethesda International Consensus recommendations on the immunophenotypic analysis of hematolymphoid neoplasia by flow cytometry: optimal reagents and reporting for the flow cytometric diagnosis of hematopoietic neoplasia. Cytometry B Clin Cytom. 2007;72(Suppl 1):S14–22.

Wu JM, Borowitz MJ, Weir EG. The usefulness of CD71 expression by flow cytometry for differentiating indolent from aggressive CD10+ B cell lymphomas. Am J Clin Pathol. 2006;126:1–8.

Clinical Genomics in Oncology

10

Hugo M. Horlings, Thalia A. Farazi, and Marc J. van de Vijver

Contents

10.1 Introduction 336
10.1.1 Clinical Genomics 336
10.1.2 Human Genome Project 336
10.1.3 Methodologies and Applications of Genomics 336
10.1.4 Limitations 336

10.2 DNA Microarray 337
10.2.1 Microarray Experiment 337
10.2.2 Isolation and Amplification of RNA 338
10.2.3 Different Labeling Methods of Amplified mRNA 338
10.2.4 Quantifying the RNA Expression 338

10.3 Gene Expression 339
10.3.1 Dynamic and Systematic 339
10.3.2 Gene Expression Profiling 339

10.4 Data Analysis of Gene Expression Data 340
10.4.1 Normalization of Gene Expression Data 340
10.4.2 Statistical Analysis 340
10.4.3 Unsupervised Classification 342
10.4.4 Supervised Classification (Knowledge Driven) 343
10.4.5 Classifying Tumors Based on "Functional" Gene Expression Signatures 345
10.4.6 Predictive Genomic and Clinicogenomic Decision Tree Models 345

10.5 Application of Gene Expression Profiles in Diagnostics of Clinical Oncology 346
10.5.1 Clinical Genomics in Breast Cancer 346
10.5.2 Clinical Genomics in Hematologic Malignancies 350
10.5.3 Clinical Genomics in Prostate Cancer 351
10.5.4 Clinical Genomics in Gastrointestinal Tumors 351
10.5.5 Clinical Genomics in Carcinoma of Unknown Primary 352

10.6 Interpretation of Gene Lists 352
10.6.1 Combining Clinicopathological and Genomic Risk Predictors 353

10.7 Points of Attention in the Design of Microarray Experiments 353
10.7.1 The Sample 353
10.7.2 The Microarray 353
10.7.3 The Statistics 353

10.8 High-Throughput Genomic Sequencing 353
10.8.1 DNA Sequencing Technologies 354
10.8.2 Sample Preparation 360
10.8.3 Second-Generation Sequencing 360
10.8.4 Third-Generation Sequencing 361

10.9 Systematic Search and Catalogues for Genetic Alterations in Cancer 362

10.10 Next-Generation Sequencing Applications 362

H.M. Horlings, MD, PhD (✉)
Academic Medical Center, Amsterdam, The Netherlands

T.A. Farazi, MD, PhD
Laboratory of RNA Molecular Biology, The Rockefeller University, New York, NY, USA

M.J. van de Vijver, MD, PhD
Head, Division of Diagnostic Oncology, NKI, Amsterdam

Head, Department of Pathology, NKI, Amsterdam

Pathologist, Department of Pathology, Leiden University Medical Center, The Netherlands

Professor of Pathology, Department of Pathology, The Netherlands

Cancer Institute/Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_11,

10.10.1 Application of High-Throughput Sequencing in Diagnostics of Clinical Oncology 363

10.11 Challenges of Using High-Throughput Sequencing Technologies 364
10.11.1 Bioinformatic Obstacles 364
10.11.2 Sufficient Amount of Tissue 365
10.11.3 Ethical Considerations 365

10.12 Conclusions and Future Directions 365

Further Reading 366

10.1 Introduction

10.1.1 Clinical Genomics

- Genomics is the study of all nucleotide sequences, including structural genes, regulatory sequences, and noncoding DNA segments, in the chromosomes of an organism
- Clinical genomics can be defined as the application of large-scale, high-throughput genomic technologies in clinical settings, such as clinical trials or primary care of patients

10.1.2 Human Genome Project

- The field of clinical genomics has grown enormously by the elucidation of the full sequence of the human genome and the characterization of all 25,000 human genes as well as the ability of large-scale surveys of gene expression, DNA sequence, single-nucleotide polymorphisms, and DNA copy numbers using microarray-based and high-throughput sequencing technologies

10.1.3 Methodologies and Applications of Genomics

- DNA microarray analysis is the most commonly used method to measure gene expression levels
- This method has been most successfully applied to characterize human cancers in order to predict clinical outcomes and define clinically relevant subgroups of tumors
- The completion of the human genome project and advances in sequencing technologies have greatly accelerated genomic cancer research

10.1.4 Limitations

- Most analyses have used gene expression levels to define broad group differences. As a consequence, there remains a considerable

diversity within these groups, and thus predictions often fall short for individual patients
- An extremely important step in obtaining reliable prognostic or predictive gene expression signatures is validation in an independent, sufficiently large cohort of patients
- Implementation of high-throughput sequencing technologies into the clinic requires an in-depth understanding of the pitfalls and biases of these technologies and a great degree of bioinformatic expertise

10.2 DNA Microarray

- A microarray is any arrangement of microscopic spots attached to a solid surface such as glass, plastic, beads, or silicon chip (Fig. 10.1)
- An array may contain thousands of spots, and the corresponding probes attached to the solid support can be the complementary DNAs (cDNA), oligonucleotides of varying length, or genomic sequences
- Most laboratories use fluorescent labeling with one or two dyes, Cy3 and Cy5 (excited by a green and red laser)
- In dual-label experiments, two samples are hybridized to a microarray, one labeled with each dye. This allows the simultaneous measurement of two samples (tumor sample vs. reference sample; that is, commercially available platform from Agilent [Santa Clara, CA] [Table 10.1])
- In single-label experiments, only one sample is hybridized to the arrays labeled with one dye. Affymetrix developed this type of hybridization (Santa Clara, CA) (Table 10.1). For each gene, several different oligonucleotides are present on the array. The hybridization is done with RNA from the sample to be analyzed without the use of a reference RNA. Instead, oligonucleotides containing one mismatch in their sequence are used to correct for background hybridization

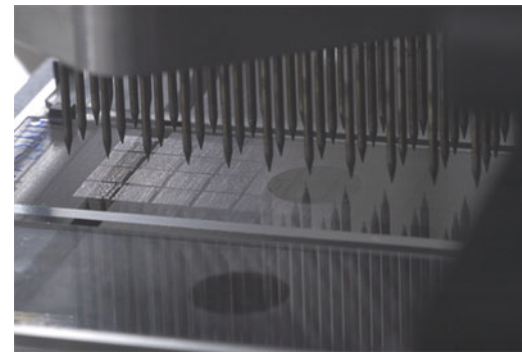


Fig. 10.1 Printing of a glass slide with 25,000 spots (CMF, NKI, The Netherlands)

10.2.1 Microarray Experiment

- Summary of the different steps of a microarray
  - Prepare a microarray chip by choosing probes (sequences)
  - Isolate and amplify RNA from tumor and reference sample (Fig. 10.2), and if a reference is used, the reference is usually prepared from a mixture of cell lines, tumor samples, or normal tissues
  - Generate a hybridization solution containing fluorescently labeled targets. Antisense cRNA is labeled in the presence of red fluorescent (Cy5) nucleotides and can then be mixed with a green fluorescent-labeled reference (Cy3) (Fig. 10.3a, b)
  - The mixture is added to the microarray, and the labeled antisense cRNA can be hybridized to the probes on the microarray
  - Detect probe hybridization
    - Scan the array using a fluorescent scanner and store output as an image
    - Quantify each spot
    - Subtract background
    - In a dual-label experiment, the level of fluorescence is digitized, and for each probe, the level of gene expression, relative to the reference, is determined and transferred to a database (Fig. 10.4)
    - Normalize data
    - Export a table of fluorescent intensities for each gene

Table 10.1 Examples of different types of microarrays

	Platforms			Platforms			
Types of microarray	Array spots	Target	Year	Array spots	Target	Year	Manufacturer
cDNA	50–6,400	cDNA clones	1996	Use is decreasing			
Nylon nitrocellulose	9,600	cDNA clones	1996	Use is decreasing			
Oligonucleotide	64,000	20 mer oligos	1998	500,000	25 mer oligos	2005	Affymetrix
Oligonucleotide	22,000	60 mer oligos	1997	44,000	60 mer oligos	2006	Agilent
Exon array	–	–	–	One million exon clusters		2005	Affymetrix
Bead array	–	–	–	46,000	50 mer oligos	2005	Illumina
Bacterial artificial chromosome array	1–2,000	Large insert clones	1992	350,000	45–85 mer oligos	2005	Nimblegen
microRNA multi-species array	–	–	–	2,708	34–44 mer oligos	2006	Invitrogen
Protein array	–	–	–	5,000	4,800 unique human proteins	2006	Invitrogen

Adapted with permission from Abdullah-Sayani et al. (2006)

- Analyze exported data
- A more detailed description of each step will follow next

10.2.2 Isolation and Amplification of RNA

- RNA is isolated totally from tumor sample (fresh frozen, cell culture, or paraffin-embedded tissue) and (if a reference is needed) reference sample (4 μg of total RNA) (Fig. 10.2)
- First-strand synthesis to generate first-strand cDNA by reverse transcriptase with cRNA-specific primer
- The cDNA is used as template to generate a double-stranded DNA molecule using exogenous primers
- After generation of the double-stranded DNA, it is used as template for an in vitro transcription reaction to generate many copies of amplified antisense cRNA, which are complementary to the sequence of the original mRNA species
- The amplified antisense cRNA is ready for labeling. If a large amount of sample RNA is used, amplification is not required

10.2.3 Different Labeling Methods of Amplified mRNA

- Incorporation of modified nucleotides in cDNA or cRNA can be done using different enzymes (Fig. 10.3a, b)
- It is also possible to directly, chemically label DNA, RNA, or nucleotides using universal labeling system (ULS™, Amsterdam, The Netherlands) by forming a covalent bond on the N7 position of guanine
- Labeled tumor sample (Cy5) and reference sample (Cy3) are hybridized to a microarray (Fig. 10.4a)

10.2.4 Quantifying the RNA Expression

- After hybridization of the labeled RNA, a laser scanner is used to excite the hybridized

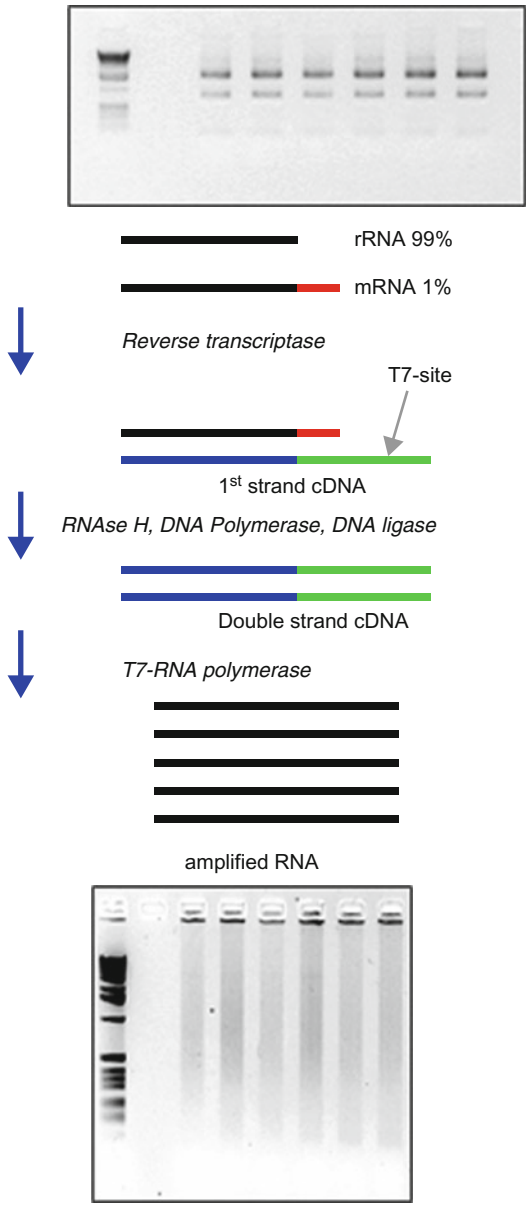


Fig. 10.2 Isolation and amplification of messenger RNA (Image from Kerkhoven RM)

array at the appropriate wavelength. In a dual-label experiment, the relative abundance of the two transcripts is visualized in colored image by the ratio of the red to green fluorescence intensities of each spot (Fig. 10.4a, b)

- Depending on the type of array used, the location and intensity of a color will indicate which gene is expressed
- The gene expression ratios are log-transformed and placed in a table in which each row corresponds to a gene and each column corresponds to a single hybridization experiment. This is represented in a so-called heat map (Fig. 10.5)
- In a single-label experiment (Affymetrix), a reference RNA is not required

10.3 Gene Expression

- Gene expression is a broad term used to describe the transcription of information encoded within DNA sequences (exons) into mRNA. It is assumed that for most transcripts, there is a linear relationship with translation of the mRNA information into proteins that regulate cell function

10.3.1 Dynamic and Systematic

- The gene expression pattern in any given cell is a highly dynamic and systematic process that alters with cell metabolism, changes in environment, and the presence of disease

10.3.2 Gene Expression Profiling

- Characterization of gene expression (gene expression profiling) has been used to study infectious and immunologic disease, but the predominant focus has been on the study of cancer
- Cancer is a genetic disease where the abnormal interaction of several genes results in the development and progression of a tumor
- With gene expression profiling, it has been possible to group genes to formulate "genetic signatures" that can potentially improve the clinical management of a cancer patient or give more insight in biologic processes of cancer

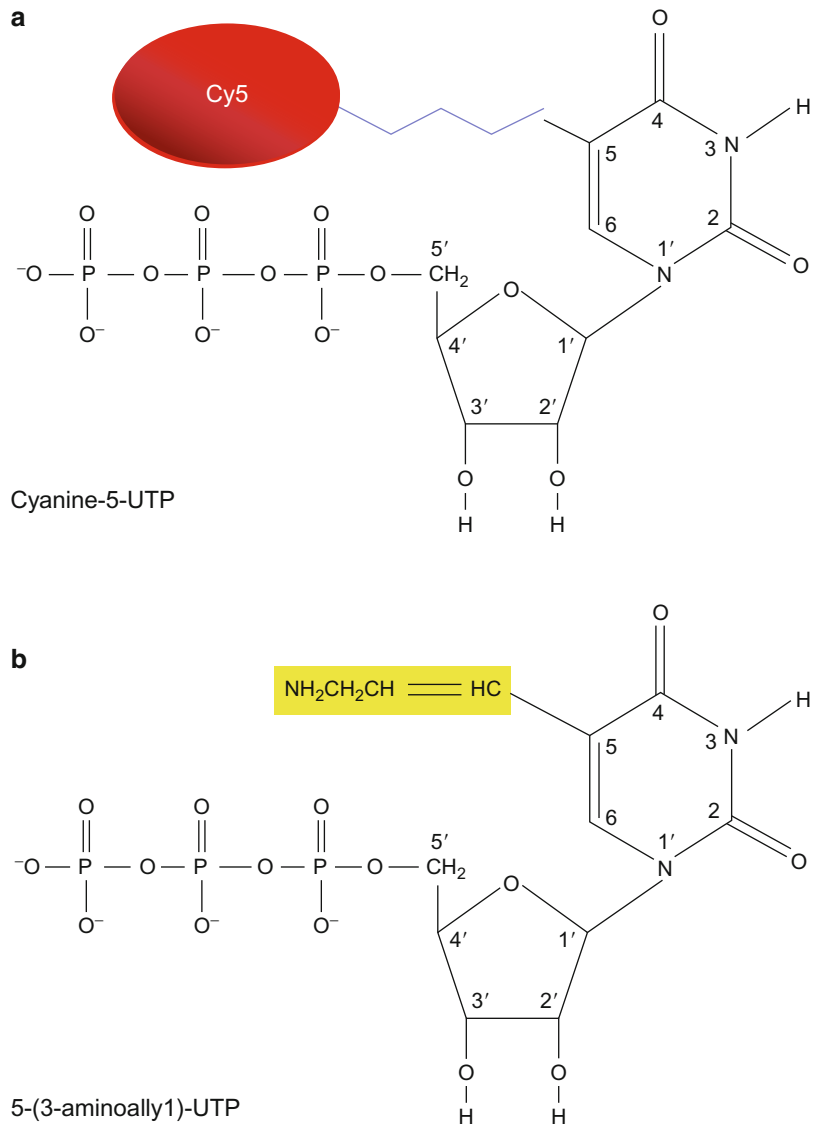


Fig. 10.3 (**a**) Cyanine-UTP. (**b**) Aminoallyl-UTP-monoreactive esters (Image from Kerkhoven RM)

10.4 Data Analysis of Gene Expression Data

10.4.1 Normalization of Gene Expression Data

- Experiment normalizations are used to standardize microarray data to enable differentiation between biologic variations in gene expression levels and variations due to the measurement methods
- A commonly used normalization method is the locally weighted scatterplot smoothing algorithm

10.4.2 Statistical Analysis

- After normalization, the purpose of any data analysis of gene expression data is to group entities on the basis of similarity of features

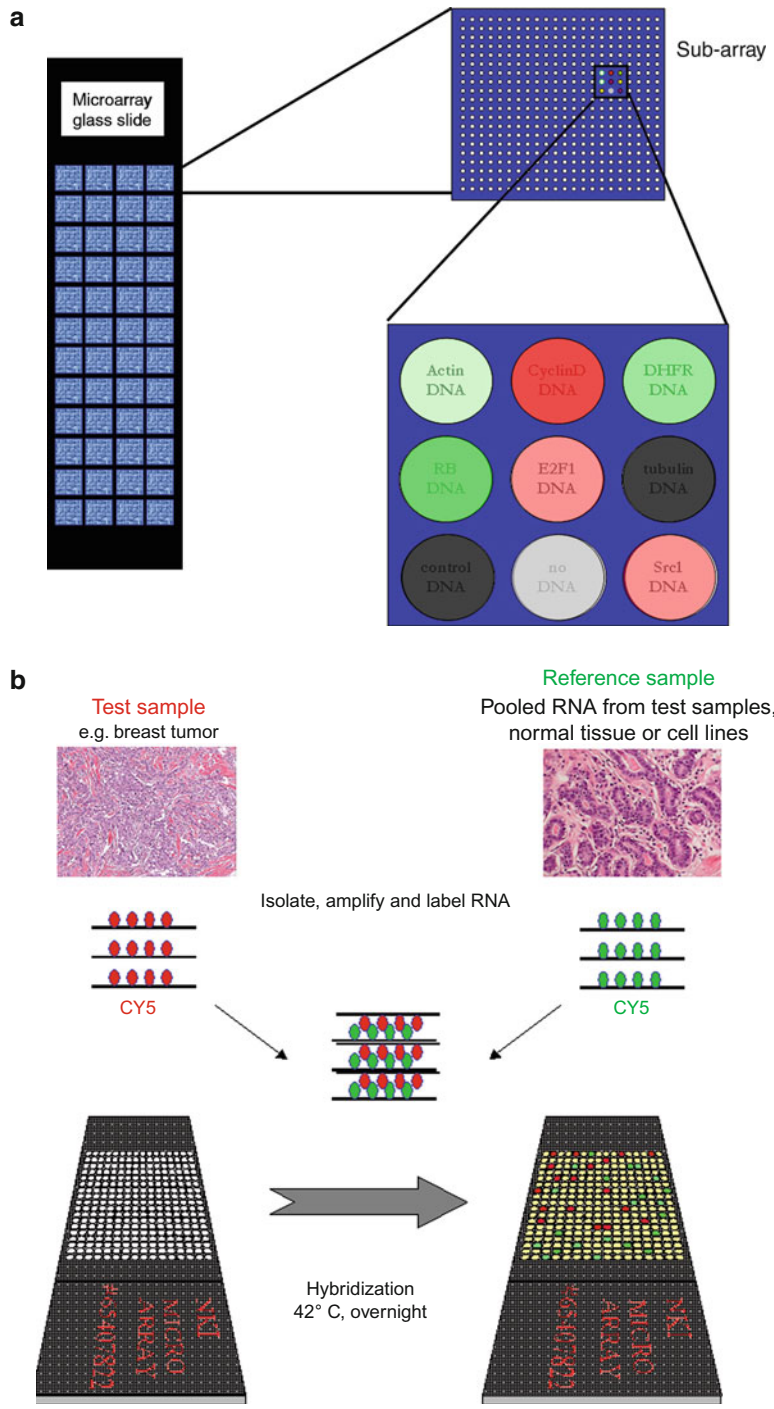


Fig. 10.4 (**a**) Microarray experiment. (**b**) Results of hybridization to microarray. *Green* indicates low expression (where the reference sample hybridized stronger to the target DNA than the tumor sample). *Red* indicates high expression (where the tumor sample hybridized more to the target DNA than the reference sample). *Black* represents a combination of control and sample DNA, where both hybridized equally to the target DNA. *Gray* represents areas where neither the control nor sample DNA hybridized to the target DNA (Adapted from image by Kerkhoven RM)

- Clustering (unsupervised and supervised) was one of the first methods to order microarray data
- Other methods to analyze gene expression data include principle component analysis, self-organizing maps, and linear discriminant

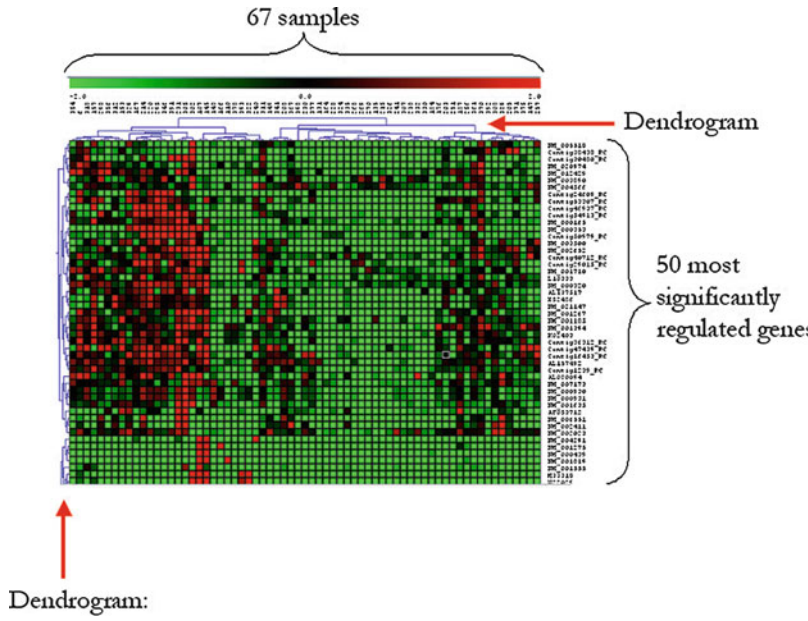


Fig. 10.5 Example of an expression data matrix. 67 samples (x-axis) and 50 genes (y-axis). Example of an expression data matrix (heat map). Hierarchical clustering of 67 samples using relative ratios (dual-label experiments) of 50 genes. Each row represents a separate gene and each column a separate sample. Shown are the relative ratios of hybridization of each sample to the reference. These ratios are a measure of relative gene expression in each experimental sample and are represented in the heap map. A color scheme is shown at the *top*. As indicated, the scale extends from fluorescence ratios of −2 to +2 log base 2 units. The dendrogram at the top indicates the similarities among the samples in their expression. The dendrogram at the *left side* indicates similarities among genes in their gene expression. *Green* indicates low expression. *Red* indicates high expression. *Black* represents a combination of control and sample DNA, where both hybridized equally to the target DNA

analysis to discover patterns of gene expression
- The main methods used to identify categories of tumors based on gene expression profiles are unsupervised and supervised classification (Fig. 10.6a, b); examples are given as follows

10.4.3 Unsupervised Classification

- One of the most commonly used unsupervised classification techniques is hierarchical cluster analysis. Hierarchical cluster analysis groups genes with similar expression patterns, with the assumption that each cluster of genes is simultaneously regulated
- Samples are clustered into groups based on overall similarity of their gene expression profiles
- Usually the result of two-dimensional clustering is represented by the combination of a "heat map," showing expression levels of each of the genes combined with a dendrogram (Fig. 10.5)
  - Example of hierarchical cluster analysis
    - Perou and Sorlie et al. showed by unsupervised classification that similar breast tumors might now be classified into five specific subtypes (basal, HER2 [HER2/neu or ERBB2], luminal A, luminal B, and normal-like) according to their distinct patterns of gene expression. Others have confirmed this classification

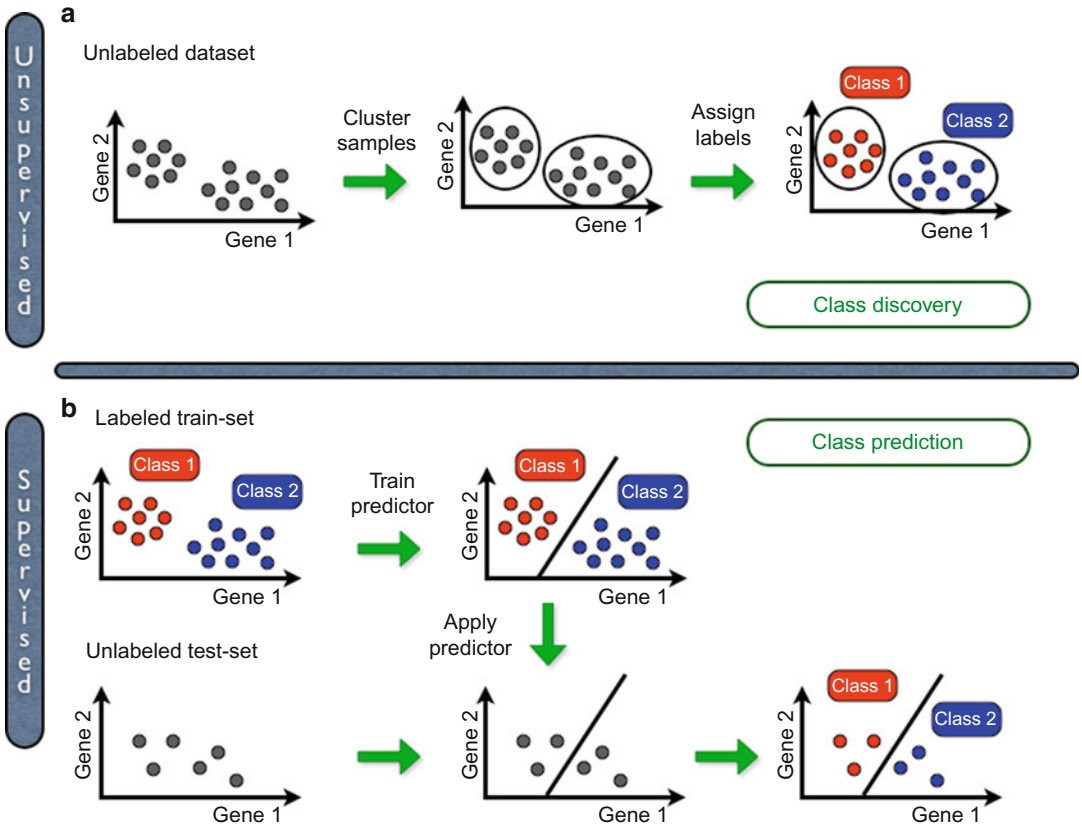


Fig. 10.6 Unsupervised and supervised clustering. (**a**) Unsupervised clustering: tumor samples are clustered into groups based on overall similarity of their gene expression profiles. This approach is useful for class discovery. (**b**) Supervised clustering: multiple tumor samples from different known classes are used to train a model capable of classifying unknown samples. This model is then applied to a tumor set for class label assignment (Image from C. Lai)

10.4.4 Supervised Classification (Knowledge Driven)

- Supervised classification is the method of choice for the analysis of gene expression profiles associated with prognosis or therapy response prediction
- Multiple samples from different known classes are used to train a model capable of classifying unknown samples. This model is then applied to a tumor set for class label assignment. There are several methods for supervised classification; examples are given as follows
  - Examples of supervised classification
- Van't Veer et al. studied invasive breast carcinomas by analyzing gene expression using an oligonucleotide array containing 25,000 probes
- DNA microarray analysis was performed on primary breast tumors of 78 young patients, and supervised classification was applied to identify a gene expression signature strongly predictive of a short interval to distant metastases ("poor prognosis" signature) in patients without tumor cells in local lymph nodes at diagnosis (lymph node negative) (Fig. 10.7)

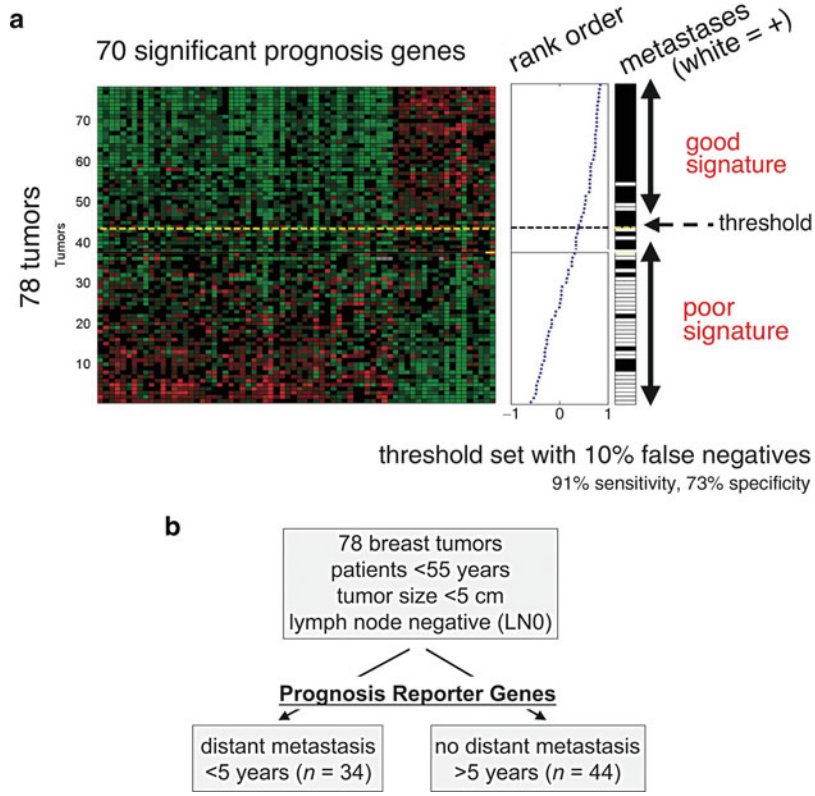


Fig. 10.7 Expression data matrix of 70 prognostic marker genes. (**a**) Expression data matrix of 70 prognostic marker genes from tumors of 78 breast cancer patients (*left panel*). Each row represents a tumor and each column a gene. Genes are ordered according to their correlation coefficient with the two prognostic groups. Tumors are ordered by the correlation to the average profile of the good prognosis group (*middle panel*). *Solid line*, prognostic classifier with optimal accuracy; *dashed line*, with optimized sensitivity. Above the dashed line patients have a good prognosis signature; below the *dashed line* the prognosis signature is poor. The metastasis status for each patient is shown in the *right panel: white* indicates patients who developed distant metastases within 5 years after the primary diagnosis; *black* indicates patients who continued to be disease-free for at least 5 years (**b**) Use of prognostic reporter genes to identify optimally two types of disease outcome from 78 sporadic breast tumors into a poor prognosis and good prognosis group. (Adapted from van't Veer et al. 2002)

- By performing supervised classification, different levels of expression have been found in the 34 patients who developed distant metastases within 5 years compared with the 44 patients with no recurrence in this period
- The 70-gene poor prognosis signature consists of genes regulating cell cycle, invasion, metastasis, and angiogenesis
- In a subsequent validation study, van de Vijver et al. investigated the prognostic power of these 70 genes in 295 patients (151 lymph node negative; 144 lymph node positive; age <53 years) (Fig. 10.8)
- Figure 10.8 shows that patients with a good prognosis signature had a <15% risk of developing distant metastases over 10 years and a <10% risk of dying. Patients with a 70-gene poor prognosis signature had a 50% risk for distant metastases and a 50% mortality rate

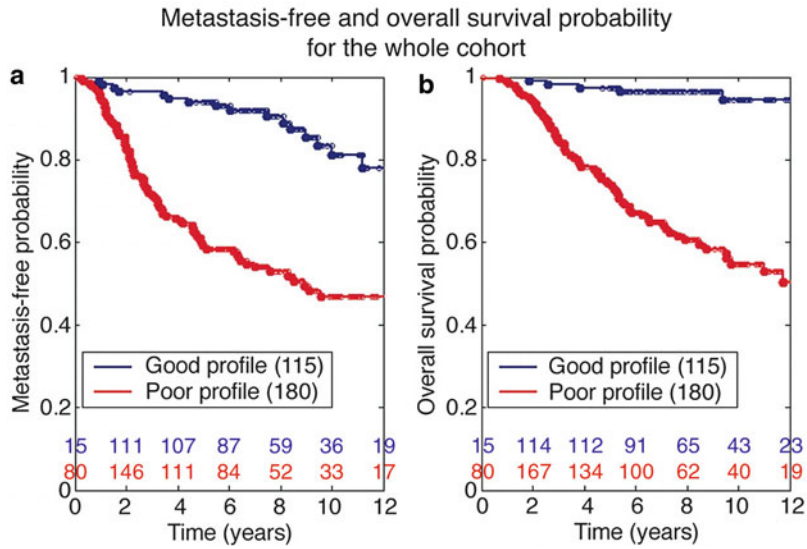


Fig. 10.8 Validation study investigating the prognostic power of the 70-gene signature. Validation study investigating the prognostic power of the 70 genes in stage I and II patients ($n = 295$). Kaplan–Meier curves of the probability that patients would remain free of distant metastases (**a**) and the probability of overall survival (**b**) among all patients (Adapted from van de Vijver et al. 2002)

10.4.5 Classifying Tumors Based on "Functional" Gene Expression Signatures

- Classifying tumors could also be based on functional annotation of gene expression algorithms; an example is provided as follows
- Chang et al. applied this strategy to identify a gene expression signature of wound response and tested its role in cancer progression
- Fifty fibroblast cultures derived from ten anatomic sites were cultured in 10% fetal bovine serum or in media containing only 0.1% fetal bovine serum
- Analysis of the global gene expression patterns, using human cDNA microarrays containing approximately 36,000 gene probes, revealed that although fibroblasts from different sites have distinctly different gene expression programs, they share a stereotype gene expression program in response to serum exposure (Fig. 10.9)
- Based on the genes in the serum-response signature, two groups of breast carcinomas could be recognized: tumors with an "activated" wound gene expression signature suggestive of active wounds and tumors with a "quiescent" gene expression signature (Fig. 10.10)
- In particular, Chang et al. found the wound signature to be an extremely strong predictor of death and metastasis in the panel of 295 breast tumors
- Given the ability of the wound signature to accurately predict metastasis, the signature further predicted which breast cancer patients would have benefited from chemotherapy (Fig. 10.11)

10.4.6 Predictive Genomic and Clinicogenomic Decision Tree Models

- Gene expression signatures can also be combined; an example is given as follows
  - Example
    - Chang et al. also developed a decision tree for combining the 70-gene signature and wound signature
    - First, patients were classified according to the 70-gene prognosis profile into the good or poor prognosis group. Subsequently, the tumors from the

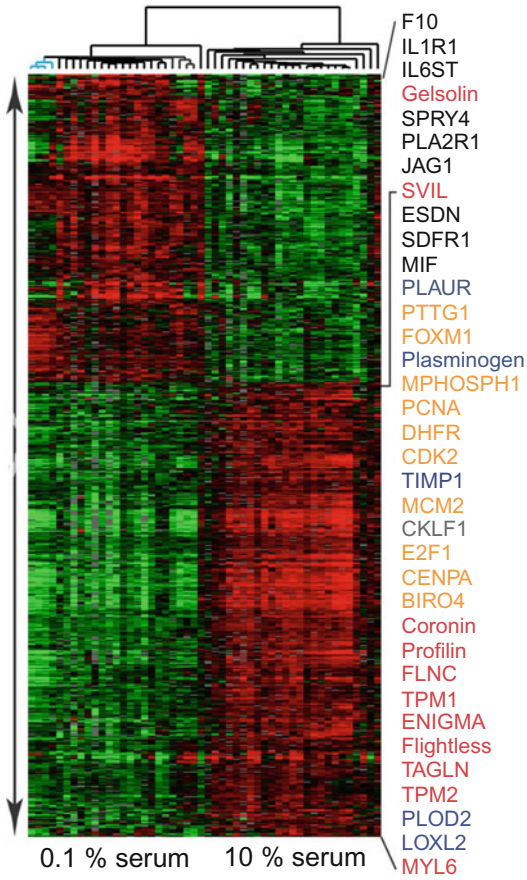


Fig. 10.9 The fibroblast common serum-response or "wound-response" signature. Genes with expression changes that demonstrate coordinate induction or repression by serum in fibroblasts is shown. *Red* indicates increased expression and green indicates reduced expression in response to serum. Representative genes with probable function in cell cycle progression (*orange*), matrix remodeling (*blue*), cytoskeletal rearrangement (*red*), and cell–cell signaling (*black*) are highlighted by colored text on the right (Adapted from Chang et al. 2004)

poor prognostic group were classified according to the wound signature as wound activated or quiescent (Fig. 10.12)

- Those patients with a poor prognosis 70-gene profile but a quiescent wound signature showed a risk similar to baseline, whereas those patients with both poor prognosis and activated wound-response signature showed a risk of metastatic disease 6.4-fold higher than baseline
- This approach shows that combining different signatures with nonoverlapping features can be used to strengthen the predictor

10.5 Application of Gene Expression Profiles in Diagnostics of Clinical Oncology

- Gene expression profiling using microarrays has been used to study several tumor types, most notably breast, ovary, colon, gastric, pancreatic, prostate, lung, melanoma, leukemia, and malignant lymphoma
- To understand the current status and relevance of gene expression profiles that have been developed, we have highlighted a few important examples as follows

10.5.1 Clinical Genomics in Breast Cancer

10.5.1.1 First-Generation Prognostic Signatures

- In breast cancer research, several gene expression profiles associated with prognosis have been identified, for example, a 70-gene classifier, a 21-gene signature, and a 76-gene expression profile. The 70-gene prognosis profile has been described in the Sect. 10.4.2; the other two gene expression signatures are described next
  - 21-gene signature
    - In 2004, the company Genomic Health in collaboration with the American National Surgical Adjuvant Breast and Bowel Project identified a "recurrence score" (RS)
    - This score of 21 genes quantifies the likelihood of distant metastasis in tamoxifen-treated patients with lymph node-negative, estrogen-positive breast cancer

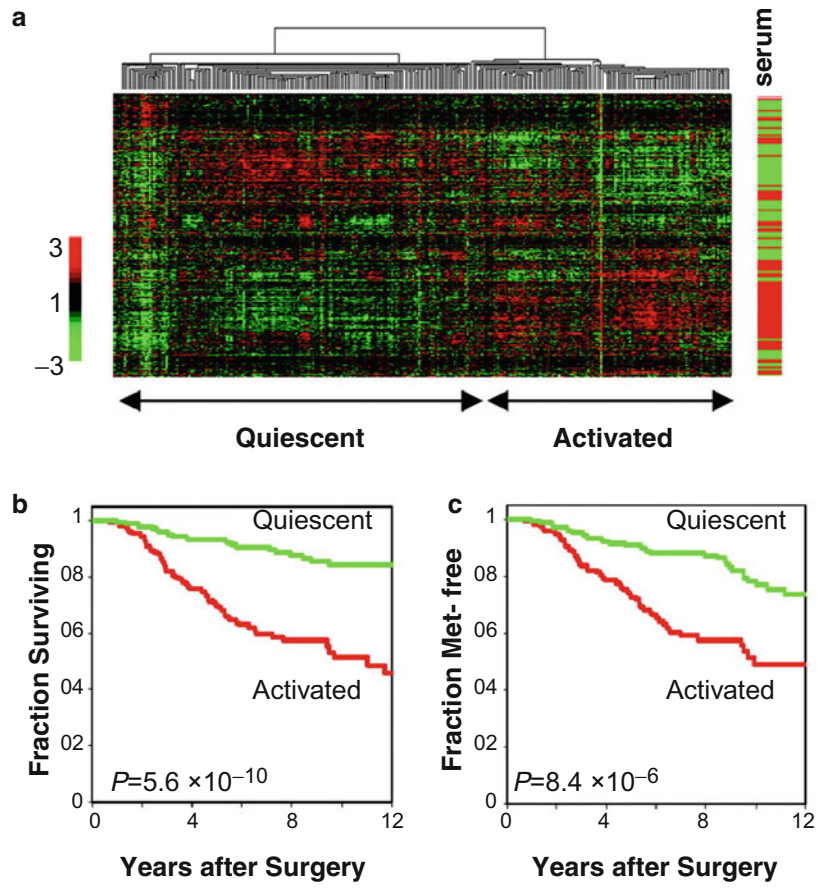


Fig. 10.10 Performance of a "wound-response" gene expression signature in predicting breast cancer progression. (**a**) Unsupervised hierarchical clustering of 295 breast cancer samples. The transcriptional response of each gene in the fibroblast serum response is shown on the right bar (*red* indicates increased expression, and *green* indicates reduced expression in response to serum). The dendrogram at the top indicates the similarities among the samples in their expression of the wound signature genes. Two main groups of tumors were observed: one group with a gene expression pattern similar to that of serum-activated fibroblasts, termed "activated," and a second group with a reciprocal expression pattern of wound signature genes, termed "quiescent." (**b** and **c**) Kaplan–Meier survival curves for the two classes of tumors. Patients with tumor expression of the activated wound-response signature had worse overall survival and distance metastasis-free survival compared with those with a quiescent wound-response signature (Adapted from Chang et al. 2005a)

- Gene expression in fixed, paraffin-embedded tumor tissue was measured as described by Cronin et al. and has resulted in the Oncotype DX assay (Genomic Health [Redwood City, CA])
- The list of 21 genes and the recurrence score algorithm were generated by analyzing the results from three independent preliminary studies involving 447 patients and 250 candidate genes found in earlier studies (including microarray-based studies)
- 16 cancer-related genes were selected primarily based on the correlation with outcome in three trials
- To test the prognostic value of the recurrence score, real-time PCR was successfully tested in 668 paraffin-embedded tumor blocks out of a larger study

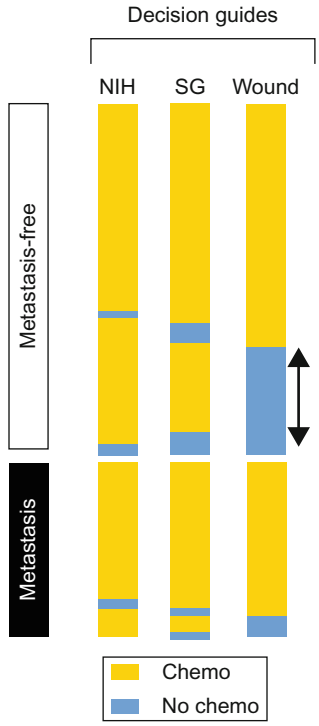


Fig. 10.11 Wound signature as a guide for chemotherapy in breast cancer. Graphical representation of a number of patients advised to undergo adjuvant systemic treatment and their eventual outcomes based on the supervised wound-response signature or the National institute of Health or St. Gallen criteria in the 185 patients in the dataset of van de Vijver 2002 (see Sect. 10.4.2) that did not receive adjuvant chemotherapy. *Yellow* indicates chemotherapy. *Blue* indicates no chemotherapy. The *bar at left* shows which patients have developed distant metastasis as first event: *black* indicates distant metastasis; *white* indicates no metastasis. Thus, *blue* in the *lower bar* indicates the potentially undertreated patients; *yellow* in the *upper bar* shows the potentially overtreated patients (Adapted from Chang et al. 2005a)

population of tamoxifen-treated patients in the B-14 study of the National Surgical Adjuvant Breast and Bowel Project

- Using this recurrence score, 338 patients (51%) had a low-risk, 22% an intermediate-risk, and 27% a high-risk profile for distant metastasis

– 76-gene expression profile

- In 2005, the Erasmus Medical Center (Rotterdam, the Netherlands) in cooperation with the American company Veridex identified a signature of 76 genes, which identifies lymph node-negative breast cancer patients at high risk of distant recurrence and eligible for adjuvant systemic therapy
- An Affymetrix chip U133a containing 22,000 genes was used to measure the level of gene expression
- Frozen samples of 286 untreated node-negative T1–T3/4 breast cancer patients of all ages were included
- This 76-gene prognostic signature was identified using a training series of 171 tumors and consists of two separate profiles, one for estrogen receptor-positive (60 genes) and one for estrogen receptor-negative breast carcinomas (16 genes)
- The gene expression levels were analyzed using log rank analysis and validated on an independent validation set of 115 tumors, without any overlap with the training set
- The distant metastasis-free survival of a "poor" profile present in 65% of the patients was after 60 months 53% and after 80 months 49%. For a "good" profile present in 35% of the patients, the disease-free survival after 60 months was 93% and after 80 months 88%
- The overall survival after 60 months is for a "poor" profile 70% and after 80 months 63% and for a "good" profile 97% and 95%, respectively
- Although the 70-gene profile from the Amsterdam group and the 76-gene profile of the Rotterdam group have only three genes in common, most genes are involved in the same regulatory pathways
- Another reason why the overlap between these two signatures may be small is that different microarray platforms were used, Agilent (dual-label experiment) and Affymetrix (single-label experiment)

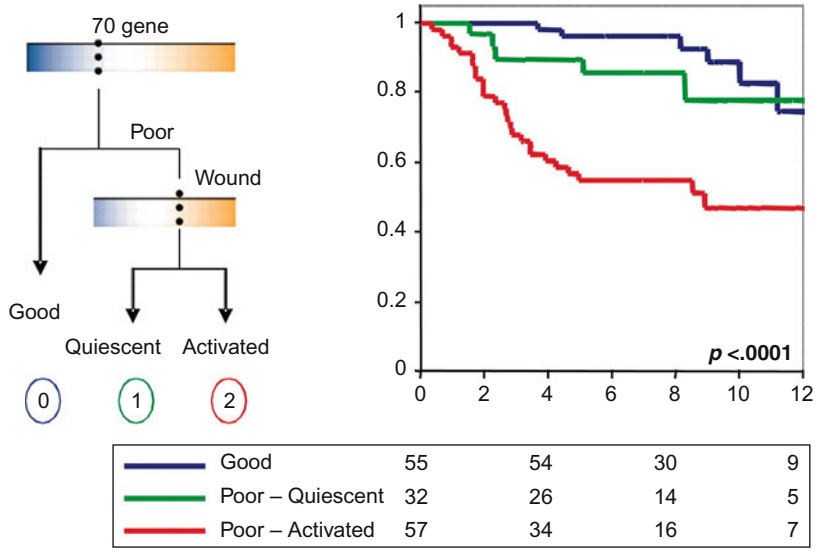


Fig. 10.12 Combining gene expression signatures. Example of decision tree analysis. At each node, the dominant risk factor in multivariate analysis is used to segregate breast cancer patients and the process is repeated in each subgroup until patients or risk factors became exhausted. Chang et al. (2005a) found that the 70-gene signature was able to identify a group of patients with very good prognosis (group 0), and then the wound signature could divide the patients called "poor" by the 70-gene signature into those with moderate and significantly worse outcomes (groups 1 and 2) (Adapted from Chang et al. 2005a)

10.5.1.2 New Additional Prognostic Signatures Are Also Being Developed

- First-generation prognostic signatures in breast cancer have not yet taken into consideration that there are major differences in the gene expression signatures between subgroups of breast cancers, such as ER-positive and ER-negative breast cancers and HER2-negative and HER2-positive breast cancers
- Most of these mRNA profiles are made up with many genes related to the cell cycle and proliferation and only predict the outcome of ER-positive disease
- The number of genes that predict the outcome of ER-negative breast carcinomas is smaller and seems to be strongly dependent on the data set analyzed
- Recent studies have shown that in ER-negative breast carcinomas, the expression of genes related to immune response also provides prognostic information
- The amount of inflammatory infiltrate as defined by immunohistochemical methods and the expression of immune metagenes (high B cell and low Il-8 metagenes) as defined by microarray-based gene expression profiling provide independent prognostic information
- The analysis of genes expressed in the tumor stroma compared with normal stroma from patients with breast carcinomas has also led to the development of stroma-related prognostic gene signatures
- For example, Finak et al. obtained a 26-gene stroma-derived prognostic predictor (SDPP), although further independent validation of their prognostic accuracy is needed
- Lehmann et al. analyzed gene expression profiles of 3,247 primary breast carcinomas from 21 breast cancer datasets, including 587 triple-negative breast carcinomas (TNBC)
- Unsupervised clustering analysis identified six TNBC subtypes exhibiting unique gene expression profiles and ontologies,

including two basal-like, immunomodulatory (IM), a mesenchymal, a mesenchymal stemlike (MSL), and a luminal androgen receptor (LAR) subtype
- BL1 and BL2 subtypes had higher expression of cell cycle and DNA damage response genes and representative cell lines responded to cisplatin
- IM and MSL subtypes were enriched for epithelial–mesenchymal transition, and cell lines exhibiting this subtype responded to NVP-BEZ235 (a PI3K/mTOR inhibitor) and dasatinib (an abl/src inhibitor)
- The LAR cell lines were uniquely sensitive to an AR antagonist (bicalutamide)
- Together these data may be useful in biomarker selection, drug discovery, and clinical trial design that will facilitate suitable targeted therapies to TNBC patients
- Expression levels of large intervening noncoding RNAs (lincRNAs) may be a powerful predictor of eventual metastasis and death in breast cancer patients. Gupta et al. hybridized RNA derived from normal human breast epithelia, primary breast carcinomas, and distant metastasis to ultradense HOX tiling arrays, which offer an interface between DNA and specific chromatin-remodeling activities. They showed that lincRNAs in the HOX loci became systematically deregulated during breast cancer progression. They found a lincRNA, termed HOTAIR, to be increased in expression in primary breast tumors and metastases, suggesting that HOTAIR expression levels in primary tumors may be a powerful predictor of eventual metastasis and death

10.5.2 Clinical Genomics in Hematologic Malignancies

- Acute lymphoblastic leukemia (ALL)
  - Yeoh et al. studied 360 cases of pediatric ALL
  - Using unsupervised hierarchical cluster analysis, they were able to correctly identify the leukemia subtypes of prognostic significance, that is, T lineage leukemia (T cell ALL), B lineage leukemia (E2A–PBX1, BCR–ABL, TEL–AML) and mixed-lineage leukemia (MLL), and hyperdiploid type
  - The same scientists could develop a gene expression profile with an accuracy of 97%
  - This gene expression profile holds great promise in clinical practice, and efforts will be made to refine the signature for general application
- Mixed-lineage leukemia
  - Gene expression profiling is not only aimed at better classification, but can also lead to the identification of novel targets for therapy
  - Mixed-lineage leukemia (MLL) cells distinctly have elevated levels of the receptor tyrosine kinase fms-related tyrosine kinase 3 (FLT3) and thereby represent an opportunity for targeted therapy
  - Armstrong et al. studied the effect of the small molecule inhibitor of FLT3 and showed that the drug stopped tumor progression
- Diffuse large B cell lymphoma
  - Gene expression profiling of diffuse large B cell lymphoma has allowed its categorization into groups based on cellular origin
  - One subgroup has gene expression characteristics of germinal center B cells, that is, "germinal" center B-like diffuse large B cell lymphoma (DLBCL), while the second group consists of genes normally induced during in vitro activation of peripheral blood B cells, that is, "activated" B-like DLBCL
  - Patients with germinal center B-like DLBCL have significantly better overall survival than those with activated B-like DLBCL, and this knowledge can allow stratification of patients for clinical management
  - To validate these results, it has been shown that the BCL6 and HGAL genes that are specifically expressed in the germinal center B cells predict overall survival in unrelated groups of patients, while other genes, such as CD10, are differentially expressed

10.5.3 Clinical Genomics in Prostate Cancer

- Prostate cancer
  - Yu et al. studied 152 prostate tissue specimens, among which were frozen tissue samples from tumor and adjacent nontumor tissue
  - Using an Affymetrix platform, they analyzed 37,777 probe elements, and using a combination of principal component analysis, supervised hierarchical clustering, and tenfold cross validation, they developed a 70-gene expression profile predictor of aggressiveness with an accuracy of 93%
  - The results were validated on a small independent group of 23 patients
  - The group of Lapointe et al. studied 121 frozen tissue samples, comprising 62 cases of prostate cancer, 41 normal tissue specimens, and 9 lymph node metastases
  - They studied 39,711 gene probes and used unsupervised hierarchical clustering to analyze the group
  - Malignant and normal tissues could be distinguished, and further, the malignant tumors could be classified into groups based on risk of recurrence
  - Using the microarray data, immunohistochemical analysis for mucin1 (MUC1) on 225 independent prostrate tumors was developed and elevated levels of MUC1 were correlated with aggressiveness of the prostate tumors in the training set and concurrent high risk of recurrence in the independent validation set ($P = 0.003$)
  - These data show that using gene expression profiling it is possible not only to identify those genes that discriminate indolent from aggressive prostate carcinomas but also highlight the power of microarrays to quickly screen for genes of interest that can be measured using immunohistochemistry or other methods
  - More recent advances in biomarker discovery for prostate cancer by microarray profiling of mRNA and miRNA expression are reviewed by Dalsgaard Sorensen et al.
  - The most promising biomarkers identified from gene expression profiling studies include AMACR, EZH2, TMPRSS2–ERG, miR221, and miR141
  - Although gene expression profiling in prostate cancer is possible, it is not yet ready for clinical use as validation is still lacking

10.5.4 Clinical Genomics in Gastrointestinal Tumors

- Eschrich et al. studied 78 colon cancer specimens and developed a 43-gene prognostic classifier that could predict with 90% accuracy the likelihood of survival at 36 months. There are only very little additional data on gene expression profiling in colorectal cancer
- Lymph node metastasis is a determinant of therapeutic strategy for patients with esophageal carcinoma. However, current methods such as computed tomography scanning and endoscopic ultrasound do not provide accurate assessments of lymph node status. Kan et al. studied 28 primary esophageal squamous cell carcinomas and applied a supervised classification technique called artificial neural networks to develop a gene expression signature that was 86% accurate in predicting lymph node metastasis. Tamoto et al. performed a similar type of analysis using 36 esophageal tumor specimens. They developed a 44-gene signature that was predictive of lymph node metastasis. Some genes present in this signature are known to have a biological role in metastasis
- Like many current microarray experiments, the obvious differences among similar types of studies are due to the small numbers of tumor samples and methodologic differences. For example, in the aforementioned studies on esophageal cancer, one pertinent difference is that Kan et al. used normal esophageal tissue as a reference RNA, while Tamoto et al. used tumor tissue

10.5.5 Clinical Genomics in Carcinoma of Unknown Primary

- Patients with metastatic carcinoma in which no primary site of malignancy can be identified, despite extensive and standardized investigation, constitute approximately 3–5% of all malignancies
- For optimal treatment decisions, it is of great benefit to identify the true nature of the process
- Several studies using gene expression microarrays have demonstrated that the expression levels of thousands of genes can be used as a "molecular phenotype" to classify a multitude of tumor types. Examples are given as follows
  - Examples
    - Using a support vector machine algorithm, Ramaswamy et al. demonstrated 78% classification accuracy in classifying 14 common tumor types
    - Bloom et al. extended the coverage of tumor types to 21 by combining multiple datasets and built a neural network classifier with 85% accuracy
    - While these studies have demonstrated that gene expression microarrays hold great promise as a powerful tool for cancer diagnosis, their survey of the human tumor universe has been rather limited (at most 21 types). They also require the use of frozen tumor biopsies, which are not readily available in the current clinical setting. To overcome these shortcomings, Ma et al. have established a comprehensive microarray database of human tumors (466 fresh frozen, 112 formalin-fixed paraffin-embedded, and 39 tumor types)
    - Ma et al. generated a 92-gene quantitative real-time PCR assay for classifying 32 tumor classes, which can use archival formalin-fixed paraffin-embedded tissues
    - MicroRNA gene expression profiling may be an even better method to improve diagnostic accuracy in carcinoma of unknown primary
    - Lu et al. used a training set of 68 more differentiated tumors, representing 11 tumor types and for which both mRNA and microRNA profiles were available in order to generate a classifier
    - This classifier was then used without modification to classify the 17 poorly differentiated tumor samples, representing carcinoma of unknown primary
    - The microRNA-based classifier established the correct diagnosis of the poorly differentiated samples with far greater accuracy (12/17) than the mRNA-based classifier (1/17)
    - Varadhachary et al. prospectively studied the performance of a microRNA-based assay to identify the tissue of origin in 104 carcinomas of unknown primary patients
    - In 62 out of 74 (84%) samples that were processed successfully, microRNA profiling resulted in an agreement with the clinicopathological presentation

10.6 Interpretation of Gene Lists

- Getting insight in biological processes from microarray data is one of the major goals
- Examples of methods that have been developed to identify biological themes are DAVID, The Database for Annotation, Visualization, and Integrated Discovery; CLENCH, a program for calculating Cluster Enrichment; eGON, a program to explore Gene Ontology (GO) terms; GOstat, a program to compute GO statistics of a list of genes selected from a microarray experiment; OM, Onto-Miner, a program to return all known information about a given list of genes; and IPA, Ingenuity Pathways Analysis
  - Examples
    - Gene set enrichment Analysis developed by Subramanian
    - "Molecular modules" underlying human malignancies

10.6.1 Combining Clinicopathological and Genomic Risk Predictors

- The next step in optimization of the prognostic and predictive risk prediction is combining well-known clinicopathological with genomic risk predictors
- For example, Tang et al. combined the 21-gene breast cancer assay recurrence score (RS) used for assessing recurrence risk and predicting chemotherapy benefit in patients with estrogen receptor (ER)-positive breast cancer with clinicopathological risk predictors (histological grade, size of the tumor, and lymph node status and age at diagnosis)
- They showed that inclusion of tumor size and nodal status in combination with the genomic test can improve the prognostic accuracy of the model but could at the same time reduce its predictive value for the chemotherapy benefit

10.7 Points of Attention in the Design of Microarray Experiments

10.7.1 The Sample

- Frozen tumor tissue is indispensable as the RNA quality remains optimal due to the rapid fixation method, but is not widely available as in most hospitals the tumor samples are directly fixed in formalin and embedded in paraffin blocks
- Problems related to making use of mRNA
  - mRNA is a very fragile molecule, which can degrade within minutes of surgical manipulation dramatically affecting the final result of microarray data
  - Likewise, subtle variations in tissue handling and method of RNA extraction from samples can result in different levels of gene expression

10.7.2 The Microarray

- Differences in platform design (single- vs. dual-label experiments), microarray design, probe annotation, methods of RNA labeling, the process of hybridization, data acquisition, and normalization make direct comparison of cross platform comparisons of gene expression studies difficult

10.7.3 The Statistics

- "Over fitting" is one of the limitations of clustering methods. This certainly explains often that the proposed association of a "genetic signature" with disease outcome is significantly stronger in preliminary studies than in subsequent research. It indicates that the number of parameters of a model is too largely relative to the cases or specimens studied
- Studies with a small sample size (<50) tend to have large variance and confidence intervals
- It is essential to validate a predictive gene expression pattern in a sufficiently large independent series of tumors/patients
- A limitation of unsupervised cluster analyses is that it provides qualitative and not quantitative information on differences between genes or classes
- Gene lists derived from a single dataset have been found to be highly dependent on the composition of the dataset

10.8 High-Throughput Genomic Sequencing

- DNA sequencing includes a variety of technologies to detect the order of the nucleotide bases in a DNA molecule – adenine, guanine, cytosine, and thymine
- "First-generation" DNA sequencing technology was first described by Frederick Sanger, Allan Maxam, and Walter Gilbert and was used to determine the first complete human genome sequence
- Next-generation sequencing (NGS) results in thousands to millions of sequence reads within the same run by either using clonally amplified templates generated from single DNA molecules, or single DNA molecule templates.

The advantage of NGS is the ability to generate an enormous amount of data cheaply, setting the goal of the "thousand-dollar genome"

- The terms second- and third-generation sequencing have been used to denote our gradual advances in sequencing technologies. Second-generation sequencing used PCR clonally amplified templates and resulted in throughput and cost per base improvement. Third-generation sequencing uses single-molecule sequencing, avoiding PCR-based amplification biases, and results in higher throughput
- There are many different applications for NGS, such as genome sequencing (DNA), transcriptome sequencing (RNA seq), and chromatin immunoprecipitation sequencing of associated DNA (ChIP seq), each one requiring a specific protocol for construction of a DNA (or cDNA) library suitable for loading onto the sequencing platform
- Successful translation of NGS high-throughput sequencing into diagnostic use will rely on standardized guidelines for sample preparation, sequencing technology used, and data processing and analysis. Each sequencing technology is characterized by its own biases and error rate, as well as biases based on the source and method of isolation of nucleic acid material used. Construction of the DNA library is the single most important determining factor for quality of the results

10.8.1 DNA Sequencing Technologies

10.8.1.1 General Principles of Next-Generation Sequencing

- Template preparation
  - Template is a segment of recombinant DNA comprised of a known region, usually a vector or an adaptor sequence to which a universal primer can bind, and the target unknown DNA sequence to be sequenced
  - Genomic DNA is randomly broken into smaller sizes from which either fragment templates or mate templates are generated
  - A fragment library is prepared by randomly fragmenting DNA using a variety of methods (see below) and requires a small amount of DNA
  - A mate-pair library is generated by circularizing the fragmented DNA, after it is size-selected for a particular length. The circularized DNA is then cut into linear DNA fragments, yielding mate-pair templates which bring together DNA ends that were not in close proximity in the genome
  - The template is attached or immobilized to a solid surface or support, with each template being spatially separated, allowing millions of sequencing reactions to take place simultaneously
  - Templates are either clonally amplified molecules or single molecules. Second-generation sequencing usually uses clonal amplification, while third-generation sequencing uses single-molecule methods
  - Clonally amplified templates accommodate technology that cannot detect single-fluorescent events (fluorescent readout being one of the main detection methods in sequencing technology). Amplification usually involves emulsion PCR, solid-phase amplification, or rolling circle replication. Biases for clonally amplified template methods include the following: (a) PCR steps create mutations that can be mistaken for sequence variants and (b) AT-rich and GC-rich target sequences can show amplification bias
  - Emulsion PCR prepares sequencing templates in a cell-free system by separating and capturing DNA into single strands under conditions that favor one DNA molecule per bead. Adaptors with universal priming sites are ligated to the target DNA ends; the DNA is separated into single strands and amplified on beads. Beads are then immobilized in a polyacrylamide gel on a standard microscope slide, chemically cross-linked to an amino-coated glass surface or deposited into individual picotiter plate wells where the NGS reaction can take place

- Solid-phase amplification produces randomly distributed, clonally amplified clusters on a glass slide (flow cell). Primers are covalently attached to the slide producing clusters with free ends to which a universal sequencing primer can be hybridized to initiate the NGS reaction
- DNA nanoball sequencing uses rolling circle replication to amplify small fragments of genomic DNA into DNA nanoballs. Each DNA fragment is reproduced in a way that connects all of the copies together in a head to tail configuration through adaptors, forming a long single molecule of connected nucleotides. Each long single molecule is then consolidated (balled up) into a small particle of DNA, called DNA nanoball (~200 nm in diameter). Each DNA nanoball contains hundreds of copies of the 70 bases of DNA to be read in each fragment
- Single-molecule templates require less starting material, involve more straightforward template preparation protocols, and lack amplification biases. Templates are usually spatially distributed and immobilized on solid supports
  - Through individual primer molecules that are then subsequently hybridized to the single-molecule templates (which are processed to include common adaptors at their ends) (one-pass sequencing)
  - By directly immobilizing the single-molecule templates (through priming and extending single-stranded, single-molecule templates from immobilized primers) and then hybridizing a common primer to the template (two-pass sequencing)
  - Through single polymerase molecules to which a primed template molecule is bound; this last approach can generate longer read lengths
- Single-molecule methods use diverse sequencing and detection technologies (see below) and include semiconductor, exonuclease DNA sequencing, quantum drop (Qdot) nanocrystals, and fluorescence resonance energy transfer (FRET)
- The term "paired-end" sequencing refers to the sequence analysis of two ends of the same DNA molecule, resulting in higher confidence when mapping the sequences to the genome. A simple modification to the standard single-read DNA library preparation facilitates reading both the forward and reverse template strands of each cluster during one "paired-end read." One can sequence one end and then turn it around and sequence the other end: the two sequences obtained are "paired-end reads." In addition to sequence information, both reads contain long-range positional information, allowing for highly precise alignment of reads
- Table 10.2 comprises a broad overview of sequencing technologies, separating them into ones that use clonally amplified and single-molecule templates

10.8.1.2 Sequencing and Imaging

- General principles
  - Various sequencing (i.e., cyclic reversible termination) and detection (i.e., fluorescence, bioluminescence) methods are being used
  - There are basic differences in sequencing technologies used for clonally amplified or single-molecule templates, leading to a different set of biases
  - Dephasing, for example, is a problem with clonally amplified templates, which results in fluorescence noise, causing base-calling errors and shorter reads. Since clonal amplification results in many identical templates, each one undergoing the sequencing reaction, the signal is the consensus of the nucleotide or probe addition for a given cycle. Incomplete extension or extra extension of the template populations results in lagging-strand or leading-strand dephasing, respectively
  - Single-molecule templates are susceptible to multiple-nucleotide or probe additions in any given cycle. For example, no signal will be detected because of the

Table 10.2 Clonally amplified versus single-molecule DNA sequencing technologies

	Type	Examples	Comments
Clonally amplified templates	Emulsion PCR	Roche/454 ABI 5500	One DNA molecule per microreactor (bead); clonal amplification occurs in beads; beads chemically cross-linked to a glass slide
	Solid-phase amplification/bridge PCR	Illumina HiSeq 2000	One DNA molecule per cluster on glass slide; clusters randomly distributed; bridge amplification and cluster growth
	PCR on DNA nanoballs	Complete Genomics	Rolling circle replication
Single-molecule templates	Primer immobilized	Helicos Biosciences	One-pass sequencing; billions of single-molecule templates
	Template immobilized	Helicos Biosciences	Two-pass sequencing; billions of single-molecule templates
	Polymerase immobilized	Pacific Biosciences	Thousands of single-molecule templates

incorporation of nucleotide probes that do not contain a fluorescent label (see below)

- Sequencing technologies
  - Cyclic reversible termination: Uses reversible terminators in a cyclic method that comprises nucleotide incorporation, fluorescence imaging, and cleavage. Each nucleotide is "marked" by pairing to a different fluorescently labeled probe, a modified nucleotide associated with a particular fluorescent dye
    - First, a DNA polymerase bound to the primed template adds or incorporates a single fluorescently modified nucleotide that is complementary to the template
    - Following incorporation, the remaining unincorporated nucleotides are washed away
    - Imaging is performed to determine the identity of the incorporated nucleotide
    - The terminating group and the fluorescent dye are then cleaved and washed away to allow the next incorporation step to take place. Four-color or one-color cyclic reversible termination can be used. There are two types of reversible terminators: 3′ blocked (requiring modified DNA polymerases) and 3′ unblocked (that can be incorporated by wild-type polymerases)
  - Sequencing by ligation: Uses DNA ligase and either one base-encoded probes (a probe where one fluorescent dye is associated with a specific base) or two base-encoded probes (a probe where the fluorescent dye is associated with a combination of two bases) in a cyclic fashion with a fluorescent readout
    - First, a fluorescently labeled probe hybridizes to its complementary sequence adjacent to the primed template
    - DNA ligase then joins the dye-labeled probe to the primer
    - Nonligated probes are washed away
    - Imaging is then performed to determine the identity of the ligated probe. The use of two base-encoded probes can improve accuracy in base calling
  - Single-nucleotide addition/pyrosequencing: Uses a nonelectrophoretic, bioluminescence method that measures the release of inorganic pyrophosphate by proportionally converting it into visible light through a series of enzymatic reactions
    - DNA polymerase adds a single deoxyribonucleotide triphosphate (dNTP) in limiting amounts, instead of using modified nucleotides to terminate DNA synthesis
    - Once the complementary dNTP is incorporated, the polymerase extends the primer and pauses
    - The addition of the next complementary dNTP reinitiates DNA synthesis

- The order and intensity of the light peaks are recorded as flowgrams, which reveal the underlying DNA sequence

– Real-time sequencing: Uses imaging of the continuous incorporation of labeled nucleotides during DNA synthesis and may involve fluorescent or other detection technology. Nucleotides do not stop the process of DNA synthesis, in contrast to reversible terminators
  - Examples that use the traditional or other fluorescent detection schemes
    – An engineered DNA polymerase with an attached fluorescent dye produces enhanced signal by FRET upon nucleotide incorporation
    – Dye-quencher nucleotides, which in their native state produce low signals but upon release and diffusion of the dye-labeled pyrophosphate analogue away from the immobilized DNA polymerase produce a fluorescent signal
  - An example using an alternative detection scheme: Single DNA polymerase molecules are attached to the bottom surface of individual zero-mode wavelength detectors (ZMW) that can obtain sequence information while phospholinked nucleotides are being incorporated into the growing primer strand

10.8.1.3 Data Processing and Analysis of High-Throughput Sequencing

- Data generated from NGS platforms can be up to 15 terabytes and likely greater by the time of this publication
- Example of data generated from fluorescent detection methods: image data comprising color-space base transitions, which are then converted into strings of bases representing sequence reads (base calling)
- Sequence reads are filtered based on a quality control score
- Sequence reads are then aligned to a reference genome to generate the genomic coordinates in the reference genome
- The sequence reads are also usually visualized as to how they map to the genome in a browser viewable format to ensure data quality and assess proper coverage across example locations (this step will vary depending on the application, i.e., small RNA sequencing can yield precise mapping of reads to the entire mature small RNA or small RNA genomic location (pre-miRNA) with each read, whereas for transcriptome sequencing, reads are usually assembled, scattered across the length of the entire transcript)
- The sequence reads are usually reduced to nonredundant (unique) reads and counted. Once reads are mapped to the genome, a file is produced with the number of sequence reads corresponding to each gene, isoform, etc. Further downstream applications include quantification (i.e., expression levels), mutation/SNP identification, copy number analysis, and analysis of structural variations. For example, for mRNA RNA sequencing expression analysis can be conducted similarly to data derived from microarray expression analysis data (i.e., evaluation of expression differences or gene ontology analysis)
- Integration and visualization of data can be done using genome browsers, such as the University of California Santa Cruz (UCSC) browser, NCBI map viewer, Ensemble, or platforms such as Galaxy
- Galaxy, from Penn State University, is an open web-based platform for interactive large-scale genome analysis. Whether on the free public server or a local computer, one can perform, reproduce, and share complete analyses of NGS data

10.8.1.4 DNA Sequencing

- Genome sequencing
  – *De novo sequencing*: The initial generation of the primary genetic sequence of an organism (reference genome). A detailed genetic analysis of any organism is possible only after de novo sequencing has been performed
  – *Targeted resequencing*: Resequencing experiments allow identification of causative mutations within populations to

advance both basic and clinical researches, once the reference sequence for an organism is determined. Targeted resequencing involves the comparative analysis of candidate genes or regions and requires a high level of accuracy to identify low frequency SNPs and structural variants. Traditional capillary electrophoresis methods provide the highest accuracy and are best suited for analyzing a limited set of genes in a large number of patient samples. However, when analyzing a large number of genes, exome capture methods coupled with NGS are required such as in-solution or solid-phase hybridization
  - *Whole-genome resequencing*: Once the reference sequence for an organism is determined, comparative sequencing can characterize the genetic diversity within an organism or between closely related species
  - Whole-genome or targeted resequencing has been used to recognize SNPs, mutations (either somatic or germline), insertions, deletions, inversions, copy number, and structural variations (i.e., translocations) in the human genome (see following section for examples). Mate-pair libraries of varying insert sizes (0.6–6 kb) allow the detection of a variety of structural changes
  - Targeted "*exomic resequencing*": Focuses only on the protein-coding genes, given that they bear the majority of known disease-related mutations that have traditionally been targeted therapeutically
  - *Metagenomics*: Refers to the genomic analysis of microorganisms by extraction of DNA from uncultured accumulation of microbes in environmental and clinical samples
- Modified DNA NGS approaches
  - Modified NGS protocols allow studying epigenetic mechanisms of gene regulation by assessment of histone modifications, DNA methylation, and related chromatin structures (chromatin immunoprecipitation (ChIP seq), methylation (Methyl seq), DNase seq)
  - To fully understand the regulation of the human genome defined by the human genome project, we need to define the human epigenome. The epigenome is more variable than the genome itself, so it could modify more variable and transient type interactions, such as environmental influences. Epigenetic mechanisms involve DNA methylation patterns, posttranslational modification of histones, the interaction between transcription factors and their targets, and nucleosome positioning
  - Epigenomic mapping is highly complex, given that the epigenome is much more variable than the genome. The epigenome varies with age, developmental stage, tissue type, environment, and between health and disease. Thus, understanding the epigenome as a function of all these factors can help us better understand the mechanisms of disease and the process of aging
  - *ChIP seq*: Refers to genome-wide sequencing profiles of immunoprecipitated DNA–chromatin complexes. It determines the genomic localization (DNA binding sites) for a particular protein of interest. This technique gives a picture of the protein–DNA interactions that occur inside the nucleus of living cells or tissues. Determining how proteins interact with DNA to regulate gene expression is essential for fully understanding many biological processes and disease states
  - *Methyl seq*: Refers to the characterization of DNA methylation patterns. To do this for specific regions, bisulfite sequencing on automated capillary electrophoresis instruments has been utilized in a lower-throughput fashion. For genome-wide DNA methylation analysis, other nonsequencing techniques were initially employed, such as restriction landmark genomic scanning, which involves labeling DNA with radioactive phosphorus 32. Recently, NGS approaches have been developed which involve enrichment-based methods that use antibodies or

proteins to pull down methylated cytosines and sequence-associated DNA
  - *Dnase seq*: Refers to DNase I hypersensitivity to identify the location of regulatory regions, based on the genome-wide sequencing of regions characterized by accessibility to cleavage by DNase I. FAIRE Seq (formaldehyde-assisted isolation of regulatory elements) is a variation of this method that can be more easily applied to a variety of cell types

10.8.1.5 RNA Sequencing

- Whole-transcriptome sequencing
  - RNA sequencing (RNA seq) is a method to determine the sequence content and abundance of coding and noncoding RNAs
  - RNA seq yields assessment of gene expression in a quantitative fashion, allows for determination of nucleotide variations (i.e., candidate mutations and recurrent fusion genes), absolute quantitation of input RNA by addition of calibrator oligonucleotides, and offers the opportunity of reannotation in case the reference gene assembly is improved
  - RNA seq can be used to detect novel transcripts and splice variants
  - Current methods for transcriptome sequencing include (1) whole-transcriptome analysis, (2) gene expression profiling, and (3) small RNA analysis
  - *Whole-transcriptome analysis* could potentially allow identification of expression of all coding and noncoding RNAs, alternative splicing, SNPs/mutations, translocations/fusions, and allele-specific expression patterns. Some methods conserve the strand information while others do not. Other variations entail further purification of specific RNA fractions (i.e., only select poly(A) mRNA or entire RNA population after depletion for specific types, such as rRNA)
  - *Gene expression profiling* or otherwise referred to as SAGE (serial analysis of gene expression) provides a highly sensitive method for quantifying gene expression levels on a genome-wide scale. It defines mRNA expression levels by sequencing unique sequence tags isolated from the 3′ ends of mRNAs. It is highly reproducible and has a great dynamic range, much more sensitive than microarrays (from less than a copy per cell to over 100,000 copies per cell). Moreover, it can detect both known and novel mRNAs
  - *Small RNA analysis* focuses on the expression of specific RNA subsets, such as microRNA (miRNA), short interfering RNA (siRNA), piwi-interacting RNA (piRNA), repeat-associated siRNA (rasiRNA), and other small RNA species. Usually the small RNA is purified or enriched from total isolated RNA
  - Farazi et al. used in-house-developed barcoded Illumina sequencing to profile miRNAs in 11 normal breast tissues and 17 noninvasive and 151 invasive breast carcinomas. Clustering and comparative analysis of miRNA read frequencies showed that normal breast samples were separated from most noninvasive ductal carcinomas in situ and invasive carcinomas by increased miR21 and multiple decreased miRNA families (including miR98/let-7), with most miRNA changes already in the noninvasive carcinomas. Given the increasing sequencing depth and small number of miRNAs, this barcoded technology can now be combined with multiplexing to allow cost-effective identification of multiple samples in parallel
- Modified RNA NGS approaches
  - Modified NGS protocols can also allow study of posttranscriptional gene regulation by assessment of RNA bound to RNA-binding proteins (RBPs) or specifically to AGO2, the main component of the miRNA effector complex to identify miRNA targets. Similarly to the epigenome, posttranscriptional gene regulation also varies with age, developmental stage, tissue type, etc., and can further our understanding of health and disease

10.8.2 Sample Preparation

10.8.2.1 DNA Sequencing

- The starting material for NGS is double-stranded DNA, for example, isolated genomic DNA, or DNA immunoprecipitated with chromatin
- Quality control and quantification of the template DNA are of the utmost importance since sequencing instruments are highly affected by the quality and molar concentration of the sequencing library. An accurate approximation of the average fragment length and quantity of double-stranded DNA is essential to convert mass concentration into molar concentration measurements. Working with low amounts of nucleic acids, it is essential to work with low-adsorbing plasticware to diminish adsorption to tube walls
- In order to convert the genomic DNA into a suitable DNA library for sequencing, it is randomly fragmented into either fragment or mate-pair templates, as described earlier
- Methods for DNA fragmentation include (1) nebulization using a throwaway nebulizer driven by pressurized air, (2) ultrasound device, (3) enzymatic digestion, or (4) transposon insertion
- The DNA can be size-selected, ligated to an adaptor, purified, and quantified resulting in a DNA library of short fragments flanked by adaptors of known sequence (these steps and their order differ based on the sequencing technology used)
- As described earlier, amplification can lead to biases leading to loss of specific regions of the template DNA as other regions are more efficiently amplified, resulting in a skew of precise quantification, or can lead to misleading calling of mutations and SNPs

10.8.2.2 RNA (cDNA) Sequencing

- In order to sequence RNA, it first needs to be converted into a cDNA library
- The majority of the total cellular RNA represents ribosomal RNA. Sequencing mRNA, miRNA, tRNA, and other long noncoding RNAs requires preparation methods that enrich for these RNA populations (i.e., size selection for 19–24 nucleotide fragments for miRNAs, methods to isolate poly(A) RNA to select for mRNAs)
- For longer RNAs (i.e., mRNAs), the RNA, similarly to the DNA, needs to be fragmented. This is usually accomplished using high temperature, ions, or enzyme digestion
- The RNA fraction of interest is converted into a double-stranded cDNA library
- cDNA library construction can involve use of oligo-d(T) and thus selection of RNAs with poly(A) tails, random hexamers, or sequence-specific amplification of adaptors that are initially ligated to the RNA of interest in a strand-directed fashion (directional sequencing – see below)
- Strand information is not always maintained, depending on the method used; approximately 3,000 genes that overlap on opposite strands are measured incorrectly
- Directionality can be introduced by using directional adaptor ligation or by labeling one cDNA strand for destruction

10.8.3 Second-Generation Sequencing

- Second-generation sequencing is based on clonal amplification methods to generate clusters of an adaptor-modified random collection of DNA fragments at high densities
- These fragments are subsequently sequenced by recurrent cycles of polymerase-based nucleotide extension or by cycles of oligonucleotide ligation
- Second-generation sequencers are available from Roche (454 Ti Roche™), Illumina (Hiseq™), Applied Biosystems (SOLID, ABI 5500), and Complete Genomics

10.8.3.1 Examples of Second-Generation Sequencers

- Illumina approach (Hiseq™)
  - This platform carries out DNA amplification by attaching single-stranded DNA fragments to a solid surface known as a single-molecule array and conducting

solid-phase bridge amplification of single-molecule DNA templates
  - One end of a single DNA molecule is attached to a solid surface using an adaptor. The molecules subsequently bend over and hybridize to complementary adaptors. This results in clonal clusters, which are denatured and annealed with a sequencing primer and exposed to sequencing by synthesis using three chemically inactive labeled nucleotides, which incorporates a single nucleotide at each step
  - It employs four-color cyclic reversible termination. As described earlier, briefly, the labels are fluorophores that differentiate among the four different bases at a given sequence position, which are imaged after each base incorporation cycle
- The Roche 454 Genome Sequencer FLX system (454 Ti Roche^{TM})
  - One DNA fragment ligated to a short adaptor is attached to a single amplification bead, and emulsion PCR proceeds a highly efficient in vitro DNA amplification method. This results in each bead generating clonal copies of the given DNA molecule. The beads are loaded into picotiter plates, and the clonally related templates are analyzed using a pyrosequencing reaction. The picotiter plate allows hundreds of thousands of pyrosequencing reactions to be carried out in parallel
  - The GS FLX system is based on sequencing by synthesis with pyrophosphate chemistry (pyrosequencing). As described earlier, pyrosequencing measures the release of inorganic pyrophosphate (PP_i) by chemiluminescence. In summary, the template DNA is immobilized, and solutions of dNTPs are added one at a time. The release of PP_i whenever the complementary nucleotide is incorporated is detectable by converting it into a chemiluminescent signal
- Applied Biosystems (SOLID)
  - This sequencing by ligation-based procedure begins with an emulsion PCR step similar to that used in the 454 technique. The PCR products are attached to a glass slide where sequencing occurs by sequential rounds of hybridization and ligation with 16 dinucleotide combinations labeled by four different fluorescent dyes (two base-encoded probes). Using the four dye-coding schemes, each position is effectively probed twice, and the identity of the nucleotide is determined by analyzing the color that results from two successive ligation reactions

10.8.4 Third-Generation Sequencing

- Third-generation sequencing platforms sequence DNA at the level of a single molecule
- This circumvents biases from PCR amplification prior to DNA sequencing
- Examples of third-generation sequencers include Helicos (HeliScope), Pacific Biosciences, Ion Torrent PGM, Life Technologies FRET, and Oxford Nanopore Technology Ion Torrent
- They can be categorized into different classes
  - Sequencing technologies where single molecules of DNA polymerase are observed as they synthesize a single molecule of DNA
  - Nanopore sequencing technologies where individual bases are identified as they pass through a nanopore coupled to an exonuclease
  - Direct imaging of individual DNA molecules using cutting-edge microscopy technologies

10.8.4.1 Examples of Third-Generation Sequencers

- The Helicos (HeliScope) uses dye-labeled nucleotides admixed with nucleic acids. The dye is subsequently attached to the nucleotide by a chemical cleavage group that enables stepwise sequencing
- Pacific Biosciences developed single-molecule real-time sequencing. It detects a single molecule of DNA polymerase as it

synthesizes a strand of DNA, directly influencing the speed and process of this enzyme
- The Ion Torrent PGM sequencer uses semiconductor sensors on a high-density array of microwells that measure the release of hydrogen ions that are generated by the process of template-directed DNA polymerase synthesis
- Life Technologies developed a real-time sequencer using FRET. It uses a Qdot nanocrystal as the FRET donor coupled to a DNA polymerase enzyme. Adding a nucleotide leads to energy transferred from the Qdot to an acceptor fluorescence moiety on each labeled base. Light emission occurs only from labeled nucleotides as they are incorporated
- Oxford Nanopore technologies developed electronic-based single-molecule sequencing. It measures the current change as sequentially cleaved bases pass through a nanopore coupled to an exonuclease, identifying the DNA base on the molecule in sequence

10.9 Systematic Search and Catalogues for Genetic Alterations in Cancer

- Multiple coordinate endeavors are taking place for a comprehensive and systematic search for somatic alterations that underlie human cancer. Some examples are given below
- In 2004 the Wellcome Trust Sanger Institute started a catalogue of genes mutated in human cancer at a higher frequency than expected by chance. In 2009 the number of candidate cancer genes increased by 40% from 291 to 410
- The Cancer Genome Project of the Wellcome Trust Sanger Institute surveys somatic alterations in human tumors and tumor-derived cell lines. This is a continuing project of resequencing 4,000 candidate cancer genes derived from a variety of solid tumors. In addition, 800 tumor-derived cell lines are being analyzed for copy number alterations
- Another initiative of the Wellcome Trust Sanger Institute is the Catalogue of Somatic Mutations in Cancer (COSMIC). This is a catalogue of somatic mutations in benign and malignant tumors and tumor-derived cell lines
- The Cancer Genome Atlas (TCGA) catalogues systematically the genomic alterations within the major types of human cancer. Their goal is to identify somatic alterations in 20–25 major tumor types by 2014
- The Therapeutically Applicable Research to Generate Effective Treatments (TARGET) initiative aims to identify therapeutic targets for childhood cancer by analyzing cancer genomes and transcriptomes
- The International Cancer Genome Consortium (ICGC) was formed to standardize the approaches by which genomic alterations are identified in human cancers. It aims to produce comprehensive catalogues of the genomic alterations of up to 50 clinically significant cancer types and subtypes over the next 10 years
- The Mitelman Database of Chromosome Aberrations in Cancer systematically catalogues chromosomal alterations and clinical associations, as reported in the literature
- The Genetic Alterations in Cancer database curates mutations associated with exposure of chemical, physical, and biological agents implicated in tumorigenesis

10.10 Next-Generation Sequencing Applications

- NGS gives the researcher the ability to rapidly (1) identify mutations or SNPs in key disease genes, (2) identify structural rearrangements (i.e., translocations) and genomic copy number alterations, (3) evaluate gene expression, and (4) identify novel transcripts or transcript variants in both coding and noncoding genes
- The DNA of the first human whole genome from James D. Watson was sequenced using Roche 454 technology
- The first cancer genome sequenced using NGS was from a patient with acute myeloid leukemia and was compared with the germline DNA from the same patient

- NGS applications can involve sequencing either at the DNA or the RNA level

10.10.1 Application of High-Throughput Sequencing in Diagnostics of Clinical Oncology

10.10.1.1 General

- It will only be a matter of time before NGS will become part of routine clinical genetics
- To understand the current status and relevance of NGS in clinical oncology diagnostics that have been achieved, a few important examples are highlighted below
- NGS has already begun to be used to personalize cancer treatment. For example, a patient with advanced gemcitabine-resistant pancreatic cancer had a remarkable clinical outcome after being treated with DNA-damaging agents. This patient's carcinoma was sequenced as part of an effort to sequence the pancreatic cancer genome by Jones et al. and revealed biallelic inactivation of the PALB2 gene, a DNA repair gene, loss of which mechanistically suggests sensitivity of the patient's cancer to DNA-damaging agents. Based on these sequencing results, experiments using a personalized xenograft generated from the patient's surgically resected tumor confirmed significant activity of this class of drugs

10.10.1.2 Breast Cancer

- Before NGS, Wood et al. studied the genomes of 11 breast and 11 colorectal cancers using the conventional Sanger sequencing approach. In these 22 samples, they determined the sequences of exons representing 20,857 transcripts from 18,191 distinct genes. They found that the genomic landscapes of breast and colorectal cancers are composed of a handful of commonly mutated gene "mountains" and a much larger number of gene "hills" that are mutated at low frequency
- Shah et al. used whole-genome DNA NGS to catalogue the genetic events that took place during the development and progression of a lobular breast carcinoma by comparison to the germline. They found that only 5 of the 32 nonsynonymous coding mutations present in the metastasis were also present in the primary tumor
- Ding et al. studied the genome of a triple-negative (ER-negative, PR-negative, HER2-negative) breast carcinoma using whole-genome end-paired DNA NGS. They sequenced the primary tumor, a brain metastasis, a xenograft derived from the primary tumor, and a germline tissue from the same patient. The metastasis contained two de novo mutations and a large deletion not present in the primary tumor and was enriched for 20 mutations. The xenograft retained all primary tumor mutations, but its overall genetic landscape better resembled the metastasis rather than the primary tumor. These results suggest that a metastasis may arise from a minority of cells within the primary tumor
- Stephens et al. used end-paired DNA NGS to identify somatic rearrangements in 24 breast carcinoma genomes. These samples were investigated by sequencing both ends of 65,000,000 randomly generated ~500 bp DNA fragments. Firstly, they found more rearrangements than previously appreciated. Secondly, breast cancer subclasses defined on the basis of expression of estrogen receptor, progesterone, and HER2 showed correlation with patterns of genomic alterations. Breast carcinomas with many tandem duplications were usually ER and PR negative, while carcinomas with few rearrangements or with rearrangements within amplicons were ER and PR positive

10.10.1.3 Acute Myeloid Leukemia

- Ley et al. used whole-genome sequencing (DNA NGS) to study the acute myeloid leukemia genome and its matched normal counterpart obtained from the patient's skin. Ten genes were discovered with acquired mutations; two were previously described, while eight new mutations were identified.

In a followup study, they sequenced the exons of one of the found acquired mutations (DNMT3a), in 280 additional patients. In a total of 62 of 281 patients (22.1%), a mutation in DNMT3a was found
- Mardis et al. used DNA NGS to obtain a high level of coverage in a cytogenetically normal, de novo genome for AML. They identified recurrent mutations in IDH1
- Ramsing et al. reported the complete characterization of the microRNAome in a patient with acute myeloid leukemia. They combined miRNA and mRNA NGS and found a somatic mutation in the 3′ UTR of TNFAIP2, a known target of the PML–RARA oncogene

10.10.1.4 Diffuse Large B Cell Lymphoma

- Pasqualucci et al. combined DNA NGS with DNA copy number analysis and showed that the coding genome of six untreated individuals with de novo DLBCL contains, on average, more than 30 clonally represented gene alterations. They also identified a new mutation in MLL2, which was not previously implicated in DLBCL pathogenesis

10.10.1.5 Melanoma

- Harbour et al. used exomic NGS to define metastasis-related mutations in highly metastatic uveal melanomas. They found a recurrent inactivating somatic mutation of BRCA1-associated protein 1 (BAP1) on chromosome 3p21.1 in 26 of 31 (84%) metastasizing tumors

10.10.1.6 Lung Cancer

- Lee et al. used DNA NGS to identify genome-wide mutations in a primary lung tumor and adjacent normal tissue with a 60-fold coverage. Comparing these two genomes, they identified more than 50,000 high-confidence single-nucleotide variants. They validated 530 somatic single-nucleotide variants (including KRAS) and identified 391 variations in other coding regions, as well as 43 large-scale structural variants
- Kan et al. characterized somatic mutations in 441 tumors including lung, breast, ovarian, and prostate cancers. They used mismatch repair detection technology to identify somatic mutations across the coding exons and flanking splice sites of 1,507 genes comprising known cancer and druggable genes. They found 77 mutated genes including protein kinases and G-protein-coupled receptors
- Pleasance et al. used NGS to catalogue the somatic mutations associated with tobacco smoking, by sequencing a small-cell lung cancer cell line (NCI–H209). They identified a tandem duplication that duplicates exons 3–8 of CHD7
- Campbell et al. used end-paired DNA NGS to identify somatically acquired rearrangements in lung cancer by comparing the tumor cell line genomes from two individuals with lung cancer to the germline reference human genome. They characterized more than 300 germline structural variants and 103 somatic rearrangements in the tumor cell lines

10.10.1.7 Pancreatic Cancer

- A comparable study of primary pancreatic tumors and their metastatic counterparts was also performed by Campbell et al. They used end-paired DNA NGS to identify somatically acquired genomic rearrangements in 13 patients with pancreatic adenocarcinoma and explored clonal relationships with their metastatic counterparts. They identified 381 somatic rearrangements, of which only a proportion were also found in metastatic tissue, suggesting genetic heterogeneity among metastasis initiating cells and driver mutations beyond those required for primary tumors that may lead to tumor cell metastatic seeding

10.11 Challenges of Using High-Throughput Sequencing Technologies

10.11.1 Bioinformatic Obstacles

- Currently, the amount of sequencing data produced is more than can easily be stored

- NGS has resulted in the growth of data acquisition from megabytes per day to terabytes per day on a single sequencer
- Standard analysis of NGS requires 1–2 days on a computing cluster and several weeks on a typical workstation
- The innovation in production of genomic data is much higher than the rate of innovation within genomic informatics, known as Moore's law
- A solution is to enclose computation into transportable objects that can be run on the computers where the data is stored (bringing computation to the data instead of bringing the data to the computation)
- Given the amount of data generated, elimination of false positives and false negatives is challenging. False positives can be identified using conventional PCR and Sanger sequencing, whereas elimination of false negatives is much harder to correct
- Mutations with a very low prevalence (i.e., 1–0.1%) may be missed, obscured by similar sequencing error rates

10.11.2 Sufficient Amount of Tissue

- Obtaining sufficient amounts of DNA from the germline counterpart of test sample for integrated genomic analysis is a limitation; however, the amount of DNA required for successful NGS is going down with time

10.11.3 Ethical Considerations

- NGS data provide information not only regarding the DNA makeup of the individual but also of the individual's relatives and related population groups
- The traditional informed consent focus on the individual does not take family, community, and population groups into account; alternative methods are needed to allow individuals to exercise their autonomy while respecting the privacy of their related population groups
- No guidelines as to the correct approach in the cases when an individual requests to have their whole genome sequenced because of a family history of a given genetic disorder or what to do with findings from research participants
- NGS data can provide researchers with clear-cut, verifiable clinical consequences for research participants. Currently, the consensus is that, where there is a serious treatable disorder, researchers have a moral obligation to share this information with research participants
- In case where findings are of uncertain consequence or involve untreatable disorders, the possible benefits of informing participants need to be weighed against the participant's right not to know
- The consent process for participants in whole-genome sequencing studies should reflect a thorough explanation of these scenarios
- Storage of sequencing data should be secure and anonymous. The ethic committees of the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA) have suggested ways to ensure this

10.12 Conclusions and Future Directions

- Although genomic technologies have been successfully applied to characterize human cancers and have the ability to predict clinical outcomes, several challenges need to be overcome before these techniques will be implemented in patient management
- Until a consensus is reached among all laboratories for standardization of microarray-type technology and NGS approaches, inter-experimental variability will remain commonplace inhibiting the transfer of microarrays and NGS technologies from the bench to the bedside
- Validating the results of these experiments is an important challenge for the future
- Lack of frozen tumor material and normal germline DNA from patients with adequate clinical pathologic annotation, including outcome data, is so far a major limiting factor. It will be an advance to incorporate the

collection of frozen tumor and normal tissue samples into prospective randomized clinical trials
- The ethical aspects need to be thoroughly considered before next-generation sequencing can be applied to a large number of patients
- Ultimately, genomic technologies, including DNA sequencing technologies and gene expression profiling, will improve our ability to precisely define disease subtypes enabling accurate treatment tailored for each individual patient

Further Reading

Abdullah-Sayani A, Bueno-de-Mesquita JM, van de Vijver MJ. Technology Insight: tuning into the genetic orchestra using microarrays – limitations of DNA microarrays in clinical practice. Nat Clin Pract Oncol. 2006;3:501–16.

Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403:503–11.

Armstrong SA, Kung AL, Mabon ME, et al. Inhibition of FLT3 in MLL. Validation of a therapeutic target identified by gene expression based classification. Cancer Cell. 2003;3:173–83.

Ascano M, Hafner M, Cekan P, et al. Identification of RNA–protein interaction networks using PAR–CLIP. Wiley Interdiscip Rev RNA. 2012;3:159–77.

Bell DW. Our changing view of the genomic landscape of cancer. J Pathol. 2010;220:231–43.

Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 2006;439:353–7.

Bloom G, Yang IV, Boulware D, et al. Multi-platform, multi-site, microarray-based human tumor classification. Am J Pathol. 2004;164:9–16.

Buckhaults P, Zhang Z, Chen YC, et al. Identifying tumor origin using a gene expression-based classification map. Cancer Res. 2003;63:4144–9.

Campbell PJ, Yachida S, Mudie LJ, et al. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature. 2010;467:1109–13.

Chang HY, Sneddon JB, Alizadeh AA, et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS Biol. 2004;2:E7.

Chang HY, Nuyten DSA, Sneddon JB, et al. From the cover: robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proc Natl Acad Sci USA. 2005a;102:3738–43.

Chang JC, Hilsenbeck SG, Fuqua SA. The promise of microarrays in the management and treatment of breast cancer. Breast Cancer Res. 2005b;7:100–4.

Chuaqui RF, Bonner RF, Best CJ, et al. Post-analysis follow-up and validation of microarray experiments. Nat Genet. 2002;32(Suppl):509–14.

Cronin M, Pho M, Dutta D, et al. Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. Am J Pathol. 2004;164:35–42.

Darnell RB. HITS–CLIP: panoramic views of protein–RNA regulation in living cells. Wiley Interdiscip Rev RNA. 2010;1:266–86.

Dennis JL, Vass JK, Wit EC, et al. Identification from public data of molecular markers of adenocarcinoma characteristic of the site of origin. Cancer Res. 2002;62:5999–6005.

Dennis J, Hvidsten T, Wit E, et al. Markers of adenocarcinoma characteristic of the site of origin: development of a diagnostic algorithm. Clin Cancer Res. 2005;11:3766–72.

Diamandis M, White NM, Yousef GM. Personalized medicine: marking a new epoch in cancer patient management. Mol Cancer Res. 2010;8:1175–87.

Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010;464:999–1005.

Ebert BL, Golub TR. Genomic approaches to hematologic malignancies. Blood. 2004;104:923–32.

Eisen M, Spellman P, Brown P, et al. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA. 1998;95:14863–8.

Eschrich S, Yang I, Bloom G, et al. Molecular staging for survival prediction of colorectal cancer patients. J Clin Oncol. 2005;23:3526–35.

Farazi TA, Horlings HM, Ten Hoeve JJ, et al. MicroRNA sequence and expression analysis in breast tumors by deep sequencing. Cancer Res. 2011a;71:4443–53.

Farazi TA, Spitzer JI, Morozov P, et al. miRNAs in human cancer. J Pathol. 2011b;223:102–15.

Galasso M, Elena Sana M, Volinia S. Non-coding RNAs: a key to future personalized molecular therapy? Genome Med. 2010;2:12.

Garber M, Grabherr MG, Guttman M, et al. Computational methods for transcriptome annotation and quantification using RNA-seq. Nat Methods. 2011;8:469–77.

Gupta RA, Shah N, Wang KC, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010;464:1071–6.

Hafner M, Landgraf P, Ludwig J, et al. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. Methods. 2008; 44:3–12.

Hafner M, Landthaler M, Burger L, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR–CLIP. Cell. 2010;141:129–41.

Harbour JW, Onken MD, Roberson ED, et al. Frequent mutation of BAP1 in metastasizing uveal melanomas. Science. 2010;330:1410–3.

Hausser J, Berninger P, Rodak C, et al. MirZ: an integrated microRNA expression atlas and target prediction resource. Nucleic Acids Res. 2009;37:W266–72.

Hawkins RD, Hon GC, Ren B. Next-generation genomics: an integrative approach. Nat Rev Genet. 2010;11: 476–86.

Hirst M, Marra MA. Next generation sequencing based approaches to epigenomics. Brief Funct Genom. 2010;9:455–65.

Hurd PJ, Nelson CJ. Advantages of next-generation sequencing versus the microarray in epigenetic research. Brief Funct Genom Prot. 2009;8:174–83.

Jiao Y, Shi C, Edil BH, et al. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. Science. 2011;331: 1199–203.

Jones S, Zhang X, Parsons DW, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science. 2008;321:1801–6.

Jones S, Hruban RH, Kamiyama M, et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. Science. 2009;324:217.

Jungkamp AC, Stoeckius M, Mecenas D, et al. In vivo and transcriptome-wide identification of RNA binding protein target sites. Mol Cell. 2011;44:828–40.

Kahn SD. On the future of genomic data. Science. 2011;331:728–9.

Kan T, Shimada Y, Sato F, et al. Prediction of lymph node metastasis with use of artificial neural networks based on gene expression profiles in esophageal squamous cell carcinoma. Ann Surg Oncol. 2004; 11:1070–8.

Kan Z, Jaiswal BS, Stinson J, et al. Diverse somatic mutation patterns and pathway alterations in human cancers. Nature. 2010;466:869–73.

Kaye J, Boddington P, de Vries J, et al. Ethical implications of the use of whole genome methods in medical research. Eur J Hum Genet. 2010;18:398–403.

Kishore S, Jaskiewicz L, Burger L, et al. A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat Methods. 2011;8: 559–64.

Lapointe J, Li C, Higgins JP, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proc Natl Acad Sci USA. 2004; 101:811–6.

Lee W, Jiang Z, Liu J, et al. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature. 2010;465:473–7.

Ley TJ, Mardis ER, Ding L, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature. 2008;456:66–72.

Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010;363: 2424–33.

Linnarsson S. Recent advances in DNA sequencing methods – general principles of sample preparation. Exp Cell Res. 2010;316:1339–43.

Lonning PE, Sorlie T, Borresen-Dale AL. Genomics in breast cancer-therapeutic implications. Nat Clin Pract Oncol. 2005;2:26–33.

Lossos IS, Jones CD, Warnke R, et al. Expression of a single gene, BCL-6, strongly predicts survival in patients with diffuse large B-cell lymphoma. Blood. 2001;98:945–51.

Lossos IS, Alizadeh AA, Rajapaksa R, et al. HGAL is a novel interleukin-4-inducible gene that strongly predicts survival in diffuse large B-cell lymphoma. Blood. 2003;101:433–40.

Lu J, Getz G, Miska E, et al. MicroRNA expression profiles classify human cancers. Nature. 2005;435:834–8.

Ma XJ, Patel R, Wang X, et al. Molecular classification of human cancers using a 92-gene real-time quantitative polymerase chain reaction assay. Arch Pathol Lab Med. 2006;130:465–73.

Mardis ER, Ding L, Dooling DJ, et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009;361:1058–66.

Maxam AM, Gilbert W. A new method for sequencing DNA. 1977. Biotechnology. 1992;24:99–103.

Metzker ML. Sequencing technologies – the next generation. Nat Rev Genet. 2010;11:31–46.

Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet. 2010;11:685–96.

Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics. 2008;92:255–64.

Muers M. Technology: getting Moore from DNA sequencing. Nat Rev Genet. 2011;12:586.

Natrajan R, Reis-Filho JS. Next-generation sequencing applied to molecular diagnostics. Expert Rev Mol Diagn. 2011;11:425–44.

Nevins JR, Huang ES, Dressman H, et al. Towards integrated clinico-genomic models for personalized medicine: combining gene expression signatures and clinical factors in breast cancer outcomes prediction. Hum Mol Genet. 2003;12:R153–7.

Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004;351:2817–26.

Pasqualucci L, Trifonov V, Fabbri G, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet. 2011;43:830–7.

Pavlidis N, Fizazi K. Cancer of unknown primary (CUP). Crit Rev Oncol Hematol. 2005;54:243–50.

Pawitan Y, Bjohle J, Amler L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Res. 2005;7:R953–64.

Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumors. Nature. 2000; 406:747–52.

Pleasance ED, Stephens PJ, O'Meara S, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature. 2010;463:184–90.

Quackenbush J. Microarray analysis and tumor classification. N Engl J Med. 2006;354:2463–72.

Ramaswamy S, Golab TR. DNA microarrays in clinical oncology. J Clin Oncol. 2002;20:1932–41.

Ramaswamy S, Tamayo P, Rifkin R, et al. Multiclass cancer diagnosis using tumor gene expression signatures. Proc Natl Acad Sci USA. 2001;98:15149–54.

Rebbeck TR. Inherited genetic markers and cancer outcomes: personalized medicine in the postgenome era. J Clin Oncol. 2006;24:1972–4.

Reis-Filho JS, Westbury C, Pierga JY. The impact of expression profiling on prognostic and predictive testing in breast cancer. J Clin Pathol. 2006;59:225–31.

Ross ME, Zhou X, Song G, et al. Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. Blood. 2003;102:2951–9.

Rothberg JM, Leamon JH. The development and impact of 454 sequencing. Nat Biotechnol. 2008;26:1117–24.

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA. 1977;74:5463–7.

Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. Hum Mol Genet. 2010;19:R227–40.

Schuster SC. Next-generation sequencing transforms today's biology. Nat Methods. 2008;5:16–8.

Segal E, Shapira M, Regev A, et al. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet. 2003;34:166–76.

Slonim DK. From patterns to pathways: gene expression data analysis comes of age. Nat Genet. 2002;32 Suppl:502–8.

Sorlie T, Perou C, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguishes tumor subclasses with clinical implications. Proc Natl Acad Sci USA. 2001;98:10869–74.

Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci USA. 2003;100:10393–8.

Statnikov A, Aliferis CF, Tsamardinos I, et al. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. Bioinformatics. 2005;21:631–43.

Stephens PJ, McBride DJ, Lin ML, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature. 2009;462:1005–10.

Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50.

Tamoto E, Tada M, Murakawa K, et al. Gene-expression profile changes correlated with tumor progression and lymph node metastasis in esophageal cancer. Clin Cancer Res. 2004;10:3629–38.

Tinker AV, Boussioutas A, Bowtell DD. The challenges of gene expression microarrays for the study of human cancer. Cancer Cell. 2006;9:333–9.

Tothill RW, Kowalczyk A, Rischin D, et al. An expression-based site of origin diagnostic method designed for clinical application to cancer of unknown origin. Cancer Res. 2005;65:4031–40.

Tsai MC, Spitale RC, Chang HY. Long intergenic noncoding RNAs: new links in cancer progression. Cancer Res. 2011;71:3–7.

van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002;347:1999–2009.

Varadhachary GR, Abbruzzese JL, Lenzi R. Diagnostic strategies for unknown primary cancer. Cancer. 2004;100:1776–85.

Villarroel MC, Rajeshkumar NV, Garrido-Laguna I, et al. Personalizing cancer treatment in the age of global genomic analyses: PALB2 gene mutations and the response to DNA damaging agents in pancreatic cancer. Mol Cancer Ther. 2011;10:3–8.

Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. Clin Chem. 2009;55:641–58.

Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet. 2005;365:671–9.

Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10:57–63.

Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell. 2002;1:133–43.

Yu YP, Landsittel D, Jing L, et al. Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy. J Clin Oncol. 2004;22:2790–9.

van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002;415:530–6.

Zhang C, Darnell RB. Mapping in vivo protein–RNA interactions at single-nucleotide resolution from HITS–CLIP data. Nat Biotechnol. 2011;29:607–14.

Zhou X, Ren L, Meng Q, et al. The next-generation sequencing technology and application. Protein Cell. 2010;1:520–36.

Clinical Proteomics in Molecular Genetic Pathology

11

David H. Geho, Virgina Espina, Lance A. Liotta, Emanuel F. Petricoin, and Julia D. Wulfkuhle

Contents

11.1 Introduction to Clinical Proteomics . . . 370
11.1.1 Protein-Building Blocks . . . 370
11.1.2 Protein Structure . . . 370
11.1.3 Tools Used for Protein Studies: An Overview . . . 370
11.1.4 Physical Detection Systems . . . 371
11.1.5 Mass Spectrometry . . . 371
11.1.6 Surface Plasmon Resonance . . . 371
11.1.7 UV Spectroscopy . . . 371
11.1.8 Circular Dichroism . . . 371
11.1.9 Electrophoresis . . . 371
11.1.10 Capillary Electrophoresis . . . 372
11.1.11 Chromatography . . . 372
11.1.12 Specific Affinity Techniques for Protein Detection . . . 372
11.1.13 Hybrid Technologies . . . 372
11.1.14 Limitations of Classical Protein Detection Tools as Clinical Tools . . . 372

11.2 Protein Microarray-Based Clinical Proteomics . . . 372
11.2.1 Clinical Proteomic Diagnostics: Protein Microarrays . . . 372
11.2.2 Specimen Procurement . . . 373
11.2.3 Laser Capture Microdissection . . . 373
11.2.4 Protein Microarrays . . . 373
11.2.5 The Architecture of a Protein Microarray . . . 374
11.2.6 Forward-Phase Arrays . . . 374
11.2.7 Reverse-Phase Microarrays . . . 375
11.2.8 Array Surfaces . . . 375
11.2.9 Arrayer Technologies . . . 375
11.2.10 Contact Printers . . . 375
11.2.11 Noncontact Printers . . . 375
11.2.12 Additional Considerations . . . 375
11.2.13 Antibodies . . . 376
11.2.14 Microarray Reporter Technologies . . . 376
11.2.15 Chromogenic Reporter Technologies . . . 376
11.2.16 Fluorescent Reporter Technologies . . . 376
11.2.17 Image Acquisition . . . 376
11.2.18 Data Analysis . . . 376
11.2.19 Analysis Software . . . 377
11.2.20 Downstream Analysis . . . 377
11.2.21 Protein Microarray Conclusions . . . 377

11.3 Mass Spectrometry-Based Proteomics . . . 377
11.3.1 Methods of Ionization . . . 378
11.3.2 Solid Tissue Mass Spectrometry . . . 378
11.3.3 Body Fluid Mass Spectrometry . . . 378
11.3.4 Working Model for the Genesis of the Serum Peptidome . . . 378
11.3.5 High-Abundance Versus Low-Abundance Blood Proteome . . . 379
11.3.6 A New Mass Spectrometry–Based Proteomics Work Flow for Body Fluid–Based Biomarker Discovery and Characterization . . . 379
11.3.7 Specimen Procurement and Preservation . . . 379

Further Reading . . . 380

D.H. Geho, MD, PhD (✉)
Imaging, Merck and Company, Inc., West Point, PA, USA

V. Espina, MS • L.A. Liotta, MD, PhD • E.F. Petricoin, PhD • J.D. Wulfkuhle, PhD
Center for Applied Proteomics and Molecular Medicine, George Mason University, Manassas, VA, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_12,

11.1 Introduction to Clinical Proteomics

- Clinical proteomics is a rapidly maturing research discipline
- The fundamental questions of clinical proteomics are the following:
  - What proteins or protein isoforms are present in a disease process
  - How do those proteins interact
  - What are the relative abundance and activation states of disease-related proteins
- This chapter reviews technologies and experimental procedures that enable clinical proteomics research

11.1.1 Protein-Building Blocks

- Proteins are polymers of amino acids, linked together by peptide bonds. A peptide bond is the amide linkage between a carboxyl group in one amino acid and amino group in another amino acid
- Vast numbers of potential amino acid sequences are possible
- However, the human genome is made of about $3–4 \times 10^4$ genes. Therefore, the proteins produced in the human system are a small subset of the theoretical protein complement
- Clinical proteomics focuses on the relatively limited population of clinically relevant proteins transcribed from the genome

11.1.2 Protein Structure

- The amino acid sequence of a protein is called its primary structure. Out of this sequence arise the intrinsic properties of the protein, such as surface shape, size, and charge. These are important characteristics of a protein, which determine the ultimate function(s) of a protein
- Based on the primary structure, a linear chain of amino acids coalesce and fold to form a series of secondary structural elements such as α-helices, β-pleated sheets, and random coils
- Tertiary structure is comprised of higher-order arrangements of secondary structural motifs. The secondary structural elements fold in such a manner so as to assume a thermodynamically stable conformation
- Cysteines can be linked via disulfide bonds, which provide further structural stability
- Quaternary structure of a protein refers to the higher-order arrangements of tertiary structures
- Historically, structural changes in proteins have been linked to diseases. As one example, sickle cell anemia results from of a single amino acid change in an otherwise unaltered primary amino acid sequence. This substitution gives rise to a protein with a different shape than normal hemoglobin and altered higher-order protein structures. The aggregation of sickle hemoglobin results in the formation of rigid fibers which causes red blood cells to become abnormally sickle-shaped
- Amino acid side chains within proteins are sites for the covalent addition of molecules such as phosphates, sugars, and lipids. These modifications occur after the protein has been translated from mRNA and are termed post-translational modifications (i.e., phosphorylation, glycosylation, and lipidation)
- The study of posttranslational modified proteins represents a vast and important area of disease pathophysiology research. Certain proteins are phosphorylated or dephosphorylated on specific residues in response to cellular signals. These signaling cascades play an important role in orchestrating cellular growth, migration, and apoptosis, among other functions. In general, posttranslational modifications are not detected using genomic approaches

11.1.3 Tools Used for Protein Studies: An Overview

- Broadly speaking, two classes of protein characteristics have been used to study protein functions

- Intrinsic physical properties. These are broadly applicable across a range of protein types and provide information about the protein's mass, charge, or structure. Examples of this type of technology are mass spectrometry, surface plasmon resonance, electrophoresis, ultraviolet (UV) spectroscopy, and chromatography
- Protein-specific properties. These techniques are protein-specific and provide information about posttranslational modifications as well as presence or absence of specific proteins in complex mixtures. They are usually derived from previously well-characterized individual proteins, such as antibody-based detection systems. Examples of approaches used to study protein-specific properties include flow cytometry, immunohistochemistry, enzyme-linked immunosorbent assay, and Western blots

11.1.4 Physical Detection Systems

- Within a protein, the overall sum of amino acids, the intrinsic qualities, and relative abundance of the amino acids provide physical elements suitable for detection by a number of complementary technologies, including mass spectrometry, chromatography, electrophoresis, surface plasmon resonance, and circular dichroism

11.1.5 Mass Spectrometry

- Mass spectrometry takes advantage of the behavior of a charged molecule in magnetic fields in order to classify proteins based on mass/charge ratios. This is discussed in further detail later

11.1.6 Surface Plasmon Resonance

- The biomolecular interactions of unlabeled proteins can be studied using surface plasmon resonance. Proteins are immobilized onto a thin metal film. If ligand binding occurs, the refractive index changes and these changes are detected by an optical sensor. Analyte association/dissociation rate constants may be calculated

11.1.7 UV Spectroscopy

- By measuring the absorbance of UV light by aromatic side chains of constituent amino acids within proteins, UV spectroscopy enables protein detection and quantitation

11.1.8 Circular Dichroism

- Circular dichroism spectroscopy uses circularly polarized light of one direction as a quick, low-resolution method for determining protein structure. Within a protein, the relative abundance of secondary structural elements such as α-helices and β-pleated sheets can be measured

11.1.9 Electrophoresis

- Standard one-dimensional (1D) gel electrophoresis provides a means for protein separation. The movement of proteins through a solution in a polymeric matrix based on the application of an electrical field represents the primary technology. Differences in migration are based primarily on size of the protein
- In two-dimensional electrophoresis, another level of separation, isoelectric point, is used to further separate protein species in addition to size-based separation. In the first step, a pH gradient permits movement of the constituent proteins until they reach their isoelectric point (pKa). In the second step, the separation occurs at right angles to the first dimension and is based on protein size as in 1D electrophoresis

11.1.10 Capillary Electrophoresis

- Electrophoresis can also be performed on protein samples within a capillary tube, providing high-resolution protein separations when voltage is applied to the system

11.1.11 Chromatography

- Proteins can be separated by passing them over a resin that partitions the molecules between the liquid or gas phase and the bound resin. Types of chromatography include size exclusion, ion exchange, and high-pressure liquid chromatography

11.1.12 Specific Affinity Techniques for Protein Detection

- A molecule or macromolecular structure that selectively interacts with a protein can be utilized as an affinity system for specific protein detection. Affinity reagents that can be used to isolate specific proteins include metals, carbohydrates, proteins, and nucleic acids
- A commonly used type of affinity reagent is an antibody that binds to a specific protein. Validation of a reagent's sensitivity and specificity must be performed for each antibody
- Antibody reagents may also be used for detection of specific proteins within clinical samples by immunohistochemistry, immunoflow cytometry, Western blots, and enzyme-linked immunosorbent assays

11.1.13 Hybrid Technologies

- Many times, a system for protein characterization and detection represents the fusion of several technologic approaches. A physical detection method, such as electrophoresis, can be paired with an antibody probing step, as is the case with a Western blot. Affinity chromatography using antibodies immobilized to a chromatography resin represents another example

11.1.14 Limitations of Classical Protein Detection Tools as Clinical Tools

- Proteomic tests in a clinical setting must be rapidly performed using very limited clinical material. Classical proteomic tools have had limited applicability due to the need for relatively large sample volume and time constraints

11.2 Protein Microarray-Based Clinical Proteomics

- Clinical proteomics is a new field that combines components of classical protein detection technologies with new technologies to create high-throughput assays that effectively utilize the proteomic information available in limited sample volumes
- Patient tissue specimens contain a wealth of potential diagnostic molecular descriptors
- Of particular interest are proteins that are involved in cell signaling, metabolism, migration, cell division, immunity, and epigenetic processes within diseased cells
- Many of the proteins of interest undergo posttranslational modifications, such as phosphorylation and cleavage. For example, many cell signaling pathways involve cascades of phosphorylated signaling proteins
- A genomics-only approach will not detect this posttranslationally defined molecular information. Assays with high sensitivity are required, because proteins cannot be copied/amplified via a polymerase chain reaction as is the case with DNA and RNA

11.2.1 Clinical Proteomic Diagnostics: Protein Microarrays

- Critical phases in a protein microarray evaluation include the following: specimen

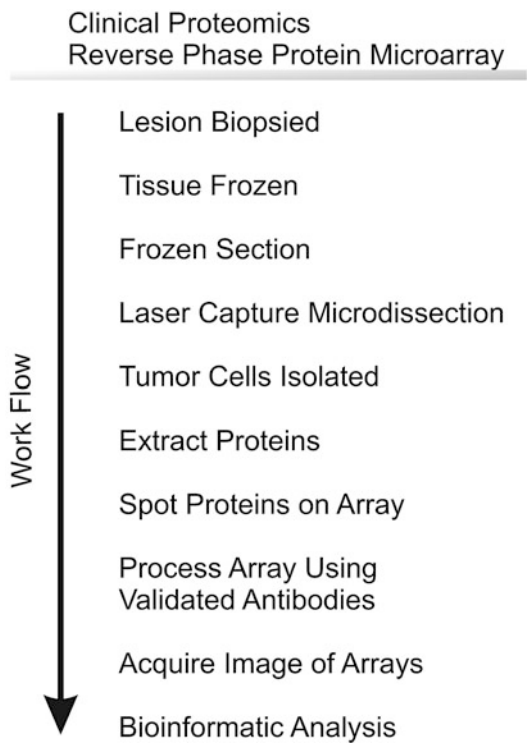


Fig. 11.1 Clinical proteomics work flow for reverse-phase protein microarrays. Immediately after harvesting, the tissue specimen must be frozen in order to preserve labile molecular end points. A frozen section is later prepared and stained. A targeted cell population, such as tumor cells, is isolated using laser capture microdissection. Proteins are extracted from the isolated cells and printed onto an array substrate using a robotic, high-throughput protein microarrayer. The proteins of interest are detected and quantified on the microarray by applying a series of chemical reagents, including validated antibodies, onto the array surface, followed by a reporter detection system (such as streptavidin–HRP). The microarray is then scanned, an image is saved, and bioinformatics analysis and clinical correlation is performed

procurement, specimen preservation, specimen processing, molecular characterization, data analysis, and clinical correlation (see Fig. 11.1)

- Key technologies for protein microarray analysis: laser capture microdissection, a reliable array printing device, array substrates (surfaces), an image acquisition system (scanner), and bioinformatics analysis

11.2.2 Specimen Procurement

- Integration of a clinical proteomics assay into pathology practice requires the development of a comprehensive program that involves nurses, physicians, tissue procurement specialists, tissue processing facilities, and sufficient storage space
- For clinical proteomics assessments, the tissues must be rapidly processed shortly after procurement. Typically, tissue is frozen at −80 °C or placed in liquid nitrogen. Rapid processing limits protein degradation and phosphatase activity
- Protease and phosphatase inhibitors may be added to the samples
- Formalin fixation is the gold standard for preserving tissue histomorphology, but because it penetrates tissue slowly, it is unsuitable for stabilizing phosphoproteins
- New biomarker and histology preservative (BHP) chemistries that stabilize signaling protein phosphorylation and retain formalin-like tissue histomorphology with equivalent immunohistochemistry are being developed that provide the convenience of formalin-type fixation plus the preservation of highly labile biomolecules

11.2.3 Laser Capture Microdissection

- Many distinct cell types are present within a tissue specimen, including cells of epithelial, endothelial, hematopoietic, or mesenchymal derivation
- For this reason, it is important to isolate specific cell populations, such as tumor cells, from a biopsy for clinical proteomic assessments
- Laser capture microdissection enables pure cell populations, such as tumor cells, to be isolated from the complex tissue microenvironment

11.2.4 Protein Microarrays

- Protein microarrays provide a means for measuring the levels of disease-related proteins extracted from patient tissues

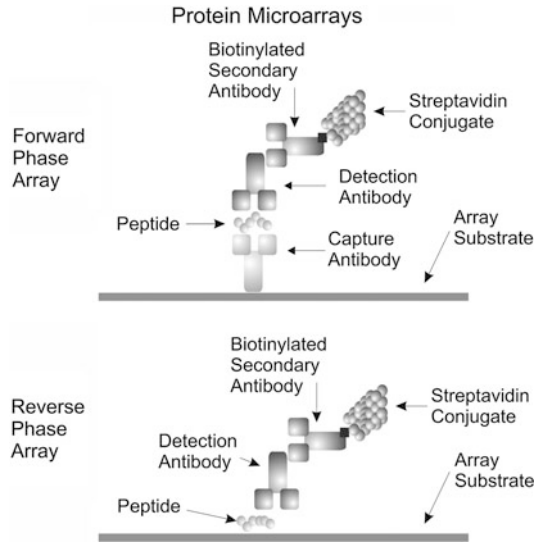


Fig. 11.2 Comparison of forward-phase arrays and reverse-phase arrays. For forward-phase arrays, an array surface is coated with bait molecules, such as an antibody, to capture a particular protein analyte. The remainder of the array surface is blocked to decrease nonspecific binding interactions. The array is then incubated with a mixture of analytes. Analytes bound to the bait molecule are detected using a second, distinct antibody. Alternatively, the analytes may be directly labeled. For the antibody detection system, two distinct epitopes, or antibody-binding regions, must be available on the same analyte. Steric hindrance between two antibodies on the same small analyte may interfere with detection. The binding of the second antibody is detected using a biotinylated detection antibody. For fluorescence detection, streptavidin-conjugated fluorophores are applied, or streptavidin linked to HRP is incubated on the array followed by the addition of a chromogenic substrate such as DAB (diaminobenzidine) for colorimetric detection. For reverse-phase arrays, the cellular proteins are directly arrayed onto a substrate. The remainder of the array surface is blocked to decrease nonspecific binding interactions. An antigen-specific antibody is then incubated on the array surface. The presence of bound detection antibody is in turn probed using a biotinylated secondary antibody. Streptavidin linked to HRP is incubated on the array followed by treatment with a chromogenic substrate such as DAB for colorimetric detection. Streptavidin-conjugated fluorophores may be applied after the secondary antibody for fluorescence-based detection

11.2.5 The Architecture of a Protein Microarray

- Certain technologies are required for protein microarrays as an assay class: an arraying device, a substrate that functions as the array surface, antibodies, or some other detection probe, a reporter system, a detection device, and bioinformatics analysis
- Varying microarray architectures have been employed. There are two fundamental types of arrays: forward-phase arrays and reverse-phase arrays (see Fig. 11.2). The difference lies in how the analytes, or proteins of interest, are captured for study
- Once captured, the analytes are detected using a specific probe, such as an antibody

11.2.6 Forward-Phase Arrays

- An array surface is coated with bait molecules, such as an antibody, to capture a particular analyte
- Capture moiety (bait molecules) can be arrayed in dilution curves on the same array
- Multiple bait molecules can be spotted on the same array
- The array is then incubated with a mixture of analytes
- Analytes bound to the bait molecule are detected using a second, distinct antibody. Alternatively, the analytes may be directly labeled
- For the antibody detection system, two distinct epitopes or antibody-binding regions must be

available on the same analyte. Steric hindrance between two antibodies on the same small analyte may interfere with detection

11.2.7 Reverse-Phase Microarrays

- In the reverse-phase architecture, the cellular proteins are printed directly onto a substrate in an array format
- Multiple samples are immobilized on the array substrate
- Denatured or nondenatured analytes may be arrayed
- Analytes may be printed in a dilution curve that enables the linear component of the antibody–analyte interaction to be evaluated

11.2.8 Array Surfaces

- An array surface must be capable of binding proteins. One example is nitrocellulose-coated glass slides
- The substrate must have a high surface area and low intrinsic background signal

11.2.9 Arrayer Technologies

- A reliable arrayer is an important technology for clinical proteomics microarrays. These devices must have reliable, reproducible delivery (printing) of proteins onto the array surface
- Two primary forms of printing are available commercially: contact and noncontact devices

11.2.10 Contact Printers

- Contact printers deposit samples via direct contact of the print head with the substrate and utilize solid pins or split pins with a built-in microreservoir to transfer sample to the substratum
- Pin diameter and the fluid properties of the sample dictate the volume deposited and the spot size
- Solid pin printers require repeated dipping into a sample reservoir to print replicate spots onto the array
- Variations such as the pin and ring assembly, ink jet technology, and split pin configurations allow multiple spots to be printed from each sample without repeated replenishment from the main sample reservoir
- Volumes deposited are generally in the nanoliter range

11.2.11 Noncontact Printers

- Noncontact printing devices utilize a sensor for depositing sample from above the substratum
- These arrayers may involve a piezoelectric crystal sensor or employ syringe-solenoid technology
- Piezoelectric printing technology utilizes a piezoelectric crystal closely apposed to a fluid reservoir with a capillary tip. Changes in electrical voltage control crystal deformation such that sample is ejected through the glass capillary
- Syringe-solenoid technology utilizes a syringe pump coupled to a solenoid valve to aspirate sample into a reservoir, which is then dispensed at high pressure upon opening of the solenoid valve
- Volumes dispensed from these devices are in the picoliter to nanoliter range

11.2.12 Additional Considerations

- Control of temperature and humidity conditions are critical to printing quality and reproducibility in that they affect sample evaporation from the reservoir and/or water condensation on the instrument or the substratum
- Pin washing between samples is critical for reproducibility and prevents sample carryover.
- Pin calibration for the degree of contact with the substrate is an important consideration

with contact printing devices affecting spot size and quality
- Dust and debris on the printing substratum can lead to printing irregularities, damage to the substrate, or cause interference with spot detection

11.2.13 Antibodies

- For reverse-phase arrays, after analyte spotting, the array surface can be probed with molecules, such as antibodies that specifically recognize proteins of interest within the lysate
- The specificity of an antibody must be validated using a Western blot assay
- For forward-phase arrays, antibodies are used to both capture and detect the analytes of interest
- A limitation of proteomic arrays is the availability of sensitive and specific antibodies that detect posttranslational modifications such as phosphorylation
- A widespread effort, including academic and industry laboratories, is underway to produce and characterize antibody libraries to meet this critical need

11.2.14 Microarray Reporter Technologies

- The detection of an analyte by an antibody probe must be transformed into a detectable signal. Reporter technologies enable this transformation either through chromagen deposition, fluorophore deposition, or a radioactive tag

11.2.15 Chromogenic Reporter Technologies

- Chromogenic detection systems provide reproducible, sensitive signals
- As in immunohistochemistry, a biotinylated antibody is bound by streptavidin linked to an enzyme such as horseradish peroxidase (HRP)
- Enzymatic signal amplification strategies can be used to increase the sensitivity of the assay, such as biotinyl tyramide amplification
- With chromogenic detection strategies, utilization of a printed sample dilution curve to identify and analyze the linear range of detection for the antibody–analyte interaction in each sample is critical due to the catalytic nature of the deposition reaction

11.2.16 Fluorescent Reporter Technologies

- Fluorescence-based detection systems also provide reproducible, sensitive signals
- Enzymatic signal amplification strategies, such as biotinyl tyramide amplification, can be used in combination with streptavidin-conjugated fluorophores for analyte detection
- Printing samples in serial dilution curves is not critical when using fluorescent reporter molecules because current image acquisition instruments allow the user to control exposure time or gain level to prevent signal saturation

11.2.17 Image Acquisition

- For chromogenic arrays, a flatbed scanner can be used to acquire an image of an array
- For fluorescent arrays, a CCD camera or a laser scanning system (e.g., Tecan PowerScanner, Tecan Group Ltd., Männedorf, Switzerland) can be used to detect signals from arrays

11.2.18 Data Analysis

- Programs capable of quantifying pixel intensity are essential for the analysis of protein microarray data

- Additionally, statistical software packages for downstream analysis of data and correlations with clinical parameters are essential for individualized therapy

11.2.19 Analysis Software

- Array analysis software that provides automated spot detection and local area background subtraction can greatly increase efficiency and reduce operator bias during analysis
- Subtraction of negative control spot (secondary antibody alone) intensities may also be necessary when working with samples that have unusually high levels of endogenous peroxidases, avidin, or biotin
- For data comparisons across samples, normalization of the data to total protein values or to a housekeeping protein is necessary
- Addition of high/low control samples and reference standard dilution curves varying by known amounts of relative analyte concentration can serve as quality control elements for the technical aspects of array processing and allow for interpolation of sample intensity values to standardized reference units that can then be compared between study sets separated in time on different sets of arrays

11.2.20 Downstream Analysis

- Final data output can be used to compare samples or groups with a variety of univariate and multivariate statistical methods
- Unsupervised hierarchical clustering of the entire dataset can identify differences/similarities in protein(s) levels/activity that leads to the formation of subgroups within the dataset
- Additional tests, such as principal component analysis, can help to identify key analytes underlying differences between groups of interest
- Spearman rho correlations provide information regarding proteins that may be linked in a signaling network

11.2.21 Protein Microarray Conclusions

- Numerous protein array studies have used patient specimens to define a specific protein's contribution to disease processes such as ovarian cancer, prostate cancer, breast cancer, and lymphoma
- In order for a proteomic tool to be applied to clinical proteomics research, it must meet high thresholds for sensitivity and reproducibility, as the molecular information that is queried is often of low abundance
- Challenges for the clinical application of protein microarray technologies include standardization of the procedure across laboratories and technologic improvements in arrayers, substrates, reporter technologies, and imaging systems

11.3 Mass Spectrometry-Based Proteomics

- Mass spectrometry, a highly sensitive proteomics tool, is becoming widely used as a tool to discover and catalog disease-related proteins in solid tissues and body fluids
- Mass spectrometry has been used as a method for physical analysis of molecules for many years
- The basis for the technique is the behavior of charged particles in magnetic fields
- The molecule to be studied can be charged either through the addition or removal of protons or electrons
- The charged molecule is introduced into a chamber, which is under a vacuum, and hurtled through a high electromagnetic field
- Mass spectrometry analysis utilizes some type of mass analyzer, including ion trap, time of flight, quadrapole, or Fourier transform ion

cyclotron resonance. These analyzers enable a mass/charge (m/z) value to be assigned to the molecules undergoing analysis
- Using protein digestion methods (i.e., trypsin-based fragmentation), the identity of proteins can be determined by coupling the fragment patterns with computer bioinformatics algorithms

11.3.1 Methods of Ionization

- With electrospray ionization, molecules in a solution are sprayed through an electric field. The field introduces a charged state within the fluid. The solvent is evaporated and the charges become associated with the molecules
- Matrix-assisted laser desorption ionization (MALDI) is another type of ionization wherein a matrix containing molecules of interest is broken up in the presence of laser light. This process ionizes the molecules for further analysis
- Surface-enhanced laser desorption and ionization is a type of MALDI, wherein the matrix has surface properties that enable binding of a subset of proteins (i.e., weak cations, strong anions, and so on)

11.3.2 Solid Tissue Mass Spectrometry

- Solid tissues can be evaluated in a variety of ways
- One method utilizes protein lysates derived from laser capture microdissected cells. The lysates are then studied using mass spectrometry
- Direct tissue mass spectrometry is another method, wherein a frozen tissue section is applied directly onto a mass spectrometry substrate

11.3.3 Body Fluid Mass Spectrometry

- The presence of disease-related proteins in body fluids, including plasma, serum, vitreous humor, and cerebrospinal fluid, has been examined using mass spectrometry
- Current mass spectrometry technologies, despite high performance and recent advancements, still do not possess the inherent analytical sensitivity of conventional immunoassays
- Current methods to increase proteomic coverage and measure low-abundance proteins have been constrained to multiplexed fractionation that, while increasing the number of analytes measured in any run, do not effectively provide a means to measure and detect low-abundance analyte (<ng/ml) concentration
- Effective measurement and detection of low-abundance protein analytes require upfront analyte concentration steps

11.3.4 Working Model for the Genesis of the Serum Peptidome

- Many cell types contribute to the formation of a disease microenvironment
- For a cancer, the cells that coalesce to form a disease microenvironment include the neoplastic cells and surrounding stromal cells
- All of these cells produce proteins that are released into the interstitium
- Proteins may also be released as a cell dies and degrades. Resident enzymatic proteins cleave the proteins, resulting in an array of proteolytic fragments from both high- and low-abundance proteins
- The fragments enter the vascular compartment, which filters molecular information from the interstitium through the endothelial cell wall barrier
- These shed, fragmented molecules provide a molecular portrait of ongoing normal and pathologic processes within a tissue
- The peptide fragments are protected by association with larger, high-abundance proteins in the blood, such as albumin
- Further degradation of the proteins in blood may occur once the blood specimen is collected, contributing to the array of protein fragments available for study

11.3.5 High-Abundance Versus Low-Abundance Blood Proteome

- Most of the proteins in blood are high-abundance, high-molecular-weight molecules (99% of the mass) such as albumin, transferrin, and immunoglobulins
- It is in the low-molecular-weight (LMW) fraction of the blood (<40 KDa) where significant attention is currently focused for discovery of disease-specific proteomic information
- The LMW proteome is of intense interest due to the ability of this archive to naturally pass from the tissue interstitium into the vasculature, thus sampling directly the disease microenvironment
- Experimental evidence indicates that the LMW molecules in blood are noncovalently bound to high-abundance proteins like albumin that exist in millions to billion-fold molar excess compared to low-abundance biomarkers
- A current challenge for blood-based proteomics is the development of technologies that release the peptide fragments from proteins such as albumin, so they can be measured using mass spectrometry while concomitantly excluding the isolation of albumin and other high-abundance and high-molecular-weight proteins
- Peptidomic and LMW proteomics require both concentration and size exclusion simultaneously

11.3.6 A New Mass Spectrometry–Based Proteomics Work Flow for Body Fluid–Based Biomarker Discovery and Characterization

- New core-shell hydrogel nanoparticles have been developed for one-step, high-throughput LMW biomarker concentration, size filtration, and preservation
- These new biomarker harvesting nanoparticles can be easily incorporated into any mass spectrometry top-down or bottom-up workflow
- The LMW proteome is preserved once analytes are bound within the nanoparticles, thus minimizing degradation following procurement
- Often, tryptic digestion of the captured and fractionated proteins is performed during proteomic analysis
- The resulting polypeptide fragments are separated using a device such as high-performance liquid chromatography
- The fragments are then ionized and detected using mass spectrometry systems such as electrospray ionization (ESI) or MALDI or multiple reaction monitoring (MRM) mass spectrometry
- Bioinformatics programs enable the polypeptide fragments to be identified
- Bioinformatic interrogation of mass spectrometry data can also be used to identify sets of peptides/proteins that are uniquely found in disease or healthy patients
- Once biomarker candidates are identified, verification and validation can be performed using antibody-based approaches such as protein microarrays or ELISAs or other immunoassays or antibody-independent techniques such as MRM using a triple quadrupole mass spectrometer

11.3.7 Specimen Procurement and Preservation

- Formalin fixation is a slow process (>1 mm/h); thus, tissue proteins, mRNA, and metabolites are rapidly changing as a consequence of the tissue being alive for hours after acquisition
- Tissues and body fluid biospecimens must be processed according to a standardized protocol
- Specimens should be snap-frozen within 5–15 min following removal from a patient
- Specimens should be handled in the same manner for all of the patients in a particular study

- Frozen biospecimens should be stored at −80 °C
- Proteomic analysis of tissue and body fluids should avoid freeze–thaw cycles
- New tissue preservation fixatives have been recently developed for rapid preservation of molecular information and stabilization of labile biomarkers such as phosphoproteins

Further Reading

Anderson NL, Anderson NG. The human plasma proteome: history character, and diagnostic prospects. Mol Cell Proteomics. 2002;1:845–67.

Blume-Jensen P, Hunter T. Oncogenic kinase signalling. Nature. 2001;411:355–65.

Celis JE, Gromov P. Proteomics in translational cancer research: toward an integrated approach. Cancer Cell. 2003;3:9–15.

Chaurand P, Caprioli RM. Direct profiling and imaging of peptides and proteins from mammalian cells and tissue sections by mass spectrometry. Electrophoresis. 2002;23:3125–35.

Deutsch EW, Eng JK, Zhang H, et al. Human plasma peptide atlas. Proteomics. 2005;5:3497–500.

Emmert-Buck MR, Bonner RF, Smith PD, et al. Laser capture microdissection. Science. 1996;274: 998–1001.

Espina V, Mueller C. Reduction of preanalytical variability in specimen procurement for molecular profiling. Methods Mol Biol. 2012;823:49–57.

Espina V, Wulfkuhle JD, Calvert VS, et al. Laser capture microdissection. Nat Protoc. 2006;1:586–603.

Espina V, Wulfkuhle J, Calvert VS, et al. Reverse phase protein microarrays for theranostics and patient-tailored therapy. Methods Mol Biol. 2008;441: 113–28.

Espina V, Mueller C, Edmiston K, et al. Tissue is alive: new technologies are needed to address the problems of protein biomarker pre-analytical variability. Proteomics Clin Appl. 2009;3:874–82.

Fredolini C, Meani F, Reeder KA, et al. Concentration and preservation of very low abundance biomarkers in urine, such as human growth hormone (hGH), by cibacron blue F3G-A loaded hydrogel particles. Nano Res. 2008;1:502–18.

Fredolini C, Meani F, Luchini A, et al. Investigation of the ovarian and prostate cancer peptidome for candidate early detection markers using a novel nanoparticle biomarker capture technology. AAPS J. 2010;12: 504–18.

Gallagher RI, Silvestri A, Petricoin 3rd EF, et al. Reverse phase protein microarrays: fluorometric and colorimetric detection. Methods Mol Biol. 2011;723: 275–301.

Gallagher RI, Blakely SR, Liotta LA, et al. Laser capture microdissection: Arcturus(XT) infrared capture and UV cutting methods. Methods Mol Biol. 2012;823: 157–78.

Grubb RL, Deng J, Pinto PA, et al. Pathway biomarker profiling of localized and metastatic human prostate cancer reveal metastatic and prognostic signatures. J Proteome Res. 2009;8:3044–54.

Gulmann C, Espina V, Petricoin E, et al. Proteomic analysis of apoptotic pathways reveals prognostic factors in follicular lymphoma. Clin Cancer Res. 2005;11: 5847–55.

Gygi SP, Aebersold R. Mass spectrometry and proteomics. Curr Opin Chem Biol. 2000;4:489–94.

Herrmann PC, Gillespie JW, Charboneau L, et al. Mitochondrial proteome: altered cytochrome c oxidase subunit levels in prostate cancer. Proteomics. 2003;3: 1801–10.

Hillenkamp F, Karas M. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. Methods Enzymol. 1990;193: 280–95.

Knezevic V, Leethanakul C, Bichsel VE, et al. Proteomic profiling of the cancer microenvironment by antibody arrays. Proteomics. 2001;1:1271–8.

Liotta LA, Kohn EC. The microenvironment of the tumour-host interface. Nature. 2001;411:375–9.

Liotta LA, Kohn EC, Petricoin EF. Clinical proteomics: personalized molecular medicine. JAMA. 2001;286: 2211–4.

Longo C, Patanarut A, George T, et al. Core-shell hydrogel particles harvest, concentrate and preserve labile low abundance biomarkers. PLoS One. 2009;4:e4763.

Lowenthal MS, Mehta AI, Frogale K, et al. Analysis of albumin-associated peptides and proteins from ovarian cancer patients. Clin Chem. 2005;51:1933–45.

Luchini A, Geho DH, Bishop B, et al. Smart hydrogel particles: biomarker harvesting: one-step affinity purification, size exclusion, and protection against degradation. Nano Lett. 2008;8:350–61.

Meani F, Pecorelli S, Liotta L, et al. Clinical application of proteomics in ovarian cancer prevention and treatment. Mol Diagn Ther. 2009;13:297–311.

Mehta AI, Ross S, Lowenthal MS, et al. Biomarker amplification by serum carrier protein binding. Dis Markers. 2003;19:1–10.

Mueller C, Edmiston KH, Carpenter C, et al. One-step preservation of phosphoproteins and tissue morphology at room temperature for diagnostic and research specimens. PLoS One. 2011;6:e23780.

Nita-Lazar A, Saito-Benz H, White FM. Quantitative phosphoproteomics by mass spectrometry: past, present, and future. Proteomics. 2008;8:4433–43.

Paweletz CP, Charboneau L, Bichsel VE, et al. Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. Oncogene. 2001;20:1981–9.

Petricoin EF, Zoon KC, Kohn EC, Barrett JC, Liotta LA. Clinical proteomics: translating benchside promise

into bedside reality. Nat Rev Drug Discov. 2002;1:683–95.
Petricoin 3rd EF, Espina V, Araujo RP, et al. Phosphoprotein pathway mapping: Akt/mammalian target of rapamycin activation is negatively associated with childhood rhabdomyosarcoma survival. Cancer Res. 2007;67:3431–40.
Pierobon M, Vanmeter AJ, Moroni N, et al. Reverse-phase protein microarrays. Methods Mol Biol. 2012;823: 215–35.
Rosenblatt KP, Bryant-Greenwood P, Killian JK, et al. Serum proteomics in cancer diagnosis and management. Annu Rev Med. 2004;55:97–112.
Silvestri A, Colombatti A, Calvert VS, et al. Protein pathway biomarker analysis of human cancer reveals requirement for upfront cellular-enrichment processing. Lab Invest. 2010;90:787–96.
Stoeckli M, Chaurand P, Hallahan DE, Caprioli RM. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. Nat Med. 2001;7:493–6.
Tamburro D, Fredolini C, Espina V, et al. Multifunctional core-shell nanoparticles: discovery of previously invisible biomarkers. J Am Chem Soc. 2011;133:19178–88.
Wolf-Yadlin A, Sevecka M, MacBeath G. Dissecting protein function and signaling using protein microarrays. Curr Opin Chem Biol. 2009;13:398–405.
Wulfkuhle JD, McLean KC, Paweletz CP, et al. New approaches to proteomic analysis of breast cancer. Proteomics. 2001;1:1205–15.

Nanotechnology in Molecular Diagnostics

12

Andrew N. Young and Brad A. Kairdolf

Contents

12.1 What is Nanotechnology? 383

12.2 What Is the Likely Role of Nanotechnology for Molecular Diagnostics? 384

12.3 Types of Nanotechnology with Diagnostic Applications 385
12.3.1 Nanoparticles 385
12.3.2 Nanowires and Nanocantilevers 394
12.3.3 Nanofluidic Materials 396

12.4 Other Clinical Applications of Nanotechnology 397
12.4.1 Molecular Imaging 397
12.4.2 Drug Delivery and Therapy 397

12.5 Summary of Key Points for Nanotechnology in Molecular Diagnostics 397

Further Reading 398

A.N. Young, MD, PhD (✉)
Department of Pathology & Laboratory Medicine, Emory University School of Medicine; Grady Health System, Atlanta, GA, USA

B.A. Kairdolf, PhD
Department of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA, USA

12.1 What is Nanotechnology?

- The National Nanotechnology Initiative (http://nano.gov) states that "Nanotechnology is the understanding and control of matter at the nanoscale, at dimensions between approximately 1 and 100 nm, where unique phenomena enable novel applications. Encompassing nanoscale science, engineering, and technology, nanotechnology involves imaging, measuring, modeling, and manipulating matter at this length scale"
- Nanoscale materials that are useful for diagnostic medicine exhibit physical, chemical, and biological properties that differ from the properties of bulk materials and single atoms or molecules
- The properties of nanoscale materials can be varied in a controlled manner, depending on size, shape, and composition
  - Compared to other forms or sizes of the same material, nanoscale materials are often stronger; conduct heat or electricity better; have different magnetic properties; become more chemically reactive; have different optical properties; or interact differently with biomolecules in the same size range, such as proteins (1–20 nm) and nucleic acids (2.5 nm in diameter)
- Nanotechnology is not simply the study of naturally occurring nanoscale materials. Rather, it is the controlled production and

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_47,

manipulation of nanoscale materials, designed with specific size, shape, and composition to produce intended physical, chemical, and biological properties

- There are two main approaches to producing nanoscale materials: top-down fabrication and bottom-up synthesis
  - Top-down fabrication reduces larger pieces of material down to the nanoscale, using methods such as controlled etching, elimination, and layering, which are similar to common processes in the semiconductor industry
    - Compared to bottom-up synthesis, the top-down approach is relatively fast and inexpensive, but tends to require larger amounts of starting materials, produce excess discarded waste, and result in higher defect rates
  - Bottom-up synthesis creates nanoscale materials by building up from atomic- and molecular-scale components, through directed assembly or spontaneous self-assembly
    - Compared to top-down fabrication, the bottom-up approach is more precise, but tends to be expensive, time-consuming, and difficult to scale up for commercial applications
- Regardless of the approach to fabrication, nanotechnology results in materials with one or more dimension between 1 and 100 nanometers (nm). For example
  - Nanoparticles measure 1–100 nm in all three size dimensions. Nanospheres are equal in all dimensions, while nanorods are greater in length than width
  - Nanowires, nanotubes, or nanocantilevers measure 1–100 nm in cross-sectional diameter. Length may be larger than nanoscale
  - Nanoporous materials and nanofluidic devices contain openings or channels measuring 1–100 nm. These features may be contained within structures larger than nanoscale

12.2 What Is the Likely Role of Nanotechnology for Molecular Diagnostics?

- The aim of researchers in this field is to develop molecular diagnostic assays with novel performance characteristics, based on the unique properties of nanoscale materials. Nanotechnology has potential to improve performance of existing diagnostic platforms or create fundamentally novel diagnostic platforms
- Characteristics that can be improved by nanotechnology include diagnostic sensitivity and specificity, analytical sensitivity (including single molecule detection), speed, multiplexing, affordability, portability, and ease of use
- Nanoscale materials exhibit high surface area per volume and thus provide a large surface for chemical reactions or conjugation of biorecognition molecules, such as nucleic acids, antibodies, ligands, or aptamers
  - These features are important for rapid biochemical reactions with high sensitivity and specificity
- Nanoparticles exhibit many unique optical properties related to their size and composition
  - They can be designed to generate intense, nonoverlapping optical detection signals, such as fluorescence or Raman light scattering
  - Nanoparticles conjugated to biorecognition molecules are used for ultrasensitive, multiplexed bioassays
  - Excitation and emission in the near-infrared (IR) or IR spectrum is possible, which reduces background signal from blood, body fluids, or tissues
- Nanocantilevers or nanowires that are conjugated to biorecognition molecules transduce specific binding of biomarkers into electrical or optical signals
  - These structures are useful for label-free detection of biomarkers in platforms such as microarrays and point-of-care devices

- Nanoporous materials permit passage of specific biomolecules, based on molecular composition and shape
  - When linked to appropriate sensors (e.g., electrical), nanopores are used to detect and sequence specific DNA species or detect and probe the function of biomarkers at single molecule sensitivity
- Nanomaterials have been combined with micro- or nanofluidic channels, and electrical or magnetic sensors, to create portable "lab-on-a-chip" platforms
- Nanotechnologies have great potential for integrating molecular diagnostics with in vivo molecular imaging and therapy
  - Nanoscale materials are of appropriate size and structure for multiple functional attachments (e.g., detection signals and therapeutic agents), as well as for targeted delivery to anatomical sites
- Based on the novel features of nanoscale materials, nanotechnology is widely expected to play an important role in early disease detection and personalized, predictive medicine

12.3 Types of Nanotechnology with Diagnostic Applications

- Numerous nanotechnologies have been developed, for use in a diverse range of biomedical applications. Key examples are listed in Table 12.1. For additional information on a more complete listing of nanotechnologies useful for biomedicine, please refer to "Further Reading" at the end of the chapter

12.3.1 Nanoparticles

- Nanoparticles can be synthesized from a variety of materials, with sufficient purity and uniformity for biomedical applications. They are usually fabricated with the bottom-up approach. Nanoparticles with great diagnostic potential generate a measurable signal (such as optical, electronic, or magnetic), which is controlled by varying particle size, shape, or composition
- Diagnostically useful nanoparticles include semiconductor quantum dots, gold nanoparticles and nanorods, surface-enhanced Raman scattering (SERS) gold nanoparticles, and biobarcodes, among others

12.3.1.1 Quantum Dots (QDs)

- Structure
  - QDs are semiconductor nanocrystals used as fluorescent labels in bioassays. The electronic and optical characteristics of QDs are highly dependent on nanocrystal size; nanocrystal shape and composition can also affect these properties
  - In electronic terms, the excitons of QDs are confined in all three spatial dimensions, with sizes smaller than the exciton Bohr radius. This "quantum confinement" creates properties intermediate between bulk semiconductors and individual molecules
  - Typical QDs used in diagnostic assays are colloidal inorganic nanocrystals with diameters ranging from 2 to 8 nm. They are often produced as core/shell structures, with a cadmium selenide core capped by a zinc sulfide shell (CdSe/ZnS)
  - Unmodified QDs are neither water-soluble nor biocompatible, but they can be coated, for example, with silica or an amphiphilic substance, such as polyethylene glycol, for use in bioassays
  - The optical emission of a QD is determined by its band gap, which describes the energy difference between the excited state and resting state of an electron in the nanomaterial
    - Small nanocrystals generally display large band gaps. Thus, smaller QDs emit bluer (higher-energy) fluorescence, while larger QDs emit redder (lower-energy) fluorescence

Table 12.1 Types of nanomaterials with biomedical and diagnostic applications

Nanomaterial	Applications	Key features
Quantum dots	Detection of protein and nucleic acid biomarkers, immunofluorescence labeling of histologic specimens, single molecule imaging and tracking	Large quantum yield, broad excitation spectra, and narrow emission spectra Increase in sensitivity over conventional IHC Useful in quantitative multiplexed assays, including immunohistology
Gold nanoparticles	Detection of protein and nucleic acid biomarkers	Rapid assays Applicable for simple point-of-care devices Nontoxic
Surface-enhanced Raman scattering (SERS) nanoparticles	Detection of protein and nucleic acid biomarkers, in vivo imaging	Ultrasensitive and applicable for high-order multiplexed analysis Unique SERS spectral fingerprints based on reporting molecule adsorbed on nanoparticle Nontoxic
Dendrimers	Useful for diagnostic detection agents, imaging contrast agents, and therapeutic delivery	Near perfect monodispersity High degree of multivalency permits a large dose of detection agent or therapeutic modality to be linked to nanomaterial
Biobarcode assay	Detection of protein and nucleic acid biomarkers	Ultrasensitive and applicable for high-order multiplexed analysis Unique barcode signal is based on the specific oligonucleotide sequence that is conjugated to gold nanoparticle
Nanotubes and nanowires	Label-free detection of biomarkers	Highly sensitive multiplexed detection Easily integrated into microfluidic systems
Nanocantilevers	Detection of protein and nucleic acid biomarkers	Combined in larger arrays for high-order multiplexed analysis
Nanoelectromechanical ("lab-on-a-chip") systems	All-in-one devices for sample processing and biomarker detection	Potential for simple point-of-care testing; performed in healthcare organizations or in the field

  - Biocompatible QDs can be conjugated to antibodies, oligonucleotides, or aptamers, or coated with streptavidin, for specific binding to biomarker targets in molecular diagnostic assays (Fig. 12.1)
- Optical signal
  - QDs exhibit several optical properties that are advantageous for bioassays, compared with conventional organic fluorescent dyes (Fig. 12.1)
  - QDs emit strong fluorescence, with high quantum yields and molar extinction coefficients 10–50 times larger than organic dyes, making them much brighter in fluorescent assays
  - QD fluorescence intensity is highly stable in comparison to organic dyes, which can rapidly lose fluorescence through irreversible chemical changes. This photostability reduces the effect of photobleaching on assay performance and permits archiving of specimens
  - QDs exhibit size-tunable fluorescence emission color. QDs are manufactured under controlled conditions to specify precise sizes, resulting in different species with different fluorescent properties
  - Researchers have extended the emission wavelength into the near infrared (NIR) (650–950 nm) to take advantage of the improved tissue penetration and reduced background, which is of particular value for assays of whole blood and tissues, including histopathology

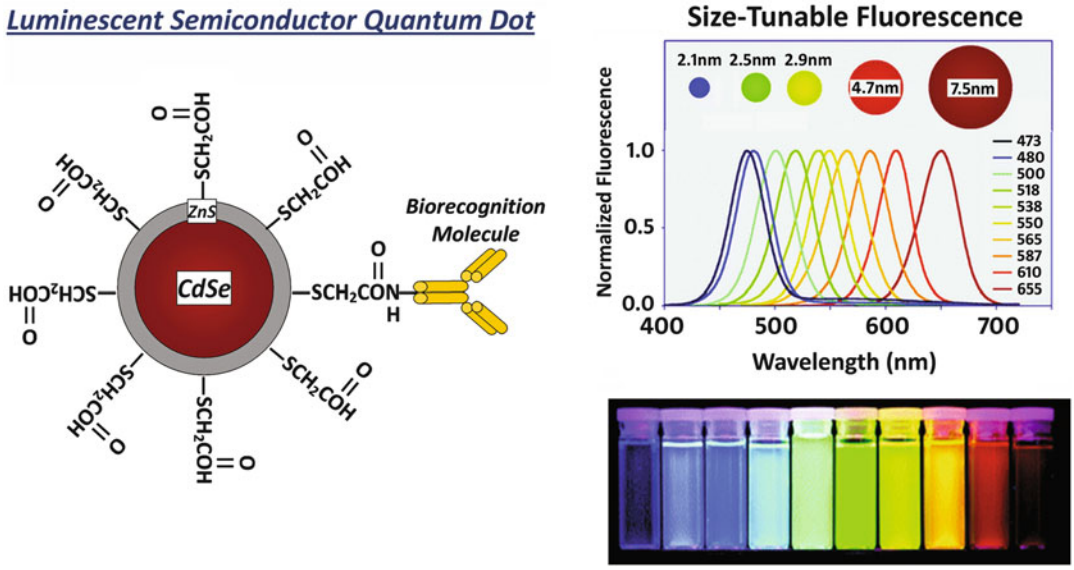


Fig. 12.1 Luminescent semiconductor quantum dot (QD). (*Left*) Structure of typical CdSe/ZnS QD. (*Top right*) Size-tunable fluorescent emission of QDs. Note narrow emission spectra. (*Bottom right*) Fluorescence from aqueous suspensions of QDs of increasing size, after excitation with a near-UV lamp (From S Nie laboratory)

  - QDs exhibit broad excitation and narrow emission spectra. Therefore, multiple QDs can be excited with a single wavelength, and their emissions can be resolved with relative ease. These properties make QDs well suited for multiplexed assays
  - QDs have been embedded in various combinations in larger structures, such as microbeads, to create optical barcodes. Theoretically, one million barcodes are possible by combining six resolvable QD emission colors and ten intensity levels for each color. Beads embedded with specific barcodes are conjugated to unique biorecognition molecules for very high-level multiplexed bioassays
- Biomedical applications
  - Applications include in vitro diagnostic assays such as immunofluorescence microscopy, flow cytometry, and multiplexed barcoding assays. Since QDs are typically composed of toxic heavy metals, the applicability to in vivo diagnostics is limited
  - Immunofluorescence microscopy of cells and formalin-fixed paraffin-embedded tissues has been pioneered by the group of Shuming Nie at Emory University, among others (Figs. 12.2 and 12.3)
    - The capacity for high-order multiplexed analysis using a single excitation source is valuable for protein expression profiling of cancer and other complex diseases
    - Signals from multiple QDs are detected with standard fluorescence microscopy and can be easily resolved and quantified with spectral imaging techniques
    - Effects of tissue autofluorescence are reduced by use of QDs that emit at red or near-infrared wavelengths
    - Data are more quantitative than standard immunofluorescence or immunohistochemistry
    - While recent advances in computerized image analysis and multispectral imaging have increased the potential

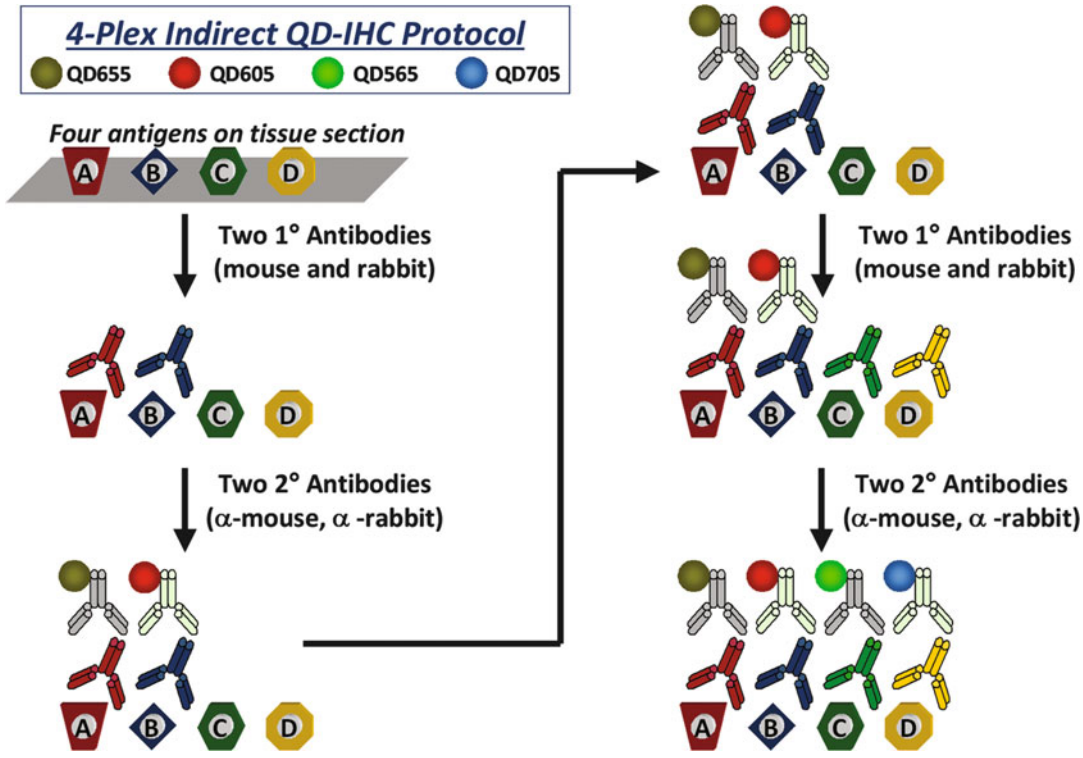


Fig. 12.2 Sequential QD staining. Two primary antibodies from two animal species are used to detect two tissue antigens. After washing, secondary antibody QD conjugates are applied. The specimen is washed and the procedure is repeated with primary antibodies for two other antigens followed by secondary antibody QD conjugates with distinct emission (Adapted from Liu et al. 2010b)

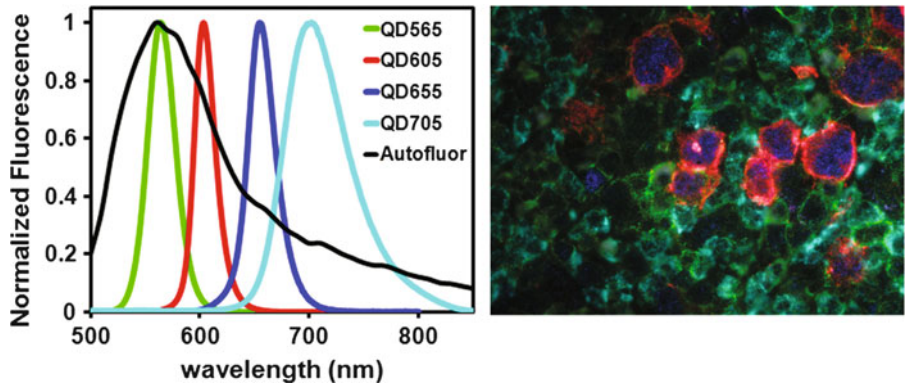


Fig. 12.3 Multiplexed QD immunofluorescence of cancer. (*Left*) Spectral imaging of immunofluorescence of formalin-fixed paraffin-embedded cancer tissue, using four primary antibodies. (*Right*) Integration of QD fluorescence with tissue morphology

of quantitative multiplexed analysis using bright-field immunohistochemistry, the capacity to develop systems for high-order multiplexed analysis with precise quantification is expected to be greater with QDs

– Table 12.2 shows a comparison of quantum dot immunofluorescence

Table 12.2 Comparison of routine immunohistochemistry with quantum dot immunofluorescence

Routine immunohistochemistry	Quantum dot immunofluorescence
One or few biomarkers per slide	Potential for many biomarkers per slide
Semiquantitative	Quantitative
Bright-field microscopy	Dark-field fluorescent microscopy

Table 12.3 Comparison of quantum dots with standard organic fluorophores

Characteristic	Quantum dots	Standard organic fluorophores
Excitation	Very broad. UV light can excite QDs of various sizes, permitting a single excitation source to be used in multiplexed studies	Narrow. Multiple excitation sources are needed in multiplexed studies
Emission band width	20–40 nm. Narrow emission spectra are advantageous for multiplexed studies	50–100 nm. Broad emission peaks tend to overlap and cause interference in multiplexed studies
Fluorescence lifetime	10–40 ns	Few nanoseconds
Photostability	Stable for at least 10 h	Photobleaching in less than 1 h, often in less than 1 min
Quantum yield	High (strong fluorescent signal per unit of light absorbed)	Orders of magnitude lower than quantum dots

with standard immunohistochemistry. In the majority of pathology departments, routine tests are designed to analyze one biomarker semiquantitatively per microscope slide
    - Multiplexed QD immunofluorescence has been used for expression profiling of prostate cancer and identifying rare Reed–Sternberg cells in lymph nodes (Liu et al. 2010a, b)
- QDs have been use as fluorescent probes for flow cytometry
    - Compared to standard organic fluorescent dyes, signals are stronger and more stable
    - Fewer excitation wavelengths can produce a larger number of resolvable fluorescent signals
        - Table 12.3 shows a comparison of QD fluorescence with the properties of standard organic fluorophores
- The Nie group achieved multicolor optical barcoding by embedding different-sized QDs into porous silica and polystyrene microbeads at precisely controlled ratios
    - The QD-tagged beads were microscopically and spectroscopically uniform and reproducible, yielding bead identification accuracies as high as 99.99%
    - When conjugated to nucleic acid probes, the beads were able to detect and distinguish DNA species in multiplexed assays (Gao and Nie 2005)

12.3.1.2 Gold Nanoparticles

- Structure
    - The synthesis of colloidal gold nanoparticles ranging from 3 to 100 nm in size is very reproducible. Particles are stable and their surfaces can be modified by a variety of chemical reactions to create specific binding or signaling properties
    - Gold nanospheres in the range of 20–120 nm appear bright red, due to efficient light scattering and surface plasmon resonance (SPR)
    - Gold nanorods have distinct properties of tissue penetration and light absorption or scattering, based on their aspect ratio (length of the major axis divided by width of the minor axis)
- Biomedical applications
    - Gold nanoparticles are used in many clinical bioassays based on colorimetry, fluorescence, and light scatter. Gold nanoparticles are nontoxic, in contrast to QDs, which could expand their potential for in vivo applications
    - Point-of-care lateral flow (immunochromatographic) tests represent an early and very

common diagnostic application for 20–120 nm gold nanoparticles
- In a standard lateral flow assay, the analyte is applied to a solid substrate (such as a nitrocellulose strip) and flows in a single direction via capillary action
- During lateral flow, the analyte is bound by a colored detection reagent (such as gold nanoparticle bound to antibody), and the complex migrates until encountering zone(s) on the substrate that are pretreated with a capture molecule (such as a capture antibody)
- Capture of the complex results in a line that can be detected visually or with an instrument
- Gold nanoparticles and dyed latex particles are the most commonly used detection reagents in lateral flow assays. Gold nanoparticles have several advantages over latex
  - They are smaller and have higher diffusion rates, allowing them to mix well with the analyte during lateral flow and readily penetrate pores in the solid substrate
  - They pack very densely at the capture line, resulting in greater visibility
  - Using appropriate instrumentation, optical signals can be quantified more precisely than by simple visualization
  - FDA-cleared point-of-care lateral flow tests based on gold nanoparticles are available commercially, such as rapid tests for pregnancy or specific IgE against common allergens

– Gold nanoparticles functionalized with nucleic acids have also been applied in commercial benchtop microarray platforms
  - Gold nanoparticles are amplified by chemical deposition of silver and detected and quantified by measuring evanescent wave-induced light scatter
  - This method was reported to be several orders of magnitude more sensitive than Cy3-based fluorescence microarray analysis
  - The gold nanoparticle microarray platform is utilized in FDA-cleared assays for respiratory viruses, inherited hypercoagulable states, and warfarin resistance (Buchan et al. 2011)

12.3.1.3 Surface-Enhanced Raman Scattering (SERS) Gold Nanoparticles

- Structure
  – SERS nanoparticles are optical detection tags consisting of Raman-active reporter molecules adsorbed on a gold nanoparticle core (40–50 nm), surrounded by a biocompatible shell (Fig. 12.4)
  – Absorption of reporter molecule on gold (or other noble metal) nanoparticle results in surface enhancement of Raman light scattering, in some cases by more than one billion times
  – The shell is usually comprised of silica or polyethylene glycol. It prevents aggregation, retains the reporter on the nanoparticle, and serves as a conjugation site for biorecognition molecules
- Optical signal
  – SERS nanoparticles offer the potential for extremely bright, photostable signals, with potential for high-level multiplexing. The theory of surface-enhanced Raman scattering is described below
  – Raman scattering is inelastic scattering of photons from a molecule, caused by interaction of incident light with molecular vibrations
  – The vibrational modes available to a molecule are influenced by its unique chemical bond structure. Thus, a molecule is characterized by a unique Raman spectral signature, comprised of multiple narrow spectral peaks (Fig. 12.4)
    - The vast majority of light scattering is elastic (Rayleigh scattering), with scattered photons characterized by the same frequency and wavelength as incident photons
    - Rare photons undergo inelastic Raman scattering, when incident light interacts

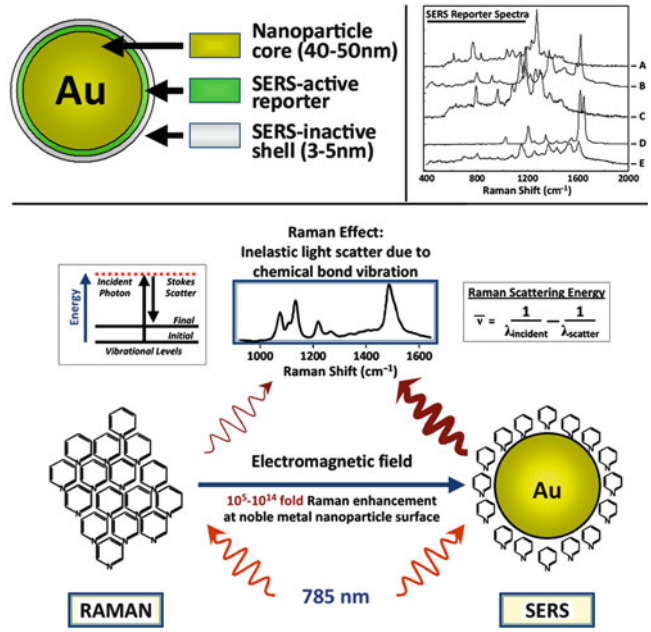


Fig. 12.4 Surface-enhanced Raman scattering gold nanoparticles. (*Top left*) Encapsulated SERS nanoparticle. (*Top right*) Raman spectra of reporter molecules. Each spectral signature (*A*–*E*) contains multiple narrow peaks with reproducible relative peak size. Total area under peaks is proportional to amount of reporter. Each spectrum has one or more unique peaks. Deconvolution of complex spectral data is capable of resolving signals in quantitative multiplexed assays. (*Bottom*) Schematic diagram of SERS phenomenon

with chemical bonds to excite molecules from ground state to a virtual vibrational energy state
    - Raman-scattered photons are characterized by shifted frequency and wavelength compared to incident light, usually of lower energy (Fig. 12.4)
    - Since unenhanced Raman scattering is a relatively rare event, signals are weak and difficult to resolve from the dominant Rayleigh scattering
  - Raman scattering is enhanced by many orders of magnitude when molecules are adsorbed on a rough noble metal surface, such as a gold nanoparticle
    - Raman enhancement is mainly due to surface plasmon resonance
    - In contrast to unenhanced Raman spectra, SERS spectra are sufficiently intense to serve as detection signals for biomedical assays
- SERS nanoparticles have many optical properties that are useful for bioassays
  - Signals are intense and stable, with a high signal-to-noise ratio, providing high analytical sensitivity
    - Surface enhancement results in optical emission that is stronger and more resistant to photobleaching than organic fluorophores or even QDs (Qian et al. 2008)
    - Excitation can be achieved with red or near-infrared wavelengths, which minimizes background autofluorescence from biological materials
  - Raman spectra are quantitative and specific to the unique chemical bonds of the reporter molecule, resulting in high specificity and capacity for multiplexing
    - Raman spectra cover vibrational energies from 300 to 5,000 cm^{-1} and are characterized by narrow line widths that can be easily resolved in multiplexed signals (Fig. 12.4)
  - SERS nanoparticle shells are readily conjugated to a variety of biorecognition molecules such as antibodies, aptamers, ligands, and nucleic acids
    - Multiplexed assays are developed using a panel of SERS nanoparticles, conjugated to distinct biorecognition molecules, and distinguished by specific Raman reporter molecules

- Clinical applications
  - SERS nanoparticles have been used for multiplexed immunoassays, tumor targeting, and detection of circulating tumor cells
    - Natan et al. developed a multiplexed lateral flow assay for respiratory viruses using silica-encapsulated SERS nanoparticles. Assay signals were resolved with a Raman spectrometer applicable for benchtop or point-of-care testing
      - SERS nanoparticles improved virus detection sensitivity, quantification, and assay reproducibility compared to standard lateral flow tests
    - The same group devised a homogeneous SERS-based cell detection assay for rapid quantification of circulating tumor cells in whole blood (Sha et al. 2008)
      - Magnetic beads and silica-encapsulated SERS nanoparticles were conjugated to antibodies against distinct cancer biomarkers, for cell capture and detection in homogenous assays
      - Near-infrared excitation permitted use of the SERS nanoparticles in whole blood
      - The reaction was quantitative, with a limit of detection of 50 tumor cells/mL
    - The Nie group has described a polyethylene glycol-encapsulated SERS nanoparticle assay for circulating tumor cells (Wang et al. 2011)
      - SERS nanoparticles conjugated to epidermal growth factor (EGF) were incubated with circulating tumor cells and admixed leukocytes, isolated from whole blood by density-gradient centrifugation
      - Cells were washed, pelleted, and assayed by infrared light excitation (785 nm)
        - Raman signals were analyzed with a handheld Raman spectrometer
      - The reported limit of detection was 5–50 circulating tumor cells/mL of whole blood
      - In a clinical trial, circulating tumor cells were identified in 17 of 19 patients with EGF receptor-positive head and neck squamous cell carcinoma
    - The Nie group targeted EGF receptor-positive tumor cells in vitro and in vivo, using polyethylene glycol-encapsulated SERS nanoparticles conjugated to anti-EGF receptor antibody (Fig. 12.5)

12.3.1.4 Dendrimers

- Structure
  - Dendrimers are branched, polymeric nanomaterials with symmetric and ordered architecture. They are synthesized in a highly controlled, sequential process to produce nanoparticles with near perfect monodispersity
  - Dendrimers can be grown using a divergent process (where all of the branches grow exponentially from the core) or in a convergent process (where individual branches, or dendrons, are grown separately and then attached to the core during a final step)
  - Dendrimer size is largely controlled by its generation, which refers to the number of branch points moving radially from the core to the periphery (G1, G2, G3, etc.)
  - A wide variety of materials can be used in dendrimer synthesis, providing versatile material properties which can be tuned for the intended application
  - One important property of dendrimers is multivalency, resulting from the highly branched architecture. Multivalency results in a number of points for the attachment of detection agents, imaging agents, and targeting molecules or drugs, for the production of multifunctional biomedical tools (Fig. 12.6)
- Biomedical applications
  - Dendrimers have been used for drug and gene delivery and can be modified with

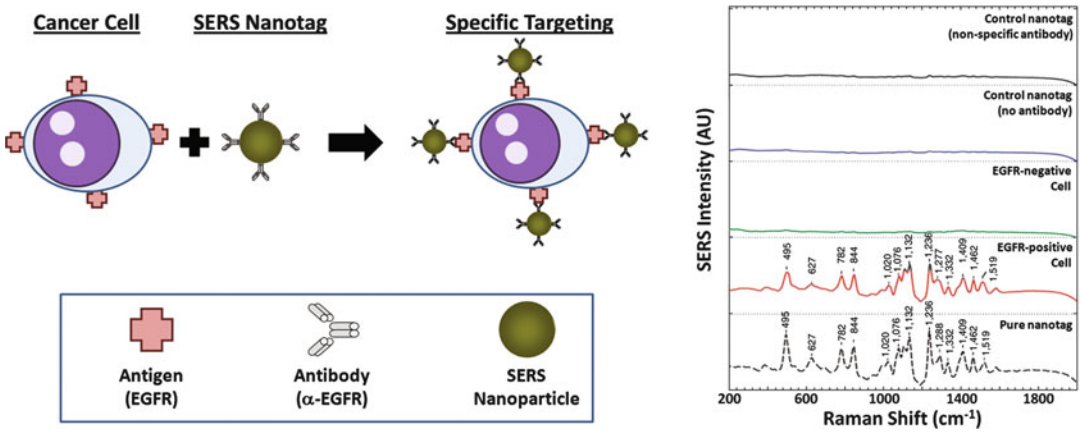


Fig. 12.5 In vitro tumor cell targeting with SERS nanoparticles. The targeted SERS agent consisted of a gold nanoparticle core, Raman reporter, and polyethylene glycol shell, conjugated to anti-EGF receptor (EGFR) antibody. Experimental tumor cells (EGFR-positive) were treated with the targeted SERS agent, washed, and analyzed by Raman spectroscopy. The SERS spectrum from treated cells was identical to the pure SERS nanoparticle (compare *red* and *black* spectra on *right*). No SERS signal was obtained from negative controls, including experimental cells treated with control SERS agent (no conjugated antibody) and control cells (EGFR-negative) treated with targeted SERS agent (*black*, *blue*, and *green* spectra on *right*). Targeting was also achieved when tumor cells were implanted in mice and SERS agents were injected in the tail vein (not shown) (From Qian et al. 2008, with permission)

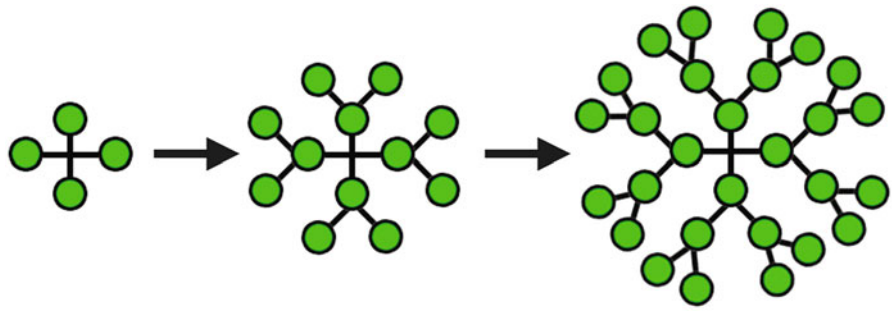


Fig. 12.6 Dendrimer. Dendrimers are branched, polymeric nanomaterials with symmetric and ordered architecture. Dendrimers are characterized by a high degree of multivalency, which provides many attachment points for detection agents or imaging contrast agents, for use in diagnostic applications

imaging moieties to serve as nanoparticle contrast agents for bioimaging

– Wiener et al. developed a new class of MRI contrast agent using dendrimers as a carrier for gadolinium
  - PAMAM dendrimers ranging from 8.5 to 140 kDa in size were produced and used to chelate gadolinium ions in a nanocomplex
  - The nanoparticle contrast agents showed up to six-fold enhancement of the relaxivity per Gd(III) ion in comparison to standard chelates and had significantly increased half-lives in vivo, making them ideal for in vivo imaging applications
– Kukowsla–Latallo et al. have synthesized a dendrimer-based carrier for the targeted delivery of therapeutics for cancer treatment
  - A G5 PAMAM dendrimer was synthesized to produce a nanoparticle <5 nm

in diameter with approximately 100 functional groups on the surface
- Folic acid and methotrexate were conjugated to the nanoparticle for targeting functionality and therapeutic effect, respectively
- Targeted nanocarriers showed increased antitumor activity and reduced toxicity in comparison to free drug

12.3.1.5 Biobarcodes

- The biobarcode assay was developed by the Chad Mirkin group at Northwestern University. It is an ultrasensitive, enzyme-free method to detect biomarkers including proteins and nucleic acids
  - The analytical sensitivity for nucleic acid detection is comparable to polymerase chain reaction, and for protein detection exceeds that of ELISA by several orders of magnitude
- Assay system
  - In a biobarcode assay, the biomarker is detected using a sandwich method
  - Biomarker is captured by magnetic microparticles conjugated to biorecognition molecules
  - Biomarker is sandwiched by gold nanoparticles conjugated to detection biorecognition molecules, as well as a high density of "barcode" DNA sequences, which are designed to specifically represent the biomarker of interest and amplify the detection signal (Fig. 12.7)
    - For nucleic acid detection, the biorecognition molecules can be complementary nucleic acid probes (unrelated to the barcode)
    - For protein detection, the biorecognition molecules can be antibodies, ligands, or aptamers
    - Biomarkers sandwiched by the magnetic microparticle and barcoded nanoparticle are separated from solution in a magnetic field
    - Barcode DNA is dehybridized from the gold nanoparticles and measured by conventional methods such as scanned microarrays or quantitative PCR
    - The presence of numerous DNA barcodes per sandwich structure amplifies the detection signal massively, resulting in extremely high detection sensitivity
    - Assays for many biomarkers can be combined for high-level multiplexed profiling, by utilizing unique biobarcodes for each biomarker
- Biomedical applications
  - The biobarcode system has been modified for simple desktop and point-of-care assays, as well as complex multiplexed analysis
  - The Mirkin group developed a biobarcode assay for PSA with a reported analytical sensitivity of 330 fg/ml, several orders of magnitude more sensitive than commercial immunoassays
    - In a clinical trial of patients following radical prostatectomy, the biobarcode assay detected minute residual PSA in all patients and identified patients with rising PSA at an earlier point than the commercial immunoassays (Thaxton et al. 2009)
  - Biobarcode assays for nucleic acids have demonstrated comparable sensitivity to PCR without enzymatic amplification of target sequence

12.3.2 Nanowires and Nanocantilevers

- Nanowires and nanocantilevers have been used as label-free signal transduction systems for biomarker detection

12.3.2.1 Nanowires

- Structure
  - Nanowires are materials with diameter in the nanoscale range and aspect ratios (length to width) of 1,000 or more
  - Nanowires have been fabricated from silicon, carbon nanotubes, conducting polymers, and other materials

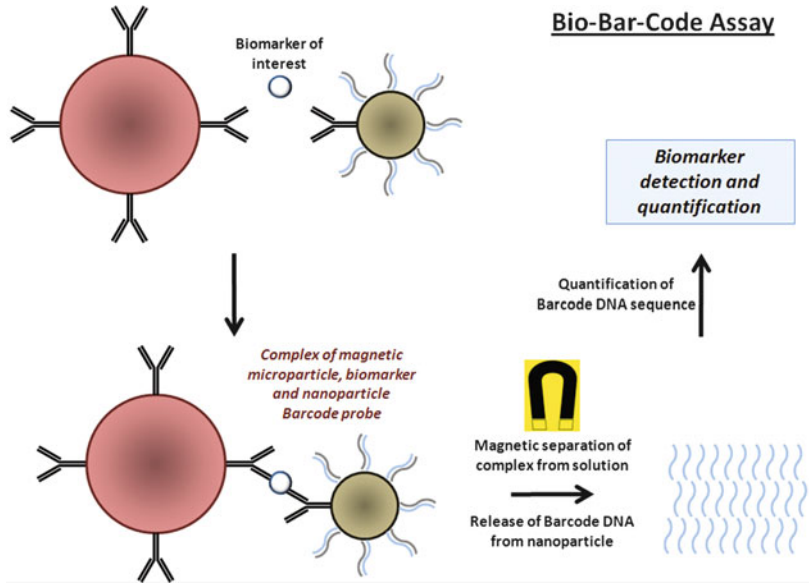


Fig. 12.7 Biobarcode assay. The biobarcode assay is useful for ultrasensitive, nonamplified detection of nucleic acid or protein targets. Two types of particle are used to detect biomarker. The first is a magnetic microparticle with recognition elements for biomarker (e.g., antibody). The second is a gold nanoparticle with a second recognition agent, which forms a sandwich around the biomarker target; this gold nanoparticle is also conjugated to a large number of oligonucleotide barcodes used for enhanced signal. After reaction with biomaterial containing the biomarker target, a magnetic field is used to isolate the sandwich structures, and a reducing agent is used to release the barcode strands. The barcode strands are identified and quantified using standard microarray or scanometric methods

  - They are functionalized for bioassays by conjugation to biorecognition molecules
  - Nanowires exhibit electrical properties that are not seen in bulk materials because electrons undergo quantum confinement in the lateral dimension
  - Silicon nanowire biosensors undergo a change in electrical conductance when bound by target biomarker because the binding event causes a change in chemical potential, producing a field-effect gate upon the nanowire
  - Detection sensitivity of nanowire biosensors depends on solution ionic strength. Desalting is required for samples with a high ionic strength to optimize sensitivity
- Biomedical applications
  - A high density of functionalized nanowires can be combined in array formats for multiplexed analysis
  - Nanowire assays have been used for multiplexed detection of cancer biomarkers at femtomolar concentrations in undiluted serum

12.3.2.2 Nanocantilevers

- Structure
  - Cantilevers function as mechanical sensors, which are tethered at one end and free at the other
  - Devices used for biomedical diagnostics typically consist of an array of cantilevers in the nanoscale, each conjugated to a specific biorecognition molecule (Fig. 12.8)
  - Nanocantilevers are deformed or deflected when bound by their target biomarker
  - This movement can be detected optically or electronically and transformed into electrical current for data analysis and biomarker measurement

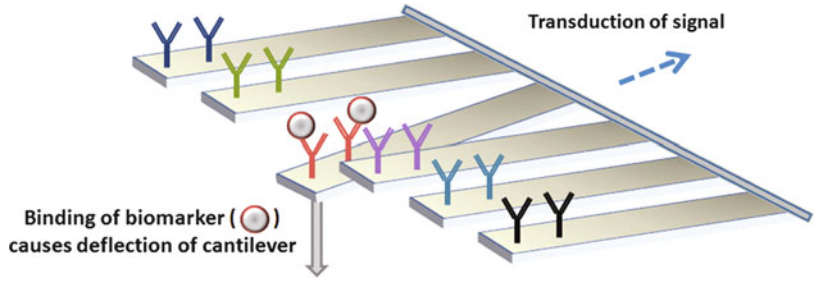


Fig. 12.8 Nanocantilever device. The cantilevers are flexible nanoscale beams built using semiconductor lithographic techniques. A large number of nanocantilevers are constructed as part of a larger diagnostic device, which can be used for high-order multiplexed detection of protein, DNA, RNA, and other biomarkers

- Biomedical applications
  - A high density of nanowire or nanocantilever biosensors can be combined in electronically addressed arrays
  - In theory, large-scale circuits can be constructed in microfluidic environments, enabling the rapid measurement of large number of biomolecules from a minute sample

12.3.3 Nanofluidic Materials

- Devices with nanoscale channels can be designed to process fluid samples for complex diagnostic analyses
- Nanofluidic devices are generally fabricated using top-down approaches, similar to those used for the machining of microelectromechanical systems (MEMS)
  - Nanofluidic devices are often referred to as nanoelectromechanical systems (NEMS). Like MEMS, they are characterized by high accuracy and precision
  - Many examples of NEMS integrate nanoelectronic transistors with mechanical pumps or motors
  - For biomedical purposes, NEMS can serve as biological and chemical sensors
  - Total analysis systems (TAS) at the micro- or nanoscale (sometimes referred to as "lab-on-a-chip" devices) are capable of performing many analytical functions in small, self-contained units applicable to point-of-care testing of minute specimens
    - Functions include sample introduction, sample processing (such as cell lysis or dilution), analyte separation (such as electrophoresis or chromatography), and analyte detection
- Compared with microfluidic devices, nanofluidic technologies permit novel analytical approaches by interacting with fluids at the molecular scale, with high surface areas for chemical processes
- By controlling size and surface chemistry, nanochannel devices can be designed to perform diagnostically useful processes, with minimal requirements for sample volume (in the picoliter range)
  - By fabricating nanochannels that are sensitive to the structure or sequence of biomolecules, the devices can designed for the following functions in diagnostic assays
    - Regulated transport of analytes
    - Highly selective filtering and enrichment of specific biomolecules
    - Directed chemical reactions between diffusing reagents
    - Detection of specific nucleic acid sequences
- Nanoporous materials can be used similarly for highly selective fractionation and binding of diagnostically important molecules from complex specimens
  - Glass and silicon nanopores have been used to selectively fractionate complex protein mixtures from biological fluids such as serum

    - In some studies, fractionated proteins were analyzed by proteomic techniques for biomarker discovery
  - Sequence-specific nanopores linked to electrical sensors have been applied for sensitive and specific detection of DNA species of interest

12.4 Other Clinical Applications of Nanotechnology

12.4.1 Molecular Imaging

- Multifunctional nanostructures for biomedical imaging combine contrast agent with functionality that targets the agent to a desired anatomical location. These agents have potential for emerging in vivo diagnostic assays
- Gold nanoparticles coated with polyethylene glycol have been described as radiographic contrast agents with high signal and long circulation time
- Paramagnetic iron oxide nanoparticles have been proposed as MRI contrast agents with uniquely high contrast at low concentrations
- Other nanoscale materials have been used as carriers for a variety of imaging agents, including buckyballs (hollow spherical carbon molecules, or fullerenes); modified recombinant adenoviruses; dendrimers (spherical, monodisperse, repetitively branched molecules with high degree of symmetry); and liposomes (spherical vesicles composed of a lipid bilayer membrane, usually containing an aqueous core)
- Nanoscale imaging reagents have several favorable features
  - High signal with low background
  - Blood circulation times that are longer than current contrast agents
  - Generally low toxicity
  - Functionalization (for example, with biorecognition molecules) that targets agent to specific tissue
  - An important limitation of nanoscale imaging agents is accumulation in the reticuloendothelial system

12.4.2 Drug Delivery and Therapy

- Similar to imaging agents, nanoscale materials have been utilized to deliver therapeutic agents to diseased tissues
- Nanoparticle albumin-bound (nab™) technology employs drugs that are encapsulated in negatively charged, monodisperse albumin nanoparticles (50–150 nm)
  - The nanoparticle drugs pass through leaky capillaries in diseased tissue and may be delivered specifically by receptor-mediated transcytosis and binding to secreted protein acidic rich in cysteine (SPARC) on the surface of target cells
  - Nanoparticle albumin-bound paclitaxel is the first agent of this class and is FDA-approved for treatment of breast cancer
- Liposomes can be designed for temperature- or pH-sensitivity by formulation with lipids of different fatty acid chain lengths, to permit controlled release of their contents
- Chemotherapy drugs loaded or attached nanoparticle serve as therapeutic agents
  - The collaborative groups of Shuming Nie and Dong Shin have developed ternary complexes comprised of a self-assembling heparin-based nanoparticle, targeting ligand (such as folate or EGF), and a chemotherapeutic or imaging agent, for treatment of head and neck and other cancers

12.5 Summary of Key Points for Nanotechnology in Molecular Diagnostics

- Nanotechnology is the understanding and control of matter at the nanoscale, at dimensions between approximately 1 and 100 nm, where unique phenomena enable novel applications
  - Compared to bulk materials and single atoms or molecules, nanoscale materials display unique optical or magnetic properties, electrical conductivity, chemical reactivity, and other features

  - Nanomaterials interact differently with biomolecules in the same size range, such as proteins (1–20 nm) and nucleic acids (2.5 nm in diameter)
  - Important examples of nanotechnology include nanoparticles, nanowires, nanocantilevers, nanopores, and nanofluidic devices
- The aim of nanodiagnostics to develop assays with novel performance, based on the unique properties of nanomaterials. Nanotechnology has potential to improve diagnostic sensitivity and specificity, analytical sensitivity, speed, multiplexing, affordability, portability, and ease of use
- Based on the novel features of nanomaterials, nanotechnology is expected to profoundly affect laboratory medicine and molecular diagnostics, playing an important role in early disease detection, genomic technologies, in vivo diagnosis, and personalized, predictive medicine

Further Reading

Azzazy HM, Mansour MM, Kazmierczak SC. Nanodiagnostics: a new frontier for clinical laboratory medicine. Clin Chem. 2006;52:1238–46.

Azzazy HM, Mansour MM, Kazmierczak SC. From diagnostics to therapy: prospects of quantum dots. Clin Biochem. 2007;40:917–27.

Buchan BW, Peterson JF, Cogbill CH, et al. Evaluation of a microarray-based genotyping assay for the rapid detection of cytochrome P450 2C19 *2 and *3 polymorphisms from whole blood using nanoparticle probes. Am J Clin Pathol. 2011;136:604–8.

Cheng MM, Cuda G, Bunimovich YL, et al. Nanotechnologies for biomolecular detection and medical diagnostics. Curr Opin Chem Biol. 2006;10:11–9.

Gao X, Nie S. Quantum dot-encoded beads. Methods Mol Biol. 2005;303:61–71.

Hu Y, Fine DH, Tasciotti E, et al. Nanodevices in diagnostics. Wiley Interdiscip Rev Nanomed Nanobiotechnol. 2011;3:11–32.

Jain KK. Applications of nanobiotechnology in clinical diagnostics. Clin Chem. 2007;53:2002–9.

Kukowska–Latallo JF, Candido KA, Cao Z, et al. Nanoparticle targeting of anticancer drug improves therapeutic response in animal model of human epithelial cancer. Cancer Res. 2005;65:5317–24.

Liu J, Lau SK, Varma VA, et al. Multiplexed detection and characterization of rare tumor cells in Hodgkin's lymphoma with multicolor quantum dots. Anal Chem. 2010a;82:6237–43.

Liu J, Lau SK, Varma VA, et al. Molecular mapping of tumor heterogeneity on clinical tissue specimens with multiplexed quantum dots. ACS Nano. 2010b;4:2755–65.

Nam JM, Stoeva SI, Mirkin CA. Bio-bar-code-based DNA detection with PCR-like sensitivity. J Am Chem Soc. 2004;126:5932–3.

Qian X, Peng XH, Ansari DO, et al. In vivo tumor targeting and spectroscopic detection with surface-enhanced Raman nanoparticle tags. Nat Biotechnol. 2008;26:83–90.

Sha MY, Xu H, Natan MJ, Cromer R. Surface-enhanced Raman scattering tags for rapid and homogeneous detection of circulating tumor cells in the presence of human whole blood. J Am Chem Soc. 2008;130:17214–5.

Thaxton CS, Nam JM, Mirkin CA. PCR-like sensitivity for proteins with bio-bar-code amplification. Discov Med. 2003;3:58–60.

Thaxton CS, Elghanian R, Thomas AD, et al. Nanoparticle-based bio-barcode assay redefines "undetectable" PSA and biochemical recurrence after radical prostatectomy. Proc Natl Acad Sci USA. 2009;106:18437–42.

Wang X, Qian X, Beitler JJ, et al. Detection of circulating tumor cells in human peripheral blood using surface-enhanced Raman scattering nanoparticles. Cancer Res. 2011;71:1526–32.

Wiener EC, Brechbiel MW, Brothers H, et al. Dendrimer-based metal chelates: a new class of magnetic resonance imaging contrast agents. Magn Reson Med. 1994;1:1–8.

Xing Y, Chaudry Q, Shen C, et al. Bioconjugated quantum dots for multiplexed and quantitative immunohistochemistry. Nat Protoc. 2007;2:1152–65.

Gene Therapy: Vector Technology and Clinical Applications

13

Sunyong Tang and Kenneth Cornetta

Contents

13.1 Introduction 399

13.2 Ethical Issues 400

13.3 Gene Transfer Techniques 400
13.3.1 Plasmid Vectors 400
13.3.2 Retrovirus Vectors and Lentiviral Vectors 401
13.3.3 Adenovirus Vectors 402
13.3.4 Adeno-associated Virus Vectors 403
13.3.5 Herpes Virus Vectors 404
13.3.6 Other Vector Systems 404

13.4 Gene Transfer Applications 404
13.4.1 Compensation for Genetic Mutations 404
13.4.2 Gene Therapy for Pharmacologic Effect 406
13.4.3 Inactivation of Harmful Genetic Sequences 406
13.4.4 Cell Engineering 408
13.4.5 Cancer Immunotherapy 409
13.4.6 Replication-Competent Viruses 410

13.5 Safety Principles and Regulatory Issues 411
13.5.1 Previous Adverse Events 411
13.5.2 General Safety Considerations 411
13.5.3 Regulatory Issues 412

Further Reading 412

S. Tang, PhD
Indiana University Vector Production Facility, Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

K. Cornetta, MD (✉)
Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

13.1 Introduction

- Gene therapy can be defined as the transfer of genetic material with therapeutic intent. Gene transfer has become clinically feasible as our understanding of the molecular basis of disease has matured, and we have developed improved techniques for manipulating genetic material
- Since the first human in vivo gene transfer study in 1989, a variety of clinical trials involving gene transfer have been initiated. A number of successes have been shown in inherited genetic diseases. Ongoing work is looking to develop gene therapy approaches to almost every disease, including cancer, heart disease, infectious disease, and neurologic disorders
- The term gene therapy vector refers to a system designed to transfer exogenous genetic material (the transgene) into a target cell. The simplest vectors are composed of naked DNA, usually in the form of plasmid DNA. Plasmid vectors are limited by low gene transfer efficiency and are generally not well suited to systemic administration. To address these limitations, several viruses have been engineered to transport genetic material. Viral vector systems and their characteristics are presented in Table 13.1
- The critical determinants for choosing a particular vector system include
  - Host range and tissue specificity of the vector

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_13,

Table 13.1 Characteristics of viral vectors

	Murine retrovirus	Adenovirus	AAV	Herpes virus	Human lentivirus
Genome	RNA	ds DNA	ss DNA	ds DNA	RNA
Transgene size (kb)	3–7	7–36	2–45	10–100	8–9
Titer (infectious units/mL)	10^6–10^8	10^{11}–10^{12}	10^6–10^9	10^4–10^{10}	10^6–10^9
Host cell proliferation	Required	Not required	Improves efficiency	Not required	Improves efficiency
Stable integration	Yes	No	Occasional	No	Yes
Clinical trials	Yes	Yes	Yes	No	Yes

ss single-stranded, *ds* double-stranded

  - Efficient transfer into the target cell
  - Desire to maintain vector in an integrated versus episomal form
  - Immunogenicity
  - The amount of exogenous DNA that must be transferred

13.2 Ethical Issues

- Gene therapy clinical trials have focused on the treatment of individuals afflicted with disease. Specifically, the target cells for gene therapy are somatic cells. Current review boards have voluntarily placed a moratorium on gene transfer procedures that introduce material into germline tissues due to ethical and scientific concerns of altering the gene pool
- The ethical issues surrounding genetic engineering were debated long before gene transfer became technically feasible
  - Transfer of genetic material for the sole purpose of treating serious disease has not generally provoked major ethical objections from the religious, political, or scientific communities. Ethicists have agreed that preclinical data must be available so that a risk to benefit ratio can be formulated and patients must provide informed consent
  - Enhancement engineering is gene transfer aimed at improving a specific phenotype in a healthy person. To date, this type of genetic modification has not been undertaken and remains an area of considerable ethical debate. For example, gene therapy that delays the onset of atherosclerosis in asymptomatic persons may be considered acceptable. The concern is that this technology will be used for performance enhancement of specific traits, such as stature
  - Eugenics, whereby complex human traits (such as personality and intelligence) are manipulated, is essentially a subject for philosophical debate until we better understand the genetics of these traits. Nevertheless, the possibility of eugenics is of concern to society and has led to regulatory oversight of gene therapy that is unique in medicine
- One of the most active areas of ethical debate in gene therapy is the use of gene transfer in fetuses. Vigorous debate on this issue continues
- A future challenge will be defining what enhancement gene therapy is. For example, some might argue that a therapy to treat erectile dysfunction should be considered enhancement therapy and should not be offered. Others would argue that a significant number of men forego potentially life-saving surgery for prostate cancer when they learn it will result in impotence. The line between enhancement and quality of life will be an area of ongoing debate

13.3 Gene Transfer Techniques

13.3.1 Plasmid Vectors

- General
  - Plasmids are derived from circular bacterial genetic elements and can be reengineered so they express a gene, or genes, of interest in eukaryotic cells

  - Several chemical and physical techniques have been developed to enhance transfer of plasmid DNA into target cells
    - Electroporation, which uses electrical current to facilitate DNA uptake
    - Precipitation of plasmid DNA in calcium phosphate, a method known as transfection
    - Approaches to enhancing plasmid gene transfer using anionic or cationic liposomes
    - Particle bombardment of plasmid-coated gold particles into cells ("gene gun" approach)
- Advantages
  - The advantages of plasmid vectors are their ability to carry large amounts of genetic sequences, they are relatively inexpensive to produce, possess a favorable safety profile, and are unlikely to integrate into germ line cells
- Disadvantages
  - The disadvantages of plasmids are low efficiency of gene transfer, rapid degradation by serum, and generally do not integrate into target cells

13.3.2 Retrovirus Vectors and Lentiviral Vectors

- General
  - Murine gammaretrovirus-based retroviral vectors were the first viral vectors developed and the first viral vector to enter clinical trials. More recently, retroviruses such as HIV-1 (subsequently referred to as "lentiviral vectors") are being used as gene therapy vehicles. These viruses are attractive gene delivery vehicles because of their unique life cycle
  - The initial retroviral vectors were based on the murine leukemia viruses, which are membrane-bound RNA viruses. The viral genes of these viruses are relatively simple and include the gag region that encodes structural proteins involved with capsid formation. The pol region encodes proteins with enzymatic functions including reverse transcriptase and integrase. The viruses also contain an env gene that encodes a membrane-associated glycoprotein that targets the particles to specific cell receptors
  - Lentiviruses also contain gag and pol regions but are generally more complex, containing a variety of accessory proteins and regulatory regions important in their lifecycle and virulence. Fortunately, the accessory regions can be deleted which simplifies vector engineering and also provides a safety feature to vector design
  - Gammaretroviral and lentiviral vectors are generated by deleting the viral protein coding sequences (gag, pol, and env), allowing the introduction of exogenous genes of interest within the viral backbone (Fig. 13.1). Deletion of the viral protein coding sequences renders the vector replication-defective
  - Vector constructs generally retain the long terminal repeats (LTR) as these sequences are required for vector integration. The LTRs also contain promoter and enhancer functions, which drive vector expression. A packaging (ψ) sequence is also needed to facilitate efficient uptake of vector RNA

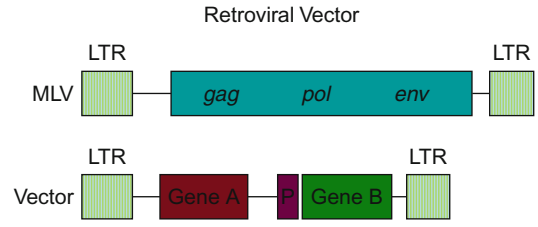


Fig. 13.1 Schematic of a murine leukemia virus (*MLV*) genome and a retroviral vector *LTR*: long terminal repeats which contain the promoter and enhancer regions and sequences required for integration into target cell DNA. The viral genes *gag*, *pol*, and *env* encode capsid protein, enzymatic proteins such as reverse transcriptase and integrase, and the envelope glycoprotein, respectively. Vectors are generated by removal of the viral genes and substitution of one or more genes of interest. In this case, gene *B* is run off an internal promoter (*P*), while gene *A* responds to the LTR promoter. Internal ribosome entry signals (IRES) may be used in place of the internal promoter

into the virion. The pol and env regions are deleted and only a small portion of the gag region is retained to maintain high vector titer
  - Infection with the wild-type retrovirus is initiated by binding of viral envelope glycoprotein (env) to a specific receptor(s) on the target cell. For example, the major receptor for the native HIV-1 envelope is the CD4 antigen, an antigen limited to only a few cell types (e.g., human T cells and macrophages). Pseudotype vectors have been engineered to express the envelope protein of a different virus on the vector particle, thus changing their host cell specificity. In the case of lentiviral vectors, the native HIV-1 env is often replaced with the vesicular stomatitis virus G envelope (VSV-G). This "pseudotyped" particle now allows the lentiviral vector to infect a wide range of cell types and species
  - As vectors are replication-defective, the vectors are produced by one of two methods
    - The simplest and quickest method is utilizing the transient transfection method in which three plasmids, encoding the vector sequence, the viral gag and pol regions, and the envelope gene are introduced into cell lines such as HEK293 or HT1080 cells. These cell lines have high transfection efficiency and are capable of producing vector at high titer, but only for 2–3 days
    - Packaging cell lines have also been generated which stably express *gag*/*pol* and an envelope plasmid. A plasmid containing the vector sequence can be introduced and clones of cells stably expressing vector can be obtained
- Advantages
  - Gammaretroviruses and lentiviruses possess many characteristics suitable for vector development. They efficiently integrate their genome into the target cell chromosome, thus passing the transgene to the transduced cell and all progeny of the cell. Infection and integration are not associated with significant cell death, chromosomal disruption, or other deleterious effects that may negatively impact the viability of the target cell. Lentiviruses have sequences that allow the double-stranded vector genome to cross the nuclear membrane allowing the vector to integrate into the genome of nondividing cells. As a result, lentiviral vectors show significantly higher efficiency in hematopoietic stem cells, as well as terminally differentiated cells. In addition to hematopoietic cells, lentiviral vectors appear to be more efficiently taken up by brain, liver, and eye tissues than the traditional gammaretroviral vectors
- Disadvantages
  - A major challenge for gammaretroviral vectors is their poor transduction of quiescent cells such as stem cells. A major safety concern for vectors derived from the murine leukemia virus (MLV) is insertional mutagenesis. While this adverse event is uncommon, it has been reported in at least three trials of MLV-based vectors. In these cases, enhancers within the LTR upregulate oncogenes near the vector insertion site that contribute to a multistep process resulting in leukemia or myelodysplasia. Lentiviral vectors have deletions in the enhancer region and to date, malignancy has not been reported but the number of patients treated is small. Vector design to decrease this adverse event remains an active area of research

13.3.3 Adenovirus Vectors

- General
  - Adenoviruses have a double-stranded DNA genome surrounded by a protein capsid, but are devoid of lipoprotein envelope. At least 47 serotypes that can infect humans and virtually all adults are seropositive for multiple serotypes
  - Adenovirus infection is initiated by binding of fiber protein to receptors on the host cell, followed by endocytosis of the virion.

Virions are released into the cytoplasm, and the viral genome is then transported to the nucleus where it is maintained episomally. Integration into the host cell genome does not usually occur
  - Most adenoviral vectors have been derived from human serotypes 2 and 5. First-generation vectors were rendered replication-incompetent by deletion of sequences including the E1A gene, which encodes a transcriptional activator required for replication of the viral genome. Production of replication-incompetent virions requires helper or packaging cells. The human embryonic kidney cell line 293 has been stably transfected with the E1A gene and constitutively expresses the E1A transcription factor. Thus, 293 can complement E1A-deleted, replication-incompetent adenoviral vectors
  - New-generation vectors including "gutless" adenoviral vectors that aim to limit the expression of immunogenic proteins are being evaluated. Modification in the viral capsid, and tissue-specific promoters, is also being developed to provide cell-specific infection with the aim of decreasing toxicity
- Advantages
  - Adenoviral vectors are relatively easy to purify and can be concentrated to high titers. They can infect both dividing and nondividing cells very efficiently. They do not integrate into the host cell genome and entail virtually no risk of insertional mutagenesis
- Disadvantages
  - Since adenoviral vectors do not integrate, a major limitation is the duration of transgene expression which can be brief in actively dividing cells due to dilution of the episomal form. Also, adenoviral gene products are highly immunogenic and can stimulate destruction of transgene-expressing cells by the host immune system. Intravenous administration of adenoviral vectors at high titers can lead to cytokine-mediated multiorgan failure, which can be fatal. This reaction has not been seen with local administration of vector

13.3.4 Adeno-associated Virus Vectors

- General
  - Adeno-associated virus (AAV) is small, single-stranded DNA virus of the parvovirus family that requires coinfection with adenovirus or other viruses for propagation. In humans, wild-type AAV2 demonstrates site-specific integration into chromosome 19, but this targeted integration is not seen with recombinant AAV vectors
  - AAV has a large number of serotypes that have different affinities for various tissues. An active area of research is to determine the optimal serotype for specific target tissue
  - AAV has been used clinically to target muscle and liver, but immune responses to the viral capsid protein have led to elimination of gene-transduced cells. Phase I trials are ongoing in other sites to determine their effectiveness. There is considerable interest in AAV use in protected sites, such as the central nervous system, to avoid immune recognition. The eye appears to be one such site where clinical benefit has been shown without immune response
- Advantages
  - The advantages of AAV include a broad host range (including rodents, nonhuman primates, and humans) and the ability to produce helper virus-free vector at high titer. AAV typically is not pathogenic and AAV vectors are generally believed to have a favorable safety profile
- Disadvantages
  - A major disadvantage of this vector is the small size of the vector genome, limiting the exogenous DNA insert to less than 4.5 kb of sequence. AAV gene therapy applications are thus currently restricted to diseases where the therapeutic transgene is

encoded on moderate to small cDNAs. These vectors integrate at very low efficiency so stem cells or other cells that will expand in great numbers are generally not good targets for AAV vectors. Elimination of transduced cells by immune responses to the viral capsid proteins has proven problematic in early clinical trials

13.3.5 Herpes Virus Vectors

- General
  - Herpes simplex virus type 1 (HSV-1) is under investigation as a gene transfer vector. An advantage of HSV-1 is its tropism for neurologic tissue, which is generally resistant to infection by other viral vector systems
  - HSV-1 is a large DNA virus that replicates at high efficiency. The viral genome includes more than 80 genes, 38 of which are required for production of infectious virions. The remaining accessory genes encode proteins that affect pathogenicity, host range, and immunogenicity as well as enhance viral replication and latency in nondividing cells
- Advantages
  - The relatively large genome of herpes viruses (~152 kb) suggests that herpes virus vectors will be able to transfer large amounts of exogenous DNA. Some of the viral genes are potentially advantageous if successfully incorporated into a gene therapy vector. For example, all herpes viruses produce proteins that inhibit expression of major histocompatibility (MHC) class I antigens and may blunt an immune response against the vector. The regulatory regions in the herpes virus genome may also be useful for optimizing transgene expression in the nervous system
  - Recently, vectors with deletions in the immediate early genes of the HSV-1 genome have been developed and have reduced cytotoxicity and provide long-term transgene expression. Furthermore, these vectors can be produced in relatively high titer
- Disadvantages
  - The large viruses require complex production procedures and extensive testing to insure significant mutations have not occurred in the viral sequences. Furthermore, the toxicity of the vector, and the remaining viral genes, remains to be determined. The chance of rescue of viral sequences by wild-type virus must also be considered in vector design

13.3.6 Other Vector Systems

- A variety of other systems are being evaluated as potential gene therapy vectors. Vectors based on foamy virus, simian virus 40 (SV40), bovine papillomavirus, vaccinia virus, and other poxviruses are being developed or are in Phase I clinical trials

13.4 Gene Transfer Applications

13.4.1 Compensation for Genetic Mutations

- General
  - Many genetic diseases are due to a specific mutation that results in a nonfunctional protein or a decrease in the level of a functional protein
  - Reintroduction of the deficient gene has been shown to restore a normal phenotype in a wide variety of animal models
  - The first approved gene therapy protocol sought to treat subjects with adenosine deaminase (ADA) deficiency. This disease has been considered an ideal candidate for gene therapy because
    - It is a single gene defect and the gene has been cloned
    - Most children die within the first year of life due to a severe combined immunodeficiency (SCID), so the risk-benefit ratio is suitable for a novel therapy

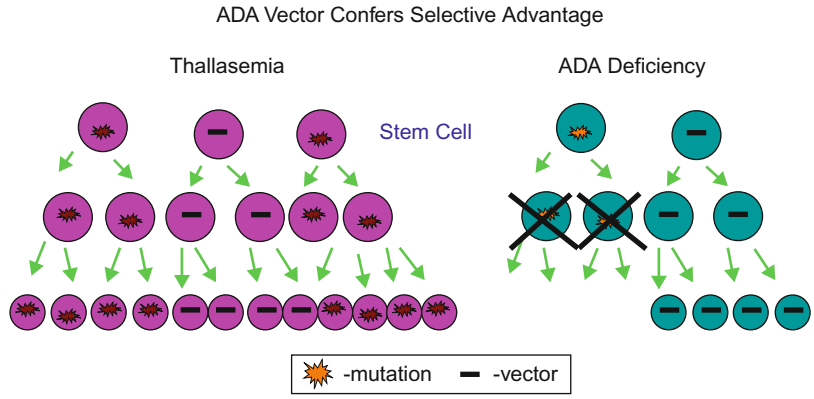


Fig. 13.2 Vectors confer selective advantage to transduced cells in adenosine deaminase deficiency (*ADA*). Hematopoietic stem cells are attractive targets for gene transfer. In the case of thalassemia, introduction of a vector which provided beta globin expression would lead to normal red blood cells, but uncorrected red cell progenitors would still be generated. Therefore, a high level of gene transfer (or elimination of competing diseased stems cells) is required. In contrast, ADA-deficient T cells die during differentiation in the thymus, so even a small corrected in hematopoietic stem cells will lead to repopulation of the T cell population as untransduced cells will die during differentiation

- A key factor is the wide range of protein expression associated with a normal phenotype. Low-level vector expression can improve the disease while overexpression should not be associated with toxicity
- As ADA is required for efficient maturation through the thymus, ADA patients have very low T cell numbers. Therefore, corrected T cell precursors should have a selective advantage (i.e., even if gene transfer is inefficient, any corrected cells should be able to repopulate the lymphoid system without having to compete with uncorrected cells)
- The target cells for ADA and many other blood cell diseases are hematopoietic stem cells, and the technology for harvesting, ex vivo manipulation, and transplantation of these cells is well defined

– As stem cells will undergo extensive cell differentiation and multiple divisions, integrating retroviral vectors has been traditionally used to target hematopoietic diseases. Studies targeting SCID, chronic granulomatous disease, Wiskott-Aldrich syndrome, thalassemia, and Fanconi anemia are currently in clinical trials. Unlike SCID, in most hematologic diseases, there is no selective advantage of vector transduced cells, so the endogenous hematopoietic cells must be eliminated by chemotherapy or other means which can alter the risk-benefit assessment (Fig. 13.2)

13.4.1.1 Retroviral Clinical Application

- In 2009, Aiuti et al. reported in the *New England Journal of Medicine* the finding in ten children treated for SCID due to ADA deficiency
- Subjects were infused with autologous CD34+ cells which were infected with a MLV-based retroviral vector expressing wild-type ADA
- Multilineage engraftment of transduced hematopoietic stem cells was obtained with repopulation of ADA expressing myeloid cells and lymphoid cells. There were no severe adverse reactions
- All patients are alive after a median 4-year followup and nine patients had immune reconstitution with increase T cell counts and normal T cell function

13.4.1.2 AAV Clinical Application

- In 2008, back-to-back articles appeared in the *New England Journal of Medicine* demonstrating improved vision in a Phase I trial for patient with Leber congenital amaurosis 2 (LCA2). The approach used an AAV vector expressing the deficient gene (see Maguire et al. 2008 and Bainbridge et al. 2008)
- LCA2 is caused by mutation in *RPE65* gene that encodes an isomerase key to the visual cycle; without *RPE65*, rod photoreceptor cells are unable to respond to light. Individuals develop impaired vision within the first few months of life and progresses to blindness in the early adulthood. Previously, there was no effective treatment
- The AAV vector was well tolerated and while vision has not been returned to normal, most patients showed sustained improvement in subjective and objective measurements of vision

13.4.2 Gene Therapy for Pharmacologic Effect

- General
  - Gene therapy can be used as a drug delivery system. Examples that are in clinical trial or under development include
    - Vectors designed to produce a secreted protein are diseases such as hemophilia. In this situation, the vector can be expressed in ectopic sites from where the protein is normally produced
    - A growing field of study is the expression of chemokines, cytokines, and growth factors. For example, nonviral and viral vectors expressing a variety of inhibitory molecules are being evaluated in cardiovascular diseases to decrease inflammation and platelet aggregation after balloon angioplasty. Vectors expressing growth factors are being evaluated alone or in combination with cellular therapy aimed at increasing collateral formation in patients with peripheral artery diseases
    - In certain metabolic disorders, vectors can be expressed in the normal site of enzyme production or in ectopic sites in order to decrease the level of toxic metabolites below that associated with disease

13.4.2.1 Adrenoleukodystrophy Clinical Application

- X-linked adrenoleukodystrophy (ALD) is a fatal disease characterized with demyelinating of the central nervous system. Most patients die before reaching adolescence
- The demyelination is caused by mutation of the ABCD1 gene, which metabolizes fatty acids, and lack of this protein leads to disruption of myelin maintenance in oligodendrocytes and microglia
- ALD progression can be halted by allogeneic hematopoietic cell transplantation (HCT) by allowing metabolism to occur in blood-derived cells, including glial cells. However, this approach is limited by donor-related constraints and carries a considerable risk of mortality due to host's immune response against donor
- In 2009, Cartier et al. reported the use of a HIV-1-based lentivirus expressing ABCD1 for ex vivo correction of autologous hematopoietic stem cell with subsequent reinfusion into two patients with ALD, aged 7.5 years and 7 years
- Nine to fifteen percent of peripheral blood mononuclear cells (PBMCs) stably express lentivirally encoded ALD protein after 2 years. With over a year of followup, the progressive cerebral demyelination appears to have been arrested

13.4.3 Inactivation of Harmful Genetic Sequences

- General
  - Gene therapy has traditionally sought to alter cell phenotype by the introduction and expression of foreign genetic material in target cells. The field of oligonucleotide

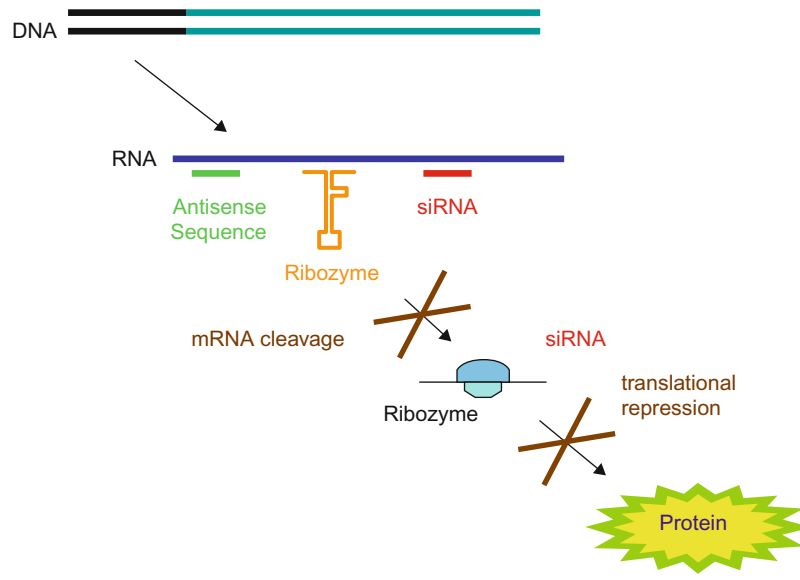


Fig. 13.3 Three strategies to preventing production of unwanted proteins. Antisense oligonucleotides, ribozymes, and short inhibitory RNAs (siRNA) all bind to RNA generating segments of double-stranded RNA which is cleaved and degraded. The siRNA molecules can also lead to translational repression as a second means of preventing protein formation

and inhibitory RNA gene therapy seeks to alter cell phenotype by eliminating expression of harmful genetic sequences. The potential applications for this approach are many, but two examples are

- Leukemias are frequently in association with chromosomal abnormalities including translocations and inversions. In chronic myelogenous leukemia (CML), the t(9;22) chromosomal translocation brings together the breakpoint cluster region (*bcr*) and proto-oncogene *abl* leading to formation of a novel fusion gene. The gene product, p210, contains amino acids from bcr and abl and is abnormally regulated and important in the disease process. Since RNA resulting from the translocation is expressed solely in the cancerous cells, targeting the unique DNA or RNA is one potential approach to selective targeting of the malignant cells. If the RNA can be eliminated before protein is generated, presumably the cell phenotype will be normal or apoptosis will occur
- This strategy appears to be most promising when applied to downregulation of viral gene expression in chronically infected cells. The first application of lentiviral vectors has used this approach as a means of treating HIV-1 by targeting viral RNA; early phase clinical trials are ongoing. This approach is being developed for other infectious diseases, including hepatitis C
  - There are currently three methods used to downregulate harmful genetic sequences (Fig. 13.3)

13.4.3.1 Antisense Oligonucleotides

- The goal of antisense therapy is the disruption of gene expression using short, sequence-specific DNA molecules. In the simplest scenario, synthesized antisense oligonucleotides bind via Watson-Crick base pairing with a complementary target RNA forming a duplex that either blocks ribosomal protein synthesis or is degraded by RNase H. Oligonucleotides are attractive because of their theoretical sequence specificity and the potential for topical application or intravenous administration. The technical challenge has been to design oligonucleotides that are not rapidly degraded after administration, can efficiently enter the target cell, and are stable within the cell while continuing to provide a disease-specific effect

13.4.3.2 Ribozymes

- An extension of the antisense oligonucleotide approach is the use of ribozymes. Ribozymes are RNA molecules that bind to target RNA molecules in a site-specific manner, cleave the target RNA, dissociate, and are then free to cleave another target RNA. Enzymatic cleavage of the target RNA means fewer ribozymes molecules per target RNA are required. The challenges for ribozymes are that they are rapidly degraded by RNase or other factors present in serum. Therefore, ribozymes have generally been expressed in the context of a plasmid or viral vector

13.4.3.3 Inhibitory RNAs

- Currently, the most promising approach is based on the recent recognition of a novel gene regulatory system based on microRNA molecules. The microRNAs are processed through a series of cellular enzymes, (including Dicer), to release inhibitory RNAs that can cause translational repression of target sequences. In addition, the inhibitory RNAs can bind with complementary target RNA sequences that are then rapidly degraded similar to that seen with antisense oligonucleotides. Large exogenous microRNAs can induce adverse cellular responses, but sequences around 20 base pairs (short inhibitor RNAs, siRNA) can provide the intended therapeutic effect without cellular toxicity. Additional modifications have generated short hairpin RNAs which are more stable and provide a more sustained effect

13.4.4 Cell Engineering

- General
  - In a sense, almost all gene therapy applications are an attempt to engineer cells and alter their phenotype. Most of the discussions above have focused on replacing missing genes or inhibiting harmful genetic sequences. Nevertheless, much of the work in gene therapy is directed at transferring sequences that are not directly involved with the disease process, per se. The following are illustrations of some approaches currently in clinical trials. As stem and progenitor cell therapies continue to grow, the ability to alter gene expression through the use of vectors will likely play an important role in improving cellular therapies

13.4.4.1 "Suicide Vectors"

- There is continued interest in vectors that can be activated leading to death of the transduced cell. The classic vectors studied contain the herpes simplex thymidine kinase gene (HS-tk). Unlike human thymidine kinases, the HS-tk will phosphorylate compounds such as acyclovir and ganciclovir that then are further processed, incorporated into DNA, and led to cell death. Vectors that target tumors but are not efficiently taken up by normal tissues have been used to deliver HS-tk in locally advanced mesothelioma, ovarian cancer, and brain tumors. Phase I studies suggest improved gene delivery is required
- These vectors are also attractive for engineering transplanted cells. Particular interest has been in engineering donor T cells in the context of allogeneic hematopoietic stem cell transplantation. If the patient developed life-threatening graft-versus-host disease (GVHD), the disease can be eliminated by the administration of ganciclovir or its analogues. Additional work is being performed to generate alternative suicide systems that do not require the infusion of an active drug but are inert except for activation of a suicide pathway in vector-containing cells

13.4.4.2 Suicide Vector Clinical Application

- Allogeneic hematopoietic stem cell transplantation (allo-HSCT) has potential to treat many types of blood cancers; unfortunately, the relapse rate for adults remains significant
- A significant number of patients who relapse after allo-HSCT can obtain a durable remission by infusing T cells from the donor to

induce a graft-versus-leukemia effect. Unfortunately, many of these patients also develop severe, often fatal GVHD

- In a trial by Bonini et al., relapsed patients were treated with donor T cells that contained a retroviral vector encoding the HS-tk gene
- All 17 patients evaluable for T cell engraftment and graft-versus-leukemia had circulating HS-tk+cells detectable beginning at a median time of 18 days. The antitumor effect tightly correlated with the in vivo expansion of HS-tk+cells
- Seven patients received ganciclovir because of severe GVHD resulting in elimination of HS-tk+cells and effective treatment of the GVHD. These data validate the feasibility, safety, and efficacy of genetically engineered HS-tk+cells in treating cancers

13.4.4.3 Drug Resistance Vectors

- The administration of chemotherapy is often limited due to cytopenia. The engineering of hematopoietic progenitors to express genes that confer resistance to chemotherapeutic agents is being evaluated as a means to allow dose intensification with decreased toxicity. This approach has been successful in protecting against the hematopoietic toxicity of methotrexate with dihydrofolate reductase; against placitaxel, doxorubicin, and vinblastine toxicity using the multidrug resistance gene-1 (MDR1); and against toxicity associated with 1,3-*bis*-(2-chloroethyl)-1-nitrosourea (BCNU) and 1-(2-chloroethyl)-3-cyclohexyl-1-nitrosourea (CCNU) with O6-methylguanine DNA methyltransferase (MGMT)

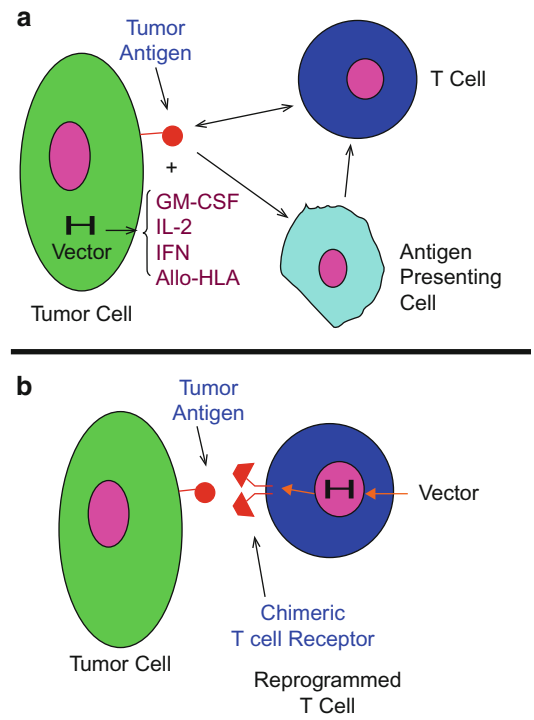


Fig. 13.4 Cancer immunology – investigational strategies. Two approaches are entering phase II clinical trials with the goal of enhancing antitumor immunity. (**a**) Autologous tumor cells are transduced with a variety of chemokines such as GM-CSF, interleukin 2 (*IL-2*), interferons (*IFN*), and allogeneic histocompatibility (*HLA*) antigens as a means of stimulating T cell immunity, either directly or through antigen-presenting cells. (**b**) Autologous T cells are reprogrammed by inserting a vector expressing a tumor-specific T cell receptor that recognizes antigens specifically expressed on the tumor being treated

13.4.5 Cancer Immunotherapy

- General
  - A very active area of basic and clinical research is the use of vector technology to induce an antitumor immune response. Much of the initial work focused on melanoma, due to its known response rate to immunologic agents such as IL-2. More recently, the field has broadened to include a variety of tumor targets. Many approaches share a common hypothesis, that tumor cells express antigens that are relatively unique and that recognition of the antigen by the immune system can lead to elimination of the cancer. Since most cancer patients die of metastatic disease, this approach is particularly attractive since the immune system will be the effector of the antitumor response
  - Tumor antigens were discovered by the identification of T cells within tumors that are reactive to tumor cells but do not

recognize normal tissue. The T cells appear tolerant in cancer patients but can be activated ex vivo and can destroy autologous tumor targets. Cloning of the T cell receptors reveal they often respond to embryonic antigens that are not normally expressed in differentiated tissues. Additionally, some receptors recognize antigens associated with the specific cell type from which the tumor cells are derived. More recently, cell surface antigens expressed on malignant and normal tissues have also been targeted, such as CD19

13.4.5.1 Inducing/Enhancing Immune Responses

- As depicted in Fig. 13.4a, a vector is introduced into autologous tumor cells, which expresses an immunostimulatory molecule. A variety of agents have been evaluated including allogeneic HLA antigens, stimulatory chemokines and cytokines, and T cell costimulatory molecules. Animal models have suggested that targeting antigen-presenting cells may be more successful than directly stimulating T cells

13.4.5.2 Engineering T Cell Receptors

- Another approach has been to reengineer T cells to express a new receptor that recognizes a specific tumor antigen (Fig. 13.4b). The receptor may be obtained by sequencing the receptor from T cells known to recognize specific tumor antigens (T cell receptor, TCR). The other approach is to combine signaling domains from the T cell receptor with protein recognition regions of antibodies (chimeric T cell receptors, CAR). These engineered receptors are then placed into gene therapy vectors and when expressed in a T cell allow the cell to recognize the target antigen. Vectors recognizing melanoma-specific antigens, carcinoma embryonic antigen (CEA, expressed in colon and liver cancers), prostate-associated antigens, and lymphoid antigens are currently in phase I/II clinical trials

13.4.5.3 Melanoma Clinical Application

- In 2006, Morgan et al. reported circulating T cells obtained from patients with melanoma unresponsive to therapy. A retroviral vector expressing a T cell receptor that recognized the MART-1 tumor antigen was inserted into the T cells. After chemotherapy administration to provide a moderate immunodepletion, T cells were then reinfused back into the patients
- T cell engraftment was noted in fifteen patients who obtained a durable engraftment at levels more than 10% of peripheral blood lymphocytes. Two patients had complete remission with sustained levels of circulating engineered cells beyond one year of treatment and appear to be long-term survivors
- These data revealed that genetically engineering approach can specifically confer tumor recognition ability to lymphocytes

13.4.5.4 Chronic Lymphocytic Leukemia Clinical Application

- In 2011, Porter et al. introduced a CAR targeting CD19 into autologous T cells from three patients with refractory CLL
- A small number of T cells were introduced but the cells expanded over 1,000 times and resulted in complete remission that had persisted for 10 months at the time of publication. The expected side effect of eliminating normal B cells with secondary hypogammaglobulinemia was observed

13.4.6 Replication-Competent Viruses

- General
  - There has been a long-standing interest in viruses as anticancer agents beginning in 1912 when De Pace noted that vaccination with a rabies vaccine led to tumor regression in a patient with ovarian cancer
  - The Newcastle virus has been of interest since this virus cannot replicate in normal tissue but can produce a lytic infection in

cells that have activation of the Ras oncogene

13.4.6.1 Replication-Competent Adenoviruses

- The major activity in current trials is with replication-competent adenoviral vectors. Normally, human cells can prevent adenoviral replication through a pathway that involves the tumor suppressor gene p53. The adenovirus has circumvented this by inactivating p53 through a viral protein E1b. To use this virus as anticancer therapy, the viruses have been engineered with deletions in E1b. Since normal cells express p53, the E1b deleted virus cannot replicate. In contrast, p53 is absent in many tumors so the adenovirus can freely replicate and leads to lysis of the tumor cells. The lysed cell then releases more E1b-deficient virus that can now infect surrounding tumor cells. Initial trials with replication-competent adenoviral vectors (Onyx-15®) were used in locally advanced head and neck tumors, and the responses were minimal. Subsequent studies in which the virus has been used in combination with chemotherapy are ongoing
- Additional viruses are also entering clinical trials, including replication-competent vaccinia, measles, and vesicular stomatitis viruses

13.5 Safety Principles and Regulatory Issues

13.5.1 Previous Adverse Events

- While there have been thousand of subjects treated on gene therapy protocols without significant adverse events, there are two significant adverse events that have altered the field
  - In September of 1999, a subject died while being treated on a phase I dose escalation study in which an adenovirus was injected intravenously. The goal was gene replacement in hepatocytes. The subject developed a cytokine-mediated adverse reaction to the adenoviral proteins, developed multiorgan failure, and died
  - The second adverse event of note was mentioned above: the development of malignancy in individuals treated with retroviral vectors. The first case occurred in 2002 when subjects participating in a gene therapy trial for X-linked SCID developed leukemia. While insertional mutagenesis has been a theoretical possibility, it had not been seen in prior human trials. The leukemia developed in four of the 11 subjects treated, with vector integration near the oncogene LMO-2 in three of the cases. Insertional mutagenesis has also been reported in a trial of chronic granulomatous disease and recently in a patient treated for Wiskott-Aldrich syndrome. Interestingly, there have been no reports of malignancy in those patients with the longest follow-up, those treated for ADA SCID. Moreover, there has been no report of this adverse event in the large number of patients with cancer immunotherapy approaches. It appears the vector regulatory sequences, the transgene itself, and the target cell all play a role in a multistep process of insertional mutagenesis. For most applications, investigators are now turning to lentiviral vectors since HIV-1 is not known to cause malignancy directly, in contrast to the murine leukemia viruses

13.5.2 General Safety Considerations

- Safety must be evaluated on a number of levels. The possible side effects of transgene expression must be considered. For example, excessive production of a factor VIII transgene could cause abnormal clotting after successful factor VIII gene transfer to patients with hemophilia. Furthermore, the manufacturing process itself must be scrutinized to insure pathogens or toxic materials have not been introduced during vector production and that the generated material has sufficient activity to confer the intended

therapeutic benefit. Finally, the potential risks of the proposed vector system must be addressed

13.5.3 Regulatory Issues

- In the United States, a variety of regulatory and advisory bodies function to evaluate the safety profile of materials used in clinical gene therapy studies. An Institutional Review Board (IRB) must approve and monitor any research study involving human subjects. Guidelines for conducting IRB activity and oversight are provided by the NIH Office for Protection from Research Risks and the US Food and Drug Administration (FDA)
- The Office of Biotechnology Activities (OBA) was formed to oversee recombinant DNA research in the United States. The office oversees clinical gene therapy at the federal level through the Recombinant DNA Advisory Committee (RAC). It also oversees local oversight through the Institutional Biosafety Committee (IBC) at the investigators' institution. Clinical trials involving use of recombinant DNA must be approved by the local IBC and reviewed by the RAC prior to treating subjects
- Currently, there are no approved gene therapy products in the USA or Europe. In the USA, clinical gene therapy studies require review by the FDA through the Investigational New Drug application (IND) process. Gendicine® is a recombinant adenoviral vector expressing p53 and is the first approved gene therapy product and is currently only available in China. A number of phase III gene therapy trials are in process in the USA, and a number of licensure applications to the FDA are anticipated in the next few years

Further Reading

Abonour R, Williams DA, Einhorn E, et al. Efficient retroviral-mediated MDR-1 gene transfer into autologous human long-term repopulating hematopoietic stem cells. Nat Med. 2000;6:652–8.

Aiuti A, Cattaneo F, Galimberti S, et al. Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med. 2009;360:447–58.

Bainbridge JWB, Smith AJ, Barker SS, et al. Effect of gene therapy on visual function in Leber's congenital amaurosis. N Engl J Med. 2008;358:2231–9.

Cartier N, Hacein-Bey Abina S, Bartholomae CC, et al. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. Science. 2009;326:818–23.

Cavazzana-Calvo M, Payen E, Negre O, et al. Transfusion independence and HMGA2 activation after gene therapy of human β-thalassemia. Nature. 2010;467:318–23.

Cornetta K, Smith FO. Regulatory issues for clinical gene therapy trials. Hum Gene Ther. 2002;13:1143–9.

Davidson BL, McCray JPB. Current prospects for RNA interference-based therapies. Nat Rev Genet. 2011;12:329–40.

DiGiusto DL, Krishnan A, Li L, et al. RNA-based gene therapy for HIV with lentiviral vector-modified CD34(+) cells in patients undergoing transplantation for AIDS-related lymphoma. Sci Transl Med. 2010;2:36–43.

Hacein-Bey-Abina S, Le Deist F, Carlier F, Bouneaud C, et al. Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. N Engl J Med. 2002;346:1185–93.

Hacein-Bey-Abina S, von Kalle C, Schmidt M, et al. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N Engl J Med. 2004;348:255–6.

Kay M. State-of-the-art gene-based therapies: the road ahead. Nat Rev Genet. 2011;12:316–28.

Levine BL, Humeau LM, Boyer J, et al. Gene transfer in humans using a conditionally replicating lentiviral vector. Proc Natl Acad Sci USA. 2006;103:17372–7.

Maguire AM, Simonelli F, Pierce EA, et al. Safety and efficacy of gene transfer for Leber's congenital amaurosis. N Engl J Med. 2008;358:2240–8.

Mingozzi F, High KA. Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. Nat Rev Genet. 2011;12:341–55.

Naldini L. Ex vivo gene transfer and correction for cell-based therapies. Nat Rev Genet. 2011;12:301–15.

Peng Z. Current status of gendicine in China: recombinant human Ad-p53 agent for treatment of cancers. Hum Gene Ther. 2005;16:1016–27.

Porter DL, Levine BL, Kalos M, et al. Chimeric antigen receptor-modified T cells in chronic lymphoid leukemia. N Engl J Med. 2011;365:725–33.

Rosenberg SA, Aebersold P, Cornetta K, et al. Gene transfer into humans-immunotherapy of patients with advanced melanoma, using tumor infiltrating lymphocytes modified by retroviral gene transduction. N Engl J Med. 1990;323:570–8.

Instrumentation for Molecular Testing 14

Jordan Laser, Miao Cui, Bruce E. Petersen, Fei Ye, and David Y. Zhang

Contents

14.1	**Instrumentation and Technology**	413
14.1.1	Cell Separation	413
14.1.2	Nucleic Acid Extraction and Purification	416
14.1.3	Spectrophotometers	421
14.1.4	Thermocyclers for Conventional PCR	423
14.1.5	Real-time PCR Instruments	426
14.1.6	Microarray Platforms	435
14.1.7	Next-Generation Sequencing	443
14.1.8	Target Sequence Capture System	446
14.1.9	Nucleic Acid Quantification	447
14.1.10	Gel Imaging Systems	449
14.1.11	Fluorescence Microscope	450
14.1.12	Slide Imaging Systems	451
Further Reading		453

J. Laser, MD (✉)
Department of Pathology and Laboratory Medicine, North Shore/Long Island Jewish Health System, Manhasset, NY, USA

M. Cui, MD • B.E. Petersen, MD • F. Ye, PhD • D.Y. Zhang, MD, PhD, MPH
Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

14.1 Instrumentation and Technology

14.1.1 Cell Separation

14.1.1.1 autoMACS® Pro Separator (Miltenyi Biotec, Bergisch Gladbach, Germany)

- General information
  - The autoMACS® Pro Separator (Fig. 14.1) is an automated benchtop magnetic cell sorter that allows sorting of more than 10 million cells per second from samples containing up to 4 × 109 total cells. The instrument allows processing of up to six samples in one run. The instrument was designed for use with any of Miltenyi Biotec's MACS Cell Separation Reagents for research applications. It is possible to choose between different cell separation strategies – from positive selection of abundant or rare cells to the isolation of untouched cells by depletion of nontarget cells or even isolation directly from whole blood. Twelve preset separation programs provide the researchers with reproducible and consistent results
- Principles of operation
  - The first step involves the highly specific magnetic labeling of cell surface antigens with MACS® MicroBeads
  - MicroBeads consist of monoclonal antibodies coupled to superparamagnetic particles of approximately 50 nm in size.

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_14,

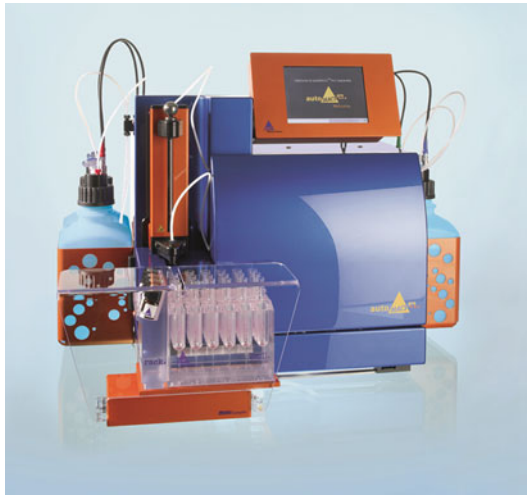


Fig. 14.1 autoMACS® Pro Separator. (http://www.miltenyibiotec.com/en/NN_1085_autoMACS_Pro_Separator.aspx)

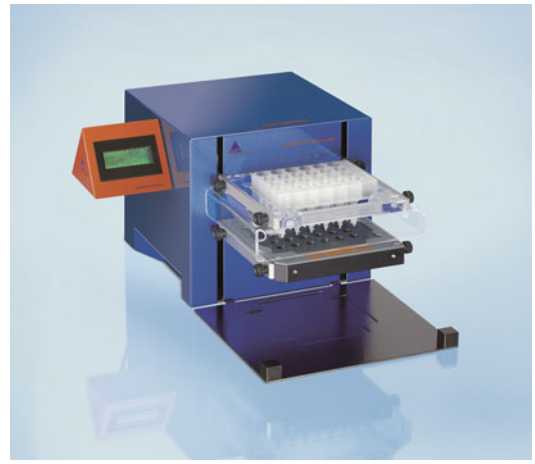


Fig. 14.2 MultiMACS™ Cell24 Separator. (http://www.miltenyibiotec.com/en/NN_1086_MultiMACS_Cell24_Separator.aspx)

MicroBeads do not change the cell's scatter properties in flow cytometry or cause cell activation or receptor capping, which is commonly observed with larger magnetic beads
  - After magnetic labeling, the cells are passed over the separation column placed in the magnetic field of the separation unit, where the small ferromagnetic spheres of the column matrix produce high magnetic gradients
  - The ferromagnetic structures generate magnetic forces 10,000-fold greater than in conventional geometries, allowing retention of the labeled cells within the column
  - Nonlabeled cells flow through and can be collected. Labeled cells can be released after removing the column from the magnet. Thus, both labeled and nonlabeled cell fractions can be efficiently isolated
- Applications
  - Cell isolations with excellent purity, high yield, and viability are obtained from the following sources:
    - PBMCs
    - Whole blood
    - Bone marrow
    - Buffy coat
    - Cell cultures
    - Dissociated tissues
- Advantages and limitations
  - Walkaway cell isolation from up to six samples (including autolabeling)
  - Chill racks maintain samples at 4 °C
  - True whole blood cell isolation: no red blood cell lysis and density gradient centrifugation required
  - MicroBeads do not alter structure, function, or activity status of labeled cells, and they are not known to interfere with subsequent experiments
  - Since columns are reused, potential contamination of previous sample

14.1.1.2 MultiMACS™ Cell24 Separator

- General information
  - The MultiMACS™ Cell24 Separator (Fig. 14.2) is used for simultaneous multisample magnetic cell separations (24 samples each run) based on MACS® Technology. The MultiMACS Cell24 Separator can be employed as a semiautomated benchtop instrument or integrated into fully automated liquid-handling platforms. A strong 24-well

magnet is used to retain magnetically labeled cells within multi-24 columns; nonlabeled target cells are collected as flow through from the column

- Principles of operation
  - The MultiMACS Column Holder holds a multicolumn block with multi-24 columns
  - The MultiMACS 24 Magnet can be moved vertically together with or independent from the MultiMACS Column Holder. This allows lowering of the columns into the magnet wells – and thus applying the magnetic field – or, similarly, lifting of the columns out of the MultiMACS 24 Magnet
  - The MultiMACS 24 Magnet generates a high-gradient magnetic field to retain magnetically labeled cells within the columns
  - During a cell separation process with the MultiMACS Cell Separator, liquids such as samples, wash, and elution buffers are dispensed onto the columns either manually, e.g., by a multichannel pipette, or automatically by a liquid-handling workstation
  - Target cells are collected in a 24-deep well block or in 24 5-mL tubes
- Procedure
  - Cells are magnetically labeled for positive selection or depletion by using one of various MACS® Cell Separation Reagents
  - The separation program to be executed is selected on the process management screen of the MultiMACS Cell24 Separator
  - The process is started, and the magnet moves to the start position
  - The multicolumn block is inserted into the MultiMACS Column Holder; the 24-deep well block is placed into the instrument
  - The samples are applied to the columns; nonlabeled cells run through. The column is washed to collect all nonlabeled cells
  - The plate containing untouched cells is removed, and samples can be used for downstream applications
  - Magnetically labeled cells retained in the column can be collected in a subsequent step using vacuum elution
- Applications
  - In combination with MACS MicroBeads and MACS cell isolation kits, the MultiMACS Cell24 Separator can be used to isolate a wide variety of cell types or to deplete specific target cells from up to 24 samples in a single run
- Advantages
  - Multisample benchtop system for parallel cell isolation
  - Allows fast, high-quality cell separation
  - Can be employed as a semiautomated benchtop instrument or integrated with liquid-handling systems for full automation
  - Contact-free pipetting eliminates cross contamination

14.1.1.3 RoboSep (StemCell Technologies Inc., Vancouver, British Columbia, Canada)

- General information
  - RoboSep is an automatic cell separation system that uses immunomagnetic cell isolation principle. Using column-free EasySep® technology, the RoboSep® performs all steps necessary to magnetically label and separate different cell types by positive or negative selection. The RoboSep minimizes sample handling, eliminates cross contamination, and reduces "hands-on" time. EasySep™ can isolate highly purified cells in about 25 min
- Principle of operation
  - It uses tetrameric antibody complexes (TACs) and dextran-coated magnetic particles to select or deplete cells of interest
  - The TACs cross-link cells expressing specific surface antigens to the magnetic particles
  - When the sample is placed in EasySep™ Magnet, cross-linked cells are pulled to the sides of the magnet
  - Any remaining untouched cells can then be removed, resulting in two distinct cell fractions
    - Positive selection: up to four different simultaneous separations and up to 8×10^9 total cells (four samples of up to 2×10^9 cells each)

    - Negative selection: one or two different simultaneous separations and up to 2×10^9 total cells (two samples of up to 10^9 each)
    - Negative and positive selections can be executed simultaneously (e.g., one negative selection and two positive selections)
    - Sample volume: from 250 μL to 8.5 mL for each sample
    - 90–95% purity of isolated cells
- Applications
  - Isolates lymphocytes (CD3 and/or CD19) and myeloid cells (CD33 and/or CD66b) for chimerism analysis
  - Isolates T cells, B cells, and lymphocytes for HLA analysis
  - Isolates epithelial cells in blood using EpCAM and MUC1 beads
- Advantages and limitations
  - Uses disposable tips and eliminates cross contamination
  - Flow-compatible magnetic nanoparticles, allowing downstream flow cytometry analysis
  - Limited to the availability of antibodies and its affinity and specificity
  - Low throughput and four samples at a time
  - Slow process and at least 20 min for four samples

14.1.2 Nucleic Acid Extraction and Purification

14.1.2.1 COBAS® AmpliPrep (Roche Diagnostics)

- General information
  - The COBAS® AmpliPrep instrument uses solution-phase magnetic bead capture in the automated extraction of nucleic acids. It is used for large-scale preparation of DNA and RNA samples
- Principles of operation
  - The COBAS® AmpliPrep instrument purifies RNA or DNA targets for PCR in six simple steps: loading the sealed sample input tubes, lysis and hybridization, capture, washing, resuspension, transfer of prepared samples to output tubes, and then their return to their original positions on the sample rack
  - Processing of each sample takes place in a separate, self-contained, single-use, disposable sample processing unit (SPU)
  - Sample volume can range from 250 uL to 1,000 uL
  - Up to 72 samples and four different assays can be loaded on the system simultaneously
- Applications
  - High-throughput extraction and purification of nucleic acids
  - The purified RNA or DNA can be used for qualitative and quantitative molecular diagnostics
- Advantages
  - High throughput
    - Capable of continuous operation
    - 72-sample capacity
    - The first 24 samples are processed in 2 h
    - Each subsequent set of 24 samples can be processed in 1 h
    - Can process up to 144 samples per 8-h shift
    - Samples can be run overnight (20-h onboard stability)
  - Can automatically add PCR master mix and internal control/quantitation standard (IC/QS)
  - Manual steps are minimized (loading and unloading only)
  - Machine has automated decapper; thus, capped specimens may be loaded, decreasing risk of contamination
  - Reagents are loaded as a unit (no mixing of lots, increased reproducibility)
  - Continuous access for loading additional samples, reagents, and disposables
  - Onboard barcode scanner reads reagent and sample barcodes, eliminating transcription errors
  - Pipetting error is minimized by pipetting integrity check and clot detection
  - Instrument can directly load K-tubes for analysis on COBAS TaqMan Analyzers
  - Instrument inventories reagents and disposables prior to run

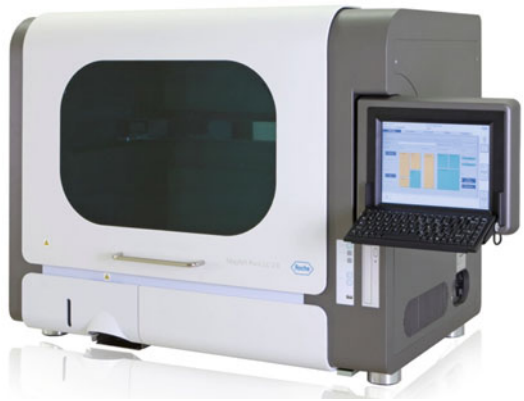


Fig. 14.3 MagNA Pure LC 2.0. (Source: http://www.roche-applied-science.com/PROD_INF/MANUALS/napi_man/pdf/chapter7/page_202-204.pdf)

- Limitations
  - Only for use with plasma or serum samples
  - Limited to total nucleic acid extraction

14.1.2.2 MagNA Pure® LC Instrument (Roche Diagnostics)

- General information
  - The MagNA Pure LC instrument (Fig. 14.3) is an automated system for purification of nucleic acids following prior cell lysis. The instrument can process a wide range of sample types including blood, blood cells, culture cells, plasma, serum, sputum, stool, and body fluids
- Principles of operation
  - Total nucleic acid is bound to magnetic glass beads
  - Beads are then aspirated by a pipette and drawn against the wall of the pipette tip by a magnet (disposable nuclease-free pipette tips also serve as reaction vials or "reaction tips")
  - Cellular debris is removed by extensive washing
  - Nucleic acid is eluted/separated from magnetic beads at high temperature
  - The process for purification of mRNA is similar to the above; however, nucleic acid binding is mediated by biotinylated poly-dT oligonucleotides, which interact with streptavidin-coated magnetic beads
  - Up to 32 nucleic acid isolations can be performed in a single run of 60–180 min
    - Sample volume can range from 10 ul to 300 ul
    - Elution volume ranges from 25 ul to 100 ul
    - Recovery rate: up to 80%
  - The instrument also has the capability of preparing PCR master mix solutions
  - The instrument can directly load sample and master mix into LightCycler® capillaries, 96-well plates, A-rings, or tubes
- Applications
  - Nucleic acids may be isolated from a wide variety of sources using various kits available from Roche Diagnostics
    - Genomic DNA: from whole blood, blood cells, cultured cells, and tissue
    - Total RNA: from whole blood, blood cells, cultured cells, and tissue
    - mRNA: from whole blood, blood cells, cultured cells, and tissue
    - Bacterial and fungal DNA: from BAL fluid, sputum, CSF, urine, swabs, and bacteria cultures
    - Viral nucleic acids: from serum or plasma
- Advantages
  - High level of automation provides walk-away performance
    - Automated clot and tip loss detection
    - Sample tracking
  - Can process a wide range of sample types (see above)
  - Reduced chance of cross contamination compared to manual methods Safeguards include
    - Uses of piston-driven pipetting head instead of vacuum pumps and tubing
    - Built-in UV lamp and HEPA filter
    - Reactions performed in disposable reaction tips
  - Automated loading of isolated nucleic acids and master mixes into standard PCR reaction tubes, LightCycler® capillaries, 96-well plates, or A-rings
- Limitations
  - Does not automatically add QS template
  - Does not perform specimen lysis
  - Low throughput and slow process

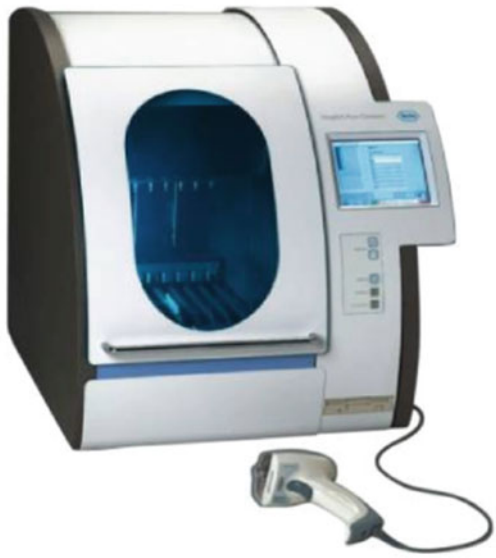


Fig. 14.4 MagNA Pure Compact (Source: www.roche-diagnostics.com)

14.1.2.3 MagNA Pure® Compact Instrument (Roche Diagnostics)

- The MagNA Pure Compact Instrument (Fig. 14.4) is similar to the MagNA Pure LC instrument but allows for smaller sample volumes. Up to eight isolations of DNA or RNA can be performed in only 20–40 min

14.1.2.4 Autopure LS (QIAGEN Systems)

- General information
  - This system combines the advantages of automation with Puregene chemistry to purify DNA from large samples to meet the need for high-volume DNA purification in demanding areas such as linkage analysis, clinical trials, and population-based genetic association studies
- Principles of operation
  - This instrument automates Puregene chemistry (see above)
  - 8–16 samples of 1–10 ml of whole blood or 80 samples of 150 million cultured cells can be processed in 8 h
  - The typical yield is routinely greater than 90% with a minimum yield of 70%
    - From 10 ml of whole blood, 350 μg of high-molecular-weight DNA that is 100–200 kb in size can be expected
    - A_{260}/A_{280} ratios of 1.7 to 1.9
    - DNA stable at 4°C for many years
- Applications
  - Ideally suited for PCR, Southern blotting, restriction digestion, and SNP analysis
  - Long-term DNA banking/archiving
- Advantages
  - Accommodates a wide range of samples
    - Buccal swabs and mouthwash
    - Tissue homogenates
    - Amniotic fluid
    - Blood and blood spots
    - Bone marrow
    - Cultured cells
  - Complete sample tracking with barcodes provides complete chain of custody for sample tube, input tube, output tube, and storage tube and also minimizes clerical errors for accurate data tracking

14.1.2.5 BioRobot Universal System (QIAGEN)

- General information
  - The BioRobot Universal System provides walkaway automation of sample preparation for applications in clinical laboratories. The instrument performs nucleic acid isolation from blood and cell-free body fluids for 48–96 samples in parallel
- Principles of operation
  - Purification is performed by magnetic separation
  - Pipet tips function as separation chambers
  - Up to 96 samples can be processed per run, eight samples at a time
  - The workstation has high-precision positioning for accurate liquid handling through eight channels and uses disposable filter tips to eliminate carryover
  - Automated vacuum processing eliminates centrifugation steps, allowing walkaway automation and fast sample processing
  - A sample tracking system identifies and tracks barcode-labeled labware for fully traceable results

  - The BioRobot Universal System is supplied with ready to run QIAamp (QIAGEN) protocols as well as capability for user-designed isolation of genomic DNA from blood and viral DNA and RNA from plasma and serum
  - The yield is about 30–60 µg DNA from 1 ml of blood
  - Built-in UV light can be used for decontamination between runs
- Application
  - Purification of RNA from cells, tissues, and blood
  - Purification of DNA from swabs, blood, and forensic samples
  - Purification of plasmid DNA from bacteria
  - DNA cleanup from amplification reactions
  - RT-PCR, PCR, sequencing reaction, and forensic assay setup

14.1.2.6 QIAcube (QIAGEN)

- General information
  - The QIAcube is an automated sample preparation system that enables to use well-established QIAGEN spin-column kits. The QIAcube system controls integrated components including a centrifuge, heated shaker, pipetting system, and robotic gripper. This enables the QIAcube to fully automate more than 40 QIAGEN spin-column kits using a simple lyse, bind, wash, and elute procedure
- Principle of operation
  - Uses spin column to separate total DNA or RNA or protein
  - 1–12 samples per run
  - 3–8 ug DNA yield from 200 ul of blood
  - Purification of DNA, RNA, or proteins
  - Require 2 h of processing
- Applications
  - Protocols available for purification of RNA, genomic DNA, plasmid DNA, viral nucleic acids, and proteins, plus DNA and RNA cleanup
  - Sequencing/sequencing analysis
  - Gene expression analysis
  - Genotyping
  - Proteomics

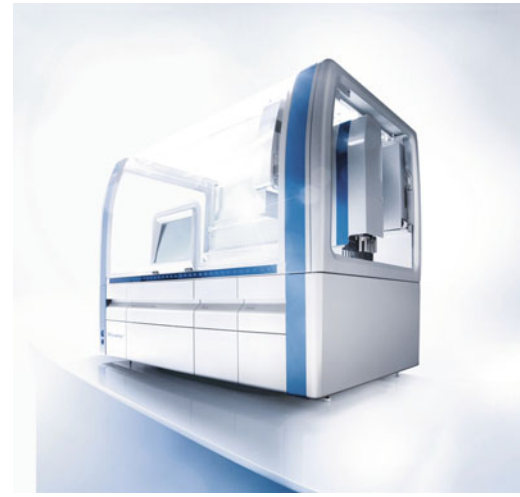


Fig. 14.5 QIAsymphony SP/AS instruments (http://www.qiagen.com/Products/QIAsymphonySP.aspx)

- Advantages and limitation
  - More than 40 well-established QIAGEN spin-column kits available
  - Low throughput (maximum of 12 samples per run)

14.1.2.7 QIAsymphony SP/AS Instruments (QIAGEN)

- General information
  - QIAsymphony technology (Fig. 14.5) combines the speed and efficiency of silica- or Ni–NTA-based purification with the convenient handling of magnetic particles. With a dedicated range of QIAsymphony Kits, the QIAsymphony SP enables sample preparation of DNA, RNA, bacterial and viral nucleic acids, and 6xHis-tagged protein from a wide range of samples. The QIAsymphony AS extends the capabilities of the QIAsymphony SP by integrating automated PCR assay setup which, in combination with the Rotor-Gene Q and QIAGEN real-time and endpoint PCR kits, enables you to complete your automated PCR workflow and maximize your efficiency
- Principle of operation
  - Processing 1–96 samples in batches of 24
    - Blood: 192 × 200 µl with the QIAsymphony DNA Mini Kit and

Table 14.1 DNA yields adequate for most downstream analysis

Sample type	Sample amount or volume	Elution volume (µl)	Typical DNA yield (µg)
Whole blood	1000 µl	500	15 – 45
	400 µl	400	8 – 24
	200 µl	200	4 – 12
Buffy coat	400 µl	400	24 – 72
	200 µl	200	12 – 40
Rat tail	50 mg	200	20 – 40
Muscle	50 mg	200	5 – 15
Spleen	25 mg	200	40 – 80
Liver	25 mg	200	25 – 50
Kidney	25 mg	200	15 – 30
Lung	25 mg	200	10 – 25
Jurkat cells	1.00E + 07 cells	200	60 – 80

144 × 400 µl or 96 × 1000 µl with the QIAsymphony DNA Midi Kit
    - Buffy coat: 192 × 200 µl with the QIAsymphony DNA Mini Kit and 144 × 400 µl with the QIAsymphony DNA Midi Kit
  - Sample preparation consists of four steps: lyse, bind, wash, and elute. Samples are lysed under denaturating conditions in the presence of proteinase K. Lysates are transferred to sample prep cartridges, and DNA binds to the silica surface of magnetic particles. Contaminants are removed by washing, and pure DNA is eluted in a user-specified volume of either modified TE buffer or water
  - Sample volumes of up to 1 ml (20–1000 µl) and up to 50 mg of tissue or 10^7 cells can be processed
  - Elution volumes on the QIAsymphony SP range from 60 µl to 500 µl
  - DNA yields are adequate for most downstream analysis (Table 14.1)
  - RNA yield is 15 ug from 1 × 10e6 HeLa or Jurkat cells
- Applications
  - Purified DNA is ready for use in downstream applications, such as multiplex ligation-dependent probe amplification (MLPA), array comparative genomic hybridization (aCGH), SNP analyses, Southern hybridizations, and STR profiling
  - Sample types: aspiration, swab, BAL, sputum, nasal secretions, and blood
    - Plasma, serum, and CSF samples can be processed using the virus cell-free protocols of the QIAsymphony Virus/Bacteria Kits
    - Respiratory samples (sputum, BAL, aspirates, dried swabs, transport media) and urogenital samples (urine, urogenital swabs, transport media) can be processed using the pathogen complex protocols

14.1.2.8 NorDiag Arrow (NorDiag)

- General information
  - Automated benchtop instrument for purification of nucleic acids and cell isolation with magnetic beads
  - Unique pipetting system that allows for the development of a wide range of applications and protocols for different sample types and minimizes the risk of contamination and with limited maintenance requirements
  - See Fig. 14.6
- Principles of operation
  - Processes 1–12 samples in parallel
  - Magnetic bead-based system with a single-use cartridge for nucleic acid isolation or cell purification
  - Different protocols and reagents for automated nucleic acid extraction from difficult clinical samples including stool, urine, and swabs
  - Automated processing of from 1 to 12 samples, 30 –60 min depending on the protocol
- Applications
  - DNA/RNA from blood, tissue, saliva, buccal swabs, and paraffin-embedded tissue
  - Stool cartridges and protocol for purified NA from stool samples
  - Unique BUGS'n BEADS technology for extraction of bacterial, viral, and fungal nucleic acids from urine, stool, sputum, and swabs
  - Viral protocol for sensitive detection of viruses

Fig. 14.6 NorDiag Arrow (http://www.isogen-lifescience.com/home/automation/automated-sample-preparation/nordiag-arrow)

  - Molzym Blood Pathogen for isolation of blood pathogens including yeast and fungi
  - CellSep cartridges and protocol for isolation of targeted cell populations using magnetic beads with specific monoclonal antibodies
- Advantages
  - Economical instrument with many features including UV sterilization, touch screen display, and barcode reader
  - A multiple-purpose instrument for both cell isolation and nucleic acid purification
  - Simple design with no internal pumps or tubing and minimal service needs
  - New protocols easily loaded via USB port
- Limitations
  - Cannot process more than 12 samples at a time
  - Sample input not more than 1 ml
  - Batch processing of samples

14.1.2.9 NorDiag Bullet

- General information
  - Automated instrument for purification of nucleic acids for up to 96 samples, from primary tube to amplification setup
  - Flexible liquid-handling instrument that has the potential for customization of amplification setup for specific applications
  - See Fig. 14.7
- Principle of operation
  - Uses magnetic bead chemistry with a rapid liquid handler. NorDiag Bullet can be customized to specific applications
  - Processes up to 96 samples from primary tubes to amplification setup in less than 2 h
- Applications
  - Pathogen detection with BUGS'n BEADS magnetic bead chemistry using urine, stool, swabs, and sputum
  - Viral magnetic beads and protocol for serum and plasma
  - Blood protocol for 65 μl of blood
- Advantages
  - Integrated barcode reader
  - Rapid processing of the samples from input to amplification setup. Time for processing 96 samples in the range of 1.5–3.0 h
  - Protocols can be customized for specific applications
- Limitations
  - Limit of 96 samples per setup
  - No random access to processing of samples

14.1.3 Spectrophotometers

14.1.3.1 NanoDrop® 2000 (Thermo Scientific)

- General information
  - The NanoDrop® 2000 Spectrophotometer (Fig. 14.8) is used in the quantification of nucleic acid, protein, and cell suspension concentrations, as well as determination of DNA, RNA, and protein purity. Models differ in terms of sample volume required, recoverability of the sample, and range of linearity. Nonetheless, these instruments all operate on the basis of common principles of spectrophotometry (see above)
- Principles of operation
  - Light source: xenon flash lamp
  - Detector: 2048-element linear silicon CCD array

Fig. 14.7 NorDiag Bullet (http://www.isogen-lifescience.com/home/automation/automated-sample-preparation/nordiag-bullet)

Fig. 14.8 NanoDrop 2000 (http://www.nanodrop.com/Productnd2000overview.aspx?gclid=CMKx2baozK0CFaQRNAodISWKhQ&AspxAutoDetectCookieSupport=1)

- Wavelength range – 220–750 nm
- Absorbance of each sample is measured at two different path lengths (0.2 mm, 1.0 mm), allowing for a very wide range of detection (5–11,300 ng/ul dsDNA) without dilution
- Absorbance calculation: absorbance = −log [intensity (sample)/intensity (blank)]
- Concentration calculation
  - Fluorescent dye concentration is calculated using the general form of the Beer–Lambert equation: $A = Ebc$, where A = absorbance, E = extinction coefficient (liter/mol–cm), b = path length, and c = molarity
  - Nucleic acid concentration is calculated using a modified form of the Beer–Lambert equation: $c = Ae/b$, where e = extinction coefficient (ng–cm/ul)

- Procedure
  - Blanking cycle is performed, prior to running test samples, using the same solvent or buffer solution that is present in the test samples
  - One to two microliters of sample is pipetted onto a measurement pedestal which houses a fiber-optic cable
  - The sample arm is lowered over the sample and slightly compresses the droplets, and a sample column is formed. Surface tension alone holds the samples in place

  - A spectral measurement and quantification of each sample are made based on the tightly controlled path length
  - After measurement is completed, the sample arm is raised, and surfaces that were in contact with the sample are wiped clean. The sample may be recovered with a pipette
- Applications
  - Measurement of nucleic acid concentration and purity
  - Measurement of fluorescent dye-labeling density of nucleic acid microarray samples
  - Analysis of the purity of protein, up to 100 mg/ml (BSA)
  - Expanded spectrum measurement and quantitation of fluorescent dye-labeled proteins, conjugates, and metalloproteins
  - Bradford assay analysis of protein
  - BCA assay analysis of protein
  - Lowry assay analysis of protein
  - Cell density measurements
  - General UV and visible light range spectrophotometry
- Advantages
  - Requires very small sample volume (0.5–2 ul)
  - Very wide dynamic range (5–11,300 ng/ul dsDNA)
- Limitations
  - Microvolume samples are subject to rapid evaporation; replicate measurements require reloading of fresh sample
  - There is risk of sample carryover if the instrument is not adequately cleaned between samples
  - DNA samples must be homogeneous; due to microvolume sampling, heterogeneity will substantially affect reproducibility
  - Low throughput: one sample at a time

14.1.3.2 NanoDrop 8000 (Thermo Scientific)

- General information
  - The eight-sample retention system used by the NanoDrop 8000 Spectrophotometer allows measurements of eight samples simultaneously. It uses the same technology as NanoDrop 2000
- Principle of operation
  - Measures nucleic acid and protein concentration of samples as small as 1 μL
  - Direct, easy measurements from tubes or plates
  - Single sample mode or up to eight samples at a time
  - Analysis of 96 samples in less than 6 min
  - Full spectral output
  - Calculates sample purity ratios (260/280)
  - Preconfigured methods for common applications such as nucleic acid, protein A280, microarray, proteins and labels, Bradford, and BCA
  - No consumables
  - Sample position illuminator which increases efficiency and reduces error by keeping track of the samples to be measured

14.1.4 Thermocyclers for Conventional PCR

- Polymerase chain reaction (PCR) is a process which employs a heat-stable DNA polymerase (Taq DNA polymerase) in the exponential amplification of a target DNA sequence. Three temperature-dependent steps (denaturation, annealing of sequence-specific primers, and elongation) are repeated in a series of cycles. Theoretically, the copy number of the target sequence doubles with each cycle. The process is described in detail in the Methodology chapter
- A thermocycler instrument produces the necessary temperature changes between denaturation, annealing, and elongation phases in a series of preprogrammed steps. Reaction tubes, containing target nucleic acid and all PCR reagents, are fitted within a temperature-controlled block. In order to verify reliable operation, the temperature of each well within the block should be tested at least twice per year, using an external probe that has been calibrated against a temperature standard
- The original PCR instrument was developed by Cetus Instrument Systems in 1985. Since

then, various modifications have been made to increase amplification efficiency, specificity, and sensitivity. Currently, many manufacturers produce PCR machines, all of which are based upon the same fundamental principles. The most common model of thermocycler used in clinical laboratories is discussed below

14.1.4.1 GeneAmp PCR System 9,700 (Applied Biosystems, Foster City, California, USA)

- General information
  - This instrument offers many of the features of the GeneAmp PCR System 9600 (Fig. 14.9) which is no longer available and has the added advantages of a smaller footprint and interchangeable blocks. Networking software is available for single-source control of multiple instruments
- Principles of operation
  - 0.2-ml MicroAmp reaction tubes in a 96-position sample tray (MicroAmp tray)
    - PCR volume range – 10–100 μL
  - The instrument contains a programmable heating and cooling block designed to heat and cool up to 96 PCR samples in a rapid and uniform manner
    - 3 interchangeable blocks available: aluminum, silver, and gold-plated silver
  - Temperature range of block – 4.0–99.9 °C
    - Displayed sample temperature matches average true temperature +/−0.75 °C
    - Max block ramp rate – 2.3–3.5 °C/s
    - Temperature accuracy – ±0.25 °C (35–99.9 °C)
    - Temperature uniformity – ≤0.5 °C (30 s after reaching 95 °C)
  - A heated cover is positioned over the sample block
    - Ensures tubes fit tightly into wells
    - Prevents condensation on top surface of tubes
  - Coolant flows through 17 holes within the block
    - Eight are used for rapid cooling of the block (ramp cooling)
    - Nine are used for cold biasing the system

Fig. 14.9 GeneAmp PCR System 9700 (http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/PCR/thermal-cyclers-realtime-instruments/thermal-cyclers/geneamp-pcr-system-9700.html)

  - Rapid heating is provided by a Kapton heater beneath the block
    - Power density at the edges is greater than at the center in order to compensate for heat loss at the periphery
  - A keypad is used for creating, storing, editing, and running PCR programs
  - Indicator lights show when block is heating, hot, or cooling
- Applications
  - DNA amplification for sequencing, genotyping (allele-specific PCR, restriction fragment length polymorphism analysis, microsatellite studies), identification of viral and bacterial pathogens, and downstream cloning applications
  - Reverse transcription PCR (RT-PCR) for expression analysis
  - Multiplex PCR reactions (use of multiple primer sets to simultaneously amplify multiple targets)
- Advantages
  - Reproducibility of cycle times
  - Uniformity of PCR yields

  - Thin-walled reaction tubes allow efficient heat transfer
  - Rapid heating and cooling
  - Oil-free operation
  - Low cost
- Limitations
  - Product detection requires post-PCR processing, which is often lengthy and increases the potential for contamination
  - Since PCR reactions are generally carried through to the plateau stage of amplification, accuracy of original product quantification is limited

Fig. 14.10 The COBAS AMPLICOR® Analyzer (Source: www.roche.com.ua/. . ./cobas_amplicor_ct_ng.jpg)

14.1.4.2 Veriti™ Thermal Cycler (Applied Biosystems)

- This instrument has similar features as GeneAmp PCR System 9700. However, it offers faster block ramp rate (5.0C/s) and possible six independent blocks for PCR (VeriFlex)

14.1.4.3 COBAS AMPLICOR® Analyzer (Roche Diagnostics)

- General information
  - The COBAS AMPLICOR® Analyzer (Fig. 14.10) is an automated batch analyzer system combining conventional PCR with post-PCR product detection on a single instrument. Although it is used for molecular testing in clinical microbiology, it is gradually replaced by real-time TaqMan Analyzer instrument
- Principles of operation
  - The instrument incorporates five components
    - Thermal cyclers (two thermal cycler units: TCA and TCB)
    - Automatic pipettor
    - Incubator
    - Washer
    - Photometer
      - Reads signals at a single wavelength – 660 nm
      - Light source: pulsed light-emitting diode (LED)
      - Detector: photodiode
  - PCR is performed using biotin-labeled primers
  - Amplification is followed by alkaline denaturation of the amplicon, followed by hybridization with oligonucleotide capture probes bound to magnetic microparticles and then multiple washing steps. During washing, the microparticles are held in place by a magnet
  - Colorimetric detection
    - Bound amplicons react with avidin-conjugated horseradish peroxidase (HRP), taking advantage of the extremely high affinity between avidin and biotin molecules
    - Additional washing is performed
    - Reaction with tetramethylbenzidine (TMB) substrate produces color
    - Photometer performs absorbance readings at 660 nm
      - Forty eight-sample capacity per run
      - Multiplexing capability: up to 6 different detections per sample
      - Incorporation of dUTP and AmpErase® (uracil-N-glycosylase) in reaction mixture prevents contamination of by-products of previous PCR reactions
      - Results calculation
    - Qualitative: test result is reported as absorbance value corrected for reagent blank [test result (A660) = A660 sample – A660 reagent blank]
      - Test result is compared to a preprogrammed test-specific absorbance range

  - Result may be positive, negative, or equivocal (gray zone)
  - Quantitative: titer can be calculated on the basis of a quantitation standard
- Applications
  - Available tests
    - Virus: HIV1, HCV detection and quantitation, HBV, and CMV
    - Bacteria: *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, *Mycobacterium tuberculosis*, *Mycobacterium avium*, and *Mycobacterium intracellulare*
- Advantages
  - Fully automated system
  - High throughput (up to 144 tests per day)
  - Specimens do not have to be handled manually for post-PCR processing
    - Increases efficiency of procedure
    - Less opportunity for sample contamination
  - Large sample scale
- Limitations
  - Endpoint detection and limited accuracy of quantitative results

14.1.5 Real-time PCR Instruments

- In real-time PCR, product detection and quantitation are based on measurements made during the amplification process. This differs from conventional endpoint PCR, in which products are detected in separate steps following the completion of amplification. The general features of real-time PCR, as well as its applications and advantages, are discussed
- Real-time PCR methodologies rely on the use of fluorescent reporters to produce detectable signals, the intensity of which is quantitatively related to amplicon production. Fluorescent molecules employed in real-time PCR are described in the Methodology chapter and include SYBR green, hybridization probes, and hydrolysis probes (e.g., TaqMan® probes). Real-time PCR instruments generally have specific calibration protocols specified by the manufacturers

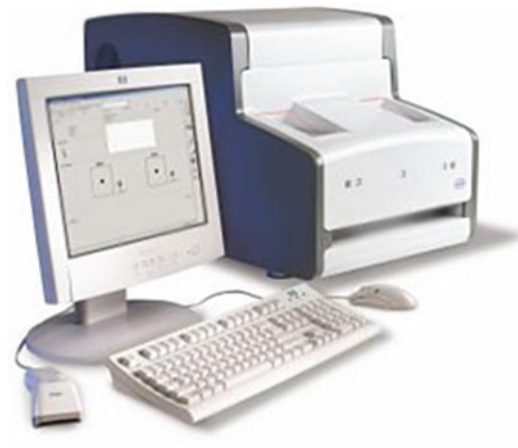


Fig. 14.11 COBAS TaqMan 48 Analyzer (Source: Operator's Manual, COBAS TaqMan 48 System)

14.1.5.1 COBAS TaqMan 48 Analyzer (Roche Instrument Center AG, Switzerland)

- General information
  - The COBAS TaqMan 48 Analyzer is a homogeneous real-time PCR-based system for the amplification, detection, and quantitation of DNA or RNA from clinical specimens (Fig. 14.11)
  - Detection is based on hydrolysis probes which exploit the inherent 5′ exonuclease activity of Taq polymerase. The mechanism of TaqMan® probes is briefly summarized below (see Chap. 8, Methodology for a detailed description)
    - A target-specific oligonucleotide probe conjugated with a fluorescent reporter dye and closely adjoining quencher dye (dual-labeled fluorescence probe) hybridizes with a target DNA sequence or internal standard sequence between the forward and reverse primers during the annealing phase of PCR. The probe is blocked at the 3′ end to prevent extension
    - As a consequence of fluorescent resonance energy transfer (FRET) between the reporter and quencher dyes, the intact probe produces little fluorescent signal. However, during

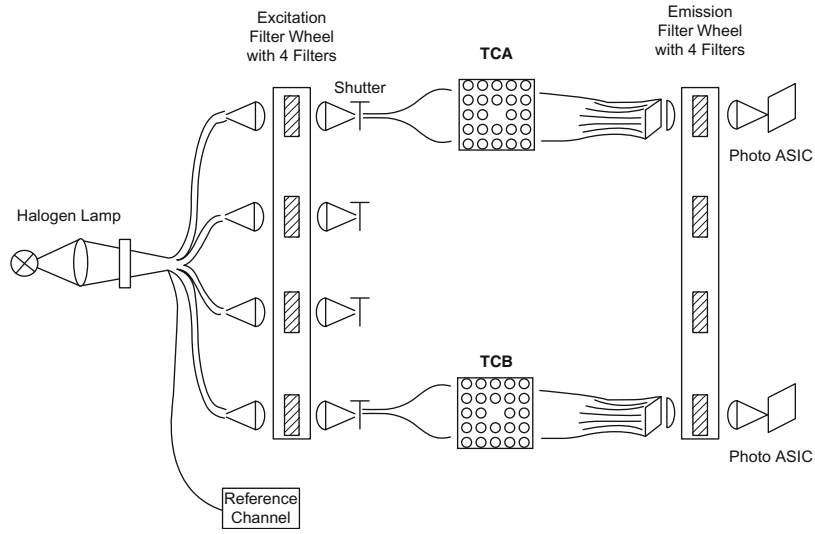


Fig. 14.12 Optical pathways of the COBAS TaqMan 48 System (http://molecular.roche.com/instruments/Pages/COBASTaqMan48Analyzer.aspx)

PCR amplification, the probe undergoes hydrolysis due to Taq polymerase 5′exonuclease activity; this causes separation of the reporter dye from the quencher dye and a corresponding increase in fluorescence
    - Fluorescence intensity progressively increases with subsequent cycles of target amplification. With successive measurements of fluorescence intensity during each annealing phase, an amplification curve is produced
    - The original product concentration can be determined on the basis of comparing cycle threshold of the target sequence with that of an internal standard
  - Inclusion of uracil-N-glycosylase (AmpErase®) and dUTP as components of the PCR master mix ensures selective amplification of the target nucleic acid
  - TaqMan methodology may be used in the quantification of RNA; however, since RNA is not an efficient substrate for Taq DNA polymerase, amplification must be preceded by a reverse transcription step in order to generate a cDNA sequence from the target RNA strand
- Principles of operation
  - Thermocyclers: the COBAS TaqMan 48 Analyzer® utilizes two independently controlled thermocyclers (TCA and TCB), each of which can process 24 samples in simultaneous independent runs (maximum of 48 samples, total)
    - During amplification, each specimen is contained within a "kinetic tube" (K-tube)
    - K-tubes are fitted in a holder (K-carrier) which is loaded into the thermocycler
  - Excitation signal (Fig. 14.12)
    - White light is produced by a tungsten halogen lamp
    - Light is transmitted through one of four interference filters arranged on a wheel, rotated by a stepper motor
    - Dedicated fiber-optic pathways convey the excitation signal to each sample position, minimizing cross talk
    - A shutter blocks the excitation light during nonreading time (preventing photolysis)
  - Fluorescent signal
    - A second set of fiber optics receives the fluorescent signal
      - These fibers are oriented at a 90° angle from the excitatory fibers to minimize interference from the excitation signal
    - Signal passes through one of four interference filters from a second filter set, also arranged on a wheel

- Signal is transmitted to one of two photo ASICS (amplification selective integrated circuits), each of which is comprised of 36 silicon photodiodes, of which 24 are utilized (one corresponding to each sample position within each thermocycler unit). Features of photo ASICs
  - Read all K-tubes simultaneously
  - Optical range – 490–730 nm
  - Large dynamic signal range
  - Tolerate stronger signals than photomultiplier tubes
- Signals from photo ASICs are integrated and amplified

– Monitoring of light source and background
- Intensity of light source is monitored by a reference channel
- When samples are not illuminated, data is collected for background (dark reading)
- Sample readings are corrected for instrument fluctuations
  - Fluorescence = (light reading–dark reading)/reference channel reading

– Multiplex PCR reactions
- Amplification and detection of multiple target sequences in the same K-tube can be performed
- Detection is achieved using multiple probes, one specific to each target sequence, each labeled with different dyes
- Dyes must be chosen carefully to minimize spectral overlap
- Fluorescent signals can be separated using different filter combinations
- Multiplex capacity allows for measurement of signal from internal standard

– Results calculation
- Precheck
  - Raw data acquired
  - Baseline slope corrected
  - Spikes removed
  - Step correction
  - Assigned fluorescence level determined
- Assigned fluorescence level (AFL) = critical threshold line: level of detection at which a reaction reaches statistically significant increase in fluorescence over background
- Threshold cycle (Ct value, crossing point, or "elbow"): fractional cycle number at which fluorescence reaches AFL
  - Ct value indicates beginning of exponential growth phase
  - Ct value is used for titer calculation
  - Assuming 100% amplification efficiency, a tenfold change in concentration changes Ct value by 3.3 cycles
- Internal quantitation standard (QS)
  - A synthetic construct of DNA or RNA designed to closely resemble the length and sequence content of the actual target and therefore amplify with the same efficiency as the target
  - Has sufficient sequence differences from the target such that it hybridizes to a separate specific probe with a distinctive fluorescence signal
  - Incorporated into each sample in a precisely known quantity in TaqMan tests designed to quantify target nucleic acids
  - Carried through the sample preparation and amplification/detection steps along with the target nucleic acid sequence
  - The difference between Ct values of QS and target used is in the determination of target quantity
  - Can also correct for instrument, chemistry, and sample variances
- Five possible results
  - No target detected
  - Titer < titer minimum
  - Sample within dynamic range of assay (titer given)
  - Titer > titer maximum
  - Invalid result

- Applications
  - Detection of viral, bacterial, and parasitic pathogens
  - Determination of viral DNA copy number

- Quality control/assay validation
- Quantitation of gene expression
- Genotyping/detection of single nucleotide polymorphisms using allele-specific probes
- Verification of microarray results

• Advantages
- See above for advantages of real-time PCR over conventional PCR

• Limitations
- High cost of instrument and reagents in comparison with conventional PCR

14.1.5.2 LightCycler 1.0 (Roche Applied Science, Indianapolis, Indiana, USA)

• General information
- The LightCycler is a real time PCR instrument with 3 fluorescence detection channels. This system is compatible with SYBR Green 1 experiments, as well as single or dual color hybridization probe experiments. The LightCycler instrument consists of a thermocycler component and a fluorimeter component

• Principles of operation
- Thermocycler
  - Samples and reagents are contained within glass capillaries which are loaded into a carousel; the carousel fits into a thermal chamber
  - The carousel can accommodate up to 32 capillaries, thus the instrument can process a maximum of 32 samples per run
  - Reaction temperature is regulated by circulation of heated or ambient air through the thermal chamber; air temperature is determined by the voltage across a heating coil
  - The temperature is graphically displayed with LightCycler software
  - An autocorrection function compensates for differences in heat capacity between air and water
  - A small reaction volume (10–20 uL) and very high surface area to volume ratio of capillaries ensures rapid thermal transfer, allowing for fast temperature transition times, with about 15–20 s required for each cycle. Use of air as the heat transfer medium also facilitates high speed cycling
- Fluorimeter
  - Excitation signal
    - Energy source: blue light emitting diode (LED)
    - Signal passes through interference filter
    - Median wavelength of excitation signal: 470 nm
    - Fluorescein absorbance peak (maximum excitation) = 493 nm
    - SYBR Green 1 absorbance peak (maximum excitation) = 497 nm
    - Signal is focused onto individual glass capillaries as they are sequentially positioned over the fluorimeter optics by rotation of the carousel
  - Fluorescence detection
    - Dichroic mirrors divert light emitted from sample into one of three detection channels
    - Each channel is equipped with an interference filter with specific bandpass

      Channel Bandpass – Emission maxima of relevant dyes

      Channel 1:530 +/– 20 nm SYBR Green1(521 nm) Fluorescein (525 nm)

      Channel 2:640 +/– 20 nm LightCycler-Red 640 (640 nm)*

      Channel 3:710 +/– 20 nm LightCycler-Red 705 (705 nm)*

      *LightCycler-Red 640 and LightCycler-Red 705 are FRET partners of fluorescein
    - Signal is received by photohybrid type detectors
    - 32 capillaries can be measured in approximately 5 s
- Fluorescence acquisition modes
  - Single: fluorescence is measured once per sample at end of selected temperature segment

  - SYBR Green 1 format: measurement at end of elongation phase
  - Hybridization probes format: measurement at end of annealing phase
- Continuous: fluorescence of all samples is measured continuously from first sample to last
  - Used for melting curve analysis
- Step: fluorescence of all specimens is measured between stepwise changes in temperature
  - Used for melting curve analysis

- Dual color detection
  - Simultaneous detection of two target sequences in one sample
  - Two different acceptor dyes (LC-Red 640 and LC-Red 705)
  - Fluorescein serves as the FRET donor dye for both LC-Red 640 and LC-Red 705
  - Dual color detection is used in the analysis of internal controls, interpretation of duplex PCR runs, and extended mutation analysis
  - Crosstalk between channels occurs due to overlap between emission spectra of LC-Red 640 and LC-Red 705
    - Crosstalk is corrected by color compensation function
- SYBR Green 1
  - Fluorescence is measured at the end of each elongation phase
  - Melting curve analysis increases sensitivity and specificity
- Sequence specific hybridization probes (introduced in section III)
  - Very sensitive and specific
  - Can detect single copy sequences in complex DNA templates
  - Two separate dye-labeled probes are used
  - The upstream probe has a fluorescence energy donor dye bound to the 3′ end (Fluorescein)
  - The downstream probe has an acceptor dye bound to the 5′ end (LC-RED 640 or LC-RED 705)
  - While in solution, the distance between the probes disallows energy transfer between them
  - When both probes hybridize with the target in head to tail fashion, the donor and acceptor dyes are brought into close proximity permitting energy transfer from donor to acceptor dye (FRET), with emission of a specific fluorescent signal from the acceptor dye
  - Measured fluorescence signal is proportional to the amount of product present in the reaction
  - Fluorescence signal is acquired with each annealing phase
  - Probes are displaced by Taq polymerase during elongation
  - The 3′ end of the acceptor probe is phosphorylated to prevent extension
- Determination of copy number
  - A standard curve is used in the determination of original copy number
    - The standard curve can be generated on the basis of standards, of known concentration, included with each run
    - Alternately, one can import an external standard curve produced in a previous run
    - Standard curve data points derive from the crossing points of each standard, which are plotted against the original copy number. (The "crossing point" is the first cycle at which fluorescence measurement significantly exceeds background level; this occurs in early log phase amplification)
    - The crossing point of a sample is related to sample concentration, as determined by the standard curve
    - The higher the original copy number, the lower the crossing point value
  - The standard curve is calculated by either fit points method or second derivative maximum method
    - The second derivative maximum method is preferred for most runs
    - The fit points method is preferred for runs with few standards or irregular standards

  - Melting curve analysis
    - Useful for:
      - Product identification (SYBR Green 1 experiments)
      - Identification of unwanted byproducts (e.g., primer dimers)
      - Mutation detection in hybridization probe experiments
      - (A single point mutation can drastically alter melting temperature)
      - Distinguishing between wild type, homozygous mutants and heterozygotes
    - Two separate mutation sites can be analyzed in one reaction using a dual probe assay
    - A polynomial, rather than linear method is recommended for all melting curve analyses
- Procedure modifications
  - Carryover prevention with uracil DNA glycosylase (UNG)/AmpErase®
    - UNG and dUTP can be incorporated into reaction mix to minimize carryover of contaminants from previous reactions
  - Hot start
    - Minimizes primer dimer formation
    - Kits are available containing heat-activated FastStart® Taq DNA polymerase
    - Anti taq polymerase antibody can also be employed in the hot-start technique
  - RT-PCR
    - With the appropriate kit, one-step RT-PCR can be performed using SYBR Green 1 or sequence specific hybridization probes
- Applications
  - Kits for diagnostic use
    - Factor V Leiden (Roche Diagnostics)
    - Prothrombin mutation (Roche Diagnostics)
  - Kits for research use: numerous kits are available from Roche Diagnostics for research applications in microbiology, oncology, hematology, and pharmacogenetics, including assays for detection of single nucleotide polymorphisms (SNPs)
- Advantages
  - See above for advantages of real time PCR over conventional PCR
  - Short run time (20–30 min)
  - Melting curve analysis in hybridization probe assays is exquisitely sensitive for detection of single base pair mutations
- Limitations
  - High cost of instrument and reagents in comparison with conventional PCR
  - SYBR Green 1: when there is a low number (1–100) of target sequences, signal is barely measurable above background
  - False positive due to nonspecific amplification

14.1.5.3 LightCycler 2.0 (Roche Applied Science)

- General information
  - The LightCycler 2.0 Instrument (Fig. 14.13) is an upgraded version of the original LightCycler Instrument with several added features
    - Six (vs three) detection channels (530 nm, 560 nm, 610 nm, 640 nm, 670 nm, 705 nm)
    - Ability to accommodate either 20-ul capillary tubes or 100-ul capillary tubes
    - Improved fan design for efficient heating and cooling of larger (100 ul) capillary tubes with minimal increase in run time
    - Upgraded software

14.1.5.4 Roche LightCycler 480 (Roche Applied Science)

- General information
  - The LightCycler 480 real-time PCR system (Fig. 14.14) is a fully integrated multi-well plate-based real-time PCR platform for accurate qualitative and quantitative analysis of gene expression and genetic variation
- Principle of operation
  - Thermal block cycler, i.e., heating and cooling, is achieved using Peltier elements and ensures optimal heat distribution across the plate and, as a consequence, optimized

Fig. 14.13 LightCycler® 2.0 (http://www.roche-applied-science.com/proddata/gpip/3_8_1_3_1_2.html)

Fig. 14.14 LightCycler® 480 (https://www.roche-applied-science.com/sis/rtpcr/htc/index.jsp)

well-to-well homogeneity as well as maximized interwell assay reproducibility
- The 96- and 384-well thermal block cyclers are easily interchangeable
- The choice between multiple filter combinations for excitation and emission enables optimized excitation as well as a reduction of cross talk between different dyes in the reaction mix
- The filters are placed in two filter wheels and offer five different wavelengths for excitation (440, 465, 498, 533, and 618 nm) and six for detection (488, 510, 580, 610, 640, and 660 nm) (Table 14.2)
- Different fluorescent dyes and probes can be used including SYBR Green I (for quantification, melt curve genotyping, and endpoint genotyping), hydrolysis probe (for quantification and endpoint genotyping), and hybridization probes (for quantification, melt curve genotyping, and endpoint genotyping)

- Application
  - Applications include gene quantification and genetic variation analysis
    - Absolute quantification (using external standards) and relative quantification/gene expression analysis (using internal gene)
    - Melt curve genotyping
      - Differentiation of SNPs by high-resolution melting
      - High-throughput haplotyping of adjacent SNPs by melting curve analysis with one HybProbe probe
    - Endpoint genotyping
      - Automated scatterplot analysis
    - Mutation scanning
      - Amplicon melting to screen for heterozygotes
      - Homozygote differentiation by amplicon melting
      - High-resolution melting analysis resolves complex patterns of genetic variation
      - Probe-free diplotyping of two adjacent SNPs
      - Diplotyping of adjacent SNPs using an unlabeled probe

Table 14.2 The LightCycler 480 real-time PCR system

Excitation filters		440	465	498	533			618
Emission filters		488	510	580		610	640	660
Dye		LightCycler Cyan550	SYBR ResoLight	Fluorescein FAM	HEX (VIC)	LightCycler Red 610	LightCycler Red 640	Cy5
Detection formats	Melting curve		*					
	HRM		*					
	SimpleProbe probes			*				
	HybProbe probes				*		*	*	*
	Hydrolysis probes 1–3 colors			*		*			*
	Hydrolysis probes 4 colors	*			*		*		*

- Advantages
  - Rapid cycling ensures completion of a PCR run within 40 min
  - Six colors for multiplex detection

14.1.5.5 GeneXpert® Dx System (Cepheid, Sunnyvale, California, USA)

- General information
  - The GeneXpert® Dx System (Fig. 14.15) integrates automated sample preparation and real-time PCR on one platform. All extraction, purification, amplification, and detection processes are performed within a single disposable cartridge (Fig. 14.16). This high level of integration permits largely hands-off operation with extremely rapid turnaround times. The system includes the GeneXpert instrument, personal computer, and preloaded software
- Principles of operation
  - Instrument
    - The instrument contains multiple (up to 16) "amplification modules"
      - Each module is independently controlled
      - Each module contains a dedicated fluorimeter with four excitation channels and four detection channels
      - Multiplexing capability
      - Each module performs continuous optical monitoring
      - Reaction is automatically stopped when target is detected
      - Shortens time to results
  - Cartridges
    - Are self-contained, single use
    - Can handle a range of volumes
    - Cartridges contain PCR reagents
      - Primers, probes, dNTPs, polymerase, and buffer components
    - Additional reagents must be introduced prior to use, e.g., washing, elution, and lysis reagents
    - Components of each cartridge
      - Syringe drive
      - Rotary drive
      - Sonic horn
      - Produces ultrasonic energy which lyses cells in raw specimen
      - Processing chambers
      - Contains reagents and filters and captures mechanisms for washing, purifying, and concentrating nucleic acids
      - Reaction chamber (for amplification and detection)
      - Thin chamber facilitates efficient heat exchange for rapid thermocycling
      - Optical windows permit real-time four-color detection
- Applications
  - FDA-cleared assays
    - Bacteria: Group B *Streptococcus*, MRSA, *C. difficile*, and van A

Fig. 14.15 GeneXpert® Dx System. The GeneXpert® instrument (right), shown with computer and barcode scanner. (Source: www.cepheid.com)

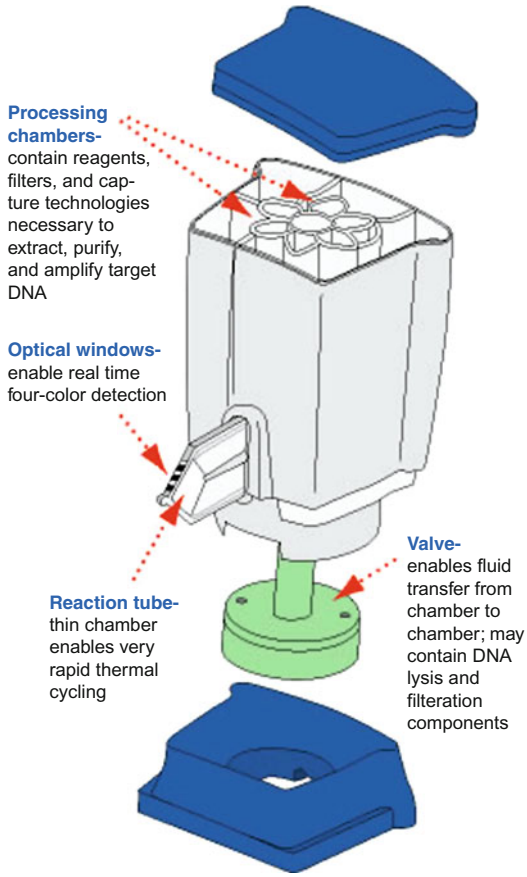


Fig. 14.16 Disposable cartridge for use with the GeneXpert® System. (Source: www.cepheid.com)

    - Virus: Enterovirus and influenza virus panel
    - Factors II and V assay
  - Many ASR products are also available
- Advantages
  - Rapid turnaround time (test results from raw sample may be available in <30 min)
  - Minimal effort required by operator
  - Cross contamination is effectively eliminated, as each sample is contained within a separate sealed disposable cartridge during all phases of preparation and amplification/detection
  - Each unit of the GeneXpert instrument is independently controlled; thus, different protocols can be run simultaneously
  - Extremely sensitive (capable of single-cell detection)
  - Instrument has small footprint and minimal power requirement, making it relatively portable
- Limitations
  - Cartridges are expensive and not reusable
  - Relatively few samples can be processed simultaneously

14.1.5.6 Fluidigm Dynamic Array System

- General information
  - Fluidigm Dynamic Array IFC chip was initially developed at California Institute of Technology. The fabrication process is called multilayer soft lithography (MSL) which uses rubber that deflects under pressure to create an effective seal. More importantly, the structures made from this material are so small that tens of thousands of them can be integrated into a dense

network of channels for regulating solutions on a micro-, nano-, or even picoliter scale. The elegantly simple valve is the foundation of Fluidigm's various integrated fluidic circuits
- Principle of operation
  - The valve consists of a membrane that deflects under pressure to pinch off the flow of fluids in a microchannel. The valve is made from two separate layers of elastomeric rubber that have been placed on a micro-machined mold. By bonding the layers together, the recesses form channels and chambers in a rubber chip
  - When pressurized gas is applied to the channels in one layer of the chip, the rubber deflects at precisely the intersection of the channels in the bottom layer. This constitutes a simple, effective valve
  - Several platforms available include BioMark™ HD System which is real-time PCR system to run up to 192 samples and EP1 System which is an endpoint PCR system to run large numbers of sample at low to mid multiplex for SNP and CNV
  - The Dynamic Array IFC chip has 12–96 samples and can carry out up to 9216 (96 × 96) reactions
- Application
  - The applications include gene expression, single-cell gene expression, SNP genotyping, targeted resequencing, sample quantitation, copy number variation, and protein crystallization
- Advantages
  - Uses as little as 1 pg of material, 5–10-nl volume, and 5–8-ul input per well
  - Reusable chip formats have also been developed to dramatically decrease costs and increase throughput
  - Digital Array integrated fluidic circuit (IFC) partitions a sample, premixed with PCR reagents, into hundreds of individual PCR reactions
  - The PCR product can be harvested out of the Access Array IFC for downstream analysis, such as sequencing

14.1.6 Microarray Platforms

- DNA microarrays are of two basic types. cDNA arrays, currently used only in research, contain ssDNA probes ranging in length from 0.6 Kb to 5 Kb, each of which may encompass an entire gene, a partial gene, or an expression sequence tag (EST). A single slide may contain from 30,000 to 50,000 different probes. By contrast, oligonucleotide arrays utilize much shorter DNA probes and have recently come into use in the molecular diagnostics laboratory. The details of oligonucleotide microarray systems are described below
- General features of oligonucleotide microarrays
  - Each chip contains thousands of different oligonucleotide probes at specified locations
    - Each location contains millions of copies of a single probe
    - Oligonucleotide probes are 25 nucleotides in length, "25-mer"
  - Specificity is enhanced by combining probes to form probe sets
    - Each set contains 16 to 25-mers
    - Probes within a given probe set are specific to different locations on a single target gene
    - Each set of perfectly matched (PM) probes for a given target has a corresponding set of mismatch (MM) probes. A "probe pair" is the combination of a PM probe with its corresponding MM probe
      - MM probes contain a single mismatched base in the middle of the sequence (usually the 13th position)
      - At low levels of target concentration, MM probes show greater sensitivity to changes in target concentration than do PM probes
      - MM probes serve as internal controls for nonspecific binding and background noise

Fig. 14.17 Affymetrix GeneChip System 3000Dx (GCS 3000Dx) (Source: http://www.affymetrix.com/products/application/clinical_applications.affx)

14.1.6.1 Affymetrix GeneChip® System 3000Dx (Affymetrix, Santa Clara, California, USA)

- General information
  - The Affymetrix GeneChip System 3000Dx (GCS 3000Dx), pictured below (Fig. 14.17), is a microarray instrument FDA cleared for use in clinical diagnostics. The system includes all necessary components for processing of Affymetrix oligonucleotide chips and consists of a scanner, autoloader, fluidics station, computer workstation, and operational software. The autoloader can hold up to 48 chips for walkaway performance. The fluidics station performs washing and staining functions; it contains four modules and thus can process four chips simultaneously. Currently, an FDA-approved cytochrome P450 chip is available from Roche Diagnostics (AmpliChip CYP450 Test) for use with the GCS 3000Dx System. Chips designed for other applications are in development. The system supports DNA analysis, expression analysis, and resequencing applications
- Principles of operation
  - Includes three components
    - GeneChip® Fluidics Station 450Dx v.2
    - GeneChip® Scanner 3000Dx v.2 with AutoLoader
      - Light source: solid-state green laser
      - Excitation wavelength: 532 nm
      - Detector: meshless photomultiplier tube
      - Capable of genotyping up to 500,000 SNPs
    - Workstation with Affymetrix® Molecular Diagnostic Software
- Applications
  - Clinical diagnostics
    - Cytochrome P450 genotyping
    - Additional tests are in development
      - Diagnosis of leukemias
  - Research
    - Expression profiling
    - DNA analysis
      - Whole-genome scanning (GeneChip® Human Genome U133 Plus 2.0 Array)
      - Cancer genetics research and analysis of loss of heterozygosity (GeneChip® Human Mapping 250 K NSP Array)
      - Copy number variation analysis (Affymetrix® Cytogenetics Whole-Genome 2.7 M Array)
    - Pathogen subtyping
      - Analysis of genes affecting pathogenicity and drug resistance
    - Pharmacogenetics
      - Analysis of molecular basis for variations in drug response (Affymetrix® DMET™ Plus Array)

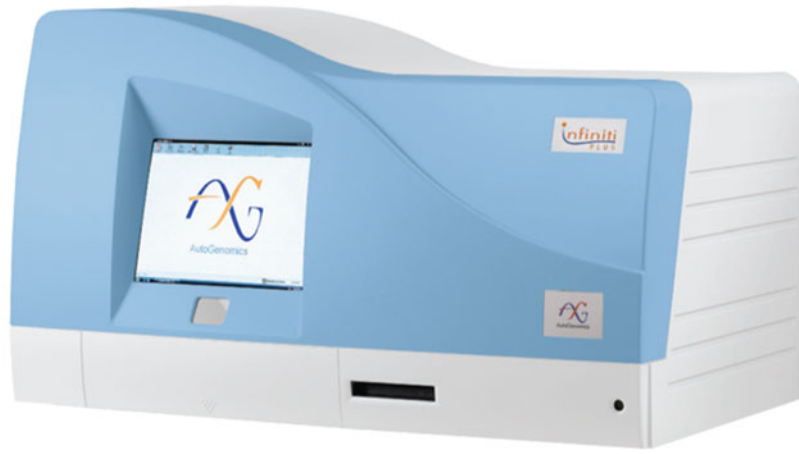


Fig. 14.18 INFINITI Plus Analyzer (http://www.autogenomics.com/infiniti_plus.php)

- Advantages (of Affymetrix oligonucleotide arrays over cDNA arrays)
  - Easy to use and has an integrated system for rapid adoption of both RNA and DNA applications
  - The most cost-effective approach for multiple applications through a single, flexible system
  - Fully integrated analysis and reporting software for a seamless interface
  - Higher density of probe pairs than in cDNA arrays; thus, more genes can be assayed on a single chip
  - Multiple independent measurements for each transcript enhance sensitivity and specificity
  - Commercially produced disease-specific oligochips are available
- Limitations
  - Only one sample can be run on one chip at one time
  - Costly

14.1.6.2 AutoGenomics INFINITI® Plus Analyzer (AutoGenomics Inc., Vista, California, USA)

- General information
  - The INFINITI® Plus Analyzer (Fig. 14.18) utilizing BioFilmChip® microarrays integrates all the discrete test processes in DNA analyses, such as sample handling, reagent management, hybridization, stringency, and optical detection into a totally self-contained system. Requiring no plumbing or tubing, the analyzer has minimal biohazardous liquid waste. It is network compatible and can be readily interfaced to laboratory information systems. The system is capable of performing qualitative determinations for a wide range of genetic and proteomic applications. The analyzer is capable of genotyping up to 240 SNPs per microarray
- Principles of operation
  - The INFINITI® Plus Analyzer has the following components
    - Optical detection – narrow bandwidth LED
    - Detector – high-resolution CCD image sensor
    - BioFilmChip® Microarray Magazine
    - Analyzes automatically
      - Performs detection primer extension
      - Hybridizes sample to microarrays
      - Washes microarrays
      - Interrogates microarrays using onboard optic system
      - Performs data analysis
  - PCR product is loaded onto the analyzer with assay-specific reagents (Intellipack® Reagent Module) and microarrays (BioFilmChip®)
    - The instruments can assay to 48 samples
    - Determines SNPs/markers present based on relative fluorescence
- Applications
  - Genetic disorders

    - Factor II (IVD), factor V Leiden (IVD), and MTHFR
    - CFTR-31
    - Ashkenazi Jewish Panel
    - FMF Panel
    - ApoE
  - Pharmacogenetics
    - Warfarin assay (IVD)
    - CYP450 2C9-VKORC1, CYP2C19 (IVD), CYP450 2D6, CYP450 3A4, and CYP450 3A5
    - TAMX3
    - MDR1
  - Oncology
    - HPV genotyping
    - KRAS–BRAF
    - EGFR
    - 5-FU
    - UGT1A1
    - NAT-2
    - CHEK-2
    - Breast cancer panel
  - Infectious diseases
    - RVP Plus
    - STD-6 Panel
    - CT–NG QUAD
    - UroGen QUAD
    - Bacterial vaginosis QUAD
    - Candida vaginitis QUAD
    - MDR-TB
    - NTM
- Advantages
  - All discrete test processes such as reagent management, sample handling, all fluidics, hybridization, stringency, optical detection, data reduction, results analysis, and reporting in one workstation
  - Capable of performing both genomics and proteomics on the same platform
  - Random access capability enabling multiple genomic or proteomic tests to be processed at the same time
  - Multiple methodologies on same platform: SNP, STR detection, and gene expression
  - "Load and go" automation: load sample plate and reagents and then go with no manual intervention
  - Data analysis and results formatting for easy interpretation
  - Benchtop plug-and-play platform with no tubing or plumbing, minimizing liquid waste
  - Runs multiple tests in the same run
  - Automated throughput yielding rapid turnaround with the first test result within 3 h and 48 results within 5 1/2 h
  - BioFilmChip® microarray provides ability to assess multiple analytes on one microarray
  - QUAD microarray processes 4 samples on 1 BioFilmChip
  - BioFilmChip provides optimal spot morphology with low intrinsic fluorescence
  - Replicate testing on each microarray ensures integrity of results
  - Accurate dispensing of samples and reagents with built-in sonic sensors
  - Built-in controls for PCR, ASPE, and hybridization
  - Spot-to-spot scanning using built-in confocal microscope yields high analytical sensitivity
- Limitations
  - Sample preparation must be performed prior to use of the INFINITI® Plus
  - Preamplification of the target is necessary
  - Currently not optimized for quantitative analyses

14.1.6.3 Luminex xMAP Technology (Luminex Corporation, Austin, Texas, USA)

- xMAP technology is a method suitable for high-density, high-throughput diagnostic applications. Polystyrene microspheres, each measuring 5.6 μ in diameter, contain two different fluorophores in varying proportions in order to produce 100 different microsphere sets, each with a unique fluorescence color signature. Microspheres of a given color set are bound to a specific probe (e.g., oligonucleotide, antibody, receptor molecule, or peptide). After binding with a fluorescently labeled analyte, the microspheres are interrogated by lasers as they pass single file through

Fig. 14.19 Luminex 100 IS System (http://www.luminexcorp.com/Products/Instruments/Luminex100200/index.htm)

a flow cell, much akin to the mechanism of flow cytometry

- In one format, multiplex PCR is performed, followed by primer extension using biotinylated dCTP. The primers used in the extension reaction are allele specific, and each is labeled with a 5′ tag sequence that does not hybridize to the target DNA. The extension product is then hybridized to oligonucleotide probes connected to microspheres. The oligonucleotides on microspheres of a given color all contain the same "antitag" sequence and therefore will only bind to DNA extension product corresponding to one allele. The bound product is then conjugated to streptavidin–phycoerythrin. The presence of specific alleles or polymorphisms in the sample material is determined by correlation of the fluorescence signal intrinsic to each microsphere with the presence or absence of a corresponding phycoerythrin signal

14.1.6.4 Luminex® 100 IS System (Luminex Corporation)

- General information
  - The Luminex 100 IS System (Fig. 14.19) is based on the principles of flow cytometry with various modifications to optimize performance with the xMAP microsphere format. These modifications include enhanced signal-processing performance, optimization of fluidics for single-size microspheres, and development of software suitable for handling multiplex analyses with up to 100 analytes. Samples are contained within 96-well plates. A full plate can be analyzed in as little as 1 h
- Principles of operation
  - Fluidics systems
    - A syringe-driven mechanism performs sample uptake and transfer to cuvette
      - Sample uptake volume – 20–200 uL
    - A second fluidic pathway delivers sheath fluid to the cuvette and sample path
  - Optical systems
    - Two solid-state lasers are used for excitation
      - Reporter laser: excites fluorophores bound to analytes associated with the microsphere surface. Wavelength = 532 nm
      - Classification laser: excites fluorophores embedded within each microsphere. Wavelength = 635 nm
      - While passing single file through the cuvette, each microsphere is simultaneously interrogated by both lasers
      - Fluorescence signals are received by multiple channels: reporter channel, two classification channels, and doublet discriminator channel

Fig. 14.20 The FLEXMAP 3D system (http://www.luminexcorp.com/Products/Instruments/FLEXMAP3D/index.htm)

    - Detection is performed by photodiodes and photomultiplier tube
    - Signals are digitized and delivered to a digital signal processor
  - xMAP microspheres
    - See above for general description
    - Calibrator microspheres
      - Used to normalize settings for each channel
    - Control microspheres
      - Used for verification of calibration and optical pathway integrity
- Applications
  - Diagnostic
    - DNA based
      - Cystic fibrosis: use of "liquid bead array" format for rapid detection of 40 different mutations implicated in cystic fibrosis
      - Cytochrome p450: multiplex detection of up to 100 different alleles in cytochrome p450 family enzymes
      - HLA: multiplex detection of up to 100 different HLA alleles
    - Antigen–antibody interaction based
      - Allergy testing
      - Autoimmune disease
      - Identification of infectious agents
  - Research
    - Cancer markers
    - Cell signaling
    - Gene expression profiling
    - Numerous other applications
- Advantages
  - Combines high-density multiplex analysis with high throughput
  - The xMAP format is very flexible. The system is readily customized by using different combinations of microspheres or custom-designing microspheres with specific probes attached
  - The system is not limited to nucleic acid analysis but is also compatible with detection formats based on antigen–antibody interactions, receptor–ligand interactions, or enzymatic reactions
  - Excellent reproducibility – high volume of xMAP microspheres in a given lot facilitates assay standardization
- Limitations
  - Miss calls can result if an unknown or unexpected polymorphism is present at PCR primer or extension primer binding sites
  - Smaller multiplexing capability compared to microarray formats (limited to 50 mutant/wild-type pairs)

FLEXMAP 3D System

- General information
  - The FLEXMAP 3D system (Fig. 14.20), in combination with xMAP® (multianalyte profiling) technology, will simultaneously

Fig. 14.21 MAGPIX (http://www.luminexcorp.com/Products/Instruments/MAGPIX/index.htm)

measure up to 500 analytes from a single sample. xMAP technology incorporates microspheres, proprietary dyeing processes, flow cytometry-based fluidics, lasers, the latest in high-speed digital signal, and advanced computer algorithms, enabling a multianalyte detection system that demonstrates excellent sensitivity and specificity

- Principles of operation
  - A process created by Luminex to internally dye same-size polystyrene microspheres with multiple fluorophores facilitates the creation of 500 distinguishable microspheres when these fluorophores are combined in varying concentrations
  - The microspheres, discriminated by color, are excited by two lasers in the FLEXMAP 3D instrument. The resulting emission is detected by avalanche photodiodes (APDs) in three classification channels (CL1, CL2, and CL3) that is then further analyzed using a separate APD in a doublet discriminator (DD) channel, which measures bead size through side scatter
  - Analytes are bound to xMAP microspheres using the same surface chemistry used with earlier Luminex instruments (LX100/200)
  - Reporters, tagged with fluorescent labels excited at a different wavelength than the internal dyes, bind to the analyte of interest and are detected by a photomultiplier tube (PMT) in a reporter channel (RP1), allowing for quantitative analysis
  - As the microspheres in a fluid stream pass rapidly through the laser beams, high-speed digital signal and computer algorithms discriminate which analyte is being carried on each microsphere and quantify the reaction based on fluorescent reporter signal
  - The results are analyzed by the system software and presented in a readable format for analysis
- Applications
  - Can be designed to perform multiplex reactions/detection. Uses include
    - Genetic testing
    - Oncology testing
    - Infectious disease testing
- Advantages
  - Runs 500-plex, MicroPlex, and MagPlex beads
  - 96-well plate runs in 20 min; 384-well plate runs in 1 h and 15 min
  - Runs 96- and 384-well plates
  - Greater sensitivity
  - Greater dynamic range
  - Autoadjusting sample probe and piercing probe
  - Easy to connect to automation and liquid handlers
- Disadvantages
  - More expensive

Fig. 14.22 Applied Biosystems 3730xl DNA Analyzer (Source: https://products.appliedbiosystems.com)

MAGPIX

- General information
  - MAGPIX (Fig. 14.21) combines a fluidics system, a mechanical system, an electronic system, and an optical system with magnetic microspheres and complex computer analysis to perform multiplex assays
- Principles of operation
  - An operator places a 96-well microtiter plate on the plate carrier, which transports the plate into the instrument
  - The sample probe descends into each well, drawing a sample for testing and drawing drive fluid from the drive fluid container. The sample moves through the fluid tubing to the optic module, transported by the drive fluid
  - In the optic module, a magnet holds the magnetic microspheres in place while first a red (classification) LED and then a green (reporter) LED illuminate them. They are imaged during each illumination
  - After the images are recorded, the magnet withdraws, releasing the microspheres for transport to the waste fluid container and to clear the way for the next sample
  - Ponent software analyzes the images, the red-illuminated images to classify the microspheres and the green-illuminated images to determine what elements of the sample have bonded to their surfaces

14.1.6.5 Capillary Electrophoresis

Applied Biosystems

- General information
  - The general principles of capillary electrophoresis are discussed in Section V (DNA separation methods). Features of representative capillary electrophoresis instruments, the Applied Biosystems 3730 and 3730xl (Fig. 14.22) DNA analyzers, are described
- Principles of operation
  - These instruments utilize large capillary arrays for high-throughput sequencing
    - 3730 uses 48 capillaries; 3730xl uses 96 capillaries
    - Capillary length – 36 cm or 50 cm
  - Capillaries are contained within a temperature-controlled oven (18–70 °C)
  - A built-in robotic loader can automatically load up to 16 microplates
    - Instrument can accommodate 96-well or 384-well plates

  - In-capillary detection is performed by dual-side illumination with argon laser
  - Fluorescence signal is detected by CCD (charge-coupled device)
  - An onboard polymer delivery pump automatically replenishes polymer after each run
  - Software performs data analysis (automated base calling, genotyping)
  - Capillary electrophoresis instruments require both spatial and spectral calibration, using reagents provided by the manufacturer
- Applications
  - De novo sequencing
  - Resequencing
  - Mutation detection
  - SNP genotyping
  - Microsatellite analysis
  - AFLP (amplified fragment length polymorphism) analysis
  - LOH (loss of heterozygosity) detection
  - Methylation analysis
  - RFLP analysis
  - BAC fingerprinting
  - SAGE® (serial analysis of gene expression)
- Advantages (over slab gel electrophoresis)
  - Smaller sample size
  - Highly economical use of reagents
  - Less hands-on time (no need to prepare gel)
  - Shorter run time
  - Longer maximum read length for sequencing applications
  - No possibility for spillover of sample into neighboring lanes
  - No difficulties with lane tracking
  - No artifacts related to inhomogeneity of gel
- Limitations
  - High cost of instrumentation

14.1.7 Next-Generation Sequencing

14.1.7.1 Roche 454 Sequencing Systems (Roche 454 Life Sciences, Branford, Connecticut, USA)

- General information
  - The 454 sequencer (Fig. 14.23) uses pyrosequencing instead of Sanger dideoxy termination sequencing. The complete sequencing workflow of the GS FLX and GS Junior Systems comprises of four main steps, leading from purified DNA to analyzed results. The GS FLX offers sequencing read lengths up to 1 kb, and GS Junior System generates sequencing reads ~400 bp in length. The clonal nature of 454 Sequencing Systems allows unambiguous allele resolution of variation in complex regions of the genome, along with quantitative detection of variants present in less than 1% of a mixture. Perform haplotyping, genotyping, rare variant detection, structural variation detection, copy number variation analysis, and heterozygote calling all using the same platform
- Principle of operation
  - The sample DNA is amplified inside water droplets in an oil solution (emulsion PCR), with each droplet containing a single DNA template attached to a single primer-coated bead that then forms a clonal colony
  - The sequencing machine contains many picoliter-volume wells each containing a single bead and sequencing enzymes
  - Pyrosequencing uses luciferase to generate light for detection of the individual nucleotides added to the nascent DNA, and the combined data are used to generate sequence readouts
  - The basic steps include:
    - Generation of a single-stranded template DNA library
    - Emulsion-based clonal amplification of the library
    - Data generation via sequencing by synthesis
    - Data analysis using different bioinformatics tools
  - Sample preparation
    - Fragment library-based projects:
    - DNA must be double stranded
    - Fragments must be >70 bp
    - If fragments are 800 bp or longer, a minimum of 10 μg of DNA is required
    - For fragments less than 800 bp, a minimum of 2 μg of DNA is required

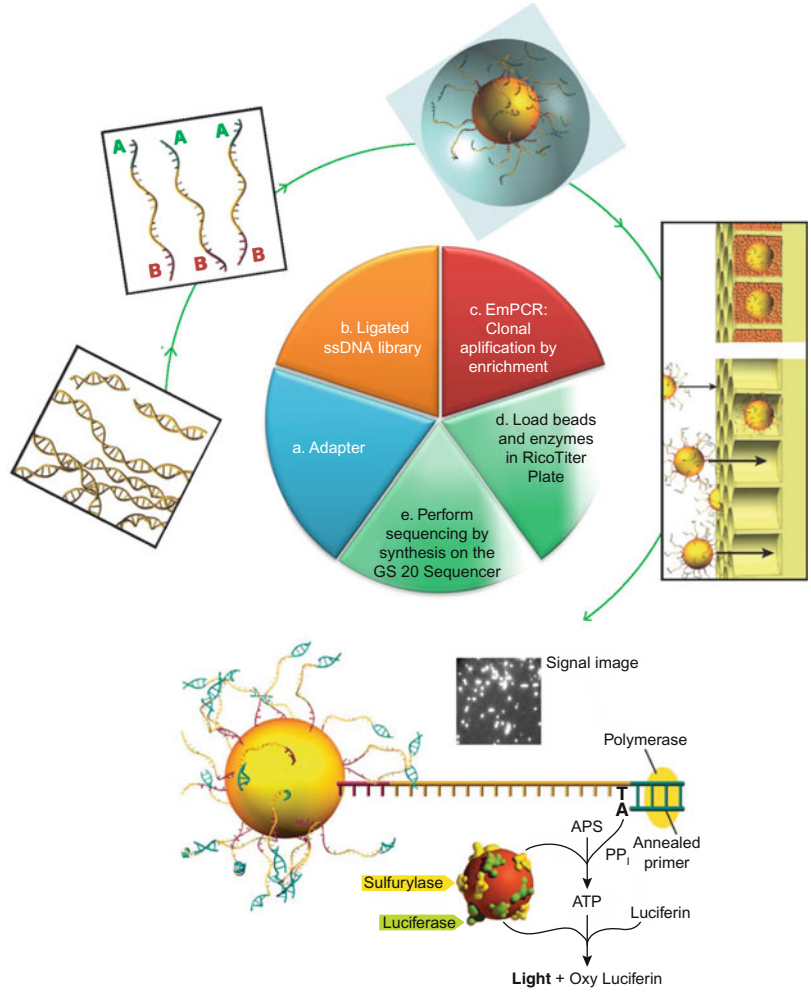


Fig. 14.23 Roche 454 Sequencing Systems (http://www.my454.com/)

- DNA should have an OD260/280 ratio of 1.8 or more
- DNA concentration should be a minimum of 50 ng/μl in TE buffer
- DNA should not be degraded nor should it contain any particulate matter

– Paired-end library-based projects:
- DNA must be double stranded
- Input DNA should be >10 kb
- Minimum of 10 μg of DNA is required
- DNA should have an OD260/280 ratio of 1.8 or more
- DNA concentration should be a minimum of 50 ng/μl in TE buffer
- DNA should not be degraded nor should it contain any particulate matter

– Amplicon library-based projects:
- DNA must be double stranded
- Input DNA should be <800 bp long
- Minimum of 500 ng of DNA is required
- DNA should have an OD260/280 ratio of 1.8 or more
- DNA concentration should be a minimum of 50 ng/μl in TE buffer
- DNA should not be degraded nor should it contain any particulate matter

• Applications
– Whole-genome sequencing

    - Performs straightforward de novo assembly to decode previously uncharacterized genomes
    - Uses shotgun reads alone or in combination with multispan paired-end reads (3 kb, 8 kb, 20 kb) to tackle large complex genome
      - Shotgun sequencing for de novo assembly of genome sequence
      - Paired-end sequencing for longer sequence reads
  - Targeted resequencing
    - Capture target sequence first and then sequence with 454
    - SNP resequencing
    - Structural variation analysis including INDELs, duplications, rearrangements, and chromosomal breakpoints
  - Metagenomics
    - Characterization of complex environmental samples and pathogen discovery
  - RNA sequencing
    - Whole transcriptome sequencing
    - Quantify and profiling novel small RNAs (i.e., profiling miRNA)
  - Methylation analysis
    - Methylation of cytosine–guanine (CpG) dinucleotides in genomic DNA, especially promoter region
- Advantages
  - Longer sequence reads (400–1000 bp), ideal for whole-genome de novo sequencing
  - Limit of detection as low as 1% which can be used as a reference method for validating low-percentage variants and is invaluable for SNP analysis in tumor samples
  - The long read lengths provided by the Roche 454 GS-FLX+ System enable to provide quantitative sequencing of complete transcriptomes
  - Paired-end libraries enable confident mapping of structural variations such as INDELs, duplications, rearrangements, and chromosomal breakpoints
- Limitations
  - Longer hours to complete the process
  - Expensive: several thousand dollars for each large titanium run for 1 sample

14.1.7.2 Illumina's HiSeq™2000 (Illumina Inc., San Diego, California, USA)

- General information
  - Illumina HiSeq 2000 platform is a widely adopted, reversible terminator-based sequencing by synthesis chemistry. HiSeq 2000 delivers a high sequencing output and fast data generation rate. It is able to process larger numbers of samples and to decode larger and more complex genomes. HiSeq 2000 takes advantage of the unique combination of read length, read depth, and flexible paired-end insert sizes which makes HiSeq data ideal for *de novo* sequencing. The system's high raw read accuracy enables confident and efficient production of high-quality, long contig assemblies
- Principle of operation
  - DNA molecules are first attached to primers on a slide and amplified so that local clonal colonies are formed (bridge amplification)
  - Four types of ddNTPs are added, and nonincorporated nucleotides are washed away
  - Unlike pyrosequencing, the DNA can only be extended one nucleotide at a time
  - A camera takes images of the fluorescently labeled nucleotides, and then the dye along with the terminal 3′ blocker is chemically removed from the DNA, allowing the next cycle
- Applications
  - Whole-genome de novo sequencing
  - Transcriptome sequencing (RNA-seq)
  - Small RNA profiling
  - Deep resequencing/targeted resequencing
  - Structural variation analysis
  - Methylation analysis
- Advantages
  - Able to generate transcript profiles in a single day
  - RNA sequencing reads can be aligned across splice junctions to identify isoforms, novel transcripts, and gene fusions
  - Flexible and efficient targeted resequencing solution in combination with

sequence capture or target enrichment options (Agilent Technologies SureSelect, NimbleGen)
- Limitations
  - Expensive instrumentation
  - Complicate data analysis

14.1.7.3 Ion Torrent Personal Genome Machine (PGM™) (Life Technologies)

- General
  - Ion Torrent Systems Inc. developed a sequencing instrument based on using standard sequencing chemistry but with a novel, semiconductor-based detection system. This method of sequencing is based on the detection of hydrogen ions that are released during the polymerization of DNA, as opposed to the optical methods used in other sequencing systems. By combining the speed and scalability of semiconductor technology with a system of measuring pH fluctuations produced during DNA replication in real time, the Ion Torrent is able to directly turn genetic information into digital data with excellent speed and uniformity. A variety of chips are available from 10 Mb to greater than 1 Gb allowing highly accurate sequence data to be generated for 16 samples per run in less than 2 h. With a raw accuracy rate 99.5% and a detection sensitivity of around 5%, the Ion Torrent is an excellent platform for SNP detection and mutational analysis
- Principle of operation
  - A microwell containing a template DNA strand to be sequenced is flooded with a single type of nucleotide
  - If the introduced nucleotide is complementary to the leading template nucleotide, it is incorporated into the growing complementary strand. This causes the release of a hydrogen ion that triggers a hypersensitive ion sensor, which indicates that a reaction has occurred
  - If homopolymer repeats are present in the template sequence, multiple nucleotides will be incorporated in a single cycle
  - This leads to a corresponding number of released hydrogens and a proportionally higher electronic signal
- Applications
  - Amplicon sequencing
  - Microbial sequencing
  - RNA sequencing (RNA-seq)
  - Targeted sequencing
  - Mitochondrial sequencing
  - Methylation analysis
- Advantages and limitations
  - Ultrafast, high-throughput workflow
  - Ideal for cross-checking other platforms and rare SNP detection
  - Alternative to RNA sequencing and transcriptome analysis. RNA sequencing reads can be aligned across splice junctions to identify isoforms, novel transcripts, gene fusions, and noncoding RNAs
  - High uniformity of coverage regardless of high GC content and homopolymer stretches
  - Sequence 16 samples per run in less than 2 h

14.1.8 Target Sequence Capture System

14.1.8.1 NimbleGen Sequence Capture (Roche NimbleGen, Madison, Wisconsin, USA)

- General information
  - NimbleGen Sequence Capture technology is designed for the enrichment of selected genomic regions from full complexity human genomic DNA in a single step. Sequence Capture was developed to eliminate the necessity of setting up thousands of PCR reactions, instead allowing for parallel enrichment of target regions in a single experiment. Targeted sequence capture can reduce the cost, improve throughput, and save time for next-generation sequencing
- Principle of operation
  - A set of oligonucleotide probes is used to capture the desired sequences from total human genomic DNA

- These captured sequences are then amplified in a single PCR reaction using common linkers or adaptors, originally attached either to the probes or to the genomic DNA, as primers. (The unwanted sequences are discarded)
- DNA "enriched" for the desired target sequences, which are then sequenced using any of the next-generation sequencing machines
- Two options are available for targeted enrichment of genomic regions
  - Solution-phase methods use probes in solution to capture the target DNA. The specific probes designed to target regions of interest from a sequencing library and an excess of probes over template are used to drive the hybridization reaction further to completion using a smaller quantity of sequencing. In-solution target enrichment can be performed in 96-well plates, using a thermal cycler, so it is more readily scalable than on-array enrichment and does not require specialized equipment:
    - SeqCap EZ Library is a solution-based capture method that enables enrichment of the whole exome or customer regions of interest in a single test tube. An automated solution for SeqCap EZ Library using the Caliper Sciclone NGS workstation can process up to 288 samples per week
    - SeqCap EZ Human Exome Library v3.0 offers the most comprehensive coverage of coding regions with genes from sequence databases plus coverage of 97% Vega, 97% Gencode, and 99% Ensembl databases
    - SeqCap EZ Human Exome Library v2.0 can capture over 30,000 genes and sequence them on a single lane
  - Solid-phase methods capture genomic DNA on probes placed on microarrays. Fragmented total human genomic DNA is ligated to an adaptor sequence and then applied to the microarray surface. The targeted genomic DNA hybridizes to the probes. Then, the hybridized sequences are eluted from the microarray and amplified in a single round of PCR. Finally, they are sequenced:
    - Sequence capture arrays enable researchers to target custom regions of interest in the human genome
    - NimbleGen Sequence Capture 2.1 M and 385 K arrays provide an ideal solution for targeted enrichment of the disease-associated regions. These customizable arrays can be made to target any region or sequence in the human genome and offer a powerful solution for large-scale targeted resequencing studies to significantly reduce time, labor, and cost while improving data quality

14.1.9 Nucleic Acid Quantification

14.1.9.1 nCounter® Analysis System

- General information
  - The nCounter Analysis System (Fig. 14.24) is a digital, midplex nucleic acid quantification platform. It is capable of counting hundreds of individual biological molecules in a single reaction without amplification. It is an integrated system comprised of a fully automated prep station, a digital analyzer, the CodeSet (molecular barcodes), and all of the reagents and consumables needed to perform the analysis
- Principles of operation
  - The technology uses fluorescent "barcodes" and single-molecule imaging for digital detection
  - Each barcode is attached to a single target-specific oligonucleotide complementary to the sequence of interest. By changing the combination of the fluors in the barcode and target oligos, hundreds of unique transcripts can be quantified in a single reaction

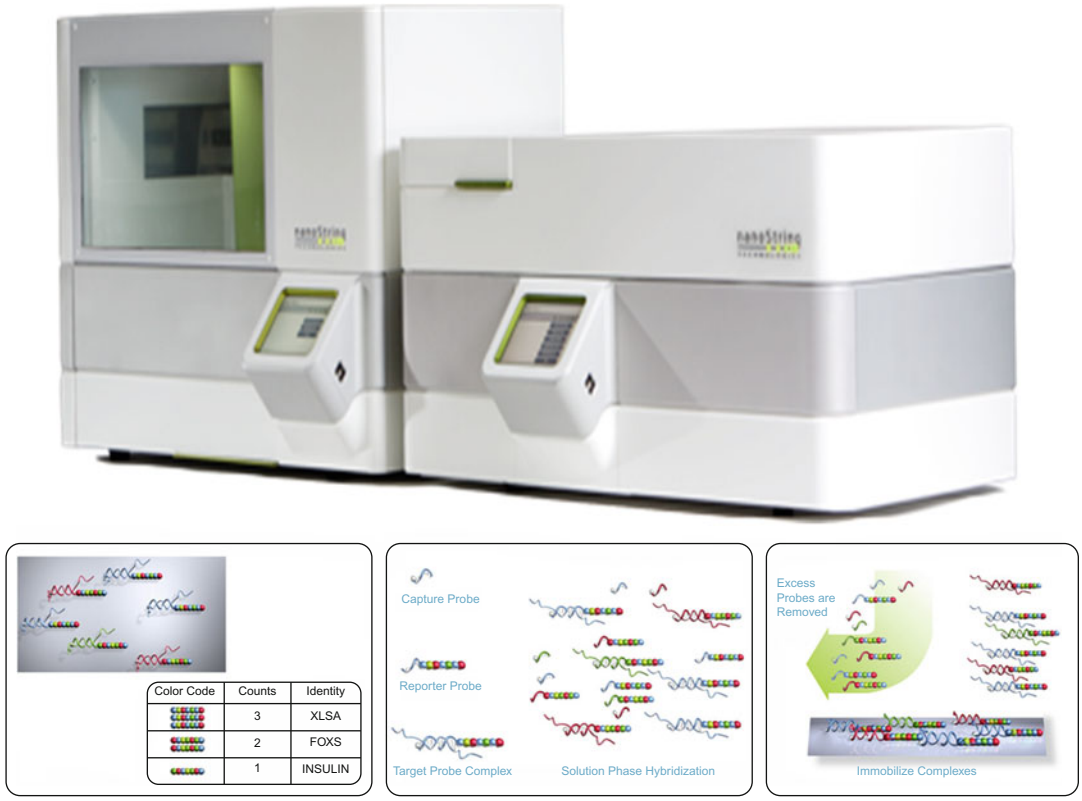


Fig. 14.24 nCounter Analysis System (http://www.nanostring.com/products/ncounter/)

- An assembled CodeSet consists of these probes plus positive, negative, and systemic controls
- Hybridization: NanoString's technology employs two ~50 base probes per target that hybridize in solution. The reporter probe carries the fluorescent barcode signal; the capture probe allows the complex to be immobilized for data collection
- Purification and immobilization: after hybridization, the excess probes are removed and the probe/target complexes aligned and immobilized in the nCounter cartridge to facilitate high-resolution imaging
- Data collection and analysis: sample cartridges are placed in the digital analyzer for data collection. Color codes on the surface of the cartridge are counted and tabulated for each target molecule
  - Data output is a comma-separated text value (.csv) file that is a list of the genes and the number of times the targets were seen (counted) in the sample. Data analysis is therefore very straightforward and can be performed in Excel or by using custom scripts or other software packages

- Applications
  - Gene expression: panel products for specific research areas as well as custom CodeSets manufactured to customer specifications
  - MicroRNA expression: panels with comprehensive coverage of human and mouse microRNA transcriptomes in a single reaction
  - Genomic copy number variation – panel products for specific research areas as well as custom CodeSets manufactured to customer specifications
- Advantages
  - Enables multiplexing of up to 800 targets in a single reaction – it is ideal for studying

focused sets of mRNA transcripts, microRNAs, or CNVs
  - Digital detection generates excellent data quality – nCounter analysis is highly sensitive (~1 copy per cell) and reproducible (interassay CVs <10%), enables fractional fold-change determination down to 1.2-fold, and offers nearly six logs of dynamic range
  - Ease of use – setting up the nCounter hybridization requires 15 min of hands-on time. Four pipetting steps are performed per sample, and the rest of the process is automated
  - Not subject to amplification bias – since nCounter does not rely on any amplification to facilitate detection, it is not subject to amplification bias
  - Minimal sample requirement – 100 ng of starting material allows one to query up to 800 targets
  - Sample type flexibility – nCounter accepts direct input of multiple sample types including total RNA, crude cell lysate, FFPE samples, and whole blood lysate
  - Data analysis is simple. Since the data is simply the number of times a target was counted in the sample, data analysis is very easy
- Limitations
  - Multiplex level (800 targets per reaction) is not of sufficiently high density for many whole-genome discovery applications
  - Requires manual transfer of cartridge from prep station to digital analyzer

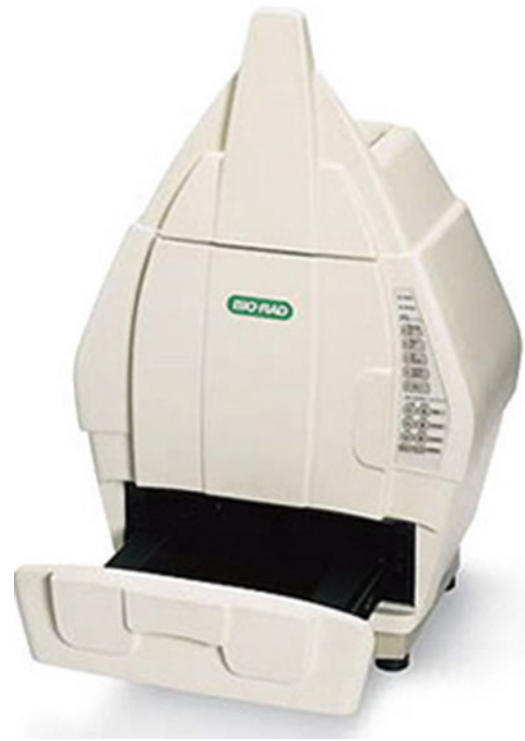


Fig. 14.25 Molecular Imager ChemiDoc™ XRS System with camera shown separately (Source: http://www.bio-rad.com)

14.1.10 Gel Imaging Systems

- A variety of manufacturers offer gel documentation systems with different features and options. Two basic methods of illumination are epi-illumination (light shines down from above the platform) and transillumination (light shines up from below the platform). The illumination source may produce UV light, white light, blue light, or light of other specified wavelengths
- These instruments have a wide range of applications, including visualization of ethidium bromide-stained gels (UV transillumination), visualization of bacterial colonies on an agar plate (white light transillumination), and visualization of bands on TLC plates or other opaque surfaces (epi-illumination). The optimal instrument depends on the imaging requirements of the laboratory. Representative systems are described below

14.1.10.1 Bio-Rad Gel Doc™ EQ, ChemiDoc™ EQ, and ChemiDoc™ XRS (Bio-Rad, Hercules, California, USA)

- General information
  - These gel documentation systems use a CCD camera to acquire images in real time, which allows for accurate positioning and focusing. Images can be analyzed, optimized, and printed using Bio-Rad

Quantity One® software (for Windows or Macintosh). The gel or other material to be visualized is placed within a light-tight hood. Visualization is accomplished using UV or white light illumination. The ChemiDoc™ XRS system is pictured below (Fig. 14.25)

- Principles of operation
  - Camera:
    - The CCD camera has a motorized zoom lens
    - Zoom, focus, and iris functions are remotely controlled
    - Optional lenses are available for low-light applications (e.g., chemiluminescence)
  - Illumination sources and filters:
    - Built-in white light epi-illumination and UV transillumination
    - A standard filter (548–630 nm) for ethidium bromide is provided with the system
    - Optional filters are available for SYBR Green, GFP, SYBR Gold, fluorescein, CY3, rhodamine, SYPRO Ruby, Texas Red, and Hoechst/coumarin
  - Optional light sources are also available:
    - White light transilluminator
    - White light conversion screen
    - 302-nm UV lamp
    - 254-nm UV lamp
    - 365-nm UV lamp
  - PCI digitizing card (included with system):
    - Converts video signal to computer image
    - Must be inserted into PCI slot of Macintosh or PC computer
  - Quantity One software (included with system) is used for:
    - Annotation of images
    - Analysis of molecular weights
    - Printing
    - Other applications
  - Two printers are offered by Bio-Rad for use with these systems:
    - Analog video printer (Mitsubishi P-91 W)
    - USB digital printer (Sony UPD895)

14.1.11 Fluorescence Microscope

- General information
  - Fluorescence is the phenomenon whereby certain chemical species can absorb light at a specific wavelength and, after a brief interval (fluorescence lifetime), emit light at a longer wavelength. Fluorescence microscopy has many research applications. In the realm of diagnostics, this technology is mainly used in FISH, in the diagnosis of autoimmune disorders, and some applications in microbiology. Although many manufactures produce different types of fluorescence microscope, the basic principle is the same. A representative example is illustrated below (Fig. 14.26)
- Principles of operation
  - A very high-intensity light source (usually a mercury or xenon arc lamp) induces fluorescence in the sample
  - The objective lens serves a dual function (epifluorescence)
    - It focuses the excitatory light beam on the sample
    - It collects light emitted from the sample
  - "Filter cube" composed of an excitation filter, an emission filter, and an intervening dichroic mirror separates the excitatory signal from the emission signal (Fig. 14.27)
    - The excitation filter selects a wavelength appropriate for excitation of a target fluorochrome
    - The emission filter is selective for the emission signal wavelength and thus blocks any contaminating light from reaching the oculars or camera tube
    - The dichroic mirror reflects the excitatory signal, such that it enters the objective lens, while transmitting (rather than reflecting) the emission signal
    - Each filter cube is optimized for a given fluorochrome, in terms of its selectivity for excitatory and emission wavelengths
    - To allow for detection of different fluorochromes, multiple filter cubes are

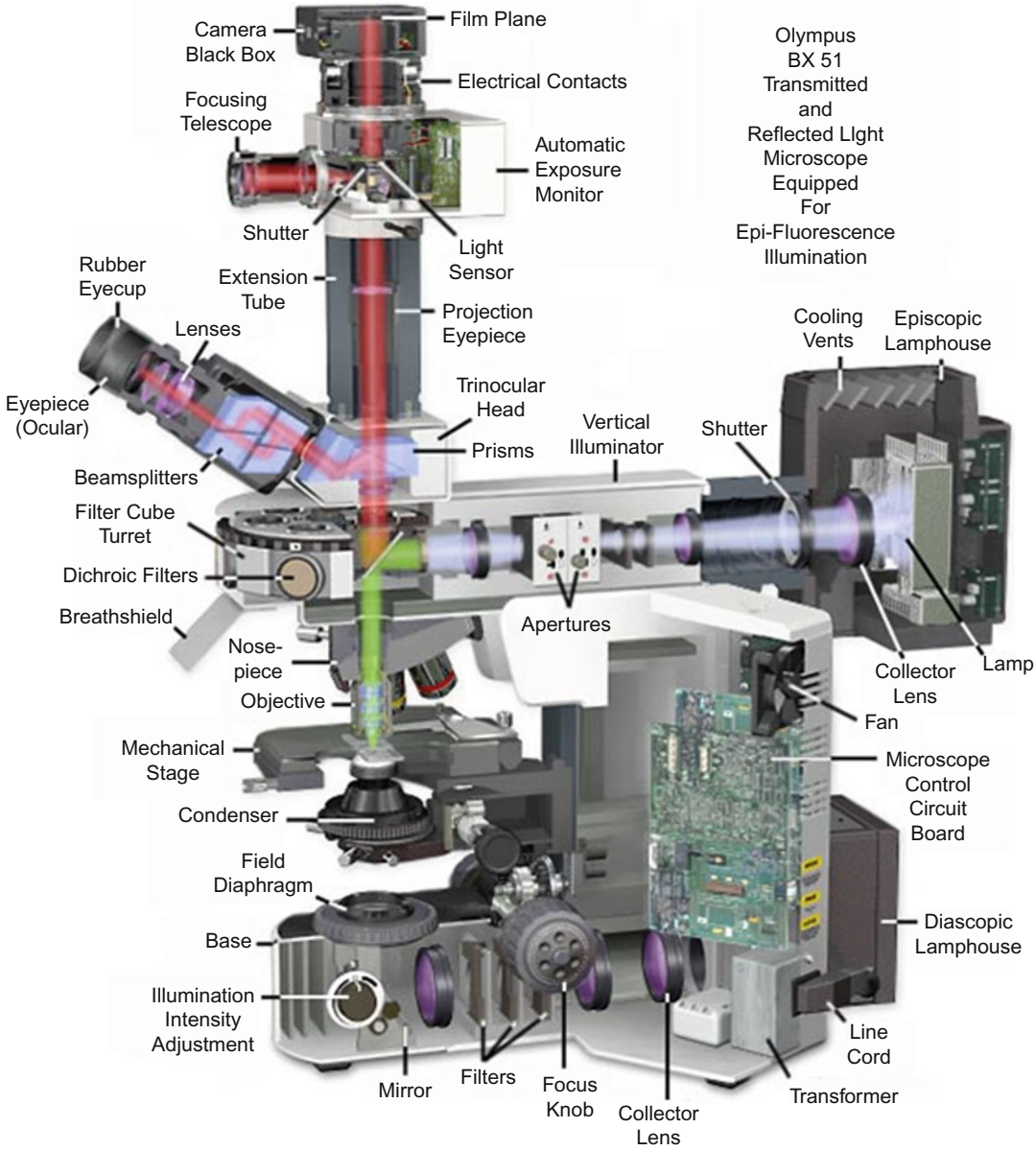


Fig. 14.26 The components of a fluorescence microscope (Olympus bx51) (Source: http://www.olympusmicro.com/primer/techniques/fluorescence/bx51fluorescence.html)

mounted on a turret. A specific filter cube is selected by manually rotating the turret

- The emission signal is focused by a projection eyepiece to produce a fluorescence image
- The image is photographed and can be processed by image analysis software

14.1.12 Slide Imaging Systems

14.1.12.1 Metafer Slide Scanning Platform (MetaSystems Gmbh, Germany)

- General information
  - The Metafer slide scanning platform (Fig. 14.28) conveniently automates a wide

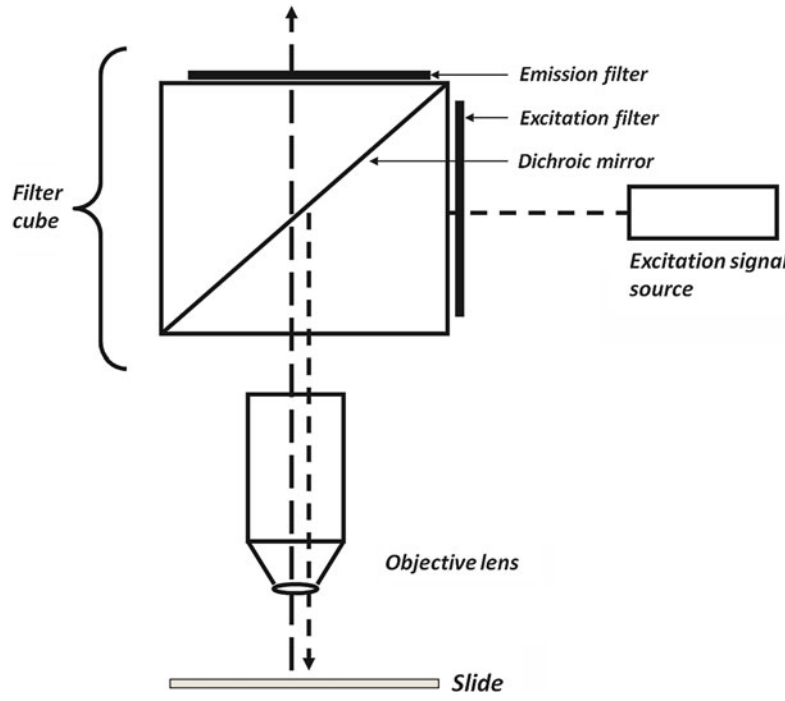


Fig. 14.27 The optics of a filter cube as configured for an inverted fluorescence microscope (Source: http://www.fluorescence.com/tutorial/fm-optic.htm)

Fig. 14.28 The Metafer slide scanning platform (http://www.metasystems-international.com/metafer)

area of microscopic image capture and analysis. Due to its modular design, it is applicable to a vast array of applications including cytogenetic diagnostics (bright field and fluorescence), hematology, pathology, toxicology, and others

– Time efficiency, automated analysis, and elimination of manual errors are the hallmark of the Metafer system. Equally important is the reproducibility achieved by applying clearly defined standards to all analysis, slide by slide, experiment by experiment, and patient by patient
– Metafer offers unlimited access to all data along with complete control over all measurement and image-processing parameters. This ensures high quality of data integrity and standardization

• Principles of operation
– Metafer slide scanning platform for cytogenetic analysis utilizes a Dell

computer, the automated Zeiss AxioImager Z2 microscope, and the following software programs from MetaSystems:
    - MSearch (metaphase search): this allows the system to scan an entire slide for all desired metaphases in bright field and fluorescence. The scan is conducted at low magnification
    - AutoCapt (autocapture): this allows Metafer to relocate to the desired images and autofocus and capture the image under higher magnification. Oil addition and barcode-reading options are available
    - Ikaros and Isis analysis software modules are available for the interactive processing of the images to final result. Virtually, an unlimited number of analysis stations can be interfaced to the Metafer server
  - The slide-loading design of Metafer accommodates use by laboratories of all sizes. The automated Zeiss microscope incorporates an eight-slide stage. The automated slide loader provides the option of holding magazines of 80 slides that can be added incrementally up to a maximum of 800 slides. This platform also includes the oil addition and the barcode reader option
- Applications
  - All types of cytogenetic analysis in bright field and fluorescence
  - Telomere analysis
  - FISH
  - Pathology slide scanning
  - Toxicology (micronuclei, chromosomal aberrations, Ames testing, comet assay)
  - Rare cell detection (CTCs)
- Advantages
  - Reproducibility
  - Digitized images that can be easily stored and sent anywhere electronically
  - Unattended operation on a platform that has been tested for over 20 years all around the world
  - Unparalleled flexibility
- Limitations
  - While Metafer can capture any slide image, for accurate analysis, slide quality is key to maximum utilization of the system

Further Reading

Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, et al. Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell. 2011;144:439–52.

C & L Instruments Inc., Tutorial – fluorescent microscope optics. C&L Instruments, Inc. Hershey, PA. http://www.fluorescence.com/tutorial/fm-optic.htm.

Garber K. Fixing the front end. Nat Biotechnol. 2008;26:1101–4.

Gaumann R, Muhlemann K, Strasser M, Beuret CM. High-throughput procedure for tick surveys of tick-borne encephalitis virus and its application in a national surveillance study in Switzerland. Appl Environ Microbiol. 2010;76:4241–9.

Heller C. Principles of DNA separation with capillary electrophoresis. Electrophoresis. 2001;22:629–43.

Heller M, Loomes KM, Cooper GJ. Synthesis of biologically active tritiated amylin and salmon calcitonin analogues. Anal Biochem. 2000;285:100–4.

Hieber L, Huber R, Bauer V, Schaffner Q, Braselmann H, Thomas G, et al. Chromosomal rearrangements in post-Chernobyl papillary thyroid carcinomas: evaluation by spectral karyotyping and automated interphase FISH. J Biomed Biotechnol. 2011;2011:693691.

Holland PM, Abramson RD, Watson R, Gelfand DH. Detection of specific polymerase chain reaction product by utilizing the 5′—3′ exonuclease activity of Thermus aquaticus DNA polymerase. Proc Natl Acad Sci USA. 1991;88:7276–80.

Hubner NC, Bird AW, Cox J, Splettstoesser B, Bandilla P, Poser I, et al. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. J Cell Biol. 2010;189:739–54.

Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. Proc Natl Acad Sci USA. 2011;108:9530–5.

Kohane IS, Kho AT, Butte AJ. Microarrays for an integrative genomics. Cambridge, MA: MIT Press; 2003.

Laakso S, Kirveskari J, Tissari P, Maki M. Evaluation of high-throughput PCR and microarray-based assay in conjunction with automated DNA extraction instruments for diagnosis of sepsis. PLoS One. 2011;6:e26655.

Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, et al. Target-enrichment strategies for next-generation sequencing. Nat Methods. 2010;7:111–8.

National Center for Biotechnology Information (NCBI) http://www.ncbi.nlm.nih.gov/About/primer/microarrays.html.

Pieprzyk M, High H. Fluidigm dynamic arrays provide a platform for single-cell gene expression analysis. Nat Meth. 2009;6. http://www.nature.com/nmeth/journal/v6/n7/full/nmeth.f.259.html.

Raggam RB, Wagner J, Bozic M, Michelin BD, Hammerschmidt S, Homberg C, et al. Detection and quantitation of Epstein-Barr virus (EBV) DNA in

EDTA whole blood samples using automated sample preparation and real time PCR. Clin Chem Lab Med. 2010;48:413–8.

Ramachandran S, Xia GL, Ganova-Raeva LM, Nainan OV, Khudyakov Y. End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: performance under optimal and suboptimal conditions. J Virol Methods. 2008;151:217–24.

Reis PP, Waldron L, Goswami RS, Xu W, Xuan Y, Perez-Ordonez B, et al. mRNA transcript quantification in archival samples using multiplexed, color-coded probes. BMC Biotechnol. 2011;11:46.

Rocha CK, Praulich I, Gehrke I, Hallek M, Kreuzer KA. A rare case of t(11;22) in a mantle cell lymphoma like B-cell neoplasia resulting in a fusion of IGL and CCND1: case report. Mol Cytogenet. 2011;4:8.

Somboonna N, Wan R, Ojcius DM, Pettengill MA, Joseph SJ, Chang A, et al. Hypervirulent Chlamydia trachomatis clinical strain is a recombinant between lymphogranuloma venereum (L(2)) and D lineages. MBio. 2011;2:e00045–11.

van de Veerdonk FL, Plantinga TS, Hoischen A, Smeekens SP, Joosten LA, Gilissen C, et al. STAT1 mutations in autosomal dominant chronic mucocutaneous candidiasis. N Engl J Med. 2011;365:54–61.

Tissue Microarrays and Biomarker Validation in Molecular Diagnostics

15

Martina Storz and Holger Moch

Contents

15.1 Tissue Microarray Technology 455

15.2 Method 456
15.2.1 Collection and Selection of Tissue Probes 456
15.2.2 Organization of a TMA 457
15.2.3 Creation of a TMA Map 457
15.2.4 Punching the Array 458
15.2.5 Cutting the Array 460

Further Reading 461

15.1 Tissue Microarray Technology

- General
  - The technology of tissue microarrays (TMAs) was developed against the background of decoding the human genome and the widespread application of high-density cDNA microarrays
  - TMAs facilitate comprehensive molecular profiling of cancer specimens with minimal tissue requirements
  - A TMA is a paraffin block composed of multiple tissue specimens
  - TMA sections provide targets for parallel in situ detection of DNA, RNA, and protein targets, allowing the rapid analysis of hundreds of molecular markers in the same set of specimens
  - Multiple specimens can be simultaneously investigated with different in situ techniques under identical laboratory conditions
- Benefits
  - TMAs are useful for rapid and high-throughput discovery and validation of biomarkers, assessing their prognostic or predictive value, and their interrelationship
  - TMAs are a cost-effective tool for quality control and standardization in immunohistochemistry (IHC)
  - TMAs can be used to evaluate sensitivity and specificity of antibodies, tissue fixation methods, and antigen retrieval methods
  - TMAs help to save reagents, manpower, and money

M. Storz, MD (✉) • H. Moch, MD
Institute of Surgical Pathology, Department Pathology, University Hospital Zürich, Zürich, Switzerland

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_15,

  - The technology is less exhausting for the original donor material because only minor tissue samples obtained from valuable materials or rare tumors are required, and the basic tissue is not destroyed
  - Collaborative studies or setup of collaborative networks is facilitated by TMAs
  - Construction of cell line TMAs is possible
  - Rapid translation of results from cell lines, xenographs, and animal models to human cancer is achieved
  - Digital images can be stored in relational databases, allowing handling the rapidly increasing amount of the generated data
  - Automatic image analysis will be more and applied for different TMA systems
  - Novel web-based database structures allow to handle clinical and pathology data for each patient in a TMA, easily facilitating intra- and interinstitutional collaborations
  - There are systems allowing secure and reliable evaluation of TMA images over the internet
- Limitations
  - Because of the small size of the individual array tissue samples (diameter usually 0.6 mm), the specimens are not totally representative of their donor tumor due to tumor heterogeneity. It is important to realize that the TMA technology has been designed to examine tumor populations, and not to survey individual tumors. The impact of tissue heterogeneity can be studied by taking multiple different punches from one tumor or by constructing replica arrays. Most, if not all, associations between molecular changes and clinical end points have been verified by TMA studies
  - Tissue spots on a TMA slide can be noninformative because of floating during the staining procedures or due to noninformative tissue (mispunching)
  - The results obtained from TMAs depend on the quality of the archival tissues. Differences in using buffered formalin or a variability in fixation can influence TMAs composed of archival tissue
  - The analysis of novel biomarkers frequently requires availability of antibodies for paraffin-embedded tissues. TMAs constructed from frozen tissue allow widespread in situ analysis of RNA and proteins. The construction of frozen TMAs is time-consuming and difficult
- Examples
  - Multitumor TMAs are composed of samples from multiple tumor types. These arrays are used to screen different tumor types for molecular alterations of interest
  - Progression TMAs have been used to study molecular alterations in different stages of one particular tumor
  - Prognostic TMAs contain samples from tumors of patients for whom clinical follow-up data and clinical end points are known
  - TMAs are constructed also for research in other fields such as inflammatory, cardiovascular, and neurologic diseases
  - TMAs can be used for cell lines, xenograft tumors, or tissues from animal model systems

15.2 Method

15.2.1 Collection and Selection of Tissue Probes

- A design of a TMA is based on the aims of the planned study. According to the aim of the study, a list of cases to be included into the TMA should be prepared
  - Collect all the hematoxylin and eosin (H&E)-stained slides and the corresponding paraffin-embedded blocks of interest
  - Make sure the H&E slides show the current state of the block or cut fresh H&E sections. This will ensure that the desired tissue observed on the H&E slide ends up on the TMA
  - Check all the collected H&E slides
  - Mark the areas of interest with a felt tip pen directly on the cover slip

Fig. 15.1 H&E slide with differentially labeled tumor areas

- Different areas on a tissue block may be of interest, for example, tumor tissue as well as normal tissue. Choose a color code and mark the different areas of interest with different colors. As example: tumor red, normal tissue blue, and other areas black
- Mark the largest possible area that contains the area of interest. It will be easier to find the desired area to be biopsied, and it enables the investigator to take more than one or two cores
- The minimum size of the circled area should be at least 3 mm^2
- To draw the areas of interest onto the H&E slides is an important part of the project. The rule is the TMA can only be punched as precisely as it has been marked (Fig. 15.1)
- Compare the circled H&E slides with the corresponding paraffin-embedded blocks. Do not include thin tissue blocks into the TMA. These cores will be missing after cutting a few sections of the TMA block
- The thickness of the designated tissue should be at least 3 mm

15.2.2 Organization of a TMA

- Arrange the corresponding block onto the H&E slide and put them on a tray by ascending numbers. This is important for an array with a large number of cases and will help later on if one has to look for a certain case
- It is recommended to include about 600 cores in a TMA. Although up to 1,000 cores can be arranged on a block, the tension will increase and can result in cracks at the edges of the TMA block and in technical problems when cutting sections
- Add normal tissues as control

15.2.3 Creation of a TMA Map

- Individual maps can be created according to the study. Arrays should be organized in different quadrants; for 640 cores, four quadrants are recommended. A space between quadrants helps to minimize the tension in the block and facilitates evaluation
- Each quadrant has 20 rows (x-axis) and eight columns (y-axis). The rows and columns get different numbers and that create a coordinate system, in which each core gets its own x and y localization
- We recommend preparing a punch file and a picture file (Figs. 15.2 and 15.3)
- The generation of the punch file is an important part in a TMA project. A mistake on this file will also appear on the TMA and on all immunostains performed on the TMA slide
- The picture file is a visual representation of the array. Different diagnoses or tissue types can be shown in colors. The marking spots on the left are extremely important for the correct orientation of the slide

Spot #	*SP number*	*Block ID*	*Tissue type*	*DX name*	*Localization x*	*Localization y*	*Coordinates*
1	B 1993.x1	2	Kidney	Chromophob	1	1	0/0
2	B 1993.x2	1	Kidney	Other	2	1	800/0
3	B 1993.x3	1	Kidney	Clear cell	3	1	1600/0
4	B 1993.x4	3	Kidney	Clear cell	4	1	2400/0
5	B 1993.x5	2	Kidney	Papillary type	5	1	3200/0
6	B 1993.x6	7	Kidney	Papillary type	6	1	4000/0
7	B 1993.x7	2	Kidney	Clear cell	7	1	4800/0
8	B 1993.x8	3	Kidney	Clear cell	8	1	5600/0
9	B 1993.x9	4	Kidney	Clear cell	1	2	0/800
10	B 1993.x10	2	Kidney	Clear cell	2	2	800/800
11	B 1993.x11	4	Kidney	Clear cell	3	2	1600/800
12	B 1993.x12	1	Kidney	Clear cell	4	2	2400/800
13	B 1993.x13	1	Kidney	Clear cell	5	2	3200/800
14	B 1993.x14	1	Kidney	Oncocytoma	6	2	4000/800
15	B 1993.x15	2	Kidney	Papillary type	7	2	4800/800
16	B 1993.x16	2	Kidney	Clear cell	8	2	5600/800
17	B 1993.x17	8	Kidney	Clear cell	1	3	0/1600
18	B 1993.x18	6	Kidney	Chromophob	2	3	800/1600
19	B 1993.x19	2	Kidney	Papillary type	3	3	1600/1600
20	B 1993.x20	1	Kidney	Clear cell	4	3	2400/1600
21	B 1993.x21	6	Kidney	Clear cell	5	3	3200/1600
22	B 1993.x22	2	Kidney	Clear cell	6	3	4000/1600
23	B 1993.x23		Kidney	Clear cell	7	3	4800/1600
24	B 1993.x24	2	Kidney	Clear cell	8	3	5600/1600

Fig. 15.2 Example of a punch file. In an Excel spreadsheet, all the information is included for punching and evaluating the stained TMA slide. The specimen number and the ID number of the donor block, tissue type and diagnosis, the *x* localization and *y* localization, and the coordinates of the micrometer screw on the manual arrayer. Other parameters can be included

- Decide on the number of cores per tumor block. For heterogeneous tumors, two to four cores per donor block are recommended. The construction of several replicates is advised if many studies are planned
- Insert all the block numbers into a punch file, prepare the picture file
- Prepare a blank recipient block using a standard tissue cassette and paraffin with a melting point of 50–60 °C. Use only recipient blocks completely filled with paraffin, discard blocks with holes or cracks

15.2.4 Punching the Array

- The following punch procedure is based on using a manual Beecher Instruments arrayer (Beecher Instruments, Sun Prairie, WI [http://www.beecherinstruments.com]) (Fig. 15.4)
  - Put the recipient block into the magnetic recipient block holder
  - Tighten the screws in the block holder carefully
  - Make sure that the block holder is precisely touching the location bars

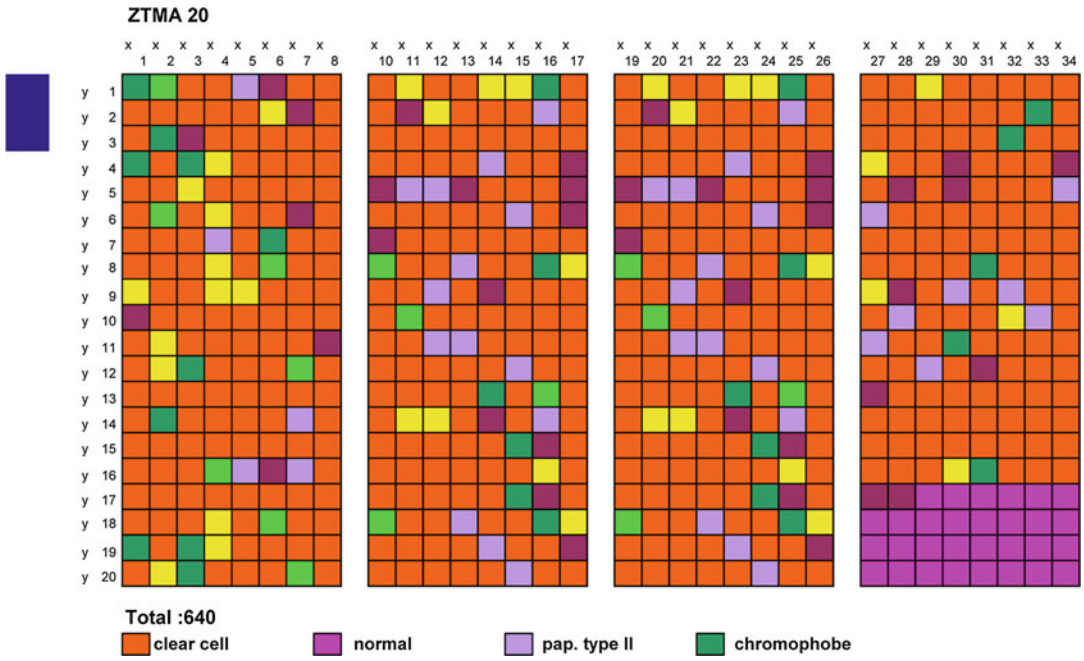


Fig. 15.3 Example of a picture file

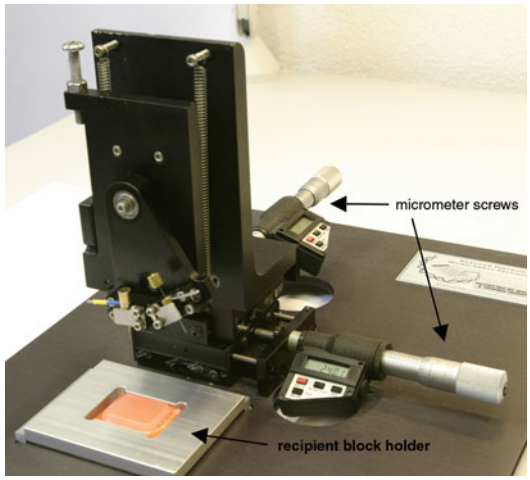


Fig. 15.4 Manual tissue arrayer

- Move the needle with the micrometer screw to the edge of the recipient block and adjust the depth to which the needle goes into the recipient block. The needle should not touch the plastic cassette
- Set up the location of the very first core x_1, y_1. Make sure that all cores will fit on the recipient block and try to maintain the same distance to all four edges of the block. Press the zero buttons on the micrometer screw; make sure that the display shows 0.000
- Punch a hole into the recipient block with the smaller needle. Put the donor block bridge over the recipient block and place the donor block on top. Punch a tissue cylinder out of the desired area using the bigger needle (Figs. 15.5 and 15.6)
- Remove the donor block bridge and push the tissue cylinder carefully into the hole of the recipient block
- Be careful not to push the tissue cylinder all the way into the recipient block, leave about 0.5 mm exposed. Place a clean glass slide on the surface of the array and gently push the remaining 0.5 mm of the tissue cylinder into the recipient block. The goal for a good array is to have all the cores at the same level. Maintaining this procedure for every single core will lead to the best result

- For the next core, move one of the micrometer screws about 800 µm when using the 600-µm needles, this results in a 200-µm space between two cores

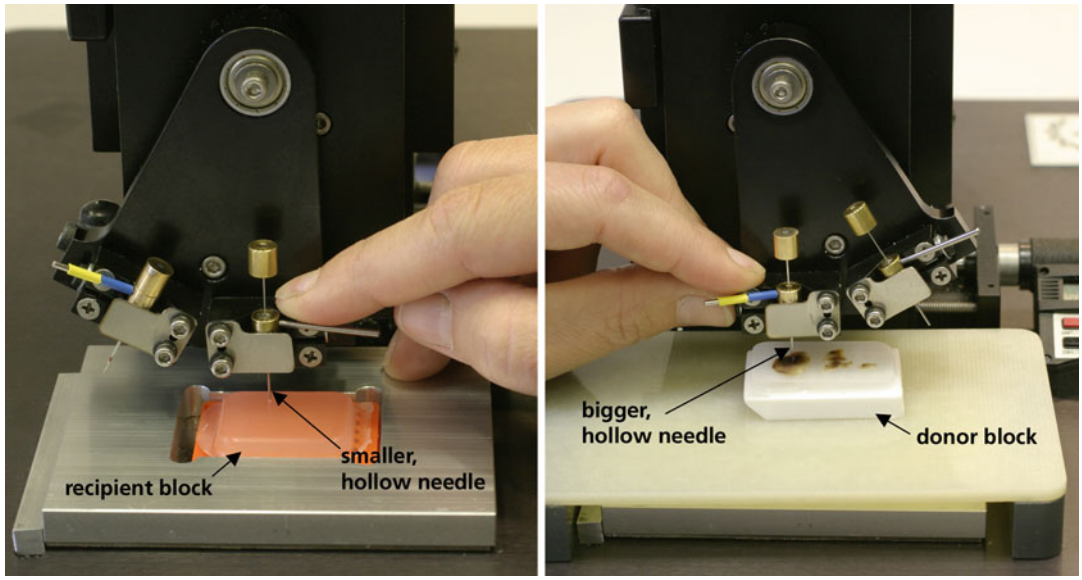


Fig. 15.5 Punching procedure with donor and recipient block

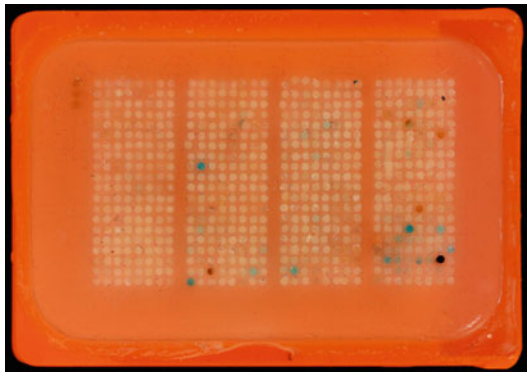


Fig. 15.6 Example of a TMA containing 640 cores

- Continue the punch procedure by following the punch file
- Put the completed array in an oven at 40–50 °C for 15–30 min. This is necessary to better adhere the tissue cylinders to the paraffin of the recipient block. Place a clean glass slide on the surface of the array and slowly press down. Be aware that this is a dangerous step. If the TMA remains too long in the oven, the paraffin will melt and the array will be damaged
- To alleviate and shorten the punching procedure, semiautomated tissue arrayer (taucherinsel-neuenburg@t-online.de) are available as shown in Fig. 15.7

15.2.5 Cutting the Array

- General
  - Cutting the array is a critical step, and only a trained person should be in charge to prevent destruction of the TMA
  - The tape transfer system from Instrumedics (Instrumedics Inc., St. Louis, MO [http://www.instrumedics.com]) is a tool to cut the array. The procedure is as follows
    - Place a tape window on the surface of the block and slowly cut the section of array and tape
    - Take a slide covered with special glue, remove the Mylar cover, and roll the section of array and tape onto the slide
    - Place the slide under an ultraviolet lamp for 30 s, put the slide in a solvent bath, and remove the tape
    - Let the slide air-dry
- Benefits
  - Each core stays in the intended location. Using a water bath could lead to a shift of the rows and columns of the array, and this could be difficult when evaluating the stained slide
  - The glue on the slide makes it possible to perform a wide range of heat pretreatment

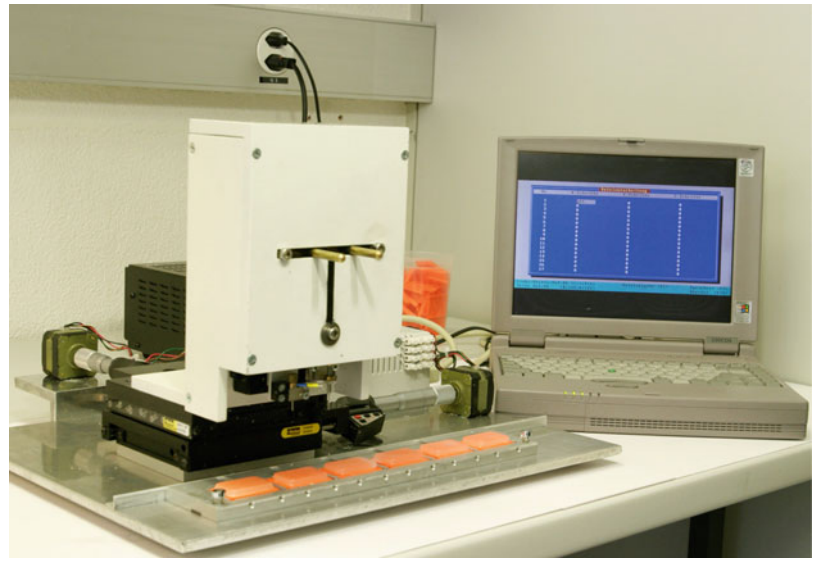


Fig. 15.7 Semiautomated tissue arrayer

steps for IHC, fluorescent hybridization, and in situ hybridization without tissue loss
  - This system allows cutting consecutive slides. No tissue is wasted trying to get a slide without crinkles and all cores present
- Limitations
  - The glue on the slide can result in background staining, but this does not affect the staining of the marker
  - The glue can interfere with procedure for fluorescence in situ hybridization
  - Daylight can polymerize the glue, and the tape will not stick on the slide anymore. Make sure to always close the lid of the slide box
  - Automated IHC systems are disturbed by the glue. Sections performed with the tape transfer system should be manually processed
- Example
  - Use the tape transfer system to cut the TMA and perform IHC manually
  - Without the tape transfer system, construct arrays with not more than 300 cores. Cut the array with the traditional method, and the application of an automated IHC system is possible
  - Store the slides at +4 °C, regardless of whether the tape system was applied or not. Use the slides within 3 months, after that the staining intensity of several antibodies will decrease

Further Reading

Berman JJ, Datta M, Kajdacsy-Balla A, et al. The tissue microarray data exchange specification: implementation by the Cooperative Prostate Cancer Tissue Resource. BMC Bioinformatics. 2004;5:19.

Bolli M, Schultz-Thater E, Zajac P, et al. NY-ESO-1/LAGE-1 coexpression with MAGE-A cancer/testis antigens: a tissue microarray study. Int J Cancer. 2005;115:960–6.

Bubendorf L, Kolmer M, Kononen J, et al. Hormone therapy failure in human prostate cancer: analysis by complementary DNA and tissue microarrays. J Natl Cancer Inst. 1999a;91:1758–64.

Bubendorf L, Kononen J, Koivisto P, et al. Survey of gene amplifications during prostate cancer progression by high-throughout fluorescence in situ hybridization on tissue microarrays. Cancer Res. 1999b;59:803–6.

Bubendorf L, Nocito A, Moch H, et al. Tissue microarray (TMA) technology: miniaturized pathology archives for high-throughput in situ studies. J Pathol. 2001;195:72–9 (review article).

Chung GG, Yoon HH, Zerkowski MP, et al. Vascular endothelial growth factor. FLT-1, and FLK-1 analysis in a pancreatic cancer tissue microarray. Cancer. 2006;106:1677–84.

Cima I, Schiess R, Wild P, et al. Cancer genetics-guided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer. Proc Natl Acad Sci USA. 2011;108:3342–7.

Dahinden C, Ingold B, Wild P, et al. Mining tissue microarray data to uncover combinations of biomarker expression patterns that improve intermediate staging and grading of clear cell renal cell cancer. Clin Cancer Res. 2010;16:88–98.

Dell'Anna R, Demichelis F, Barbareschi M, et al. An automated procedure to properly handle digital images in large scale tissue microarray experiments. Comput Methods Programs Biomed. 2005;79:197–208.

Diaz LK, Gupta R, Kidwai N, et al. The use of TMA for interlaboratory validation of FISH testing for detection of HER2 gene amplification in breast cancer. J Histochem Cytochem. 2004;52:501–7.

Fowler JM, Ramirez N, Cohn DE, et al. Correlation of cyclooxygenase-2 (COX-2) and aromatase expression in human endometrial cancer: tissue microarray analysis. Am J Obstet Gynecol. 2005;192:1262–73.

Garcia JF, Camacho FI, Morente M, et al. Hodgkin and Reed-Sternberg cells harbor alterations in the major tumor suppressor pathways and cell-cycle checkpoints: analyses using tissue microarrays. Blood. 2003;101:681–9.

Gately K, Kerr K, O'Byrne K. Design, construction, and analysis of cell line arrays and tissue microarrays for gene expression analysis. In: O'Driscoll L, editor. Gene expression profiling: methods and protocols, Methods in molecular biology, vol. 784. 2nd ed. New York: Humana Press/Springer; 2011. p. 139–53.

Gulmann C, Loring P, O'Grady A, et al. Miniature tissue microarrays for HercepTest standardisation and analysis. J Clin Pathol. 2004;57:1229–31.

Hendriks Y, Franken P, Dierssen JW, et al. Conventional and tissue microarray immunohistochemical expression analysis of mismatch repair in hereditary colorectal tumors. Am J Pathol. 2003;162:469–77.

Hewitt SM. Tissue microarrays as a tool in the discovery and validation of predictive biomarkers. Methods Mol Biol. 2012;823:201–14.

Howard BA, Zheng Z, Campa MJ, et al. Translating biomarkers into clinical practice: prognostic implications of cyclophilin A and macrophage migratory inhibitory factor identified from protein expression profiles in non-small cell lung cancer. Lung Cancer. 2004;46:313–23.

Kallioniemi OP, Wagner U, Kononen J, et al. Tissue microarray technology for high-throughput molecular profiling of cancer. Hum Mol Genet. 2001;10:657–62.

Kay E, O'Grady A, Morgan JM, et al. Use of tissue microarray for interlaboratory validation of HER2 immunocytochemical and FISH testing. J Clin Pathol. 2004;57:1140–4.

Kettunen E, Nicholson AG, Nagy B, et al. L1CAM, INP10, P-cadherin, tPA and ITGB4 over-expression in malignant pleural mesotheliomas revealed by combined use of cDNA and tissue microarray. Carcinogenesis. 2005;26:17–25.

Kim R, Demichelis F, Tang J, et al. Internet-based Profiler system as integrative framework to support translational research. BMC Bioinformatics. 2005;6:304.

Kocher T, Zheng M, Bolli M, et al. Prognostic relevance of MAGE-A4 tumor antigen expression in transitional cell carcinoma of the urinary bladder: a tissue microarray study. Int J Cancer. 2002;100:702–5.

Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med. 1998;4:844–7.

Koynova DK, Tsenova VS, Jankova RS, et al. Tissue microarray analysis of EGFR and HER2 oncogene copy number alterations in squamous cell carcinoma of the larynx. J Cancer Res Clin Oncol. 2005;131:199–203.

Kuefer R, Hofer MD, Gschwend JE, et al. Tissue microarrays. High-throughput procedures to verify potential biomarkers. Urologe A. 2004;43:659–67.

Linke SP, Bremer TM, Herold CD, et al. A multimarker model to predict outcome in tamoxifen-treated breast cancer patients. Clin Cancer Res. 2006;12:1175–83.

Liu X, Minin V, Huang Y, et al. Statistical methods for analyzing tissue microarray data. J Biopharm Stat. 2004;14:671–85 (review article).

Liu CL, Montgomery KD, Natkunam Y, et al. MA-Combiner, a simple software tool to permit analysis of replicate cores on tissue microarrays. Mod Pathol. 2005;18:1641–8 (review article).

Lusis EA, Chicoine MR, Perry A. High throughput screening of meningioma biomarkers using a tissue microarray. J Neurooncol. 2005;73:219–23.

Manley S, Mucci NR, De Marzo AM, et al. Relational database structure to manage high-density tissue microarray data and images for pathology studies focusing on clinical outcome: the prostate specialized program of research excellence model. Am J Pathol. 2001;159:837–43.

Mengel M, Kreipe H, von Wasielewski R. Rapid and large-scale transition of new tumor biomarkers to clinical biopsy material by innovative tissue microarray systems. Appl Immunohistochem Mol Morphol. 2003;11:261–8.

Mihic-Probst D, Kuster A, Kilgus S, et al. Consistent expression of the stem cell renewal factor BMI-1 in primary and metastatic melanoma. Int J Cancer. 2007;121:1764–70.

Mirlacher M, Kasper M, Storz M, et al. Influence of slide aging on results of translational research studies using immunohistochemistry. Mod Pathol. 2004; 17:1414–20.

Moch H, Schraml P, Bubendorf L, et al. High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. Am J Pathol. 1999;154:981–6.

Moch H, Kononen T, Kallioniemi OP, et al. Tissue microarrays: what will they bring to molecular and anatomic pathology? Adv Anat Pathol. 2001;8:14–20.

Nielsen TO, Hsu FD, O'Connell JX, et al. Tissue microarray validation of epidermal growth factor receptor

and SALL2 in synovial sarcoma with comparison to tumors of similar histology. Am J Pathol. 2003;163: 1449–56.

Nocito A, Bubendorf L, Tinner EM, et al. Microarrays of bladder cancer tissue are highly representative of proliferation index and histological grade. J Pathol. 2001;194:349–57.

Packeisen J, Korsching E, Herbst H, et al. Demystified tissue microarray technology. Mol Pathol. 2003;56:198–204 (review article).

Rubin MA, Mucci NR, Figurski J, et al. E-cadherin expression in prostate cancer: a broad survey using high-density tissue microarray technology. Hum Pathol. 2001;32:690–7.

Rubin MA, Dunn R, Strawderman M, et al. Tissue microarray sampling strategy for prostate cancer biomarker analysis. Am J Surg Pathol. 2002;26:312–9.

Schraml P, Kononen J, Bubendorf L, et al. Tissue microarrays for gene amplification surveys in many different tumor types. Clin Cancer Res. 1999;5: 1966–75.

Schraml P, Schwerdtfeger G, Burkhalter F, et al. Combined array comparative genomic hybridization and tissue microarray analysis suggest PAK1 at 11q13.5–q14 as a critical oncogene target in ovarian carcinoma. Am J Pathol. 2003;63:985–92.

Simon R, Sauter G. Tissue microarrays for miniaturized high-throughput molecular profiling of tumors. Exp Hematol. 2002;30:1365–72 (review article).

Simon R, Sauter G. Tissue microarray (TMA) applications: implications for molecular medicine. Expert Rev Mol Med. 2003;2003:1–12.

Simon R, Mirlacher M, Sauter G. Tissue microarrays. Methods Mol Med. 2005;114:257–68.

Staller P, Sulitkova J, Lisztwan J, et al. Chemokine receptor CXCR4 downregulated by von Hippel-Lindau tumour suppressor pVHL. Nature. 2003;425:307–11.

Struckmann K, Schraml P, Simon R, et al. Impaired expression of the cell cycle regulator BTG2 is common in clear cell renal cell carcinoma. Cancer Res. 2004;64:1632–8.

Theurillat JP, Dreier B, Nagy-Davidescu G, et al. Designed ankyrin repeat proteins: a novel tool for testing epidermal growth factor receptor 2 expression in breast cancer. Mod Pathol. 2010;23:1289–97.

Theurillat JP, Metzler SC, Henzi N, et al. URI is an oncogene amplified in ovarian cancer cells and is required for their survival. Cancer Cell. 2011;19:317–32.

Torhorst J, Bucher C, Kononen J, et al. Tissue microarrays for rapid linking of molecular changes to clinical endpoints. Am J Pathol. 2001;159:2249–56.

Tzankov A, Pehrs AC, Zimpfer A, et al. Prognostic significance of CD44 expression in diffuse large B cell lymphoma of activated and germinal centre B cell-like types: a tissue microarray analysis of 90 cases. J Clin Pathol. 2003a;56:747–52.

Tzankov A, Zimpfer A, Lugli A, et al. High-throughput tissue microarray analysis of G1-cyclin alterations in classical Hodgkin's lymphoma indicates overexpression of cyclin El. J Pathol. 2003b;199:201–7.

Varga Z, Theurillat JP, Filonenko V, et al. Preferential nuclear and cytoplasmic NY-BR-1 protein expression in primary breast cancer and lymph node metastases. Clin Cancer Res. 2006;12:2745–51.

Went P, Vasei M, Bubendorf L, et al. Frequent high-level expression of the immunotherapeutic target Ep-CAM in colon, stomach, prostate and lung cancers. Br J Cancer. 2006;94:128–35.

Laser Capture Microdissection in Molecular Diagnostics

16

Liang Cheng, Shaobo Zhang, Darrell D. Davidson, Matthew Kuhar, Mingsheng Wang, Sean R. Williamson, David Y. Zhang, and Gregory T. MacLennan

Contents

16.1 Overview of Laser Microdissection 466
16.1.1 Conceptual Importance of Procuring Pure Cell Populations 466
16.1.2 The History of Microdissection 466
16.1.3 Major Technical Platforms and Technical Basis 467
16.1.4 Principles of Major Microdissection Systems 467

16.2 Specimen Handling and Processing 469
16.2.1 Specimen Considerations and Handling ... 469
16.2.2 Major Components of Microdissection Systems 472
16.2.3 Key Characteristics of the Laser and Transfer Film 472
16.2.4 Key Characteristics of the UV Laser and Lining Foil 473
16.2.5 Expected Yield of DNA, RNA, and Proteins from Laser Microdissected Tissue 473
16.2.6 Benefits and Limitations of LCM for Cell Isolation 473

16.3 LCM Laboratory Protocols 474
16.3.1 Slide Preparation 474
16.3.2 Step-by-Step Procedures Using the Arcturus LCM System 475
16.3.3 Extraction of Genomic DNA from Formalin-Fixed, Paraffin-Embedded Tissue 475
16.3.4 Procedure and Protocol Resources 476

16.4 Other Cellular Isolation Techniques 476
16.4.1 Non-LCM Microdissection Methods 476
16.4.2 Nonmicrodissection Methods That Enrich the Target Cells 477

16.5 Methods of Analysis and Potential Application of LCM 478
16.5.1 Methods of Molecular Analysis 478
16.5.2 Potential Applications of LCM 478

Further Reading 481

L. Cheng, MD (✉) • S. Zhang, MD • D.D. Davidson, MD, PhD • M. Wang, MD • S.R. Williamson, MD
Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

M. Kuhar, MD
StrataDx & Strata Pathology Services, Lexington, MA, USA

D.Y. Zhang, MD, PhD, MPH
Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

G.T. MacLennan, MD
Department of Pathology, Case Western Reserve University, Cleveland, OH, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_16,

16.1 Overview of Laser Microdissection

16.1.1 Conceptual Importance of Procuring Pure Cell Populations

- It was originally asserted by Virchow (1821–1902) that the cell, rather than the tissue, represents the most basic unit of disease
- The myriad of cell types within neoplasm presents a major obstacle for the global analysis of such tissues. A single cell cannot be adequately analyzed without a means of reliably isolating it from adjacent cells
- Accurate and sensitive detection of molecular changes in malignant or premalignant cells usually requires that only cells of interest are examined
  - Genetic material from contaminating cells can mask a finding with conflicting data
    - Even rare unintended cells' genetic material can become amplified during polymerase chain reaction techniques
    - Using nonmicrodissected material often underestimates the actual incidence of genetic alterations
  - Human tumors are heterogeneous with admixed cell populations
    - Normal parenchyma that the tumor is invading or in which the tumor developed
    - Blood vessels supplying the tumor
    - Inflammatory cell infiltrate that often accompanies malignancy
    - Stromal cells and/or desmoplastic connective tissue in response to an invasive tumor
- LCM is capable of extremely selective sampling (Fig. 16.1)
- Microdissection permits analysis of molecular alterations within specific cell populations without interference from nontarget cell populations
- LCM is a technique developed by the National Institutes of Health. It employs a laser to dissect individual cells or small clusters of cells selected by concurrent light microscopy

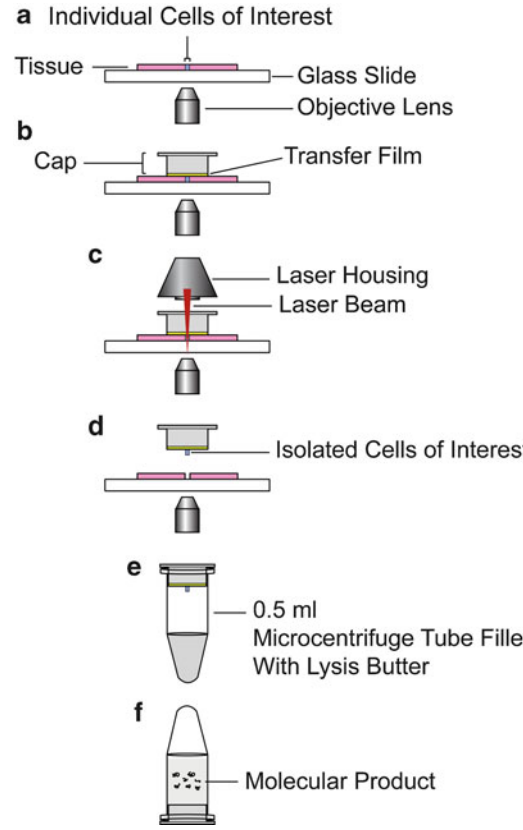


Fig. 16.1 The steps of laser capture microdissection are shown in parts (**a**–**f**). The cells are identified using the light microscope to observe the uncoverslipped slide (**a**). The cap and transfer film are then placed directly on the area of interest (**b**). Near-infrared light is passed through the cap and transfer film and cells of interest (**c**), which forms a polymer between the struck cells and the overlying transfer film. With removal of the cap (**d**), the cells adherent to the transfer film are successfully extracted. Finally, the cap is placed on a microcentrifuge tube with lysis buffer (**e**). After the lysis buffer digests the cells in the transfer film, a pure suspension is achieved (**f**)

- The objective of LCM is to obtain a pure sample comprised only of the specific cells of diagnostic, prognostic, or research interest

16.1.2 The History of Microdissection

- Early work procedure microdissects cells of interest with skilled free hand
- In 1976, Meier–Ruge et al. initiated the development of a UV laser micropreparation

Table 16.1 Traditional methods for cell enrichment

Method	Principle
Culturing of tumor cells	Selected cells expanding in culture medium
Xenograft enrichment	Serial passage of xenografted tissue through immunodeficient rodents
Cell sorting	Isolation of cells from suspension with various types of techniques, such as density gradient, fluorescence-activated cell sorting, antibody-labeled immunobeads, or affinity-labeled magnetic beads
Manual microdissection	Manual extraction of desired foci or selective ablation of unwanted regions

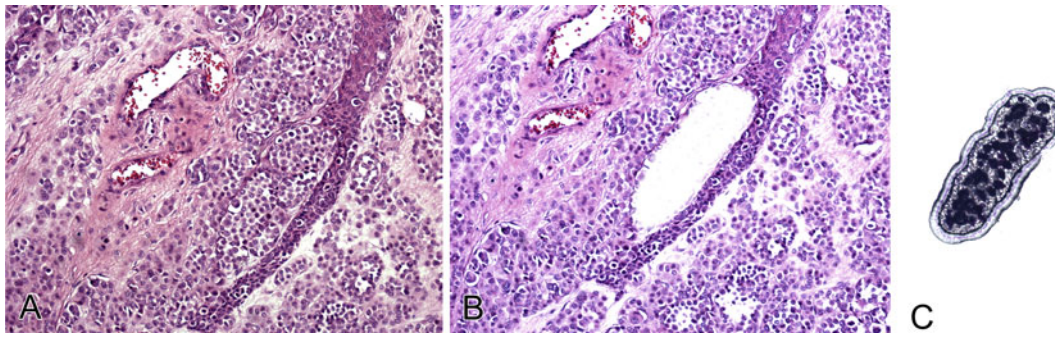


Fig. 16.2 (**a**) Cutaneous malignant melanoma with a prominent central group of melanocytes, as observed under light microscope prior to laser capture microdissection. (**b**) The central group of melanocytes has been removed without alteration to the adjacent tissue in this image following LCM. (**c**) An image of the group of isolated melanocytes following extraction

instrument using UV laser technology that was primitive but nonetheless more precise than manual techniques
- LCM was developed in 1996 at the National Institutes of Health by Emmert et al. (Table 16.1)

16.1.3 Major Technical Platforms and Technical Basis

- Two major LCM platforms are commercially available
  - Thermal-sensitive membrane and infrared laser systems
  - Laser-cutting ultraviolet laser system
- Technical basis
  - The thermal-sensitive membrane system captures target cells by a thermally activated thermoplastic ethylene vinyl acetate polymer membrane mounted on the bottom of a plastic cap
  - The laser-cutting system is a noncontact method, which harvests the target by UV laser cutting of tissue mounted onto a special slide with premounted laser-sensitive foil

16.1.4 Principles of Major Microdissection Systems

16.1.4.1 Thermal-Sensitive Membrane and Infrared Laser Systems

- The cells of interest are identified using light microscopy (Figs. 16.2, 16.3)
- A specialized centrifuge cap with an attached thin thermoplastic transfer film is placed over

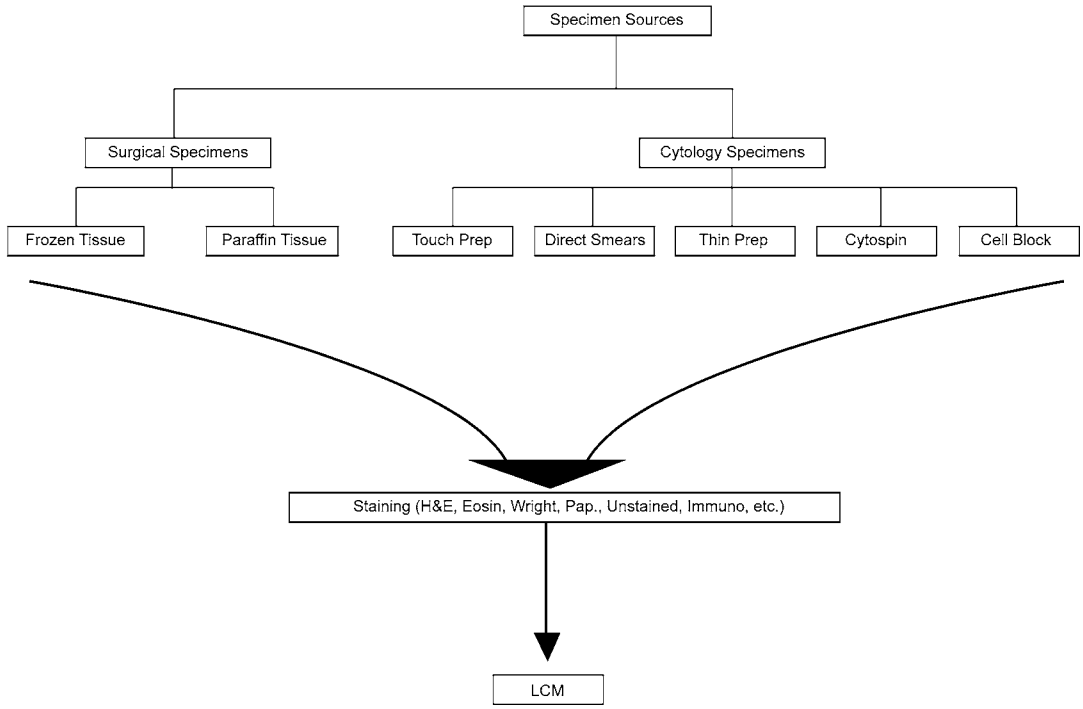


Fig. 16.3 A wide variety of specimens and stains are suitable for use with the LCM technique

the area of interest with the film coming in direct contact with the tissue on an uncoverslipped, uncharged glass slide
- A near-infrared laser with a thin focused beam is passed through the transfer film and the underlying cell(s) of interest
- The transfer film is rapidly heat activated and focally melts
- The melted transfer film soaks into the spaces in the tissue immediately after the laser shot
- The transfer film cools to form a new polymer with the selected cells
  - The laser wavelength is designed to heat the transfer film without heating the tissue
  - The heating and cooling process is complete within milliseconds
  - The new bond between the selected cell(s) and the transfer film is stronger than the bond between the glass slide and the cell
- The cell-film polymer is retracted from the surrounding tissue by lifting the cap
- The transfer film can be moved to multiple additional areas to harvest many areas of interest on a single film
- When all desired harvesting is complete, the cap and attached transfer film are removed from the setup and placed on a standard 0.5-mL microcentrifuge tube containing a lysis buffer or a digestion buffer
- The lysis or digestion buffer releases the DNA, RNA, or proteins from the film on the underside of the cap, providing a pure sample suspension with the target molecules

16.1.4.2 The Laser-Cutting Ultraviolet Laser System

- The technology requires special slides with premounted, laser-sensitive foil (Fig. 16.4)
- Tissue sections are mounted onto the ultraviolet laser-absorbing foil and the cells of interest are microdissected by laser cutting, which creates an edge surrounding the cells
- The major advantage of laser-cutting technology is that there is no physical contact between the cells and the instrument
  - Noncontact reduces the potential risk of contamination by nontarget cells

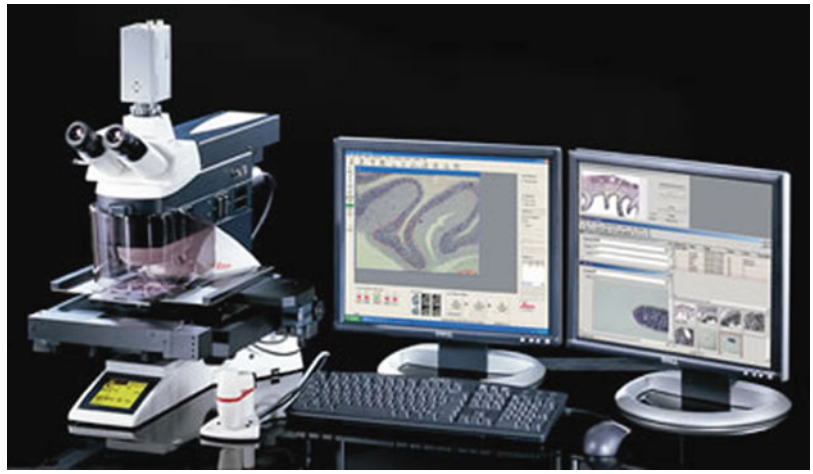


Fig. 16.4 Leica laser capture microdissection system (LMD6000) (Leica Microsystems, Wetzlar, Germany)

  - Potential chemical modification of molecules is avoided
- The ultraviolet-cutting systems are particularly useful for microdissection of tissue sections up to 200 μm in thickness

16.1.4.3 Laser Microdissection Using Combining Technology

- The laser microdissection systems offered by Arcturus, the ArcturusXT and the Veritas, combine infrared laser-enabled LCM and ultraviolet laser cutting in a single platform (Fig. 16.5)
  - The infrared LCM captures single cells or small numbers of targeted cells with great precision
  - Ultraviolet laser delivers unprecedented speed and precision optimal for dense tissue structures or large numbers of cells. The unique combination of infrared laser capture and ultraviolet laser cutting permits the use of different slide types and sample preparations

16.1.4.4 Alternative Nonlaser Technology

- The piezo-power microdissection system offered by Eppendorf is a nonlaser low-cost alternative (Fig. 16.6)
- The instrument cuts tissue using an ultrasonic metal microchisel
  - A piezo-powered stepper converts frequency (25–60 kHz) and amplitude (0–1.5 μm) into mechanical force that oscillates the fine microchisel metal tip to cut samples
  - Cells can be collected without heating or exposure to intense light
- The microdissector provides precise dissection, especially in tough tissues such as plants, cartilage, tooth, and elastic arteries as well as in live tissue
- Because tissue pretreatment is not necessary, both living and archived tissue can be used for cell isolation using piezo-power microdissection (Table 16.2)

16.2 Specimen Handling and Processing

16.2.1 Specimen Considerations and Handling

16.2.1.1 Types of Specimens

- Paraffin-embedded specimens (Fig. 16.7)
  - Formalin-fixed
    - Most common archival tissue in surgical pathology
    - Neutral buffered formalin fixation is acceptable; however, it causes extensive cross-linking of nucleic acids and protein, which makes polymerase chain reaction (PCR) amplification more difficult

Fig. 16.5 Artcurus laser capture microdissection system (PixCell IIe) (Arcturus Bioscience Inc., Mountain View, California, USA)

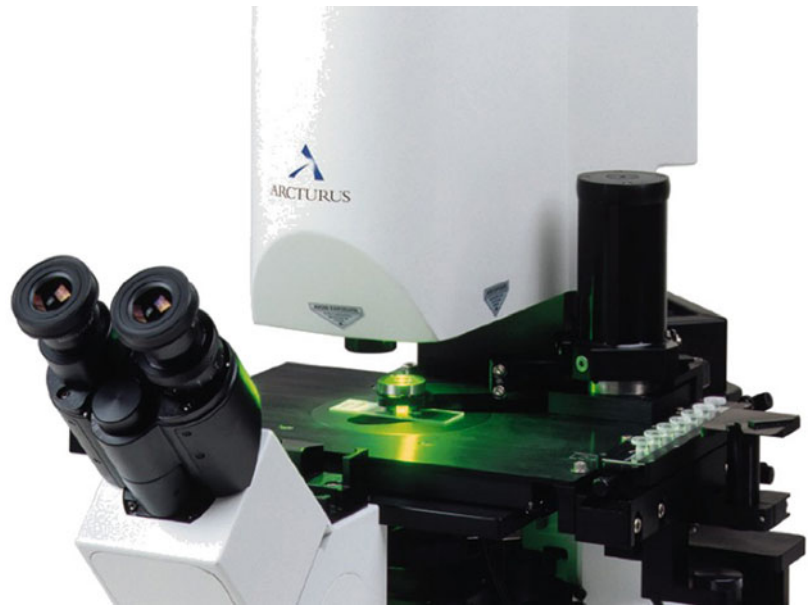


Fig. 16.6 In this close-up of the Arcturus PixCell IIe, the cap can be seen resting on the glass slide (Arcturus Bioscience Inc., Mountain View, California, USA)

  - Typical DNA fragments range from 100 to 1,500 bp
  - RNA quality may be too poor to be useful, although recent advances have overcome some technical difficulties
  - Overfixation worsens macromolecule quality
    - Biopsies should fix <12 h
    - Large specimens should fix <48 h
  - Metal salt-based fixatives should be avoided
  - Bouin and B5 fixatives are even more damaging than formalin and should be avoided whenever possible
  - Alcohol fixation is the most desirable method for paraffin-embedded specimens
- Frozen specimens
  - Specimens frozen and embedded in optimal cutting temperature compound (OCT)

Table 16.2 Characterization of commonly used microdissection techniques

Technology	Laser capture microdissection	Laser pressure catapulting	Ultrasonic piezo (piezo-power microdissection)
Power source	Infrared laser	Ultraviolet laser	Ultrasound
Principle	Thermal-activated film-contacting tissue	Ultraviolet laser cut around target tissue on foil-premounted slide	Ultrasonic motion of metal tip on target tissue
Minimum size	7.5 μm	1 μm	1.5 μm
Specimen	FFPE, frozen tissue, CytoPrep, live cell	FFPE, frozen tissue, CytoPrep, organelles	FFPE, frozen tissue, CytoPrep, hard tissue
Slide type	Noncharged	No preference	No preference

FFPE formalin-fixed, paraffin-embedded, *CytoPrep* cytology preparation

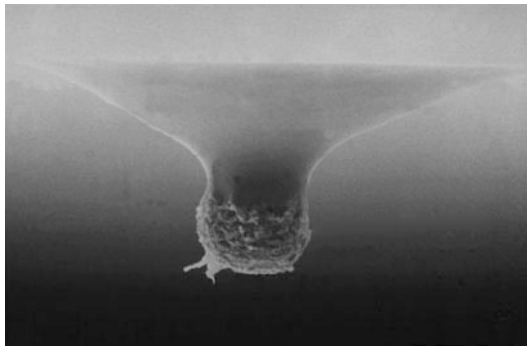


Fig. 16.7 The isolated cells are seen adherent to the transfer film in this scanning electron micrograph (Arcturus Bioscience Inc., Mountain View, California, USA)

and then sectioned with a microtome provide the best genomic and molecular preservation
  - Requires foresight since a different processing procedure is used
  - Microscopic morphology is more problematic in frozen specimens than in all other processing techniques
    - Identifying the cells of interest is more difficult
  - Specimens must be stored at −80 °C until the moment that the LCM procedure begins
- Cytology specimens
  - Isolating cells is relatively easy since most cells of interest are either already single or in small clusters comprised of identical cells
  - Fixed with methanol or ethanol
  - Archival specimens
  - Provides excellent genomic and molecular preservation, particularly since the nuclei are intact (i.e., not sectioned by a microtome or cryostat)
- Archival stained and coverslipped slides
  - A wide variety of archival specimens may be procured using LCM, including immunolabeled and FISH-hybridized cells
  - The coverslip must be removed before the tissue can be used

16.2.1.2 Specimen Processing Considerations

- For all specimens, freezing or fixation should occur as quickly as possible after cessation of perfusion to avoid degradation by endogenous RNases, nucleases, and proteases
- RNase-free techniques should be employed for RNA extraction
  - Disposable gloves
  - RNase-eliminating cleaner on instruments, including microtome, between each sample
  - Use RNase-free solutions, glassware, and plastics

16.2.1.3 Postprocessing Specimen Considerations

- Histologic sections should be 5–10 μm thick
- Cut onto uncharged, uncoverslipped glass slides
- Staining (also *see* LCM Laboratory Protocol section)
  - May be stained with hematoxylin and eosin (H&E)

Fig. 16.8 The Arcturus Veritas Series (Arcturus Bioscience Inc., Mountain View, California, USA) is capable of performing both LCM and LMM/LC technique

- H&E does reduce PCR efficiency
- H&E needs to be light and balanced
- Using only eosin (i.e., withholding hematoxylin) improves results

– Papanicolaou and Wright stains yield good results
– Tissues may be unstained, but this makes visual morphologic identification of the cells of interest vastly more difficult
– Tissues may also be stained with immunohistochemical markers
  - This can be incredibly useful in locating the cells of interest in selected studies
    – Precursor lesions
    – Borderline neoplasms
    – Early, low-grade neoplasms
    – Neoplasms with multiple components
    – Multifocal cancers
    – Other circumstances when morphology alone is not conclusive
  - Can be useful in identifying cells during LCM since their morphologic appearance can be distorted due to the lack of a coverslip
  - May be used to stain for possible contaminating cells in order to avoid their inclusion in the research sample
    – An example is using CD34 to identify endothelial cells within a tumor
  - May require an altered immunohistochemical staining protocol to decrease RNA damage

16.2.2 Major Components of Microdissection Systems

- Light microscope
- Laser source with laser control unit
- Electrically or manual operated slides stage
- CMOS/CCD camera capable of relaying real-time microscopic video and save images
- Computer system with color monitor

16.2.3 Key Characteristics of the Laser and Transfer Film

- The transfer film is composed of ethylene vinyl acetate and is 100 μm thick
  – The film has a diameter of approximately 6 mm
  – It is mounted upon an optically clear cap for ease of withdrawing it from the specimen after polymerization (Fig. 16.8)
  – The film does not transfer heat well to adjacent tissue, and the near-infrared light is closely matched to the absorption spectrum of the dye-impregnated transfer film
    - Therefore, almost all energy is absorbed by the transfer film with minimal heating (and resultant possible damage) of the underlying tissue
      – The tissue under the transfer film reaches a maximum temperature of approximately 90 °C for only a few milliseconds

  - The various biologic macromolecules are capable of enduring the process without alteration
  - Only the film directly struck by the laser photons melts, deforms, and polymerizes with the underlying cells (Fig. 16.9)
- The laser diameter (and therefore diameter of ethylene vinyl acetate that can be melted) can be adjusted from 7.5 to 30 μm
  - This can accommodate cells of various sizes or even small clusters of cells

16.2.4 Key Characteristics of the UV Laser and Lining Foil

- The foil is mounted on the upper side of the slide with glue; there is a space of 1 μm between the glass and membrane
- The foil is UV absorbing, and it is vaporized on the cutline upon laser activation
- The microdissected tissue drops by gravity or laser power

16.2.5 Expected Yield of DNA, RNA, and Proteins from Laser Microdissected Tissue

- Each human cell contains about 7 pg of genomic DNA
- Considering the DNA preservation and extraction recovery, 600–1,000 cells will provide enough DNA for several PCR amplifications
- A typical PCR needs 50–100 pg of DNA template
  - The yield of intact DNA that covers a gene from formalin-fixed, LCM-prepared samples is roughly only 1%
  - A typical mammalian cell contains 10–30 pg of total RNA, of which 5% represents mRNA
    - The quality of RNA is closely related to the tissue preservation, preparation, and ribonuclease inactivation during processing
    - Gloved hands and ribonuclease-inhibitor treated buffers are essential for protecting the RNA from degradation
    - Frozen section samples fixed with alcohol followed by acetone provide improved RNA yields
    - It is often necessary to microdissect many more cells than one would calculate in order to ensure capture of a sufficient amount of RNA
    - The protein content of cells is organ dependent and tissue specific
    - Protein degradation occurs not only due to the activity of proteolytic enzymes but also due to disruption and denaturation caused by cross-links between peptides, which often greatly affect the protein analysis
    - Generally, at least 10–100 times more cells are needed for a protein analysis than for DNA or RNA analysis

16.2.6 Benefits and Limitations of LCM for Cell Isolation

- Benefits
  - The most important advantages of laser microdissection are cell purity, speed, precision, versatility, and the ability to capture cells in their natural state
  - Overall, LCM is the least time-consuming method of cell isolation
  - Possible to document both captured and residual cells before and after microdissection
  - Different types of cells, such as normal, premalignant, and neoplastic cells, can be microdissected sequentially since LCM does not destroy adjacent tissues
  - Especially useful for isolating scattered cells admixed with other cell populations
  - Can be easily integrated into procedures for molecular genetic diagnosis
  - Can be applied to different cell or tissue preparations
  - Laser-cutting system allows contact-free collection of tissue fragments, including large intact fragments with retained histologic morphology

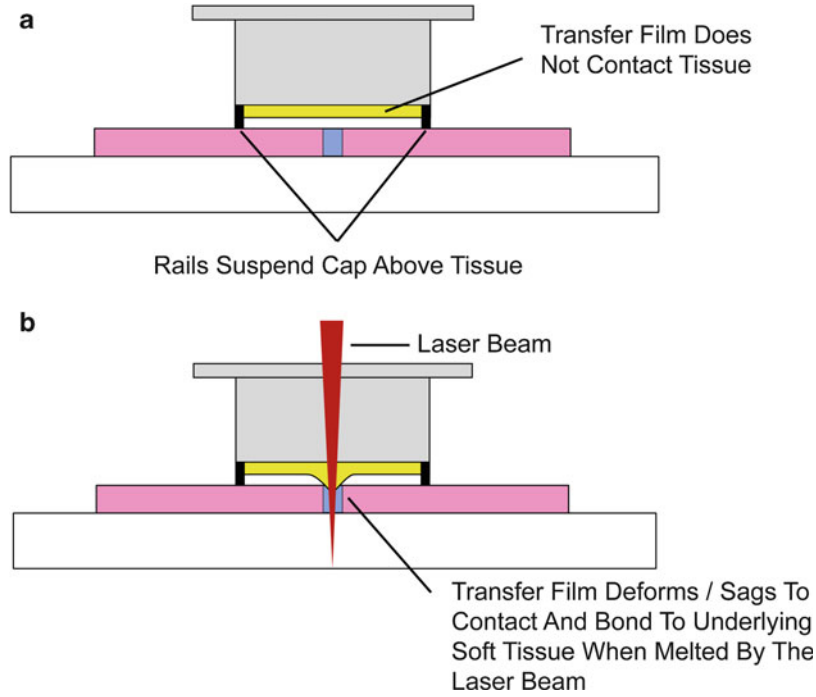


Fig. 16.9 Advanced cap/transfer films use built-in rails to suspend the transfer film just above the tissue (**a**). When the film is struck by the laser beam, the transfer film melts and deforms to contact and bond to the underlying tissue (**b**). This cap minimizes the possibility of contamination by nonspecific binding of unintended cells to the transfer film

  - Cells are attached firmly to transfer film with little risk of loss
  - Allows for examination of cells removed from their natural, in vivo surroundings
  - Less advanced lesions (e.g., dysplasia) that do not form tumors can be sampled
- Limitations
  - Difficulties of microdissection due to decreased optical resolution of stained and dehydrated tissue sections without a coverslip
  - If a tissue is mounted on positively charged slides, it may not be possible to remove the selected cells from the slide
  - Does not work well on hard tissues such as bone and cartilage
  - It is difficult to isolate cells smaller than 7.5 μm without contamination from adjacent cells
    - If a single small cell is targeted, cells adjacent to it may also be struck by the laser beam and may adhere to the transfer film
  - More expensive than manual microdissection methods
  - Potential for contamination by nonspecific weak attraction of the transfer film to cells not heated by the laser
    - The newest commercially available cap/transfer films have no direct contact with underlying tissue, thereby minimizing the possibility of contamination (Fig. 16.8)

16.3 LCM Laboratory Protocols

The following are protocols used in our laboratory

16.3.1 Slide Preparation

- Uncharged slides yield the best results
- Tissue sections should be centered on slides
- 4–5-μm sections are typically used

16.3.1.1 Formalin-Fixed, Paraffin-Embedded Slides

- See Table 16.3

Table 16.3 Formalin-fixed paraffin-embedded slides

Deparaffinize/hydrate	
Xylene	10 min
Xylene	10 min
100% ETOH	2 min
100% ETOH	2 min
85% ETOH	1 min
70% ETOH	1 min
H_2O	1 min
Stain	
Hematoxylin	30 s
H_2O rinse	
Eosin	30 s
H_2O rinse	
Dehydrate	
70% ETOH	1 min
85% ETOH	1 min
100% ETOH	5 min
100% ETOH	5 min
Xylene	5 min

Table 16.4 Frozen section slide preparation

Fix	
Cold 100% methanol or acetone at −20 °C for 10 min	
Distilled water rinse twice for 2 min	
Stain	
Hematoxylin	30 s
DEPC H_2O rinse	
Eosin	30 s
DEPC H_2O rinse	
(DEPC, diethyl pyrocarbonate, a nuclease inhibitor)	
Dehydrate	
70% ETOH	1 min
85% ETOH	1 min
100% ETOH	5 min
100% ETOH	5 min
Xylene	5 min

16.3.1.2 Frozen Section Slides

- See Table 16.4

16.3.2 Step-by-Step Procedures Using the Arcturus LCM System

- Turn on the machine using the main power switch
- Ensure that the monitor is on
- Center the joystick to make sure that the joystick is vertical and perpendicular to countertop surface
- Place a slide on the stage
- Find the area of interest, either by looking at the monitor or through the ocular lenses
- Initiate the vacuum to immobilize the slide on the stage for cell harvesting
- Align a cap on the right side of the stage on the load line
- Swing the transfer arm to the cap and firmly pull arm up to load the cap on the arm
- Swing the arm toward the specimen, and it will automatically lower over the selected area of the slide
- Enable the laser (a small red indicator light will appear in the center of the monitor screen)
- Use the laser focus wheel to focus the laser into a small, sharp, and bright red dot
- After focusing the laser, the machine is ready to harvest cells
- Choose appropriate laser size using the small arm located above the laser focus wheel
- Continue using handheld control and joystick to move around the selected field, harvesting the cells of interest until enough cells have been collected for appropriate analysis
- Lift the swing arm off the slide and place the cap into the cap holder located to the right of the loader
- Use the cap transfer tool to pick up the cap from the cap holder, and insert the cap into a 0.5-mL tube containing 50 μL appropriate digested solution
- Be sure to label the cap and the tube

16.3.3 Extraction of Genomic DNA from Formalin-Fixed, Paraffin-Embedded Tissue

- Secure pure cell population of interest using LCM
- Put the tissue into 0.5-mL Eppendorf tube containing 50 μL of digestion solution
  - *Note*: digestion solution:
    - Tris-HCl – 20 mM

- KCl – 50 mM
- $MgCl_2$ – 5 mM
- EDTA – 1 mM
- Proteinase K – 5 mg/mL

- Incubate at 37 °C overnight
- Boil the tube for 10 min to inactivate the proteinase K
- Cool down the tube with ice for 2 min
- Take 2 μL of the solution from the tube as a genomic DNA template for PCR reactions
- Estimated yields from formalin-fixed paraffin-embedded tissues
  - DNA – 7 pg per cell
  - RNA – 10–30 pg cell (1–5% of which is mRNA)
  - Protein – 10–100 pg per cell

16.3.4 Procedure and Protocol Resources

- Detailed procedures and protocols for the handling and processing of specimens as well as protocols for molecular analysis can also be found at:
  - http://dir.nichd.nih.gov/lcm/lcm.htm
  - http://www.arctur.com
  - http://www.leica-microsystems.com

16.4 Other Cellular Isolation Techniques

16.4.1 Non-LCM Microdissection Methods

16.4.1.1 Manual Extraction of Cells Under Direct Light Microscope Observation

- Concept is to manually extract small portions of tissue (often <1 mm in size) from paraffin-embedded tissue using modified Pasteur pipettes or tungsten wire needles under direct observation
  - This requires fairly advanced malignancies since sizable clusters of cells need to be present to manually obtain fairly large clusters en bloc
  - Works better for mass lesions (e.g., carcinoma) than precursor lesions (e.g., dysplasia or carcinoma in situ)
    - Malignancies are increasingly detected at far smaller sizes due to improved computed axial tomography technology and serologic studies including carcinoembryonic antigen, cancer antigen 125, and prostate-specific antigen
    - Genetic characteristics of cells in large mass lesions may be significantly different from cells in smaller lesions
    - Cannot be used to study malignancies like Hodgkin lymphoma since Reed–Sternberg cells are rare and scattered in an abundant reactive inflammatory cell infiltrate
- No expensive procurement equipment needed
- Very tedious and time consuming
- Depends on personal skills
- Susceptible to operator variability and room airflow
- Sample is not completely pure since admixed cells may be included

16.4.1.2 Selective Ultraviolet Radiation Fractionation (SURF)

- Concept is to use ultraviolet (UV) radiation to destroy all the unwanted tissue in the paraffin-embedded block
- Achieved by protecting the areas of interest with ink dots
- Cells that are not protected from the radiation then have damaged DNA, which is unsuitable for PCR
- The technique is able to isolate cells of interest; however, the remaining tissue is completely destroyed
  - Cells of interest cannot be readily compared with surrounding stromal cells
  - No tissue archiving possible

16.4.1.3 Laser Microbeam Microdissection (LMM)/Laser Cutting (LC)

- The main modern rival to LCM
- Uses a highly focused UV laser beam to photoablate a thin rim of tissue around the cells of interest

- After the cells of interest are cut from the adjacent cells, they are transferred into their media using either a needle tip or laser pressure catapulting
  - This step has increased potential for loss of the specimen and requires more technical skill than LCM
- Advantage of LMM/LC compared with LCM
  - Suitable for collection of tissue fragments, including large intact fragments with retained histologic morphology
  - Has the advantage of working on hard tissues, which have stronger intercellular bonds
    - Bone
    - Cartilage
  - Avoids the potential for nonspecific cell adhesion to the transfer film that is possible with LCM
  - Faster than LCM for isolating relatively large areas of tissue
- Disadvantages of LMM/LC compared with LCM
  - Less adept at isolating single cells
  - Destroys adjacent tissue
  - More technically challenging
  - Typically requires more time
- Some instruments are capable of performing both LCM and LMM/LC techniques (Fig. 16.9)

16.4.2 Nonmicrodissection Methods That Enrich the Target Cells

16.4.2.1 Xenograft Enrichment

- Serial passage of tissues through immunodeficient rodents to obtain human tumor cell populations whose nonmalignant cells are rodent in origin
  - Essentially limitless numbers of cells can be obtained
  - Self-propagating
  - Extensive expertise is required
  - An animal facility is required
  - The cell population requires up to 6 months to establish
  - The tumor cells are susceptible to additional genetic changes over time
    - A subset of cells with evolutionary advantage (i.e., more aggressive) may propagate
- Rodent cells may still contaminate the human tumor cells
  - The in vivo microenvironment is significantly different from the artificial in vitro microenvironment
    - Different exogenous factors are available for the cell to utilize
    - Different interactions may cause various upregulation and downregulation of genes with resultant changes in protein expression profiles

16.4.2.2 Cell Cultures (Not Cell Lines, Which Are Permanent and Monoclonal)

- Allowing human cells to mitotically divide in vitro in culture media
  - Many cells can be obtained
  - Takes time to establish
  - A fair degree of expertise is required
  - The tumor cells are susceptible to additional genetic changes over time
  - Most cell lines are derived from late stage malignancies
    - Genetic alterations in these cells may be significantly different from earlier, more diagnostically difficult lesions or lesions whose diagnosis would have prognostic impact on the patient

16.4.2.3 Cell Sorting

- Using cell density or ability to bind specific labeled antibodies to sort the cells by their characteristics
- Using flow cytometry to sort cells labeled with fluorescence conjugated antibody
- Cell sorting requires cell suspension
  - Works relatively well for hematologic disorders
  - In order to analyze solid tumors, it is necessary to dissociate the tissue with enzymes, such as collagenase, trypsin, or papain. Some mechanical dissociation may also be needed

16.5 Methods of Analysis and Potential Application of LCM

16.5.1 Methods of Molecular Analysis

- A wide variety of techniques are available for molecular analysis (Fig. 16.10)
  - DNA, RNA, and proteins are the macromolecules of analytic interest
    - Each can be altered in different disease states
    - PCR can be employed to amplify the desired DNA and RNA

16.5.1.1 DNA Analytic Techniques

- DNA array and dideoxy fingerprinting
- Loss of heterozygosity
- X chromosome inactivation
- Microsatellite instability
- DNA methylation
- Direct DNA sequencing
- Single-strand conformation polymorphism
- Restriction fragment length polymorphism
- Comparative genomic hybridization
- Southern blot
- Others

16.5.1.2 RNA Analytic Techniques

- Complementary DNA (cDNA) array and expression profiling
- cDNA sequencing
- Northern blot
- Real-time PCR
- Genetic cloning
- MicroRNA fingerprinting
- Others

16.5.1.3 Protein Analytic Techniques

- Immunoprecipitation for protein purification
- Antibody screening and drug discovery
- Western blotting
- Peptide sequencing
- Mass spectrometry
- UV spectrometry
- Chromatography
- Others

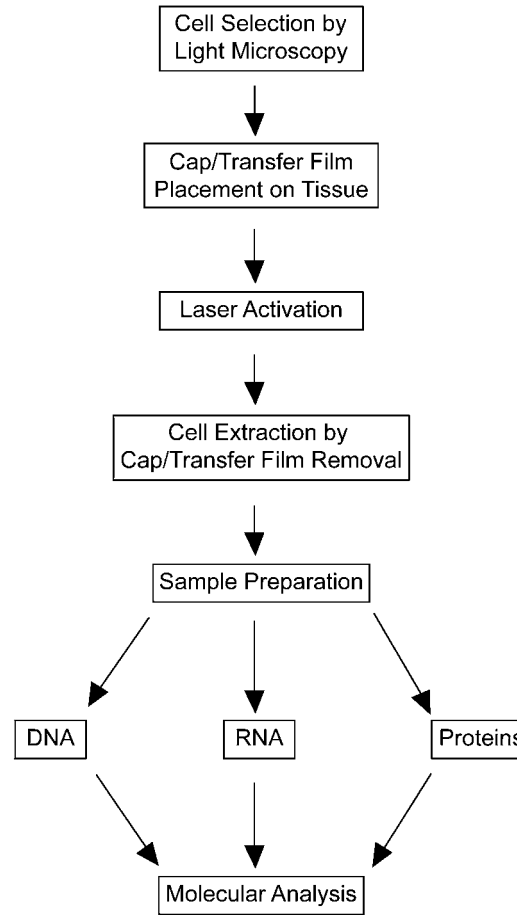


Fig. 16.10 Steps of LCM and molecular analysis of macromolecular isolates

16.5.2 Potential Applications of LCM

See Table 16.5

16.5.2.1 Diagnosis

- Discriminating between precursor lesions and early malignancy
  - Malignancy has typically been defined based upon an array of cytologic and architectural abnormalities observed under a light microscope; yet, before these morphologic manifestations, there are cellular and genetic alterations, which may prove to be a better definition of malignancy

Table 16.5 Major LCM downstream applications

Application	Examples
DNA based	Genetic mutation
	Epigenetic alteration
	Loss of heterozygosity
	X chromosome inactivation
	Restriction fragment length polymorphism
	Single-strand conformation polymorphism
	Primer extension preamplification
	Whole genome amplification
	Comparative genomic hybridization
	Microarrays
RNA based	Gene expression profile
	Representation difference analysis
	Serial analysis of gene expression
	Expressed sequence tag
	Gene chips
Protein based	Western blotting
	Mass spectrometry
	Surface-enhanced laser desorption ionization
	Peptide sequencing
	Two-dimensional electrophoreses (2D-PAGE)
	Protein microarrays
microRNA based	Posttranscription regulation of gene
	Tissue specific markers
	Tumor developmental lineage signature
	Pathogenesis, diagnosis, and prognosis
	Chemosensitivity and treatment
Special analysis	Gene mutation analysis on immunohistochemistry or FISH labeled cells

  - Carcinogenesis is a multistep process, often with a sequential order of genomic alterations appreciable in precursor lesions
- Diagnosing lesions with challenging morphologic criteria
  - Examples include histologically deceptive and benign-appearing variants of carcinoma in various anatomic sites such as bladder, skin, and soft tissues
  - Some tumors are currently classified by size criteria, such as papillary adenoma of the kidneys, which is distinguished from papillary renal carcinoma purely based on an arbitrary size cutoff
    - Conceptually, if cancer is a clonal proliferation derived from a single cell, carcinoma could exist at a very small, even single cell, level
    - Genetic profiling may be more accurate in tumor classifications
  - Many hematologic malignancies are very difficult to classify based upon light microscopy, and molecular techniques can prove essential
    - Also, using LCM can avoid the difficulties in establishing cell line cultures, which frequently requires complicated karyotyping
- Ascertaining tumor origin of unknown primary tumor site
  - Not infrequently, metastases become clinically apparent before their primary cancers are diagnosed

- Characterizing certain primary cancers can be quite difficult
- Comparing the genetic/expression/miRNA profile of the unknown tumor against that of a well-characterized database of various cancers may facilitate a diagnosis, e.g., http://bioinformatics.oxfordjournals.org, http://www.mirbase.org/

16.5.2.2 Prognosis

- Predicting how aggressively a malignancy is likely to behave
  - For some malignancies, aggressive behavior is known to be correlated with specific translocations or ploidy
    - Examples include t(9;22) BCR–ABL in some hematopoietic malignancies and ZAP70 in chronic lymphocytic leukemia/small lymphocytic lymphoma
    - Those chromosomal translocations are detectable by PCR from isolated cell populations
    - DNA sequencing for detecting genetic mutations predicts the clinical outcome and response to targeted therapy, such as *FGFR3* mutation in urothelial cancer and *EGFR* mutation in lung adenocarcinomas
  - Morphologic grading of tumor differentiation is quite subjective, and therefore not consistent
    - The degree of macromolecular derangement may prove to be a more accurate and more objective measure
- Assessing future malignancy risk
  - In some circumstances, the patient's risk of developing additional malignancies may be surmised once the nature of the initial presenting malignancy is known
    - Examples include multiple endocrine neoplasm (MEN) syndromes, some of which are associated with specific gene mutations
  - Familial risk for developing a malignancy may also be better assessed by genetic changes
    - Examples include *BRCA1* and *BRCA2* in breast, ovarian, and prostate carcinoma
- Assessing and predicting the utility of targeted therapy
  - *EGFR* mutations at exon 18, 19, and 21 are associated with response to EGFR tyrosine kinase inhibitors, whereas mutations at exon 20 are associated with resistance to the therapy
  - *KRAS* mutation is predictive of resistance to targeted therapy in patients with colorectal cancer
  - Tests for specific protein (over)expression may aid in directing therapy
    - An example includes HER2 overexpression in breast carcinoma
- Determining the nature of multiple tumors
  - In some instances, multiple morphologically similar tumors cannot be definitively classified using light microscopy as representing multiple independent primary tumors or intraglandular metastasis

16.5.2.3 Drug Discoveries

- As a larger molecular database is collected, we may be able to find targets for additional drug mechanisms of action
  - HER2, EGFR, and c-kit are just a few early examples of molecular characteristics that allow tumors to receive tailored treatment modalities

16.5.2.4 Scientific Inquiry

- Identifying the molecular differences between normal, premalignant, and malignant tissues by evaluating various macromolecular traits increases our overall knowledge of disease and will contribute to future diagnostic, prognostic, and treatment modalities
  - The National Cancer Institute is currently working on the Cancer Genome Anatomy Project, which is attempting to achieve this goal by creating a cDNA library based on cell protein expression profiles
    - http://wwwncbi.nlm.nih.gov/projects/CGAP/

Further Reading

Armstrong AB, Wang M, Eble JN, et al. TP53 mutational analysis supports monoclonal origin of biphasic sarcomatoid urothelial carcinoma (carcinosarcoma) of the urinary bladder. Mod Pathol. 2009;22:113–8.

Ashley EA, Butte AJ, Wheeler MT, et al. Clinical assessment incorporating a personal genome. Lancet. 2010;375:1525–35.

Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. Cell. 2011;145:30–8.

Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer. 2011;11:726–34.

Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. Nat Med. 2011;17:320–9.

Bonner RF, Emmert-Buck M, Cole K, et al. Laser capture microdissection: molecular analysis of tissue. Science. 1997;278:1481–3.

Cheng L, Song SY, Pretlow TG, et al. Evidence of independent origin of multiple tumors from prostate cancer patients. J Natl Cancer Inst. 1997;90:233–7.

Cheng L, Gu J, Eble JN, et al. Molecular genetic evidence for different clonal origin of components of human renal angiomyolipomas. Am J Surg Pathol. 2001;25:1231–6.

Cheng L, Gu J, Ulbright TM, et al. Precise microdissection of human bladder carcinomas reveals divergent tumor subclones in the same tumor. Cancer. 2002;94:104–10.

Cheng L, MacLennan GT, Zhang S, et al. Laser capture microdissection analysis reveals frequent allelic losses in papillary urothelial neoplasm of low malignant potential of the urinary bladder. Cancer. 2004;101:183–8.

Cheng L, Jones TD, McCarthy RP, et al. Molecular genetic evidence for a common clonal origin of urinary bladder small cell carcinoma and coexisting urothelial carcinoma. Am J Pathol. 2005;166:1533–9.

Cheng L, Zhang S, Wang M, et al. Molecular genetic evidence supporting the neoplastic nature of stromal cells in "fibrosis" after chemotherapy for testicular germ cell tumours. J Pathol. 2007;213:65–71.

Cheng AL, Huang WG, Chen ZC, et al. Identification of novel nasopharyngeal carcinoma biomarkers by laser capture microdissection and proteomic analysis. Clin Cancer Res. 2008;14:435–45.

Cheng L, Zhang S, MacLennan GT, et al. Laser-assisted microdissection in translational research: theory, technical considerations, and future applications. Appl Immunohistochem Mol Morphol. 2013;21:31–47.

Chin L, Andersen JN, Futreal PA. Cancer genomics: from discovery science to personalized medicine. Nat Med. 2011a;17:297–303.

Chin L, Hahn WC, Getz G, Meyerson M. Making sense of cancer genomic data. Genes Dev. 2011b;25:534–55.

Chowdhuri SR, Xi L, Pham TH, et al. *EGFR* and *KRAS* mutation analysis in cytologic samples of lung adenocarcinoma enabled by laser capture microdissection. Mod Pathol. 2012;25:548–55.

Emerson RE, Wang M, Roth LM, et al. Molecular genetic evidence supporting the neoplastic nature of the Leydig cell component of ovarian Sertoli-Leydig cell tumors. Int J Gynecol Pathol. 2007;26:368–74.

Emmert-Buck MR, Bonner RF, Smith PD, et al. Laser capture microdissection. Science. 1996;274:998–1001.

Espina V, Heiby M, Pierobon M, et al. Laser capture microdissection technology. Expert Rev Mol Diagn. 2007;7:647–57.

Finak G, Bertos N, Pepin F, et al. Stromal gene expression predicts clinical outcome in breast cancer. Nat Med. 2008;14:518–27.

Frohling S, Dohner H. Chromosomal abnormalities in cancer. N Engl J Med. 2008;359:722–34.

Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481:306–13.

Green ED, Guyer MS. National human genome research I. Charting a course for genomic medicine from base pairs to bedside. Nature. 2011;470:204–13.

Haber DA, Gray NS, Baselga J. The evolving war on cancer. Cell. 2011;145:19–24.

Hamburg MA, Collins FS. The path to personalized medicine. N Engl J Med. 2010;363:301–4.

Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74.

Hanson JC, Tangrea MA, Kim S, et al. Expression microdissection adapted to commercial laser dissection instruments. Nat Protoc. 2011;6:457–67.

Harris TJ, McCormick F. The molecular pathology of cancer. Nat Rev Clin Oncol. 2010;7:251–65.

Hernandez S, Lloreta J. Manual versus laser microdissection in molecular biology. Ultrastruct Pathol. 2006;30:221–8.

Jones T, Carr M, Eble JN, et al. Clonal origin of lymph node metastases in bladder carcinoma. Cancer. 2005a;104:1901–10.

Jones TD, Eble JN, Wang M, et al. Molecular genetic evidence for the independent origin of multifocal papillary tumors in patients with papillary renal cell carcinomas. Clin Cancer Res. 2005b;11:7226–33.

Jones TD, Wang M, Eble JN, et al. Molecular evidence supporting field effect in urothelial carcinogenesis. Clin Cancer Res. 2005c;11:6512–9.

Jones TD, Wang M, Sung MT, et al. Clonal origin of metastatic testicular teratomas. Clin Cancer Res. 2006;12:5377–83.

Katona TM, Jones TD, Wang M, et al. Molecular evidence for independent origin of multifocal neuroendocrine tumors of the enteropancreatic axis. Cancer Res. 2006;66:4936–42.

Katona TM, Jones TD, Wang M, et al. Genetically heterogeneous and clonally unrelated metastases may arise in patients with cutaneous melanoma. Am J Surg Pathol. 2007;31:1029–37.

Kernek KM, Ulbright TM, Zhang S, et al. Identical allelic losses in mature teratoma and other histologic components of malignant mixed germ cell tumors of the testis. Am J Pathol. 2003;163:2477–84.

Kum JB, Grignon DJ, Wang M, et al. Mixed epithelial and stromal tumors of the kidney: evidence for a single cell of origin with capacity for epithelial and stromal differentiation. Am J Surg Pathol. 2011;35:1114–22.

Lander ES. Initial impact of the sequencing of the human genome. Nature. 2011;470:187–97.

Loeb LA. Human cancers express mutator phenotypes: origin, consequences and targeting. Nat Rev Cancer. 2011;11:450–7.

Macconaill LE, Garraway LA. Clinical implications of the cancer genome. J Clin Oncol. 2010;28:5219–28.

Maitra A, Wistuba II, Virmani AK, et al. Enrichment of epithelial cells for molecular studies. Nat Med. 1999;5:459–63.

Makino H, Uetake H, Danenberg K, et al. Efficacy of laser capture microdissection plus RT-PCR technique in analyzing gene expression levels in human gastric cancer and colon cancer. BMC Cancer. 2008;8:210.

Manolio TA, Green ED. Genomics reaches the clinic: from basic discoveries to clinical impact. Cell. 2011;147:14–6.

McCarthy RP, Zhang S, Bostwick DG, et al. Molecular genetic evidence for different clonal origins of epithelial and stromal components of phyllodes tumor of the prostate. Am J Pathol. 2004;165: 1395–400.

McCarthy RP, Wang M, Jones TD, et al. Molecular evidence for the same clonal origin of multifocal papillary thyroid carcinomas. Clin Cancer Res. 2006;12: 2414–8.

McDermott U, Downing JR, Stratton MR. Genomics and the continuum of cancer care. N Engl J Med. 2011;364:340–50.

Mokhtar NM, Ramzi NH, Yin-Ling W, et al. Laser capture microdissection with genome-wide expression profiling displayed gene expression signatures in endometrioid endometrial cancer. Cancer Invest. 2011;30:156–64.

Patel V, Hood BL, Molinolo AA, et al. Proteomic analysis of laser-captured paraffin-embedded tissues: a molecular portrait of head and neck cancer progression. Clin Cancer Res. 2008;14:1002–14.

Potti A, Schilsky RL, Nevins JR. Refocusing the war on cancer: the critical role of personalized treatment. Sci Transl Med. 2010;2:28cm13.

Rini BI, Atkins MB. Resistance to targeted therapy in renal-cell carcinoma. Lancet Oncol. 2009;10:992–1000.

Roychowdhury S, Iyer MK, Robinson DR, et al. Personalized oncology through integrative high-throughput sequencing: a pilot study. Sci Transl Med. 2011;3:111ra21.

Schmid MW, Schmidt A, Klostermeier UC, et al. A powerful method for transcriptional profiling of specific cell types in eukaryotes: laser-assisted microdissection and RNA sequencing. PLoS One. 2012;7:e29685.

Standaert DG. Applications of laser capture microdissection in the study of neurodegenerative disease. Arch Neurol. 2005;62:203–5.

Stratton MR. Exploring the genomes of cancer cells: progress and promise. Science. 2011;331:1553–8.

Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458:719–24.

Suarez-Quian CA, Goldstein SR, Bonner RF. Laser capture microdissection: a new tool for the study of spermatogenesis. J Androl. 2000;21:601–8.

Sung MT, Wang M, Maclennan G, et al. Histogenesis of sarcomatoid urothelial carcinoma of the urinary bladder: evidence for a common clonal origin with divergent differentiation. J Pathol. 2007;211: 420–30.

Sung MT, Zhang S, MacLennan GT, et al. Histogenesis of clear cell adenocarcinoma in the urinary tract: evidence of urothelial origin. Clin Cancer Res. 2008;14:1947–55.

Tangrea MA, Mukherjee S, Gao B, et al. Effect of immunohistochemistry on molecular analysis of tissue samples: implications for microdissection technologies. J Histochem Cytochem. 2011;59:591–600.

Velculescu VE, Diaz LA, Jr. Understanding the enemy. Sci Transl Med. 2011;3:98ps37.

Visvader JE. Cells of origin in cancer. Nature. 2011;469:314–22.

Wang H, Owens JD, Shih JH, et al. Histological staining methods preparatory to laser capture microdissection significantly affect the integrity of the cellular RNA. BMC Genomics. 2006;7:97.

Wang X, Wang M, MacLennan GT, et al. Evidence for common clonal origin of multifocal lung cancers. J Natl Cancer Inst. 2009;101:560–70.

Weinstein IB, Joe A. Oncogene addiction. Cancer Res. 2008;68:3077–80; discussion 80.

Zheng S, Cummings OW, Saxena R, et al. Clonality and TP53 mutation analysis of focal nodular hyperplasia of the liver. Am J Clin Pathol. 2010;134:65–70.

Human Tissue Biorepository

17

William E. Grizzle, Katherine C. Sexton, and Walter C. Bell

Contents

17.1 Models of Human Tissue Repositories ... 484

17.2 Aids in Developing Tissue Repositories ... 484

17.3 Design of Repository ... 485

17.4 Challenges in Collecting Specimens ... 485

17.5 Issues in Processing of Specimens ... 486

17.6 Long-Term Storage of Biospecimens ... 487

17.7 Informatics and Records of Collecting, Processing, Storage, and Distribution of Tissue Specimens ... 487

17.8 Types of Specimens Provided by a Tissue Repository ... 487

17.9 Distribution of Specimens to Investigators ... 487

17.10 Quality Assurance Program ... 488

17.11 Education and Training in Issues Related to Tissue Repositories ... 490

17.12 Matching Available Tissues to Investigator Needs for Tissues ... 490

17.13 Equitable Distribution of Tissues ... 491

17.14 Annotation of Tissue Specimens ... 491

17.15 Bias Affecting Usefulness of Tissues ... 492

17.16 Administrative Issues ... 492

17.17 Shipping ... 493

17.18 Regulatory and Ethical Issues Affecting Tissue Repositories ... 493

17.19 Cost Recovery ... 494

17.20 Safety in Tissue Repositories ... 494

17.21 Informatics ... 495

17.22 Challenges in Meeting Specimens for Genomic Studies ... 495

17.23 Challenges Leading to New Directions in Tissue Repositories ... 496

Further Reading ... 496

W.E. Grizzle, MD, PhD (✉) • W.C. Bell, MD
Department of Pathology, The University of Alabama at Birmingham School of Medicine, Birmingham, AL, USA

K.C. Sexton, MBA
Comprehensive Cancer Center, The University of Alabama at Birmingham, Birmingham, AL, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_17,

17.1 Models of Human Tissue Repositories

Human Tissue Biorepository – an organization that is involved in any aspect of collecting, processing, storing, banking, or distributing human tissues for research, education, or clinical uses. This chapter focuses only on human tissue biorepositories supporting research.

- "Catch as catch can" – disorganized approach to collection, processing, storage, and distribution. The repository collection may not be very useful to support research
- Banking model – general tissues or specific tissues are collected according to a standard operating procedure (SOP)
  - Emphasis is usually on banking rather than distribution
  - Tissues are immediately available
  - Outcome and clinical information may be immediately available
  - Fresh specimens usually are not available
  - Increased storage requirements
  - Specimens may not meet investigator needs
  - Specimens may not be used
- Clinical trial model – subtype of banking model in which remnant tissues from clinical trials are stored
  - Specimens usually consist of cases and controls, both of which have extensive clinical information and outcome
  - Some specimens may not be able to be used for general research due to limitations of informed consent
  - Some specimens may not meet investigator needs
  - Some specimens may not be used
- Prospective model – human tissue specimens are collected to specifically meet each investigator's specific requests
  - Specimens are collected with SOPs which are established by the investigator
  - Most specimens are used, so operation is very efficient and cost effective
  - Decreased storage requirements
  - Fresh specimens are available
  - Specimens are collected prospectively so they are not immediately available when requested
  - Clinical outcome and other data are not available when specimens are requested. Outcome may take many years depending upon the disease process
- Combination of banking and prospective models – best of both approaches

17.2 Aids in Developing Tissue Repositories

- Best practices
  - International Society of Biological and Environmental Repositories (ISBER), third edition has been published in Biopreservation and Tissue Banking
  - National Cancer Institute, Best Practices, second edition
- Comprehensive and special references on repository operations
- General issues to be considered in development and operations of a human tissue biorepository
  - Quality assurance (QA) in operations
    - All operations based on a QA program
    - Quality control of tissue specimens
  - Repository science – knowledge of factors affecting quality of tissues provided to support research
  - Safety
    - Repository personnel
    - Investigators and their personnel
  - Regulations and ethics, IRBs, privacy board
    - Repositories
      - Institutional Review Board (IRB) approves repository – usually requires full review
      - Repository meets Health Insurance Portability and Accountability Act (HIPAA) requirements for protected health care information and standards for security of informatics program
    - Investigators
      - IRB may classify study as nonhuman research
      - IRB may classify study as exempt

  - IRB may approve study
- HIPAA may not apply if
  - None of 18 personal identifiers as defined by HIPAA are utilized in patient data
  - Limited dataset, which includes dates, can be used with a research agreement
- IRB may classify study as nonhuman research
  - De-identified cases (coded) – investigators have no access to the identities of tissue sources
  - Anonymized – neither repository nor investigator has access to identities of tissue sources
- IRB classifies study as exempt based on de-identified or anonymized tissue specimens
- IRB approves study – expedited or full review

17.3 Design of Repository

- Goals
  - Who are the customers and what do these investigators want?
    - What do they need?
    - What is available?
  - What types of tissues (e.g., frozen tissue) to collect and provide?
  - What is the model of a repository which will be used?
  - What not to collect?
    - Consider if specimens infected with blood-borne pathogens (e.g., HIV) will be collected
  - What not to do – do not perform the investigators research for them
- Resources to support the repository (see Box 17.1)

Box 17.1. Resources to support the repository

- Initial resources available
  - Space
  - Equipment
  - Funds
    - Personnel
- Long-term support
  - Long-term commitment – is it worth the effort to start the repository?
  - Cost recovery for all work performed for investigators

17.4 Challenges in Collecting Specimens

- Identification of cases of interest
  - Collecting bodily fluids at clinic visit prior to the operation
    - Fasting versus nonfasting
    - Efficiency of collection
    - Avoiding a separate blood draw
  - Operative schedules to identify cases of interest
  - Distributing sterile collecting containers to operating room (OR)
  - Rapid transfer from OR to pathology for diagnostic review
  - Obtaining specimens not needed for diagnosis directly from pathology
- Limitations to collection of specimens
  - Neoadjuvant therapy – therapy that occurs prior to surgery so that the tissue is "affected" before being collected
    - Population of cells may be lost from the tumor or the tumor and/or metastases may be destroyed
    - This is a great challenge to tissue biorepositories independent of the model
  - Size that requires all specimens to be submitted diagnostically
    - More effective screening/detection/diagnosis is resulting in smaller tissue specimens, especially tumors
  - Preinvasive neoplastic specimens (e.g., ductal carcinoma in situ of the breast) completely submitted to identify any areas of invasive disease

  - Newer approaches to diagnosis
    - Imaging rather than biopsy to identify metastases and smaller tumors
    - Needle aspirations to identify larger tumors and metastasis resulting in no excess tissues being available to support research
  - Warm ischemia during surgery affecting tissue available for research
- When surgery is undertaken, blood vessels are completely or partially compromised and a period described as warm (body temperature) ischemia may affect specimens
  - Ischemia causes loss of rapidly degraded molecules including mRNAs and phosphoproteins
  - Blood vessels usually are not compromised when biopsies are obtained
  - There may be differences between rapidly degraded molecules in biopsies versus resections
  - Because many rapidly degraded molecules have been lost during warm ischemia, remaining molecules seem much more stable when tissue removed from body
  - Time for collection/processing – speed the processing to limit cold ischemia
- After the removal of tumors from body, changes may occur less rapidly than changes with the body during ischemia
  - Resources may limit ability to rapidly process specimens
  - Flow of specimens from the operating room (OR) should be considered
  - Pathology administrative requirements (e.g., when to enter the specimen into clinical informatics system)
  - Quickly processed aliquots versus longer processed aliquots from same tissues
  - Keep specimen at 4 °C until processed
  - If a variable cannot be controlled, record the variable
- New approaches to collection of samples/research
  - Nitrocellulose blots to collect samples of DNA/RNA/protein without affecting tumor samples
  - Blots of tissue on microscope slide
  - Smaller aliquots to support novel methods of research
  - Expanded use of paraffin blocks
  - Tissue microarrays
    - A statistical sample of paraffin blocks
    - Design to include controls

17.5 Issues in Processing of Specimens

- Processing of specimens should be rapid and matched to the specimens
  - Personnel should be flexible in their training so they can perform multiple tasks
  - The quality assurance program should cover processing
  - SOPs should be developed for each step of processing
    - SOPs for each bodily fluid (e.g., serum, urine)
    - SOPs matched to banking of solid tissue
    - SOPs matched to preparation of each solid tissue/bodily fluid (e.g., paraffin block)
    - SOPs matched to investigator needs for prospective collection of solid tissue/bodily fluids
  - Times of processing should be documented in informatics system
    - Monitor and document parameters of fixation and tissue processing
    - Reversing effects of fixation and tissue processing
  - Sample aliquots should be obtained for quality control (QC)
  - Samples should be uniquely labeled
  - Processing of specimens should be matched to investigator needs
- Importance of correct labeling
  - Errors may occur in reading label
  - Labels may separate from the specimen
  - Bar coding (a method to limit problems with specimen identification)
    - Unique to smallest aliquot and recognizes any child samples

    - May be linked to extensive information in the associated informatics system
- Age, race, sex, history, prior therapy, outcome
- Specimen characteristics – size, times of collection and processing
- Storage information and sites on each aliquot
  - Personnel errors may occur in identifying specimens

17.6 Long-Term Storage of Biospecimens

- Options from liquid nitrogen to room temperature
- To maintain cell viability over time usually requires at least storage in liquid nitrogen vapor and no freeze-thaw cycles
- For frozen tissues, we know that −20 °C (non-self-defrost) is probably okay for 1 month, but after several months, tissues begin to lose molecular features compared to storage at −80 °C
- Differences between −80 °C and liquid nitrogen vapor storage are not obvious at the protein level, and in a study by the CHTN −80 °C storage was better at mRNA level

17.7 Informatics and Records of Collecting, Processing, Storage, and Distribution of Tissue Specimens

- There are many variables in collection, processing, storage, and distribution of specimens which cannot be accurately controlled
- Information on each aliquot and the associated variables are linked by a code which frequently is in the form of a bar code
- When variables cannot be controlled exactly, information as to times and conditions of processing to distribution should be documented in the informatics system
- Location at which each unique aliquot is stored is also included in the informatics system
- Detailed information is a critical component of the quality assurance program as is the quality control diagnosis
- When specimens are removed from storage, information should be documented as to their distribution or destruction
- Sometimes specimens may be removed from storage and further aliquoted
  - These new aliquots (child samples) should be uniquely labeled and identified in informatics system
- The information permits reconstruction of the history of the aliquot

17.8 Types of Specimens Provided by a Tissue Repository

- Tissue repositories can provide a narrow and/or wide range of specimens of tissue
- Many banks focus on the collection of samples of blood/blood products
- Other tissue repositories may provide only paraffin blocks or paraffin sections of tissue
- Many prospective or combined models of repositories can provide fresh specimens, frozen specimens, blood products, paraffin material, tissue microarrays, mRNA, microRNA, and/or DNA

17.9 Distribution of Specimens to Investigators

- Internally, investigators should arrange for specimen pickup at tissue repository
- External distribution should rely on reliable commercial couriers who permit and understand the transfer of biological specimens (e.g., Federal Express)
- Air shipments must meet the standards of the International Air Transport Association (IATA)
  - This includes training of at least one employee involved in shipping as to IATA requirements
  - Meeting IATA requirements is the responsibility of the biorepository, not the courier

- Shipments scheduled to avoid problems and identification of problems (e.g., Monday to Thursday)
  - Special arrangements should be made for shipments on Friday or Saturday
- Problems with shipments should be identified rapidly by investigators (postcard) and corrected
- The informatics system should track the distribution of specimens

17.10 Quality Assurance Program

- Quality assurance (QA) is the overall process by which the uniformity and the quality of all operations of an organization are optimized with an emphasis on quality
- QC is an aspect of QA that ensures that a process or a product meets defined standards
- Good manufacturing practices should be followed as described by the International Standardization Organization (ISO) (ISO9001)
- Standard operating procedures (SOPs) are standardized approaches to performance of specific operations of an organization (e.g., collecting serum from patients with idiopathic pulmonary hypertension)
- SOPs are written detailed descriptions of each activity performed by a tissue repository so that the activity could be performed reliably by personnel who have not performed the activity previously
  - SOPs are reviewed annually and revised if necessary
  - A change in an SOP is made by authorized designated supervisors, and the change is dated as to when made and identified as to who made the change
  - Replaced SOPs maintained as archival SOPs so prior activity can be understood and reconstructed
  - SOPs organized into a procedure manual
- Equipment standardization, maintenance, and operations are an integral component of QA
- Audits are written periodic evaluations of operating procedures and infrastructure of the tissue repository
  - Conducted regularly by QA employees
  - Adherence of all employees to the QA program
  - Submitted to upper management and chief executive officer of the tissue repository
  - Examples of audits are provided in Box 17.2
- Audits evaluate all aspects of the QA program
- QA personnel and administrative issues
  - Personnel of the QA program should be independent of management of operations, sales, etc.
  - Head of QA reports to head of the organization or designates
  - Audits reported to head of the organization and problems with operations or QC should be corrected
  - QA personnel should aid with SOPs
- Surveys document user satisfaction
  - Short term – how specimens are distributed/shipped (Is dry ice adequate? Did it arrive on time?)
  - Long term – satisfaction of investigators/users with interactions with the biorepository and quantity, quality, and usefulness of specimens
  - Reviewed by QA personnel
- Quality control of solid tissues – quality control monitors the process and the usefulness of the tissue specimens provided to investigators
  - Verification of diagnosis of actual specimens provided to investigators or stored in tissue banks

Box 17.2. Audits that may be incorporated into a QA program

- Adherence to SOPs
- Storage of specimens
- Equipment repair and maintenance
- Equipment monitoring (e.g., freezer temperatures)
- Current training of personnel
- Quality of specimens (QC)
- Survey of investigators as to their satisfaction with products and problems with products

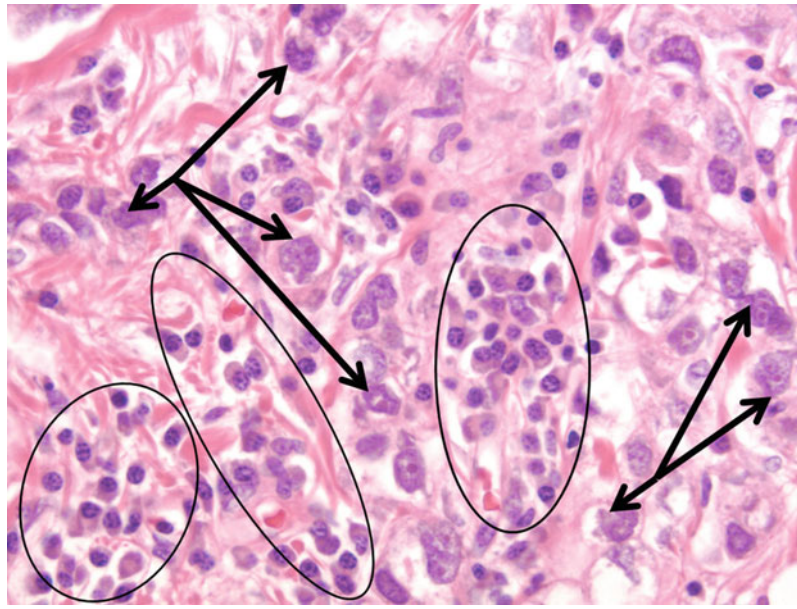


Fig. 17.1 This figure demonstrates a breast cancer (H&E, X630) in which the cancer cells, examples demonstrated by the *black arrows*, are extensively intermixed with inflammatory cells (within *ovals*). The cancer cells could be removed and analyzed as a separate cellular compartment by laser capture microdissection, a very labor intensive process. Enrichment of the number of cancer cells by micro-/macrodissection in which a paraffin or frozen section of the tumor is used as a guide to remove noncancer cells cannot be accomplished due to the close intermixture of cancer and noncancer cells. Of note, this tumor contains more inflammatory cells than cancer cells (about a 3:1 mixture), and many of the inflammatory cells are plasma cells. If the tumor were analyzed by homogenization and extraction of mRNA, microRNA, DNA, and/or protein, most of the extracted molecules would come primarily from noncancer cells. This critical limitation on the use of such specimens in molecular research is important for investigators to understand

  - Diagnosis confirmed by mirror image aliquots examined by a pathologist. This is a critical component of QC
  - For tumors, identify % tumor and of the tumor % necrosis, % fibrosis/mucin and % tumor nuclei
  - Additional quality control can be added which can include molecular quality control (e.g., RIN number via Agilent 2100 system or actual molecular analysis with RT-Q-PCR)
  - For some very infiltrative tumors such as pancreas or prostate or when there is extensive inflammation, microdissection or macrodissection may be necessary to isolate neoplastic cells
    - This can be done on frozen or paraffin embedded tissues
  - Approaches for investigators should be documented and there should be cost recovery (Figs. 17.1 and 17.2)
- Quality control in the collection of biological fluids
  - Fluids collected include fluids from fine needle aspiration, whole blood, serum, plasma, buffy coats, urine, saliva, and rare fluids such as pancreatic duct fluid ("juice") and cerebrospinal fluid
- Diagnosis is based upon overall diagnosis of the patient or sometimes a subdiagnosis (e.g., pancreatic cancer in patient with diabetes)
- Follow SOPs for collection, processing, storage, and distribution
- Speed of processing and storage after obtaining the specimens is important

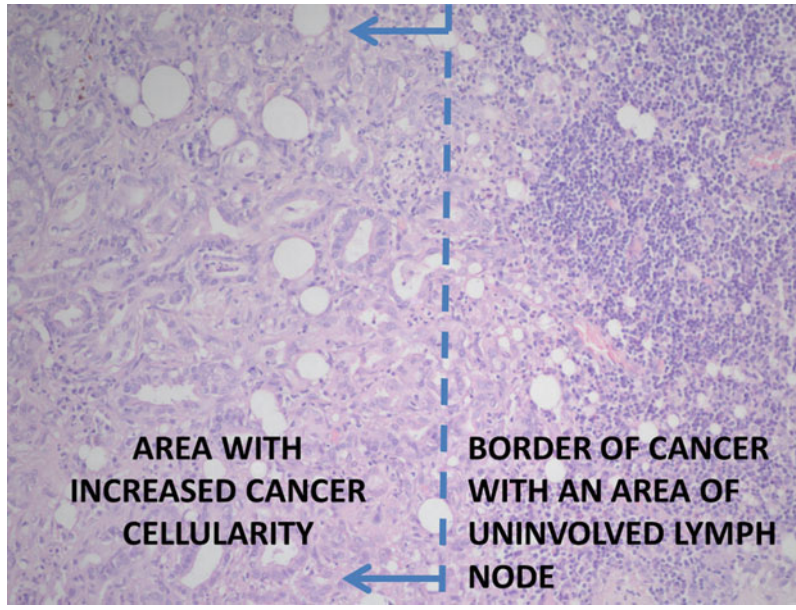


Fig. 17.2 This figure demonstrates a pancreatic adenocarcinoma partially involving a lymph node. To analyze the molecular expression of genes specifically in the cells of the pancreatic cancer, the contribution of noncancer cells (i.e., in this case, the lymphocytes and other cells of the uninvolved area of the lymph node) must be excluded. Micro-/macrodissection of the specimen to only include the area with increased tumor cellularity can be performed by removing the border area before cutting sections from the paraffin block. Such paraffin sections would only include the area indicated by the *dashed line* and *blue arrows*

- Try to complete processing and storage of blood/products, urine, and saliva within 4 h
  - Clinical followup after 6 months may be necessary to ensure a correct diagnosis

17.11 Education and Training in Issues Related to Tissue Repositories

- To produce uniform, standard specimens and to minimize bias, personnel must be carefully trained in SOPs and overall repository operations
- Potential sources of specimens and issues of repository science
- Records of training must be maintained
- To protect patients, all personnel must be trained in ethical issues (IRB) and privacy issues (HIPAA requirements)
- As discussed subsequently, all personnel must be trained in safety
- Just as important as educating repository personnel is the education of investigators using the repository
  - It is important for tissue repository pathologists to ensure investigators understand the tissues they are using in their research
  - This includes the potential for specimens to cause bias, optimization of tissues to support their research, and alternative tissues
  - Also, the investigator and all their personnel must be trained in biohazards

17.12 Matching Available Tissues to Investigator Needs for Tissues

- Pathologist needs to review each investigator's request for tissues and if appropriate, discuss the request with the investigator as well as alternative tissues

  - For example, a request for "smooth muscle" or "fat" is too general in that these tissues vary biologically depending on their source
- For example, smooth muscle from the uterus responds to estrogens and has active estrogen receptors
- In contrast, smooth muscle from arteries responds to hormones that control blood pressure
  - Some other tissues may be similar enough that additional sources can be identified
  - Sources of some tissues may be primarily met from autopsies or tissues removed for transplant
- Some tissue requests are very difficult to meet and this must be explained to investigators
  - Large numbers of rare diseases (e.g., synovial cell sarcomas)
  - Specimens with multiple requirements (male breast cancer from males less than 40 years of age who are African-Americans)
- The addition of each requirement makes the specimens more and more difficult to identify
- For example, African-Americans represent 30% of the clinical population at UAB, and breast cancers in males are uncommon and even rarer would be a breast cancer in a young male
  - Specimens that are atypical for the disease
- For example, papillary serous carcinomas of the ovary in African-Americans – a tumor that is infrequent in African-Americans
  - Specimens of primary tumors of large size in heavily screened populations in which large tumors are uncommon
  - Resection specimens processed very rapidly (e.g., 5 min)
    - Meeting this requirement would disrupt normal clinical operations
  - Determine whether or not there is scientific evidence for the need for difficult to meet requests
  - Requests for failed donor tissues or tissues removed at transplant (e.g., lungs from patients with cystic fibrosis)

17.13 Equitable Distribution of Tissues

- Tissue repositories should attempt to equitably distribute tissues to investigators
- Typically investigators are assigned priorities for specific specimens, but if specimens are not in high demand, typically specimens are assigned on a rotating basis
- Some requests are too labor intensive to be met (e.g., obtaining a whole spinal cord)
- Some requests cannot be met because of diagnostic requirements (e.g., samples of fresh ductal carcinoma in situ [DCIS] of the breast); DCIS is used totally in diagnosis to exclude microinvasion
- Some requests are so restrictive that very few specimens would meet the request

17.14 Annotation of Tissue Specimens

- Some basic information is needed on any tissue aliquot being studied including most importantly its correct diagnosis
  - Age, race, and sex are important information that usually can be readily obtained
- The need for more extensive annotation usually depends upon the use of the tissues
- Other than clinical outcome, more extensive annotation usually is project specific, so general approaches to annotation may not be adequate
- Some information needed for annotation may not be available in charts; thus, for some studies and most epidemiology studies, the extent of specific annotation may need to be designed with some of the data obtained directly from patients (e.g., family histories, environmental exposures)
  - In such cases, the study would have to provide extra resources to facilitate such annotation
- In many cases, annotation of specifics may not be needed. In these cases, it may be more efficient to collect the annotation after the specimens and matching annotations are requested

17.15 Bias Affecting Usefulness of Tissues

- Bias results from differences in cases and controls unrelated to the process or diseases being studied
- Bias causes an incorrect interpretation that experimental differences between cases and controls are due to the disease; instead, the differences are due to variables in the way the tissues are obtained and/or experiment is performed
- Bias cannot be detected by routine statistical analysis
- Bias typically is detected when there is failure to confirm a previous result using a different independent population
- The chance that bias will be important in a study increases with the number of variables being studied
- Bias may be introduced into research projects due to samples of tissue when there are differences between cases and controls due to varying conditions in collecting, processing, storing, or distributing tissues as well as differences in the populations from which samples are obtained
- Use of SOPs aid in reducing the chance of bias
- Specific specimen factors causing differences between samples and controls due to bias. These include the issues in Box 17.3
- Specific population differences: examples that may cause experimental differences in case versus controls due to populations are listed in Box 17.4
- Avoiding bias: the chance of bias can be minimized by ensuring cases and controls match (e.g., # of diabetics in cases is equivalent to # of diabetics in controls)

Box 17.3. Potential causes of bias Due to tissues

- Number of freeze-thaw cycles
- Procedures for collection, processing, and storage
- Collection sites
- Type of specimens (e.g., EDTA plasma vs. citrate plasma)
- Differences in specimen containers
- Failure to use SOPs or SOPs that vary between cases and controls
- Site-specific bias (unidentified)

Box 17.4. Potential causes of bias due to population

- Demographics (e.g., age, race, sex)
- Comorbid conditions (e.g., diabetes)
- Donor stress
- Homeostasis (fed vs. fasting)
- Diurnal variations (time when sample is drawn)

Box 17.5. Issues to be considered in an MTA

- Transfer to third parties
- Use of specimens
- Indemnification
- Commercial uses
- Intellectual property

17.16 Administrative Issues

- The transfer of tissues or tissue products to extramural investigators usually should require a material transfer agreement (MTA) between the institutions/organizations sending and receiving the specimens
- An MTA is unidirectional and addresses issues, for example, in transferring tissues from a collecting site to a site where the investigator is located
- MTAs actually should facilitate the transfer of specimens among institutions
  - Examples of issues that should be addressed in MTAs are listed in Box 17.5

17.17 Shipping

- Transfer of specimens outside the organization collecting, processing, storing, and/or distributing tissue specimens
- Use of reliable courier which is equipped to ship human specimens
- International shipping is very difficult and is problematic unless specimens are shipped by specialized couriers; international shipments are very expensive
- Even though assigned to a courier, the shipping site is responsible for meeting shipping requirements, specifically the requirements of the IATA
  - Someone at the shipping site must be trained in IATA requirements
- Shipping of flammables (e.g., specimens in alcohols) is limited in quantity by IATA

17.18 Regulatory and Ethical Issues Affecting Tissue Repositories

- Ethical issues – the IRB
  - IRB approval should be obtained for tissue biorepository; all IRB approvals, exemptions, and decisions as to a classification of "nonhuman research" are local
    - Collection of remnant tissues not needed for diagnosis
  - Specimens in the biorepository are identified as to the identification of donor in the repository – full review by IRB may be required as well as informed consent from the source of the tissue
  - All specimens are anonymized – could be classified by IRB as "nonhuman research" or IRB may require an expedited review
- The need for informed consent could be waived
  - IRB approval of investigators using a tissue repository
    - If specimens are provided as anonymized, research may be classified by the IRB as "nonhuman research." IRB approval may be necessary
    - If specimens are de-identified, research may be classified by the IRB as "nonhuman research" or IRB approval may be required
    - If specimens are identified, research will usually require IRB review and approval
    - For any situation, review by the IRB is necessary to determine how the work of the biorepositories is classified

Box 17.6. List of HIPAA 18 identifiers

- Names
- All geographical subdivisions smaller than a State, including complete zip codes (see HIPAA discussion on website) (http://www.hhs.gov/ocr/privacy/hipaa/understanding/summary/index.html)
- All elements of dates (except year) for dates directly related to an individual (see HIPAA description on website), for ages over 89, the elements may be aggregated into a single category of age 90 or older
- Phone numbers
- Fax numbers
- Electronic mail addresses
- Social Security numbers
- Medical records numbers
- Health plan beneficiary numbers
- Account numbers
- Certificate/license numbers
- Vehicle identifiers and serial numbers, including license plate numbers
- Device identifiers and serial numbers
- Web Universal Resource Locators (URLs)
- Internet Protocol (IP) address numbers
- Biometric identifiers, including finger and voice prints
- Full face photographic images and any comparable images
- Any other unique identifying number, characteristic, or code (note this does not mean the unique code assigned by the investigator to code the data)

- Privacy issues – HIPAA
  - HIPAA specifies 18 HIPAA-designated identifiers (Box 17.6)
- If all these 18 identifiers are excluded from the identification of the tissue specimen, HIPAA requirements do not apply to these tissue specimens
  - Some of the 18 HIPAA identifiers are dates (e.g., birth date)
- Age can be substituted for birth date in most research
  - If dates are required, a limited dataset can be constructed which excludes other HIPAA identifiers except dates
- If a limited dataset is used by an investigator and the investigator signs a data use agreement with the source of the clinical information or tissue, HIPAA also does not apply
- Informed consent
  - May be waived by local IRB for de-identified and/or anonymized specimens
  - There is no perfect time or place to obtain informed consent
    - At clinic – space and time limited
    - In the operating room – area, space, and time limited, patient is under stress and is fasting
    - After surgery (patient may require pain medication)
    - Trying to obtain consent after discharge is usually unsuccessful
  - Specimens obtained specifically for research (e.g., extra solid tissue, bodily fluids) usually require informed consent from patients
  - Informed consent from patients frequently requires significant resources because tissue repositories have no direct relationship with patients
  - Obtaining informed consent from some patients to use tissues in biomedical research may be a problem due to religious, social, or ethnic considerations/views
- Return of research results to patients/repositories
  - Research results may be wrong or their interpretation may be incorrect
  - Research results may not be applicable to all subpopulations
  - Research results returned to patients or potentially involved in medical decisions in the USA must be from CLIA-approved laboratories
  - Significant liability may occur if research results cause harm to patients
  - Research results may be accidentally released
  - Return of research results requires significant infrastructure and is expensive

17.19 Cost Recovery

- It is illegal in many countries to sell human tissues; however, recovery of costs associated with collection, processing, storage, and distribution of tissues is appropriate. Such cost recovery may be necessary to permit continued operation of a human tissue biorepository
- Some form of cost recovery aids in ensuring appropriate respect for and efficient utilization of human tissues
- Cost recovery is an appropriate component of long-term financial stability of tissue repositories
- The extent of cost recovery is usually based upon resources supporting a biorepository

17.20 Safety in Tissue Repositories

- A tissue repository has many potential dangers, especially biohazards and chemical hazards
  - Potential biohazards include blood-borne pathogens (e.g., HIV, hepatitis B), tuberculosis, antibiotic resistant bacteria, and prions
  - All tissues, even paraffin blocks, must be handled with universal precautions
  - Chemical hazards include formaldehyde, alcohols, and xylene
  - Physical hazards include slippery wet and paraffin coated floors, as well as dangers to employees from other individuals
  - If volatile reagents (e.g., xylene) are used or stored, fire is a potential danger

  - Electrical hazards are represented by improperly grounded equipment as well as personal appliances (e.g., radios)
- Safety program – the safety program of a tissue repository may be partially independent or a part of the safety program of the umbrella organization
  - A safety program should have a safety committee and a safety officer
  - A safety plan is developed to minimize the chance of injury including the use of engineering practices (e.g., good ventilation and drains), which includes the use of safety equipment (e.g., safety glasses, hoods)
  - The safety plan includes review of the potential dangers of each employee/visitor based upon not only the work performed but also the area of work
  - A safety training program is established for personnel that include special training focused on blood-borne pathogens, chemical hazards, and if appropriate, formaldehyde
- The safety program applies to personnel working in the work area where such chemicals are present or stored; updated yearly
  - The safety plan is reviewed yearly and when incidents occur, the safety plan is modified to prevent their recurrence
  - Vaccinations for hepatitis B should be offered to all employees who are in contact with fresh or frozen human tissues
  - Because repositories provide specimens outside their organization, it is a good idea to require that outside personnel receiving these specimens be educated in biohazards including universal precautions
  - To protect against chemical hazards, material safety data sheets (MSDS) should be available to all employees using chemicals to which the MSDS apply

17.21 Informatics

- A biorepository should have an informatics program if the repository is large enough to warrant it. The informatics program should save the repository effort and money. This might include
  - A bar coding system
  - An investigator component to keep track of the details of all investigator requests
  - A donor component to record details of the donor (e.g., age, race, sex, history, clinical information such as prior therapy)
  - The donor component might be combined with a specimen component to follow the collection, processing, storing, and distributing of each unique aliquot usually via a unique bar code
  - Unique common data elements to avoid redundancy should be used
- These should typically utilize a pathology vocabulary as well as more general designations (e.g., any breast cancer)
  - Required fields should be minimized, and the navigation should follow the business plan of the repository (e.g., the order of how specimens are collected and processed)
  - If specimens are identified as to the identity of tissue sources, the system must meet HIPAA security requirements including installing the system on a secure server behind a firewall, the use of strong passwords, and the maintenance of audit trails for even viewing of identified patient information
  - It may be beneficial to divide the database into an "identified" component which is small and has very limited access and a de-identified database devoid of all 18 HIPAA identifiers with most of the donor data; this minimizes "read only" audit trails
  - In order to meet NCI and ISBER Best Practices, the informatics system should keep track of the history of specimens including the times and details of tissue processing (e.g., freeze-thaw cycles)

17.22 Challenges in Meeting Specimens for Genomic Studies

- Many tissue banks have SOPs that are not designed to meet the stringent demands of current genomic studies (e.g., 0.15 g of

a tumor with ≥80% tumor nuclear cellularity and ≤30% necrosis); see Fig. 17.1
- Most tumors, except for selected types, cannot meet this standard because such a requirement is not characteristic of the growth pattern of the tumor and supplying such specimens would require costly microdissection (see Fig. 17.2)
- Also, these and similar requirements will introduce bias into studies in that the results will only apply to a very select group of tumors and not to tumors in general
- The costs with such collections exceed the funds provided for collecting, processing, and microdissection of such cases
- Specimens for genomic studies are best collected prospectively; however, the costs of detailed prospective collections (e.g., matching samples of blood) may be over $1,000 per case (patient) due to the many collected specimens which do not meet requirements

17.23 Challenges Leading to New Directions in Tissue Repositories

- All these challenges require relatively novel approaches and significant additional resources
- Problems
  - Neoadjuvant therapy – identify cases treated with neoadjuvant therapy
- They may or may not be useful for specific research projects
  - Smaller sizes of tumor – some tumors are too small to obtain samples
  - If only in situ lesions are present, some specimens must be submitted totally to exclude microinvasion
  - Biopsies of metastatic lesions may be so small that they must be submitted totally
  - Tissues from metastatic lesions may be unavailable because imaging is used in lieu of biopsy of suspected lesions
- Potential solutions
  - Nitrocellulose blots extract RNA, DNA, and proteins from tissues
  - Proteins can be extracted from nitrocellulose blots and used for discovery; new approaches using housekeeping proteins for quantitating specific proteins may be developed for blots
  - mRNA and microRNA can be extracted from nitrocellulose and together with housekeeping genes can determine mRNA and microRNA levels using RT-Q-PCR
  - DNA can be extracted from nitrocellulose blot and analyzed
    - Blots of tumors on glass microscopic slides can be used to analyze proteins, DNA and RNA
    - For small tumors, in situ lesions, or metastatic lesions, increase the use of diagnostic paraffin blocks and construct tissue microarrays
    - For all small lesions including fine needle aspirates, try and obtain extra samples for consented patients
    - Consider using tissues from autopsies of patients with metastatic disease; perform such autopsies rapidly (warm autopsy) to make tissues more useful
    - Use of micro-/macrodissection and laser capture microdissection if enriched (Fig. 17.2) or pure (Fig. 17.1) cellular populations are needed

Acknowledgments Supported by the Cooperative Human Tissue Network (5U01CA44968), UAB SPORES in Breast Cancer (5P50CA0189019), Pancreatic Cancer (2P50CA101955) and Cervical Cancer (5P50CA098252), Comprehensive Cancer Center Core Support Grant (P30CA13148).

Further Reading

Bell WC, Sexton KC, Grizzle WE. How to efficiently obtain human tissues to support specific biomedical research projects. Cancer Epidemiol Biomarkers Prev. 2009;18:1676–9.

Bell WC, Sexton KC, Grizzle WE. Organizational issues in providing high-quality human tissues and clinical

information for the support of biomedical research. Methods Mol Biol. 2010;576:1–30.
Edgerton ME, Morrison C, LiVolsi VA, et al. A standards based ontological approach to information handling for use by organizations providing human tissue for research. Cancer Inform. 2008;6:127–37.
Frost A, Sparks D, Grizzle WE. Methods of antigen recovery vary in their usefulness in unmasking specific antigens in immunohistochemistry. Appl Immunohistochem Mol Morphol. 2000;8(3):236–43.
Gaston SM, Soares MA, Siddiqui MM, et al. Tissue-print and print-phoresis as platform technologies for the molecular analysis of human surgical specimens: mapping tumor invasion of the prostate capsule. Nat Med. 2005;11:95–101.
Grizzle W. Special symposium: fixation and tissue processing models. Biotech Histochem. 2009;22:1–9.
Grizzle WE, Aamodt R, Clausen K, et al. Providing human tissues for research: how to establish a program. Arch Pathol Lab Med. 1998;122:1065–76.
Grizzle WE, Grody WW, Noll WW, et al. Recommended policies for uses of human tissue in research, education, and quality control. Ad Hoc Committee on Stored Tissue, College of American Pathologists. Arch Pathol Lab Med. 1999;123:296–300.
Grizzle WE, Stockard CR, Billings PE. The effects of tissue processing variables other than fixation on histochemical staining and immunohistochemical detection of antigens. J Histotechnol. 2001;24:213–9.
Grizzle WE, Bell W, Sexton KC. Best practices and challenges in collecting and processing human tissues to support biomedical research. In: AACR 96th Annual Meeting, Education Book; 2005; p. 305–310.
Grizzle WE, Bell W, Fredenburgh J. Safety in biomedical and other laboratories. In: Patrinos GP, Ansorg W, editors. Molecular diagnostics. Amsterdam/Boston: Elsevier/Academic Press; 2005. p. 421–428. Chapter 33.
Grizzle WE, Fredenburgh J, Myers RB. Fixation of tissues. In: Bancroft J, Gamble M, editors. Theory and practice of histological techniques. 6th ed. Edinburgh: Churchill Livingstone; 2007. p. 53–74. Chapter 4.
Grizzle WE, Sexton KC, Bell WC. Quality assurance in tissue resources supporting biomedical research. Cell Preserv Technol. 2008;6:113–8.
Grizzle WE, Bell W, Fredenburgh J. General considerations concerning safety in biomedical research laboratories. In: Patrinos GP, Ansorg W, editors. Molecular diagnostics. 2nd ed. Amsterdam/Boston: Elsevier/Academic Press; 2010. p. 563–572. Chapter 39.
Grizzle WE, Bell WC, Sexton KC. Issues in collecting, processing and storing human tissues and associated information to support biomedical research. In: Srivastava S, Grizzle WE, editors. Translational pathology of early cancer. Amsterdam, the Netherlands: IOS Press BV; 2012. p. 531–549.
Grizzle WE, Bell WC, Sexton KC. Issues in collecting, processing and storing human tissues and associated information to support biomedical research. Cancer Biomark. 2011;9:531–49.
Hewitt R, Watson PH, Dhir R, et al. Timing of consent for the research use of surgically removed tissue: is postoperative consenting acceptable? Cancer. 2009; 115:4–9.
Otali D, Stockard CR, Oelschlager DK, Wan W, Manne U, Watts SA, Grizzle WE. Combined effects of formalin fixation and tissue processing on immunorecognition. Biotech Histochem. 2009;84: 223–247.
Pitt K, Campbell L, Skubitz A, ISBER, et al. 2008 Best practices for repositories: collection, storage, retrieval and distribution of biological materials for research. Cell Preserv Technol. 2008;6:3–58.
PRIM&R. Human Tissue/Specimen Banking Working Group Report: Report of the Public Responsibility in Medicine and Research; 2007.
PRIM&R. Tissue Working Group Report: Section II Tools for Investigators, IRBs and Repository Managers; 2008.
Tuck MK, Chan DW, Chia D, et al. Standard operating procedures for serum and plasma collection: Early Detection Research Network Consensus Statement Standard Operating Procedure Integration Working Group. J Proteome Res. 2009;8:113–7.

Pathology Informatics

18

Roy E. Lee, Long Phi Le, and John Gilbertson

Contents

18.1 Definition and Scope of Pathology Informatics 500

18.2 The Major Components of Pathology Informatics 500
18.2.1 The Study of Information 500
18.2.2 The Study of Information Systems 500
18.2.3 The Study of Processes and Workflows ... 501
18.2.4 Information, System, and Process Management 501

18.3 Pathology Informatics and the Diagnostics Laboratory 501
18.3.1 Codependency and Shared Responsibility 501
18.3.2 Future Trends 502

18.4 The Laboratory Information System and the Diagnostic Process 502
18.4.1 The Total Testing Cycle 502
18.4.2 Involvement by the LIS 502

18.5 Molecular Informatics 505
18.5.1 Unique Aspects of Molecular Informatics 505

18.6 Pathology Imaging 508
18.6.1 Specific Uses of Digital Imaging 508
18.6.2 Telepathology 509
18.6.3 Digital Pathology 509
18.6.4 Current Status, Challenges, and Potential of Digital Pathology 510
18.6.5 Nontraditional Imaging Techniques 510

18.7 Image Analysis 511
18.7.1 Goals 511
18.7.2 Working Examples 511
18.7.3 Limitations to Image Analysis on Whole Slide Images 511
18.7.4 Techniques 512

18.8 Digital Pathology, Image Analysis, and the Laboratory Information System 512
18.8.1 Image Orders and Workflow 512
18.8.2 Specimen Identification and Messaging ... 512
18.8.3 Imaging at Sign Out 513

Further Reading 513

R.E. Lee, MD (✉)
Pathology and Laboratory Medicine Institute,
The Cleveland Clinic, Cleveland, OH, USA

L.P. Le, MD, PhD • J. Gilbertson, MD
Harvard Medical School, Department of Pathology,
The Massachusetts General Hospital, Boston, MA, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_18,

18.1 Definition and Scope of Pathology Informatics

- Pathology informatics can be conveniently defined as the study and management of pathology information, information systems, and processes (Table 18.1)
- The definition is based on the observation that laboratory processes, the data those processes consume and generate, and the computer systems that manage that data and those processes are intimately related
- Informatics is not synonymous with "information technology" (IT) or "information services" (IS) (Fig. 18.1). While informatics teams work closely with IT/IS groups, the latter tend to be focused on one component of informatics – the technical implementation and maintenance of information systems

18.2 The Major Components of Pathology Informatics

- Pathology informatics implemented at the Massachusetts General Hospital is based on four main areas of study: information, information systems, processes (workflows), and management

18.2.1 The Study of Information

- What information is required for a given process
- Measurements of information quality
- Information extraction (e.g., image analysis and natural language processing)
- Annotation and meta-data
- Information models and architectures
- Programming principles (information manipulation)
- Human processing and use of information (e.g., human–computer interaction and user interface design)
- Decision support principles (Bayes theorem)
- Information exchange and standards: data standards (e.g., SNOMED, ICD) and messaging standards (e.g., HL7, DICOM)

18.2.2 The Study of Information Systems

- Laboratory information systems (LIS): architecture and operations (see below)
- Image management systems (e.g., picture archival and communication systems [PACS])
- Bioinformatics systems
- Other clinical information systems (CIS): ADT (admission, discharge, and transfer) systems, CIS, electronic medical record

Table 18.1 Simplified breakdown of major areas under medical informatics. Because these three areas are often confused with one another, it is important to define them here. Clinical informatics differs from pathology informatics mainly in that the type of data and systems cover a scope outside of pathology and laboratory medicine and places less emphasis on process optimization and redesign. Of note, bioinformatics is the only one of these three areas that explicitly includes "computational tools" (IT) in its definition. Overall, medical informatics does not necessarily imply the use of IT; however, in today's age, IT makes the most sense to use, instead of traditional filing cabinets, and pen and paper

Clinical informatics	Pathology informatics	Bioinformatics
Concerned with information use in healthcare by clinicians: clinical decision support, CPOE, system design, system implementation (e.g., electronic health records and health information systems)	Concerned with laboratory information, its systems (e.g., laboratory information system), and associated processes. Pathology informatics differs from clinical informatics in that it is the branch of medical informatics covering pathology data and the processes and systems it covers. A greater proportion is devoted to process optimization, such as Lean Six Sigma	Associated with "hard-core" number-crunching algorithms for analysis and study of data at the DNA, RNA, and protein levels and less about information flow and processes. NIH definition (2000): research, development, or application of computational tools and approaches for expanding the user of biological, medical, behavioral, or health data, including those to acquire, store, organize, archive, analyze, or visualize such data

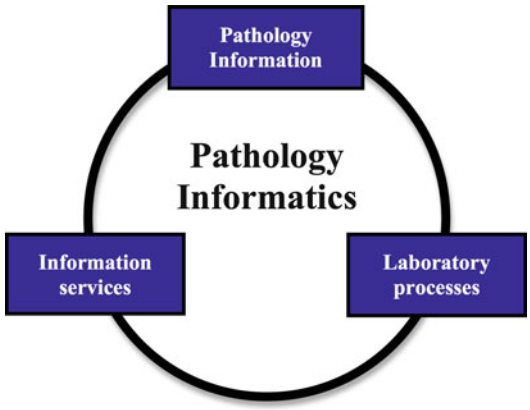


Fig. 18.1 The term "informatics" is not synonymous with information technology. In pathology informatics, these three major components of pathology information, information services/technology, and laboratory processes are intimately linked and related

(EMR)/electronic health record/patient medical record systems, and financial systems
- System interoperability and connectivity (interfaces and standards)
- System life cycle (creation, implementation, maintenance, decommission) of clinical systems
- User training

18.2.3 The Study of Processes and Workflows

- Process theory (industrial engineering)
- Process analysis, improvement, and validation: (e.g., business process management (BPM), failure modes and effects analysis (FMEA), and Lean Six Sigma)
- People working in systems (e.g., Just Culture)
- Outcomes measurement
- Process modeling and simulation
- Decision support processes

18.2.4 Information, System, and Process Management

- Pathology informatics involves large, complex, highly integrated, and interconnected systems supporting hundreds of users and processes. Successful system development and implementation and management in the pathology environment require significant communication and management skills
  - Project management
  - System management
  - Capital and accounting finance
  - Software development management
  - Validation studies management
  - System selection, implementation, and maintenance (e.g., total cost of ownership)
  - Regulatory compliance

18.3 Pathology Informatics and the Diagnostics Laboratory

18.3.1 Codependency and Shared Responsibility

- A modern diagnostic laboratory cannot function without an effective information system, and an effective information system cannot be designed, implemented, or maintained without a clear understanding of the operations and missions of the lab. There are numerous reasons for this:
  - The primary product: Accurate, actionable clinical information, in the form of finalized results or "signed out" reports, is the primary product of pathology labs
  - Required inputs: Laboratory operations depend on access to clinical, patient, specimen, and financial information
  - Management of processes: Diagnostic laboratories run complex processes that need to run efficiently, at low variance, and with limited failures. The processes must be validated and documented
  - Communications: Communication between the lab and the rest of the clinical care team (usually through formal orders and results/reports as well as ad hoc) is central to the value of the laboratory and the safety of the patient
  - Automation: Automation, standardization, machine-readable specimen identification (barcodes), and clear, human- and

machine-readable results/reports have been shown to improve efficiency, patient safety, and effective communication between the lab and the clinic
  - Data sharing and interoperability: Pathology/laboratory results are used by multiple downstream information systems (reporting systems, electronic information systems, etc.). Studies have shown that the majority of data in most EMRs originates in diagnostic laboratories and that most significant medical decisions depend on laboratory data
- Each above activity is a shared responsibility between medical staff, laboratory technical staff, the information management team, and information system developers. System developers must understand the laboratory's goals, missions, and customers, and laboratory staff should understand that the LIS is their best tool for feedback, quality assurance, and laboratory efficiency/value

18.3.2 Future Trends

- Demographic trends (national and international), the explosion of genetic and molecular testing and digital imaging, and increasing use clinical decision support systems, workflow efficiency engines, utilization and value studies, and the convergence of basic science and clinical application (all of which will require high-quality, standardized, machine-readable laboratory data) will drive increased needs for sophisticated LIS and pathology informatics for the foreseeable future

18.4 The Laboratory Information System and the Diagnostic Process

- Though normally considered a computer system, the LIS can be thought of more broadly as a set of related processes, operations, and procedures, some or all of which are supported by one or more computer systems. The primary goal of an LIS is to support the operations of a clinical laboratory. It therefore makes sense to discuss the LIS in terms of laboratory operations

18.4.1 The Total Testing Cycle

- The total testing cycle (TTC) is a classic, high-level model of the diagnostic laboratory process (Fig. 18.2). The TTC models laboratory testing as a largely linear process beginning when testing is first considered by the clinician and continues through the clinical and laboratory processes needed to order the test, collect and transport the specimen, prepare the specimen for testing, run the test, and through the interpretation of the test result first by lab staff and pathologists (which may result in further testing) and finally by clinicians. The process is modeled as a cycle because the results of one test often results in the ordering of others

18.4.2 Involvement by the LIS

- The LIS (and related information systems) is involved in every step of the TTC

18.4.2.1 Ordering and Order Entry

- Ordering is the process through which the ordering clinician chooses and orders tests or other laboratory services (such as anatomic pathology analysis of tissue specimens). Traditionally done through paper requisitions, it is increasingly done through computerized "provider order entry" systems (CPOE). By interfacing with other systems (the LIS, test formularies, lab handbooks, etc.), sophisticated CPOE systems allow clinicians to search for tests, obtain advice on appropriate test selection and cost, manage utilization, collect clinical information or clinical justification necessary for the laboratory doing the test, or provide mechanism for the ordering physician to talk directly to a pathologist
- Order communication and specimen transport
  - Orders are passed to the LIS, usually through an ordering interface, and patient

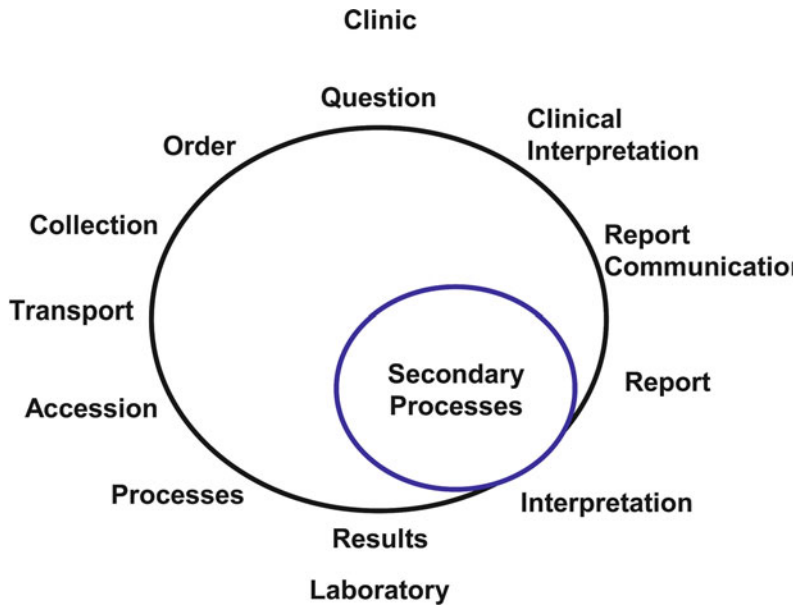


Fig. 18.2 The total testing cycle: The TTC is a model of the generic clinical testing process. It begins with a clinical question and the decision to test, through the ordering, specimen collection and transportation to the steps necessary to generate, interpret, and report results in lab and communicate them to the clinic. Information systems are involved in virtually every step in the cycle

and encounter information is sent to the LIS through an ADT (admission, discharge, and transfer) interface. In some systems, the LIS prints barcoded containers in clinic for identification of the specimens; in other cases, specimen identification is done by a separate process in the clinic. Increasingly barcoded specimens are tracked during transport from the clinic to the lab

- Accessioning
  - Accessioning is the process through which the laboratory (through the LIS) takes possession of the specimen; verifies patient, insurance, order, and specimen information; reconciles discrepancies; and associates the specimen(s) with the specific laboratory processes that will eventually lead to a diagnostic report. Laboratories often provide their services for multiple clinics and hospitals and receive specimens in many different ways, making accessions an important and complex process
- Diagnostic processes
  - Diagnostic processes (also known as protocols or workups) define the way the lab are activities that the lab employs to generate the information needed to solve the clinical question at hand. Processes can be iterative or nested. Depending of the complexity of the diagnostic question, a lab may employ multiple subprocesses (e.g., specimen dissection and description, tissue processing, embedding, tissue sectioning, staining, molecular studies, and microscope examination) involving multiple subdivision of specimens (e.g., parts, blocks, slides, and aliquots) and even different laboratories in the pathology department or at outside institutions (e.g., surgical pathology, molecular pathology, flow cytometry, and reference laboratories) to fully work up a case. A major role of the LIS is to organize and keep track of this complex flow of specimens, processing, testing, and data. LIS has several mechanisms to do this
    - Unique specimen identification, usually by barcode
    - Tracking the physical location of specimens and knowing where each specimen should be and where it should go next ("routing")
    - Protocols defining where specimens should go (routing) and how they should be processed/analyzed at each location. Protocols tend to be specific to specific specimen types: For example, a breast

biopsy may have a different route and may be subjected to different protocols (e.g., different processing, more slides, more stains, and fewer images) than a kidney resection
    - Status: Each case and tissue asset in the LIS is given a status depending on its completeness in the diagnostic cycle. Examples include "Accessioned," "In histology," "Final," and "Signed out"
- Primary or secondary protocols
  - Diagnostic processes can be iterative: The first set of protocols on a tissue biopsy may be designed to establish the presence (or absence) of cancer. If cancer is established, the lab might use other protocols to classify the cancer using tumor markers. The use of these secondary protocols can be automatically triggered based on the results of initial test results (reflex testing) or done on the order of the pathologist working up the case
- Interpretation and reporting
  - In a simple case such as the measurement of serum electrolytes, results can be accepted by the LIS from the analyzer and reported to the clinic (usually via a CIS, see below) after the test or batch passed quality assurance. Such test results are usually sent with minimal interpretive information such as normal reference ranges. For complex testing requiring an interpretive report from the lab (e.g., anatomic pathology), each report is reviewed by the primary pathologist for the case who then signs out each case (usually through electronic signature) in the LIS
- The pathology report
- In some areas of pathology (such as surgical pathology, cytology, molecular pathology, and autopsy), a clinical report is needed to communicate
  - What specimens were received
  - What processes/workup were done
  - Specific laboratory data or results (e.g., "size of tumor" or "number of lymph nodes dissected")
  - Diagnostic interpretation (e.g., "prostatic adenocarcinoma, Gleason 3 + 4 = 7")
  - Clinical communications and notes (e.g., "interpretation is limited by extensive cautery artifact")
  - In addition, the pathology report contains information on the pathologists and clinicians associated with the case, the clinical context, diagnostic and procedure codes, and information used in a wide range of operational and financial reports as discussed in the "Postdiagnostic activities" section below
  - The LIS provides a series of tools for the creation of the pathology report including the ability to create specific report templates for different types of cases, word processing including spell checking, "quick text," or "macros" for the inclusion of standard text by typing short, relatively simple codes. In many cases, LIS allows structured data entry in which information is entered in specific data elements directly into the system with a defined set of valid values and stored as discrete data elements in the database. For example, the College of American Pathologists tumor templates are often implemented through a structured data entry tool in the system, thus bypassing the use of paper
- Report and result communication
  - Laboratory reports created in the LIS are communicated to clinicians through a variety of mechanisms including printing and mailing, dedicated printers in clinical offices, faxing and, increasingly, secure Internet/web-based report distribution systems. However, the great majority of reports are communicated, via HL7 interfaces, to EMR systems and CIS
  - The LIS is usually able to display pathology reports in a much more sophisticated way than the EMR. There are two main related reasons for this: (1) The HL7 interfaces that communicate the report cannot easily communicate formatting information (such as tables or font), and (2) EMR systems tend not to support sophisticated formatting. This is a significant problem for laboratory – clinic communication – and

is slowly being mitigated by improved interfaces and support for open, standardized file formats such as the portable document format (PDF)
- Alerts, preliminary reports, addenda, and amendments
- In addition to the standard results and pathology reports, LIS provides other mechanisms for communications between lab and clinic
  - Alerts: Usually used to communicate lab results that fall in predefined "critical ranges," alerts can be implemented as font or formatting changes in the report (or EMR), automated paging or direct phone calls to the ordering or responsible physician
  - Preliminary reports: can be used for a variety of purposes, for example, to notify the clinic that a case result will be delayed
  - Addenda: an addition to a signed out report that adds new data (but usually does not change the clinical impact of the diagnosis reported in the signed out report)
  - Amendment: a mechanism for changing a signed out report with information that could have significant clinical impact. Because of the potential clinical impact, amendments are usually issued with some form of active clinical alert (vide supra)
- Postdiagnostic activities
- After result reporting or report sign out, the LIS continues to manage data, reports, and specimens. Typical postdiagnostic activities include
  - Management of cases for clinical conferences
  - Management of specimens: clinical archives for specimens, aliquot, blocks, and slides. The archive time frame can range from days to years depending on the importance of the specimen and the cost associated with storage. In many cases, specimen archiving is mandated by law
  - Management reports: The LIS in a large hospital or practice will generate hundreds of reports needed for management and control of lab operations, workload, efficiency, and finance
  - Billing reports: The LIS generates reports that include the information needed to bill for lab services. Typically, these reports are passed to billing systems that use the data to validate charges and generate actual invoices that are passed to third-party payers, lab clients, or patients
  - Reports to registries, research databases, and regulatory agencies. Examples include reports to cancer registries, enterprise research databases that provide de-identified clinical data to researchers, and regulatory and certification agencies

18.5 Molecular Informatics

- Molecular studies have become commonplace in virtually all aspects of pathology, especially in areas such as microbiology and anatomic pathology. Molecular testing includes a wide range of modalities, mostly in the form of single gene tests from commercial vendors and laboratory-developed tests. Multigene or genomic level testing and sequencing will become more widespread as the technology becomes more affordable and the practice of personalized medicine becomes more routine
- Molecular informatics as a subset of pathology informatics focused on the flow of information through a molecular laboratory, starting from test ordering through reporting, and has substantial similarities and significant differences compared to informatics in traditional clinical and anatomic pathology laboratories, especially in data storage requirements (Table 18.2)

18.5.1 Unique Aspects of Molecular Informatics

18.5.1.1 Samples, Analytes, and Testing

- Sample types of varying quantity and quality that need to be accessioned and tracked as assets
  - Blood: useful for germ line testing or blood-related diseases; typically yields sufficient quantity and high-quality material
  - Fresh/frozen tissue: similar to blood but often not routinely available

Table 18.2 Example comparisons of data storage requirements between anatomic, clinical, and molecular pathology. Numbers are in uncompressed files and, although not necessarily exact, serve to illustrate the differences in magnitude of data storage between themselves

Category	Test/entity	Data storage requirements
Anatomic pathology	Annual storage needs on AP LIS	<1 TB per year, assuming 100,000 cases per year
	Pathology imaging	Depends on compression scheme, scanning magnification, amount of tissue on slide, number of focal planes per virtual slide, and number of slides per case. A major academic institution generating over 2,000 slides per day **could** easily consume over 33 TB per month (surgical pathology only – excluding cytology)
Clinical pathology	Blood banking: transfusion service	~21 GB/year added to database (assuming 60,000 cases/year)
	Clinical labs	<1 GB/year (assuming volume of 100,000 cases per year)
Molecular pathology	aCGH	~70 MB per case (text only)
		~275–300 MB/case (if image file included)
	Sanger sequencing	~400–500 KB per amplicon (~400–600 bp lengths each, and including both *.ab1 files for forward and reverse primer reactions)
	Next-generation sequencing (exome only)	~13 GB (Illumina HiSeq platform) per run
	Next-generation sequencing (whole genome)	~500–600 GB (Illumina HiSeq platform) per run

  - Body fluids: varying quantity and quality of material
  - Formalin-fixed paraffin-embedded (FFPE) tissue: standard specimen in most anatomic pathology laboratories; varying quantity and quality affected by age of tissue and degree of formalin fixation which leads to cross-linking and fragmentation of nucleic acids
- Various analytes/substrates derived from the initial specimen which contribute different information (DNA, mRNA, noncoding RNA); one specimen may yield multiple molecular derivatives or assets, which may be challenging for tracking
- The means of testing, data analysis, interpretation, and reporting vary according to whether the target mutations are germ line (inherited as in familial diseases) or somatic (acquired after fertilization as in cancer)
- Various types of mutations to be detected which would command different testing modalities
  - Single-nucleotide polymorphisms (SNP) or single-nucleotide variants (SNV)
  - Small insertions and deletions (indels)
  - Copy number variations (CNV)
  - Structural rearrangements (translocations, inversions)
  - Epigenetics (DNA methylation)
  - Several types of data may be obtained from the same molecular derivative

18.5.1.2 Data Types

- Qualitative (e.g., base substitution/insertion/deletion, translocation between two gene partners, and positive detection of a band on a gel)
- Quantitative (e.g., fragment size, copy number change, somatic mutation at a low allele frequency in a tumor specimen, HIV viral load testing, and bcr-abl testing for minimal residual disease in chronic myelogenous leukemia)
- May be very complex with multiple mutations/changes detected for one sample

18.5.1.3 Interpretation of Results

- Various forms of results output: PCR gel image, fragment capillary electrophoresis tracings, fluorescently probed tissue slides viewed under microscopy, quantitative real-time PCR data based on interpolation of standard curves, copy number profile from genome-wide copy number analysis, and sequence data from capillary electrophoretograms or pyrograms

- Various utility of results: diagnosis, prognosis, and therapeutic significance
- Most one target (single nucleotide or single gene)-based assays are interpreted in a binary (positive or negative) fashion or quantitatively (if signal is present, then disease is present with burden relative to level of detection); these interpretations tend to be "static" in nature that is based on prior established knowledge about recurrent mutations and their functional effects
- Assays testing for the entire sequence of a gene (whole gene sequencing) or with genome-wide scope (array comparative genomic hybridization) may detect variants that have not been previously described in addition to ones that are recurrent and have known consequences; new variants require extensive literature search, cross-checking with mutation databases, in silico modeling, and population statistics to further characterize their meaning (benign, likely benign, of unknown significance, likely pathogenic, pathogenic); often, data on these new mutations are incomplete or lacking, requiring a classification of "unknown significance"; variants of unknown significance may need to be reclassified as new studies reveal their functional relevance; therefore, whole gene or genome-wide tests require "dynamic" interpretation
- Detected rare variants may be private for a certain family or subpopulation; testing of additional members may yield added information for linkage analysis and proper classification of the variants
- Molecular testing which has traditionally relied on test interpretation for one individual based on known variants will have to adopt population-based interpretation to yield the most meaning out of genetic information (e.g., what proportion of cancer patients with a particular gene mutation)

18.5.1.4 Data Curation/Archiving

- As molecular testing becomes more complex, involving many genes and genome-wide coverage, data curation/archiving is imperative
- Processes should be in place to easily capture molecular results from instruments, annotations/interpretations from molecular pathologists/geneticists, findings from clinicians (phenotype), and all other laboratory/radiological data
- A comprehensive collection of structured medical testing and clinical phenotype data affords the ability to easily query for population-based statistics and clinical genotype correlation to classify variants of unknown significance
- Structured data should also be collected throughout the test workflow for sample tracking, quality assurance, quality control, and process management

18.5.1.5 Reporting

- Reports from the molecular pathology laboratory may include both qualitative and quantitative elements
- Reporting mutations or genetic changes should be based on standardized nomenclature so that recipients from other institutions may comprehend the test results
- Results may be simple as in a single-base substitution or complex involving a multitude of mutations detected in a genome-wide manner; results may be interrelated (e.g., expression or methylation data for a gene that was also sequenced at the DNA level) and should be correlated to make sense of the observations
- Molecular data particularly those resulting from genome-wide studies will require clever visualization tools to quickly navigate through complex datasets; a simple text-based report may not be sufficient
- Reporting of variants may require documentation of references and primary literature which should be archived with the laboratory mutation database; references and interpretations may change as more functional data are accumulated

18.5.1.6 Workflow in a Molecular Pathology Laboratory Is Complex

- Molecular tests usually require multiple processes handled by multiple personnel

- Sample tracking and quality control checkpoints are needed to successfully operate a high-volume molecular laboratory, particularly for high-throughput tests
- Automation should be applied as much as possible to ensure testing accuracy and consistency and to enable the allocation of human resources for other intellectually demanding tasks
- Digitization should be implemented where possible to increase workflow efficiency (e.g., digital document management to achieve a paperless operation)

18.5.1.7 Future Issues

- The arrival of next-generation sequencing into the clinical molecular laboratories will push the limits of current molecular informatics systems
- All aspects of molecular informatics (accessioning, workflow, sample tracking, data curation, results interpretation, reporting, and so forth) will be further taxed with the increased data burden and complexity of high-throughput sequencing
- Next-generation sequencing allows the concurrent testing of many genetically barcoded samples for many targets, yielding billions of bases of information (gigabytes to terabytes of information)
- Next-generation sequencing as a single testing modality may potentially detect point mutations, indels, copy number, gene expression, structural rearrangements, and methylation status all in one assay
- Operation of a next-generation sequencing laboratory will require not only molecular informatics for routine laboratory operation but also a multitude of other informatics disciplines for successful implementation: Bioinformatics and biostatistics for analysis and interpretation, medical informatics for genotype–phenotype correlation and reporting, and information technology and information systems support to provide the infrastructure for the data and computation demands of next-generation sequencing

18.6 Pathology Imaging

- Digital imaging is used extensively in the pathology laboratory (Fig. 18.3)
- The main reasons for digitization
  - Documentation of specimens
  - Digital management of documents
  - Visualization of data
  - Communication (e.g., telepathology)
  - Computation (e.g., image analysis)
  - The ability to use multispectral (nonwhite) light for analysis

18.6.1 Specific Uses of Digital Imaging

- Documentation of gross and autopsy specimens, microbiology plates, gels, gram stains, urine sediment, and microscope fields in histopathology, cytology, hematology, and parasitology

Fig. 18.3 Current versus digital workflows: Predigitization (*blue*), digitization (*green*), and postdigitization (*yellow*) steps are shown. This picture illustrates how every step in the process is important – errors or tissue folds on the slide will affect each upcoming step

- Document management
- Display of complex data, such as flow cytometry
- Telepathology
- Diagnostic image analysis or image-based diagnostic decision support; for example, automated Pap smear screeners
- The use of multiple, specific wavelengths of light, as in multiple target FISH or quantum dots

18.6.2 Telepathology

- Telepathology is a mechanism for laboratory to laboratory communications that allows digital images created at one site to be viewed and interpreted at another. Virtually, any lab operation (from grossing to electron microscopy) can be managed through telepathology
- Fundamental requirements
  - Ability to digitize the specimen at the sending site
  - A network that can support the communication
  - Ability to display at the receiving site

18.6.2.1 Three Main Types of Telepathology

- Static
  - Relatively small, single fields are captured and sent through the network. Static telepathology is useful when the diagnostic material can be meaningfully imaged in a limited number of discrete fields and/or the when the diagnostic request is to confirm a specific finding. It has been used effectively in microbiology, parasitology, and Pap cytology (confirmation of the output of an automated screener)
- Robotic
  - A video camera is connected to a robotic microscope. An operator at a distant site operates the robotic microscope and views the video images. In some implementation, a high-resolution still camera is also attached to the microscope, allowing the operator to capture and view high-fidelity images. While robotic systems require relatively high bandwidth, they have been used successfully to support remote hospitals for both routine pathology and frozen section diagnosis
- Whole slide imaging
  - As the name implies, all of the tissue on a microscope slide is digitized at high resolution and is stored on a server as a complex, compressed "pyramid" file that saves the whole slide image at multiple resolutions (or "levels") roughly corresponding to the resolutions one would see under a microscope with 4x, 10x, 20x, and 40x objective lens and 10x oculars. Each level is divided into multiple, small "tiles" that can be independently sent from the server to the remote client system. The pathologist is initially presented with a low-resolution "thumbnail" image and "pan and zoom" into and around the image by clicking on the image, with the server presenting only the image data needed to render the area of the slide the pathologist is looking at (very much like "Google Maps"). While an entire whole slide image may be large (0.3–10 GB), the pathologist tends to do most of their examination at low resolution; therefore, a relatively small percentage of the image data (or "tiles") is sent over the network to be displayed

18.6.3 Digital Pathology

- There is increasing interest in the potential, large-scale implementation of whole slide imaging for primary diagnosis. This process, known as digital pathology, has predigitization, digitization, and postdigitization phases, each of which impacts image quality and each of which needs to be considered if large-scale digital pathology is to be successful in clinical practice

18.6.3.1 Predigitization

- In pathology, imaging begins as the physiochemical process of fixing and processing tissue, cutting tissue sections, placing them on slides, and staining them with a variety of

technologies (e.g., histochemical staining, immunohistochemical staining, and fluorescent in situ hybridization [FISH] staining). These techniques remain as the most important part of the imaging process
- Traditionally, these techniques have been highly manual, with a high variance from slide to slide that is detrimental to the downstream digitization process. Increasingly, however, automation has been developed for each step of the process (processing, embedding, staining, coverslipping – there is even work on automated tissue sectioning)

18.6.3.2 Digitization

- Over the past 10 years, automated devices have been developed to digitize (or "scan") whole slides rapidly, automatically, and at high fidelity. Current high-end devices can automatically scan a slide in approximately 1 min at a spatial sample period of approximately 1.5 μm (similar to the resolution of a microscope with a 40x lens). Image quality and system throughput (the number of slides successfully scanned per hour) are dependent on the quality of the prediagnostic image phase (vide supra)

18.6.3.3 Postdigitization

- After image capture, the images are stored on a server and prepared for viewing and interpretation as discussed in the telepathology section above. Parameters that affect image quality include the amount of compression, the quality of the pathologist's monitor display, and the viewing environment

18.6.4 Current Status, Challenges, and Potential of Digital Pathology

- Digital pathology is in the early phases of development and implementation. There are still significant challenges as well as large, potential rewards
- Imaging consumes precious time and disrupts histology workflow
  - Current high-end scanners can scan a slide in approximately one minute (a rate that has been decreasing rapidly). However, if one considers that a large lab might generate 2,500 sides per day and that the standard histology process involves a large amount batching, it is clear that imaging the majority of the lab's slide will require a large number of scanners and potentially a change to a nonbatched, continuous flow histology operation
- Cost and profitability
  - Digital pathology requires significant capital, and profitability has yet to be proven
  - Scanners, servers, networks, monitors, potential infrastructural changes to the lab, and the workstations in the pathologist office when combined in to become prohibitively expensive, and no one has yet demonstrated a model with a clear return on investment
- Regulation and validation
  - While use of WSI for second opinion consultation and intraoperative consultation seems to be established, the current opinion of the FDA is that whole slide imaging systems will likely be regulated as class three devices when used for primary diagnosis; significantly, there is no consensus of the testing necessary to prove the safety and efficacy of WSI in primary diagnosis and therefore no mechanism to get devices cleared for this application
- The potential of digital pathology
  - Despite the significant challenges, digitizing the majority of histopathology slides automatically and at high fidelity has obvious potential. It would allow laboratories to apply network connectivity and computational power – the driving forces of innovation and efficiency in the modern world – to anatomic pathology

18.6.5 Nontraditional Imaging Techniques

- In addition to WSI and digital pathology, pathology labs are beginning to examine the use of techniques developed in radiology or endoscopy

- Virtopsy (virtual autopsy): the use of CT and other radiology tools as adjunct to the autopsy
- Specimen CT: potentially to evaluate margins or large tissue volumes
- In vivo microcopy (for the epithelial/endothelial surface of hollow organs and vessels)
- Multitissue section, high resolution, and three-dimensional reconstruction

18.7 Image Analysis

18.7.1 Goals

- The goal of imaging analysis is to assist the laboratory and pathologist when making morphology-based decisions such as ruling in or out specific diagnosis, measuring the quality of the slides and stains, counting structures such as mitoses or tumor size, or standardizing tumor grade. Image analysis can also be used to improve the digital pathology viewing experience of the pathology by, for example, registering the images of two adjacent tissue sections so that they can be viewed and navigated in tandem (Fig. 18.4)

18.7.2 Working Examples

- While diagnostic image analysis on whole slide images is in its infancy, the use of image analysis in automated Pap screening is a mature technique that is used by hundreds of laboratories on millions of specimens each year

18.7.3 Limitations to Image Analysis on Whole Slide Images

18.7.3.1 Data Size

- A typical whole slide image represents gigabytes of data in uncompressed form, placing a huge data load on any computational process

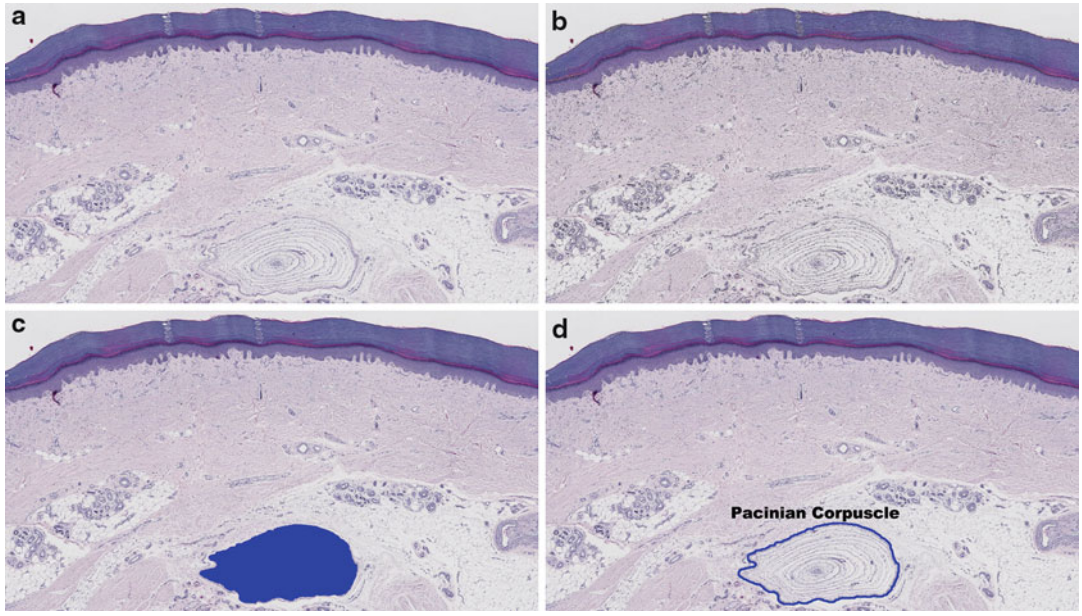


Fig. 18.4 Basic image analysis techniques: (**a**) Once the tissue section is stained and imaged, the digital slide can undergo a variety of manipulations for image analysis, such as (**b**) sharpening; (**c**) segmentation, in which an anatomical or histological structure is separated from the rest of the image via automated algorithms; and (**d**) classification, in which the segmented regions of the image is identified/associated with a specific anatomic or medical entity

18.7.3.2 Histology Artifact

- Standard, largely manual histology processes result in:
  - Tissue sections with variable thickness and that are rotated, translated, folded, and twisted
  - Staining can be highly variable

18.7.4 Techniques

- Histochemistry
  - Though not normally considered part of image analysis, histochemical, immunohistochemical, and FISH staining is an extremely powerful tool for tissue segmentation and structure classification (vide infra)

18.7.4.1 Averaging and Sharpening

- These are basic techniques, usually applied across the entire image, to remove noise (averaging) or sharpen edges. Done by applying appropriate filters to either the spatial or the color space

18.7.4.2 Segmentation

- The separation of meaningful structures in the image (usually anatomic or disease structures, such as separating bone from soft tissue); segmentation is done by applying "morphologic" or color "operators" that look for specific shapes, textures, and colors in image. Operators usually begin with a basic shape (e.g., a blue circle of a given size) and then apply "erosion," "dilation," or color variation programs to find similar structures

18.7.4.3 Classification

- Classification is the identification/association of specific objects or regions (identified through segmentation) with specific anatomic or medical entities (such as cancer cells and blood vessels)

18.7.4.4 Registration

- An image analysis technique important in digital pathology registration allows WSI from two adjacent tissue sections (which have been independently rotated, translated, and twisted in the histology process) to overlay on the same screen

18.7.4.5 Pipelining

- A typical image analysis process would involve histology staining, imaging, averaging and sharpening, segmentation, classification, and, potentially, registration in a "pipeline" and creating the completed analysis

18.8 Digital Pathology, Image Analysis, and the Laboratory Information System

- As the main workflow engine of the pathology lab, the LIS has a major role in the future implementation of large-scale digital pathology and image analysis

18.8.1 Image Orders and Workflow

- LIS has the capacity to include imaging and image analysis as part of the workflows (routes and protocols) defined in the LIS for specific specimen types and clinical questions. For example, the current protocol for a breast core biopsy might call for placing the entire core in a single block, process with a specific protocol, cut ten levels from the block and stain level 1, 5, and 10 with H&E, save the remainder for potential immunohistochemistry, and send the case to the breast pathologist on service. This could be extended to including imaging level 1, 5, and 10 at high resolution on a specific slide scanner

18.8.2 Specimen Identification and Messaging

- LIS barcodes and HL7 messages between the scanner and the LIS will manage and allow the scanner to identify slides and the appropriate imaging orders on each slide

18.8.3 Imaging at Sign Out

- Incorporation of imaging at sign out by integrating images into the standard sign out workflow

Further Reading

Baron JM, Dighe AS. Computerized provider order entry in the clinical laboratory. J Pathol Inf. 2011;2:35.

Buesa RJ. Adapting lean to histology laboratories. Ann Diagn Pathol. 2009;13:322–33.

Butte AJ. Translational bioinformatics applications in genome medicine. Genome Med. 2009;1:64.

Cankovic M, Varney RC, Whiteley L, Brown R, D'Angelo R, Chitale D, et al. The Henry Ford production system: LEAN process redesign improves service in the molecular diagnostic laboratory: a paper from the 2008 William Beaumont hospital symposium on molecular pathology. J Mol Diagn. 2009;11:390–9.

Cowan DF. Informatics for the clinical laboratory: a practical guide. New York: Springer; 2003.

Gabril MY, Yousef GM. Informatics for practicing anatomical pathologists: marking a new era in pathology practice. Mod Pathol. 2010;23:349–58.

Gilbertson J, Yagi Y. Histology, imaging and new diagnostic work-flows in pathology. Diagn Pathol. 2008;3(Suppl 1):S14.

Glenn TC. Field guide to next-generation DNA sequencers. Mol Ecol Resour. 2011;11:759–69.

Gomah ME, Turley JP, Lu H, Jones D. Modeling complex workflow in molecular diagnostics: design specifications of laboratory software for support of personalized medicine. J Mol Diagn. 2010;12:51–7.

Guidi GC, Lippi G, Solero GP, Poli G, Plebani M. Managing transferability of laboratory data. Clin Chim Acta. 2006;374:57–62.

Paszko C, Pugsley C. Considerations in selecting a laboratory information management system (LIMS). Am Lab. 2000;32:38.

Sinard JH. Practical pathology informatics: demystifying informatics for the practicing anatomic pathologist. New York: Springer; 2006.

Stephan C, Kohl M, Turewicz M, Podwojski K, Meyer HE, Eisenacher M. Using laboratory information management systems as central part of a proteomics data workflow. Proteomics. 2010;10:1230–49.

Tuthill M. Automating anatomic pathology. Implementing an AP solution that integrates with your hospital's LIS can improve workflow and productivity. Health Manag Technol. 2008;29:18, 20.

Weinstein RS, Graham AR, Richter LC, Barker GP, Krupinski EA, Lopez AM, et al. Overview of telepathology, virtual microscopy, and whole slide imaging: prospects for the future. Hum Pathol. 2009;40:1057–69.

Wilbur DC. Digital cytology: current state of the art and prospects for the future. Acta Cytol. 2011;55:227–38.

Section 3

Genetic Disorders

Prenatal Cytogenetic and Cytogenomic Diagnosis

19

Ninette Cohen and Lisa Edelmann

Contents

19.1 General 518

19.2 Prenatal Testing Modalities 518
19.2.1 Amniocentesis 518
19.2.2 Chorionic Villus Sampling 519
19.2.3 Percutaneous Umbilical Blood Sampling 519
19.2.4 Fetal Skin Biopsy 519

19.3 Indications for Prenatal Testing 520
19.3.1 Advanced Maternal Age 520
19.3.2 Ultrasound Findings 520
19.3.3 Abnormal Screening Results 520
19.3.4 Parental Chromosome Abnormality 520
19.3.5 Previous Pregnancy with Cytogenetic Abnormality 521

19.4 Fetal Abnormalities and Outcome 521
19.4.1 Structural Rearrangements 522
19.4.2 Supernumerary Marker Chromosomes 523
19.4.3 Uniparental Disomy 523

19.5 Fluorescence in Situ Hybridization (FISH) in Prenatal Testing 524
19.5.1 FISH on Direct Specimens to Screen for Common Aneuploidies 524
19.5.2 Subtelomere FISH Probes Can Be Used to Examine Fetal Chromosomes When Deletion or Duplication Is Suspected 524
19.5.3 Locus-Specific FISH Probes Can Be Used to Examine Fetal Chromosomes 525

19.6 Array Comparative Genomic Hybridization in Prenatal Testing 525
19.6.1 Prenatal Cytogenomic Diagnosis 525
19.6.2 Targeted Arrays for Copy Number Detection 526
19.6.3 Whole Genome Arrays for Copy Number Detection 526

Further Reading 526

N. Cohen, PhD (✉) • L. Edelmann, PhD
Department of Genetic and Genomic Sciences, Mount Sinai School of Medicine, New York, NY, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_19,

19.1 General

- Cytogenetic prenatal diagnosis involves examination of the fetal chromosome complement
  - Cells are obtained by various modalities during the first or second trimester and are subsequently established in tissue culture
  - Short-term in situ tissue culture methods produce sufficient cells for metaphase analysis
  - Chromosome preparations are banded and analyzed microscopically
  - Digitized images are submitted for diagnosis
- Cytogenomic prenatal diagnosis involves examination of genomic copy number at higher resolution than standard karyotype analysis
  - DNA is extracted from direct cells of the amniotic fluid specimen or tissue of the chorionic villus specimen or from cultured cells
  - Array comparative genomic hybridization (aCGH) is performed against a control
  - Specimen and hybridized to a microarray of genomic targets
  - The tiff image of the array file is analyzed by software that calculates the ratio of the signal intensity of the test specimen to the control specimen, which provides information regarding copy number

19.2 Prenatal Testing Modalities

19.2.1 Amniocentesis

19.2.1.1 Methodology

- Performed during second trimester between 16 and 18 weeks' gestation
- 20 cm^3 of amniotic fluid is withdrawn from the amniotic sac transabdominally under ultrasound guidance
  - The first 1–3 cm^3 is discarded to remove maternal cells
- Amniocytes, which are similar to fibroblasts, grow in tissue culture

19.2.1.2 Benefits

- Primary care obstetrician can perform the procedure
- Outpatient procedure in physician's office
- Procedure risk is 1 in 200 or 0.5%

19.2.1.3 Limitations

- Turnaround time is 6–10 days
- Results not available until the end of second trimester
- Interpretation is dependent upon the degree of chromosome band resolution, and high resolution is not possible

19.2.1.4 Laboratory Methodology In Situ Clonal Analysis

- Amniotic fluid specimen is centrifuged in order to obtain the fibroblast cells
- The cell pellet is suspended in tissue culture medium
- The suspension is placed on a sterile cover slip that is inside a small Petri dish
- The culture is placed in a 5% CO_2 incubator at 37 °C
- Culture medium is added after 24 h and then on a specified culture regimen
- After 5 days in culture, the cover slip is examined to determine if there is sufficient clonal activity for harvest, which occurs at 7 days
- When there is sufficient activity to produce adequate metaphases for analysis, the culture is subjected to a harvesting procedure utilizing a robotic harvester
  - Cell division is arrested by colcemid
  - Chromosomes are swollen by hypotonic treatment
  - Preparation is fixed on cover slip by acetic acid/methanol fixative
- Cover slips are transferred and fixed on slides
- Metaphase preparations are stained by a Giemsa–trypsin protocol and are then analyzed under the microscope

19.2.2 Chorionic Villus Sampling

19.2.2.1 Methodology

- Chorionic villus sampling (CVS) is performed during the first trimester at 10–12 weeks' gestation
- Chorionic villi are gently aspirated by catheter transvaginally or by needle aspiration transabdominally
- Placental tissue is dissected and enzymatically digested with a mixture containing trypsin and collagenase

19.2.2.2 Benefits

- Results are available by the end of the first trimester
- Turnaround time is 4–7 days with adequate specimens
- Direct analysis produces preliminary results in 8–24 h

19.2.2.3 Limitations

- Turnaround time is dependent upon the size of the initial sample
- Long technically proficient learning curve requiring specialist to perform procedure
- Procedure risk is 1%
- Interpretation is dependent upon the degree of chromosome band resolution, and high-resolution analysis is not possible
- Separation of villi from maternal decidua is critical for analysis
  - 1–2% risk for maternal cell contamination
- 1–2% incidence of confined placental mosaicism

19.2.2.4 Laboratory Methodology

- Cell pellet is obtained after above digestion
- Cells are suspended in media
- The same methodology as amniotic fluid cells (see Sect. 19.2.1.4) is employed to produce adequate metaphases for analysis
- *Note*: the tissue does not produce individual clones, and generally the cultures are ready to be processed after 3–5 days

19.2.3 Percutaneous Umbilical Blood Sampling

19.2.3.1 Methodology

- Needle is inserted into fetal umbilical vessel under ultrasound guidance
- Performed at 18–20 weeks' gestation
- Requires maternal fetal medicine specialist to do the procedure

19.2.3.2 Benefits

- Chromosome analysis results in 48–72 h
- Direct fetal lymphocyte diagnosis
- High-resolution chromosome analysis may be possible

19.2.3.3 Limitations

- Technically proficient specialist performs the procedure
- Risk of maternal blood contamination
- Procedure risk comparable with CVS
- Late second trimester diagnosis

19.2.3.4 Laboratory Methodology

- See Chap. 8

19.2.4 Fetal Skin Biopsy

19.2.4.1 Methodology

- A full-thickness skin biopsy is obtained from the fetus under ultrasound guidance or fetoscopy
- Performed during second or third trimester depending upon diagnosis
- Requires specialist to do the procedure
- Specific diagnoses are made by electron microscopy or immunohistochemical analysis of the tissue

19.2.4.2 Benefits

- No other methods can be used to make the diagnosis of rare dermatologic genetic disorders
  - Some examples are epidermolysis bullosa, congenital ichthyosis, and oculocutaneous albinism

Table 19.1 NT: nuchal translucency, β-hCG: maternal serum free β-human chorionic gonadotropin, PAPP-A: pregnancy-associated plasma protein A, msAFP: maternal serum α-fetoprotein, uE3: unconjugated estriol, inhA: inhibin A, ↑: high, ↓: low, √: included in screening

	Mat. age	NT	Free β-hCG	PAPP-A	msAFP	uE3	hCG	inhA
First trimester combined screening markers	√	√	√	√				
In trisomy 21		↑	↑	↓				
In trisomy 18		↑	↓	↓				
In trisomy 13		↑	↓	↓				
Second trimester quad screening	√				√	√	√	√
In trisomy 21					↓	↓	↑	↑
In trisomy 18					↓	↓	↑	Normal
In trisomy 13 – no screening available								

19.2.4.3 Limitations

- Most invasive procedure
- Useful for specialized diagnoses only
- Less invasive procedures are utilized for cytogenetic and molecular genetic diagnoses

19.3 Indications for Prenatal Testing

19.3.1 Advanced Maternal Age

- ACOG recommendation over the age of 35
  - Some practitioners present option as a consideration over the age of 30
- Risk of having a chromosome abnormality at 35 is 1 in 200
  - Gradually increasing risk as age advances
- Risk/benefit analysis determines patient preference methodology

19.3.2 Ultrasound Findings

- Structural abnormalities
  - Single or multiple congenital abnormalities
  - 25% of structural defects are associated with a chromosome abnormality
- Nuchal fold translucency
  - Cystic hygroma
  - Increased measurements of fetal neck

19.3.3 Abnormal Screening Results

- First trimester screening
  - Pregnancy-associated plasma protein A
  - Free β-human chorionic gonadotropin
  - Nuchal translucency measurements with crown-rump length to establish gestational age (assesses the thickness of soft tissues of the nape of neck of the fetus)
- Combined screening
  - Above three values plus maternal age
  - False-positive rate is 5%
  - Detection rate 80–90%
- Second trimester screening
  - α-fetoprotein
  - Unconjugated estriol
  - Inhibin A
  - Chorionic gonadotropin
- Integrated screening
  - Results from first and second trimester screening plus maternal age
- False-positive rate 2–3%
- Detection rate 80–90%
- See Table 19.1

19.3.4 Parental Chromosome Abnormality

- Structural abnormality
  - Balanced translocation carrier

    - Can produce unbalanced offspring
    - Viability determined by nature of translocation and reproductive history
    - Risks dependent on mechanism of meiotic separation (adjacent 1 segregation most likely to produce offspring with unbalanced karyotype)
  - Robertsonian translocation carrier
    - Involves acrocentric chromosomes 13, 14, 15, 21, and 22
    - Most common translocation, i.e., 13/14 associated with multiple miscarriages
    - Combinations with chromosome 21 results in translocation or familial Down syndrome
    - May be at risk for uniparental disomy (see Sect. 19.4) if translocation involves chromosomes 14 and 15
  - Pericentric inversion carrier
    - Unbalanced recombinant chromosomes may result from crossing over in a recombination loop during meiosis
    - More likely to occur if the loop is large due to greater chromosome distance
    - Duplication or deficiency of chromosomal material will produce phenotype effects
    - Viability is possible depending upon size and what chromosomes are involved
    - Small pericentric inversions have been noted as population variants, particularly involving chromosomes 9 and 2
      - No clinical consequences in these cases
  - Paracentric inversion carrier
    - During meiosis, acentric and dicentric chromosomes may result
    - Associated with early pregnancy loss
    - No risk for an abnormal live-born
- Mosaicism
  - Constitutional mosaicism
    - Different percent of normal versus aneuploid cell lines in different tissues
      - Phenotypic effects determined by percentage of aneuploid cells
  - Gonadal mosaicism
    - Mosaic cell line is present only in gonadal tissue
      - Gamete production is affected
      - Risk of abnormal offspring is based on percentage of mosaicism in gonads
    - All other tissues are normal
    - No phenotypic consequences

19.3.5 Previous Pregnancy with Cytogenetic Abnormality

- Previous trisomy 21, 18, or sex chromosome aneuploidy
  - Recurrence risk 1–2%
- Balanced de novo rearrangement
  - Sporadic recurrence unless gonadal mosaicism

19.4 Fetal Abnormalities and Outcome

Aneuploidy

- Most common trisomies identified, which may survive to term delivery
  - Trisomy 21 – Down syndrome
    - Mental retardation
    - Clinical history of hypotonia
    - Prominent occiput
    - Characteristic facies consisting of oblique palpebral fissures, epicanthal folds, low-set ears, flat nose bridge, and large protruding tongue
    - Congenital heart disease, i.e., A–V canal
    - Duodenal atresia
    - Bilateral simian creases
  - Trisomy 18 – Edward syndrome
    - Intrauterine growth retardation
    - Mental retardation
    - Micrognathia
    - Low-set ears

    - Congenital heart disease, i.e., ventricular septal defect (VSD)
    - Contractures with characteristic hand position, i.e., second digit over third and fifth over fourth
    - Rocker bottom feet
  - Trisomy 13 – Patau syndrome
    - Mental retardation
    - Characteristic craniofacial abnormalities consisting of bilateral cleft lip and palate and holoprosencephaly
    - Polydactyly
    - Polycystic kidneys
    - Congenital heart disease, i.e., ASD and VSD
- Sex chromosome aneuploidies
  - Klinefelter syndrome – XXY
    - Tall stature
    - Small sclerotic postpubertal testes
      - Azoospermia
    - Gynecomastia
    - Clinical history of learning disabilities
  - Triple X syndrome – XXX
    - Variable clinical history of spontaneous abortions
    - Normal phenotype
    - Clinical history of severe learning disabilities
  - XYY
    - Tall stature
    - Prominent metopic suture
    - Clinical history of learning disabilities
    - Clinical history of behavior problems
- Monosomy
  - Turner syndrome – 45, X
    - Majority (80%) due to a paternal meiotic error
    - High incidence in spontaneous abortions
    - 99% of conceptuses spontaneously abort
    - Of those that survive to term, phenotype consists of
      - Short stature
      - Webbed neck
      - Triangular facies
      - Coarctation of the aorta (20%)
      - Structural kidney abnormalities
      - Café au lait spots
      - Clinical history of learning disabilities (spatial)
      - Ovarian dysgenesis
- Double aneuploidy
  - Identified in very early pregnancies
  - May be picked up at time of CVS
  - Very few survive to term
  - Frequent finding in spontaneous abortions
  - Involves both autosome and sex chromosomes
- Triploidy/tetraploidy
  - Three or four sets of chromosomes
    - Chromosome number 69 or 92
    - Results from dispermy event or reabsorption of one polar body
    - Identified in very early pregnancies
    - May be picked up at time of CVS
    - Survival similar to double aneuploidy

19.4.1 Structural Rearrangements

- Known familial rearrangement
  - Parental balanced translocation (see Sect. 19.3.1)
  - Parental pericentric inversion (see Sect. 19.3.1)
  - Parental paracentric inversion
    - May have history of reproductive loss
    - Balanced inversion progeny can survive and have similar reproductive history as parent
- De novo rearrangement
  - May be associated with multiple congenital anomaly/intellectual disability (MCA/ID) syndromes
  - Risk reported as high as 10%, may be higher if X chromosome is involved
  - Cannot determine if completely balanced
    - Limitation dependent upon microscopic resolution
      - Molecular techniques may further define possibility of balanced rearrangement (see Sect. 19.6)

19.4.2 Supernumerary Marker Chromosomes

- Variable size
  - Some may be "dot" like
- Banding patterns cannot be determined by routine staining
  - Nucleolus organizer region (NOR) staining or molecular probes to determine whether satellites are present
  - C-banding to determine amount of heterochromatin
- Origin usually dependent on molecular studies
  - Fluorescence in situ hybridization (FISH) panel for most common derivatives (acrocentric chromosomes, X chromosome)
    - Chromosome 15 markers, 50% of all markers
    - Known syndrome of intellectual disability when Prader–Willi syndrome (PWS) critical region including *SNRPN* is present
- Even if origin is determined, it may not be able to predict outcome and prognosis since there is a limited pool of data
  - Unknown phenotype for the majority of cases except for chromosome 15
- Familial markers
  - Parental marker may seem to be the same by routine staining and, if so, then the risk for abnormality is reduced
  - However, molecular subtelomere studies have shown that parental marker may be balanced but proband may not be (see Sect. 19.6.1)

19.4.3 Uniparental Disomy

- Uniparental disomy (UPD) arises when an individual inherits both copies of a chromosome from one parent
  - Loss of a chromosome in a trisomic zygote (trisomy rescue)
  - Duplication of a chromosome in a monosomic gamete (monosomy rescue)
  - Fertilization with two copies of a chromosome in one gamete and no copies of the chromosome in the other gamete (nullisomic and disomic gametes)
- Associated with clinical phenotypes only for chromosomes that contain imprinted genes, which are genes that are expressed from either the maternally or the paternally inherited chromosome but not both
- One copy of an imprinted gene is silenced, whereas the other is active
- A maternally imprinted gene is not expressed from the maternally inherited chromosome and vice versa
  - Chromosome 6
    - Paternal UPD 6 – transient neonatal diabetes mellitus
  - Chromosome 7
    - Maternal UPD – Russell–Silver syndrome and IUGR
  - Chromosome 11
    - Paternal UPD – Beckwith–Wiedemann syndrome
    - Maternal UPD – Russell–Silver syndrome and IUGR
  - Chromosome 14
    - Maternal UPD 14 – Intrauterine growth retardation (IUGR) and mild dysmorphic features
    - Paternal UPD 14 – hypotonia, thoracic dystrophy, and developmental delay
  - Chromosome 15
    - Paternal UPD – Angelman syndrome
    - Maternal UPD – Prader–Willi syndrome
  - Chromosome 16
    - Maternal UPD 16 – IUGR and congenital anomalies
- Diagnosis is made by molecular methods
  - Analysis of microsatellite markers of fetus and parents
  - Chromosome microarray testing with SNP genotyping markers of fetus and parents
- UPD testing should be offered
  - When mosaic or nonmosaic trisomy for any of the above chromosomes is observed on CVS specimen and if followup amniocentesis is chromosomally normal

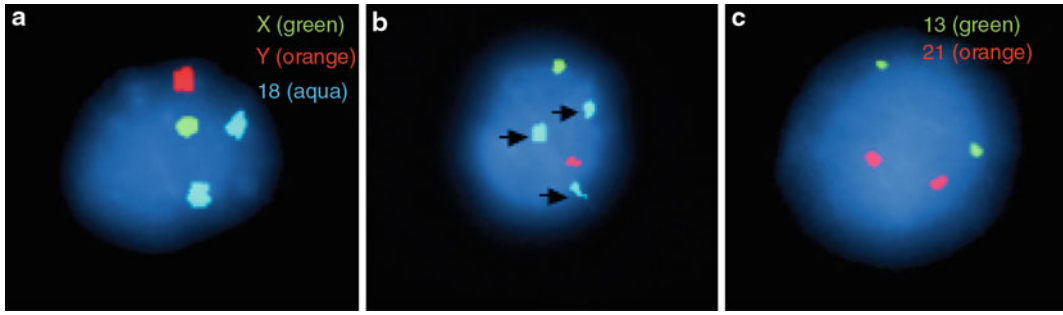


Fig. 19.1 (**a**) FISH performed on interphase nuclei from amniotic fluid using a probe mix containing the X chromosome centromere sequences fluorescently labeled in *green* fluorochrome, the Y chromosome centromere sequences fluorescently labeled in *orange* fluorochrome, and the chromosome 18 centromere sequences fluorescently labeled in *aqua* fluorochrome. This image shows a normal male hybridization pattern for chromosomes X, Y, and 18. (**b**) Probes are the same as in figure (**a**). Image of a nucleus from amniotic fluid of a male fetus with a hybridization pattern for chromosome 18 consistent with trisomy 18 (*arrows* demarcate chromosome 18 probe). (**c**) FISH performed on interphase nuclei from amniotic fluid using a probe mix containing specific loci on chromosome 13 fluorescently labeled in *green* fluorochrome and chromosome 21 fluorescently labeled in *orange* fluorochrome. This image shows a normal hybridization pattern for chromosomes 13 and 21

  - When Robertsonian translocation involving chromosomes 14 or 15 is observed in the fetal karyotype

19.5 Fluorescence in Situ Hybridization (FISH) in Prenatal Testing

19.5.1 FISH on Direct Specimens to Screen for Common Aneuploidies

- Rapid analysis on interphase nuclei of direct amniocytes or chorionic villi (8–24 h)
  - Followup for abnormal ultrasound findings
  - Cystic hygroma – abnormality most commonly associated with aneuploidy
  - Followup for abnormal first or second trimester biochemical screening
  - In cases of advanced maternal age or previous aneuploid pregnancy
- Enumeration of most common autosomal (13, 18, 21) and sex chromosome (X, Y) aneuploidies (Fig. 19.1)
- Used in conjunction with standard karyotype analysis
- Limitations
  - Difficulty interpreting mosaic findings
  - Maternal blood in specimens may confound results
  - Direct CVS specimen (trophoblast) represents different population of cells than cultured cells (villus stroma); therefore, placental mosaicism may yield FISH results that conflict with karyotype
  - Does not detect all aneuploidies

19.5.2 Subtelomere FISH Probes Can Be Used to Examine Fetal Chromosomes When Deletion or Duplication Is Suspected

- When parent is a carrier of a cryptic balanced translocation
  - Risk for partial monosomy and partial trisomy in unbalanced fetuses most likely due to adjacent 1 segregation

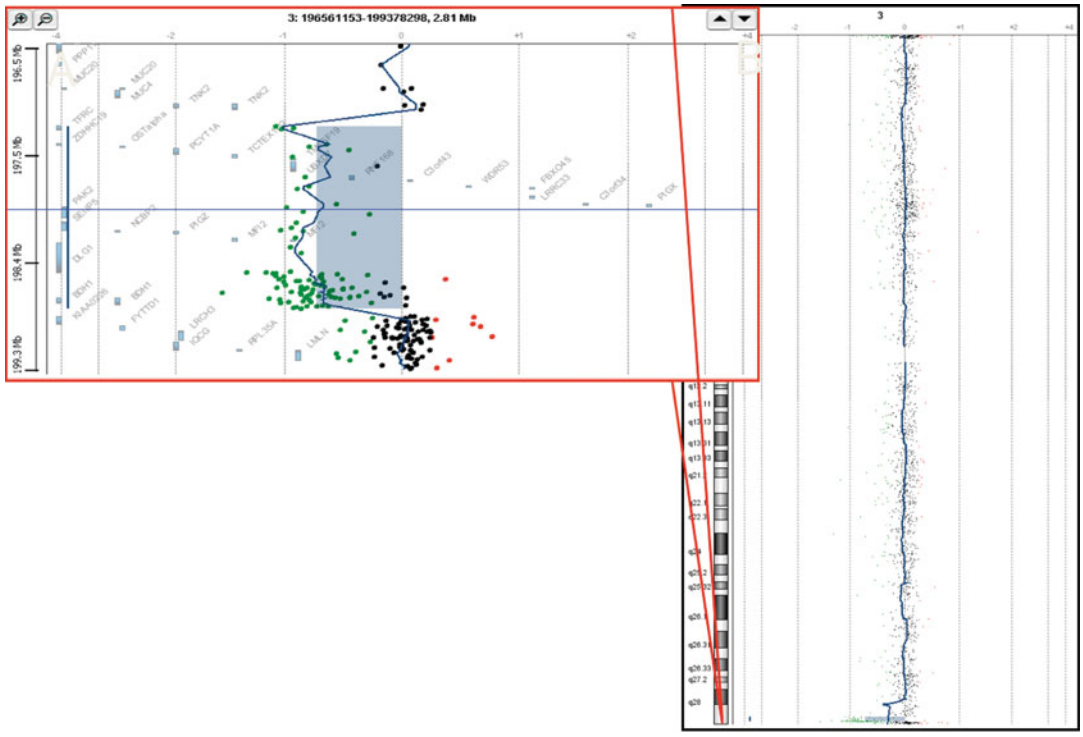


Fig. 19.2 Example of a deletion detected by aCGH. Depicted on the *left* is view of chromosome 3q29 with a 1.6-Mb deletion detected. *Green* probes are deleted with a negative log2 ratio, *black* probes have a normal log2 ratio, and *red* probes have increased log2 ratio. At least four consecutive probes must be present within an aberration. On the *right* is the whole chromosome view of chromosome 3. The deletion is present on the *bottom* of the figure and represented by the *shaded area*. This deletion is representative of 3q29 microdeletion syndrome, a known disorder with mild facial dysmorphism, ataxia, autism, microcephaly, and intellectual disability

- When de novo translocation involving at least one telomere is identified in fetal karyotype to assess whether it is balanced (see Sect. 19.6.1)

19.5.3 Locus-Specific FISH Probes Can Be Used to Examine Fetal Chromosomes

- Ultrasound abnormality indicative of a specific microdeletion/duplication syndrome
  - Cardiac defects and 22q11 FISH probe (e.g., Tuple1 or N25)
  - Lissencephaly and 17p13.3 FISH probe (e.g., Lis1)
- Parent is a carrier of a microdeletion (50% chance of transmission)
- For couples with a previous child with a microdeletion/duplication syndrome in case of gonadal mosaicism in one parent (rare)

19.6 Array Comparative Genomic Hybridization in Prenatal Testing

19.6.1 Prenatal Cytogenomic Diagnosis

- Also referred to as chromosomal microarray or molecular karyotyping
- Can interrogate the genome for copy number gains and losses (duplications and deletions)

at much higher resolution than standard karyotype analysis
- Performed on DNA extracted from direct chorionic villi or amniotic fluid cells or on cultured cells
- Involves differential fluorescent labeling of fetal DNA and control DNA hybridized together on the array to determine relative amounts of material hybridized to each probe on the array
- Usually represented as the log2 ratio of the intensities of the fetal to control DNA (Fig. 19.2) (see Chap. 10 for a detailed methodology)
- Turn around time for direct specimens <5 days
- Used in conjunction with standard karyotype analysis
  - Fetuses with structural ultrasound abnormalities (1–3% detection rate)
  - Increased NT or IUGR (detection rate unknown)
  - Family history of chromosome abnormality
  - Follow up to abnormal chromosome result
    - Balanced translocation
    - Inversion
    - Marker chromosome identification
    - Other structural rearrangement of unknown origin
  - Parental concern (no known fetal abnormality)

19.6.2 Targeted Arrays for Copy Number Detection

- Bacterial artificial chromosome (BAC) clones or oligonucleotides are printed onto glass slides or covalently linked to beads to create an array of targets
- More than 40 known microdeletion and microduplication disorders and all subtelomeric regions are generally represented
- Detects rearrangements not visible by standard karyotype analysis
- Limitations
  - Not a whole genome array – only targeted regions of the genome
  - Will not detect balanced rearrangements
  - Inherited copy number variants (CNVs) can confound results so parents may need to be tested
  - Variants of unknown significance may be detected, although risk is lower with a targeted array versus a whole genome array

19.6.3 Whole Genome Arrays for Copy Number Detection

- BACs or oligonucleotides are printed onto glass slides to create an array of targets
- Whole genome is arrayed so aberrations larger than 200–500 kb are detected
- All known microdeletion and microduplication syndromes are represented
- Useful for assessing chromosome aberrations of unknown origin and in followup of chromosome rearrangements detected by standard karyotype analysis (as described above)
- Higher detection rate than targeted arrays
- Limitations
  - Will not detect balanced rearrangements
  - Inherited CNVs can confound results so parents may need to be tested
  - Variants of unknown significance may be detected making genetic counseling difficult
  - Information regarding adult-onset disorders, including infertility, may be obtained

Further Reading

Brady PD, Devriendt K, Deprest J, Vermeesch JR. Array based approaches in prenatal diagnosis. Methods Mol Biol. 2012;838:151–71.

Crolla JA. FISH and molecular studies of autosomal supernumerary marker chromosomes excluding those derived from chromosome 15: II. Review of the literature. Am J Med Genet. 1998;75:367–81.

Evans MI, Wapner RJ. Invasive prenatal diagnostic procedures 2005. Semin Perinatol. 2005;29:215–18.

Gardner RJM, Sutherland GR, Shaffer LG. Chromosome abnormalities and genetic counseling. 4th ed. New York: Oxford University Press; 2012.

Hillman SC, Pretlove S, Coomarasamy A, McMullan DJ, Davison EV, Maher ER, Kilby MD. Additional information from array comparative genomic hybridization technology over conventional karyotyping in prenatal diagnosis: a systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2011;37:6–14.

Kagan KO, Etchegaray A, Zhou Y, et al. Prospective validation of first-trimester combined screening for trisomy 21. Ultrasound Obstet Gynecol. 2009;34: 14–18.

Kotzot D. Prenatal testing for uniparental disomy: indications and clinical relevance. Ultrasound Obstet Gynecol. 2008;31:100–5.

Ledbetter DH, Engel E. Uniparental disomy in humans: development of an imprinting map and its implications for prenatal diagnosis. Hum Mol Genet. 1995;4: 1757–64.

Nicolaides KH. First-trimester screening for chromosomal abnormalities (Review). Semin Perinatol. 2005;29: 190–4.

Randolph LM. Prenatal Cytogenetics. In: Gersen SL, Keagle MB, editors. The principles of clinical cytogenetics. Totowa: Humana Press; 2005. p. 267–321.

Shaffer LG. Risk estimates for uniparental disomy following prenatal detection of a nonhomologous Robertsonian translocation. Prenat Diagn. 2006;26:303–7.

Shaffer LG, Coppinger J, Alliman S. Comparison of microarray-based detection rates for cytogenetic abnormalities in prenatal and neonatal specimens. Prenatal Diagn. 2008;28:789–95.

Tepperberg J, Pettenati MJ, Rao PN. Prenatal diagnosis using interphase fluorescence in situ hybridization (FISH), 2-year multi-center retrospective study and review of the literature. Prenat Diagn. 2001;21: 293–301.

Van den Veyver IB, Patel A, Shaw CA, et al. Clinical use of array comparative genomic hybridization (aCGH) for prenatal diagnosis in 300 cases. Prenatal Diagn. 2009;29:29–39.

Warburton D. De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. Am J Hum Genet. 1991;49(5):995–1013.

Wenstrom KD. Evaluation of Down syndrome screening strategies. Semin Perinatol. 2005;29(4):219–24.

Molecular Medical Genetics

20

Lisa Edelmann, Stuart Scott, Liu Liu, and Ruth Kornreich

Contents

20.1 Introduction 530
20.1.1 Inheritance of Single-Gene Disorders 530

20.2 Autosomal Recessive Disorders 530
20.2.1 Cystic Fibrosis 530
20.2.2 Ashkenazi Jewish Carrier Screening 532
20.2.3 Hereditary Hemochromatosis 535
20.2.4 Spinal Muscular Atrophy 536
20.2.5 Medium-Chain Acyl-CoA Dehydrogenase Deficiency 538

20.3 Autosomal Dominant Disorders 539
20.3.1 Nucleotide Repeat Expansion Disorders 539
20.3.2 Skeletal and Connective Tissue Disorders 545

20.4 X-Linked Inheritance 550
20.4.1 Fragile X Syndrome 550
20.4.2 X-Linked Muscular Dystrophy (Duchenne Muscular Dystrophy and Becker Muscular Dystrophy) 552
20.4.3 X-linked Adrenoleukodystrophy 553
20.4.4 Rett Syndrome 554

20.5 Mitochondrial Disorders 555
20.5.1 The Mitochondrion 555
20.5.2 Defects of mtDNA 556
20.5.3 Defects of Nuclear DNA 557
20.5.4 Diagnostic Evaluation 558

Further Reading 558

L. Edelmann, PhD (✉) • S. Scott, PhD • R. Kornreich, PhD
Department of Genetic and Genomic Sciences, Mount Sinai School of Medicine, New York, NY, USA

L. Liu, PhD, FACMG
Cancer Genetics Laboratory, Baylor College of Medicine, Houston, TX, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_20,

20.1 Introduction

20.1.1 Inheritance of Single-Gene Disorders

- Mode of inheritance is important for predicting clinical status of individuals carrying mutations and for risk assessment of family members of a patient affected with a genetic disease. Modes include:
  - Autosomal recessive (AR)
  - Autosomal dominant (AD)
  - X-linked
  - Mitochondrial
- Examples will be given for each subgroup

20.2 Autosomal Recessive Disorders

20.2.1 Cystic Fibrosis

- Cystic fibrosis (CF) is one of the most common AR severe disorders in the Caucasian population with an incidence of 1 in 2,500–1 in 3,300 births in non–Hispanic-Caucasians
- Its pathophysiology results from the secretion of thick mucus by membrane epithelial cells, which interrupts the function of organs such as the lungs, pancreas, intestine, and male reproductive tract by defective chloride and sodium transport
- Clinical
  - Progressive, multisystem disorder that primarily affects the respiratory, digestive, and reproductive systems
  - Failure to thrive, meconium ileus in 10–20% of infants with CF
  - Recurrent respiratory infections lead to chronic pulmonary disease
  - Exocrine pancreatic insufficiency in 85% of patients
  - Infertility in males due to congenital absence of vas deferens
  - Treatment involves enzyme and dietary supplementation for pancreatic insufficiency and respiratory therapy with treatment of infections
  - Early death usually due to obstructive pulmonary disease
  - Mean survival ~35 years
  - Approximately 15% of CF individuals have a milder course with pancreatic sufficiency and a median survival of 56 years
- Prevalence
  - Varies depending on the ethnic group
    - Non–Hispanic-Caucasian – 1 in 2,500 births
    - Ashkenazi Jewish (AJ) – 1 in 2,300 births
    - Hispanic – 1 in 13,500 births
    - African–American – 1 in 15,100 births
    - Asian–Americans – 1 in 35,100 births
- Genetics
  - Caused by mutations in CF transmembrane regulator *(CFTR)* gene
    - Cloned in 1989 by positional cloning
    - Spans 230 kb on the long arm of chromosome 7, encodes 6,500 nucleotide mRNA, 27 exons
    - 1480-amino acid transmembrane protein predominately located at the apical membrane of epithelial cells of the lungs, sinus, pancreas, intestine, sweat and bile ducts, and vas deferens
    - Multifunctional protein
      - Chloride channel activated by cAMP
      - Transporter regulator of other channels including epithelial sodium channel
    - Five domains: two membrane spanning, two nucleotide binding that interact with ATP, and one regulatory (Fig. 20.1)
  - Over 1,900 *CFTR* mutations have been reported; most rare, but several recurring at a higher frequency (http://www.genet.sickkids.on.ca/cftr)
  - p.F508del most common with frequencies ranging from 18% to 88% in different ethnic, demographic, and racial groups and with an overall worldwide frequency of approximately 70%
  - p.G542X is the next common mutation responsible for 2.4% of alleles, and only

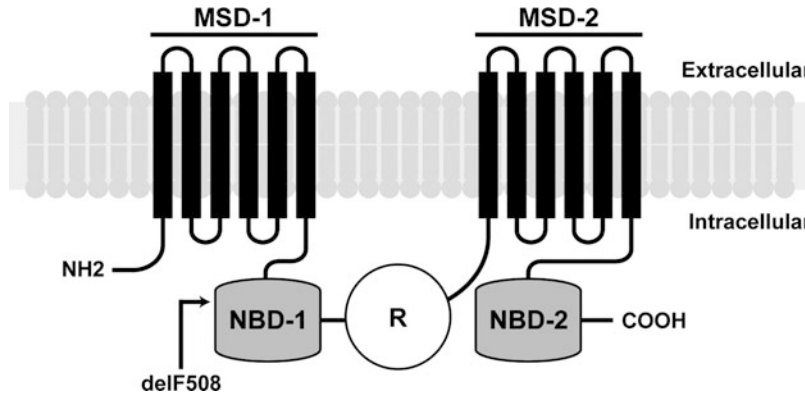


Fig. 20.1 Schematic representation of the proposed *CFTR* structure. *CFTR* is comprised of five domains: two membrane-spanning domains (MSD-1 and MSD-2) that form the chloride ion channel, two nucleotide-binding domains (NBD-1 and NBD-2) that bind and hydrolyze ATP, and a regulatory domain (R) that is phosphorylated by protein kinases A and C. The location of the most common CF-causing mutation, p.F508del, is noted

another three mutations occur above 1% in US CF patients
  - Genotype–phenotype correlations are inexact and should not be used to predict survival or level of pulmonary disease
    - Some mutations [c.3717 + 12191C > T (3849 + 10kbC > T), p.A455E, c.2657 + 5 G > A (2789 + 5 G > A), p.G85E, p.R334W] lead to milder CF phenotype with possibly pancreatic sufficiency
    - Variants reported which may lead to only limited lung involvement, sinusitis, or male infertility due to congenital bilateral absence of the vas deferens
    - Polypyrimidine tract in intron 8 variable—5T/7T/9T possible
      - 5T leads to low RNA splicing
      - 5T homozygosity or 5T with *CFTR* mutation on other chromosome not clinically significant in females but can cause congenital bilateral absence of the vas deferens in males
      - p.R117H without 5T on the same chromosome nonclassic CF allele (p.R117H with 5T on same allele, in *cis*, associated with classic CF)
- Diagnosis
  - Sweat electrolyte levels
  - Immunoreactive trypsin
  - DNA analysis
    - Most testing methods initiate with multiplexed polymerase chain reaction (PCR) amplification; however, the technologies used to differentiate wild type from mutant alleles following PCR vary
    - Commercially available kits to detect common mutations include
      - Reverse dot blot
      - Bead arrays
      - Amplification refractory mutation systems
      - Oligonucleotide ligation assay
      - Fluorescence resonance energy transfer (FRET)
    - Some laboratories develop testing in-house using homebrews, while others use commercially available analyte-specific reagents as well as specialized instrumentation
- Population carrier screening
  - National Institutes of Health (NIH) recommended in 1997 that CF mutation screening be offered to adults with a positive family history of CF, partners of individuals with CF, and couples who are pregnant or are planning a pregnancy
    - Implementation difficult due to the large number of mutations, varying distribution of mutations in different

ethnic groups, and lack of sufficient clinical correlation and educational material
- Guidelines established by joint American College of Medical Genetics (ACMG), American College of Obstetricians and Gynecologists (ACOG), and NIH committee published in 2001
  - Carrier screening should be offered to Caucasians and "made available" to other ethnic groups
  - Couple-based or sequential depending on circumstances
  - Patient literature should be provided
  - Panethnic mutation panel, which included 25 CF-causing mutations with an allele frequency of ≥0.1% in the general US population recommended (Table 20.1)
  - Advocated additional reflex testing of p.R117H carriers for the 5/7/9 T alleles in the polypyrimidine tract of intron 8 as well as reflex testing for interference from the benign variants p.F508C, p.I506V, and p.I507 V in p.F508del-positive samples when indicated
  - Panel modified to include 23 mutations in 2005
    - c.948delT (1078delT) removed because true frequency falls below the frequency threshold of ≥0.1%
    - p.I148T had a >100-fold increase in individuals having carrier screening compared with CF population
    - p.I148T deleted as it does not cause classic CF—exists as a complex allele with c.3067_3072delATAGTG (3199del6), which is the true severe CF mutation
  - Table 20.2 shows residual carrier risks in different ethnic groups after recommended screening panel

20.2.2 Ashkenazi Jewish Carrier Screening

- Ashkenazi Jews (AJ) originated from Eastern Europe and represent approximately 90% of the 5.7 million Jews in the United States today
- Over 40 disorders with a higher prevalence in this group have been described
- Screening began with Tay–Sachs disease (TSD) in the 1970s by enzymatic analysis
- ACOG recommends screening for at least Tay–Sachs, Canavan, CF, and familial dysautonomia (FD) (other disorders may be offered)
- ACMG recommends an additional five tests (Bloom syndrome, Fanconi anemia type C, Gaucher disease, mucolipidosis type IV, and Niemann–Pick disease)
- Most labs now are offering testing for at least eight "AJ" diseases and CF (Table 20.3)
- Many laboratories screen for 18 diseases or more in this population
- Screening is performed mostly by targeted mutation analysis as limited number of mutations are needed to detect >90% of carriers
- Commercial kits for the AJ panel are limited

20.2.2.1 Tay–Sachs Disease

- AR lysosomal storage disorder caused by deficiency of β-hexosaminidase A
- Mostly infantile form, which is uniformly fatal in childhood due to progressive neurologic deterioration
- Much less frequent later-onset disorder known as chronic GM2 gangliosidosis characterized by muscle weakness, ataxia, dysarthria, mild mental impairment, and psychosis
- Mostly AJ (1 in 27 carrier frequency) but also prevalent in Cajun or French–Canadian ancestry
- Tenfold lower (1 in 300 carrier frequency) in non-Jewish general population
- No treatment currently available
- Diagnostics
  - Most labs use a combination of measurement of hexosaminidase levels and mutation analysis
  - Screening can be done by targeted mutation analysis only if individual is of 100% AJ descent
  - Three mutations (c.1274_1277dupTATC, c.1421 + 1 G > C, p.G269S) account for 98% of AJ mutations, but <50% of non-Jewish alleles

Table 20.1 Recommended core panel of 25 mutations for general population cystic fibrosis carrier screening (2001)

Standard mutation panel					
delF508	delF507	G542X	G551D	W1282X	N1303K
R553X	621 + 1 G → T	R117H	1717-1 G → A	A455E	R560T
R1162X	G85E	R334W	R347P	711 + 1 G → T	1898 + 1 G → A
2184delA	1078delT	3849 + 10kbC → T	2789 G → A	3659delC	I148T
3120 + 1 G → A	–	–	–	–	–
Reflex tests					

I506V, I507, F508C: non-cystic fibrosis-causing variants test only if unexpected homozygosity for delF508 and/or delF507 (F508C has been associated with CBVAD); 5 T/7 T/9 T-test only for R117H positives
Common nomenclature is used in this table; for updated nomenclature, please see http://www.genet.sickkids.on.ca/cftr

Table 20.2 Cystic fibrosis mutation detection and carrier rates before and after testing in various ethnic groups using the recommended core panel of 23 mutations

Ethnic group	Carrier frequency	Detection rate (%)	Carrier risk after negative result
Ashkenazi Jewish	1 in 24	94	1 in 384
Non–Hispanic-Caucasian	1 in 25	88	1 in 206
Hispanic–Caucasian	1 in 58	72	1 in 203
African–American	1 in 62	65	1 in 171
African–American	1 in 94	49	1 in 183

  - Measurement of hexosaminidase A enzyme levels in serum, plasma, leukocytes, platelets, or other sources will pick up most carriers; however, 2–5% of individuals tested will be inconclusive (because of overlap between high carrier and low normal ranges)
  - Pseudodeficiency alleles exist such that individuals will have reduced enzyme activity using the in vitro screening assay, but not against the natural substrate, GM2 gangliosides (not true TSD carriers)
    - p.R247W – found in 2% of AJ and 32% of non-Jewish carriers by enzyme analysis
    - p.R249W – found in 4% of non-Jewish carriers by enzyme analysis
  - All individuals found to be carriers or who test inconclusive by enzyme analysis should have followup by molecular testing

20.2.2.2 Canavan Disease

- AR neurodegenerative leukodystrophy caused by deficiency of the enzyme aspartoacylase
- Affected infants develop normally during first few months of life and then experience a marked loss of early milestones and clinical features including megalocephaly, poor head control, and seizures
- Death usually during first decade of life
- Carrier frequencies reported to be 1 in 40–1 in 57 in AJ, significantly lower in non-Jews, but reported in many ethnic groups
- Diagnostics
  - Measurement of the substrate N-acetylaspartic acid in fluids possible, but not always reliable
  - p.E285A and p.Y231X account for approximately 97% of AJ Canavan disease (CD) mutations
  - p.A305E accounts for approximately 1% of AJ CD alleles, but 50% of non-Jewish mutations

20.2.2.3 Familial Dysautonomia

- AR neuropathy that almost exclusively affects AJ infants
- Disorder of sensory and autonomic function characterized by absence of tearing, absence

Table 20.3 Carrier screening in the AJ population

Disease	AJ carrier frequency	Number of AJ mutations commonly screened	Detectability (%)
Gaucher type I	1:15	4	95
Cystic fibrosis	1:24	23[a]	94
Tay–Sachs	1:27	3	98
FD	1:31	2	>99
Canavan	1:55	3	97
Fanconi group C	1:100	1	99
Niemann–Pick A and B	1:115	4	95
Bloom syndrome	1:134	1	99
MLIV	1:89	2	95

AJ Ashkenazi Jewish, *FD* familial dysautonomia, *MLIV* mucolipidosis, type IV
[a]Most laboratories do not test for less than the 23 ACMG recommended panel, although only five mutations are common in AJ

of papillae on the tongue, protracted vomiting, decreased discrimination to pain and temperature, and cardiovascular instability
- Carrier frequency of 1 in 31 in AJ and found almost exclusively in AJ
- Diagnostics
  - One mutation in the *IKBKAP* gene, c.2204 + 6T > C (IVS20 + 6T > C), accounts for >99% of mutant AJ alleles
  - The missense mutation, p.R696P, has also been found in small number of AJ families
  - p.P914L was found in a non-AJ individual

20.2.2.4 Type I Gaucher Disease

- AR lysosomal storage disorder caused by deficient activity of the enzyme β-glucosidase
- One of the most prevalent disorders among AJ, with a carrier frequency of 1 in 15
- The glycolipid, glucocerebrosidase, accumulates primarily in the cells of the macrophage–monocyte system
- Extremely clinically heterogeneous, ranging from early onset of severe disease to a mild course
- Asymptomatic affected individuals have been picked up in routine carrier screening
- Enzyme replacement therapy is available
- Diagnostics
  - Testing for glucocerebrosidase activity in cells is possible, but it is not accurate for carrier detection
  - Testing for four mutations [p.N409S (N370S), c.84dupG (84GG), p.L483P (L444P), and c.115 + G > A (IVS2)] in the AJ, will detect about 95% of mutant Gaucher disease alleles
  - Care must be taken when designing assays as there is a nearby processed pseudogene
- Genotype–phenotype correlations
  - Individuals homozygous for the most common AJ allele, p.N409S, will have a nonneurologic disorder with an average age of onset of about 30 years
  - Individuals with one copy of p.N409S and either c.84dupG, p.L483P, or c.115 + G > A will have the nonneurologic form of type I GD but will have a more severe course than p.N409S homozygotes
  - Affected individuals with two copies or any combination of c.84dupG, p.L483P, or c.115 + G > A will have severe disease with a neurodegenerative course

20.2.2.5 Fanconi Anemia, Group C

- AR disease characterized by short stature, bone marrow failure, congenital malformations, and a predisposition to acute myelogenous leukemia
- A single mutation in the *FACC* gene c.456 + 4A > T (IVS4 + 4A > T) accounts for nearly all AJ mutant alleles
- AJ carrier frequency of 1 in 100

20.2.2.6 Niemann–Pick Disease, Types A and B

- AR disorders resulting from the deficient activity of the lysosomal enzyme acid sphingomyelinase

- Type A disease
  - Characterized by a rapid progressive neurodegenerative course and hepatosplenomegaly in infancy with death usually occurring by 3 years of age
  - Carrier frequency of 1 in 115 in AJ
- Type B disease
  - Panethnic
  - Milder than type A with primarily hematologic and pulmonary symptomology and little if any neurologic involvement
- Diagnostics
  - Three common AJ type A mutations [p.R498L, p.L304P, and c.996delC (fsP330)] account for approximately 95% of AJ alleles
  - One recurrent type B mutation, p.R610del

20.2.2.7 Bloom Syndrome

- AR neuropathy condition that is characterized clinically by severe prenatal and postnatal growth deficiency, a sun-sensitive telangiectatic rash, and a predisposition for different malignant and benign tumors that develop in early childhood and adolescence
- Results from mutations in the *BLM* gene, a RecQ helicase
- AJ carrier frequency 1 in 134
- Diagnostics
  - A single mutation (BLM^{Ash}; c.2207_2212delATCTGAinsTAGATTC (2281del6ins7)) accounts for almost all AJ *BLM* mutations

20.2.2.8 Mucolipidosis, Type IV

- AR, neurodegenerative lysosomal storage disease characterized by a variable degree of growth and psychomotor retardation, and ophthalmologic abnormalities, which include corneal clouding and progressive retinal degeneration
- Life span may be normal, but most patients remain at a developmental level of 1–2 years
- Accumulation of cytoplasmic storage bodies but normal levels of lysosomal hydrolases present
- 80% of individuals with mucolipidosis type IV (MLIV) are AJ
- Mutations in *MCOLN1* gene are responsible
- AJ carrier frequency 1 in 89
- Diagnostics
  - Two mutations, c.406-2A > G (IVS3-2A > G) and g.511_6943del (del6.4 kb), account for 95–96% of AJ MLIV alleles

20.2.3 Hereditary Hemochromatosis

- Hereditary hemochromatosis (HH) consists of a group AR genetic disorders that can lead to tissue injury from accumulation of excess iron in the body. The most common cause of HH is mutations in the hemochromatosis (*HFE*) gene, but defects have been reported in transferrin receptor 2 (*TFR2*), ferroportin and hepcidin, and hemojuvelin (*HFE2*) genes. Must be differentiated from iron storage as a secondary complication of thalassemias, anemias, transfusions, and so on
- Clinical
  - Increased absorption of iron from gastrointestinal tract resulting in iron deposits particularly in liver, heart, pancreas, and skin
  - More males than females affected
  - Onset in fourth decade for men and fifth for women
  - Treatment by routine phlebotomies
  - If untreated, organ damage possible such as
    - Cirrhosis of the liver, hepatocellular cancer, and liver failure
    - Type 2 diabetes
    - Congestive heart failure or arrhythmia
    - Arthritis
    - Hypogonadism in males
    - Skin bronzing
- Diagnosis and prevalence
  - Typically diagnosed by combination of genetic and phenotypic findings
  - Measurement of transferrin saturation (>50% for women and >60% for men) and serum ferritin levels (>200 μg/mL)
  - Liver biopsy to measure iron concentration is the gold standard
  - Mutation analysis (*see* "HFE Mutations" section)

  - Estimated that 3–5 individuals per 1,000 in general population are clinically affected
  - Penetrance in males is higher than females before menopause
- *HFE* mutations
  - Most common cause of HH
  - *HFE* gene
    - Located on chromosome 6
    - Class I *HLA* gene
    - Binds β2-microglobulin and is associated with transferrin receptor
  - p.C282Y (c.845 G > A)
    - Major disease-causing mutation found in 80–90% of affected individuals
    - 1 in 11 Europeans are p.C282Y heterozygotes
    - Arose relatively recently (in the last 1,000–3,000 years)
    - Mostly found in Northern Europeans
    - 0.44% of non–Hispanic-Caucasian individuals are p.C282Y homozygotes—most are asymptomatic (reduced penetrance)
  - p.H63D (c.287C > G)
    - Older, more common, worldwide distribution
    - 2% of Europeans p.H63D homozygotes
    - Extremely low penetrance such that homozygotes and compound heterozygotes with p.C282Y slightly increased risk of iron overload
  - DNA diagnosis
    - Targeted mutation analysis by any number of methods including PCR followed by restriction digestion (Fig. 20.2), hybridization, fluorescence resonance energy transfer (FRET), and capillary electrophoresis
    - Polymorphisms have been reported – one is in the binding site of one of the primers used in an original report, therefore must be careful when designing primers

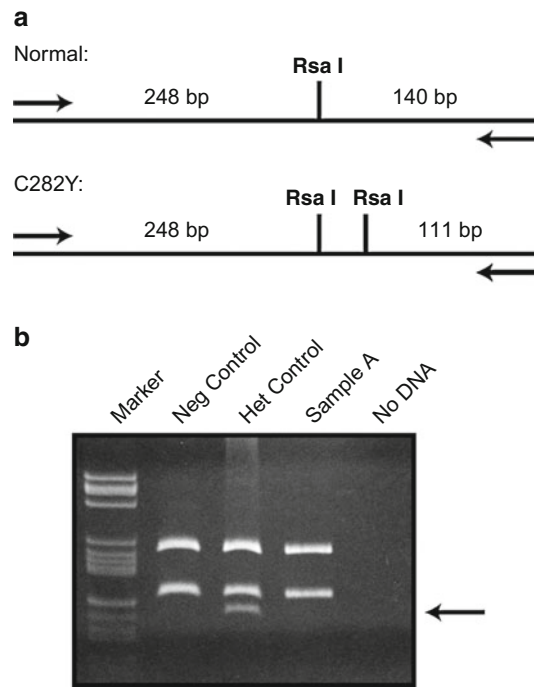


Fig. 20.2 Molecular genetic testing of HH. (**a**) The p.C282Y HFE mutation results from a c.845 G > A transition, which creates a novel *Rsa*I restriction site. PCR primers are depicted with *arrows*. (**b**) The image depicts gel electrophoresis of PCR products, which encompass the p.C282Y mutation site that is digested with *Rsa*I. In this illustration, sample A is negative for the p.C282Y mutation given the absence of the smaller *Rsa*I restriction fragment (111 bp; *arrow*). The 29-bp fragment in p.C282Y hetero- and homozygotes is typically undetectable

20.2.4 Spinal Muscular Atrophy

- Spinal muscular atrophy (SMA) is one of the most common AR diseases and is characterized by symmetric proximal muscle weakness due to the degeneration of anterior horn cells of the spinal cord.
- Clinical
  - Progressive degeneration and loss of anterior horn cells (lower motor neurons) in the spinal cord and brain stem nuclei causing symmetric muscle weakness and atrophy
  - Diagnosis is made based on poor muscle tone, symmetric muscle weakness that spares the face, and ocular muscles
  - Other signs include tongue fasciculations and absence of deep tendon reflexes
  - Normal reaction to sensory stimuli and normal intellect

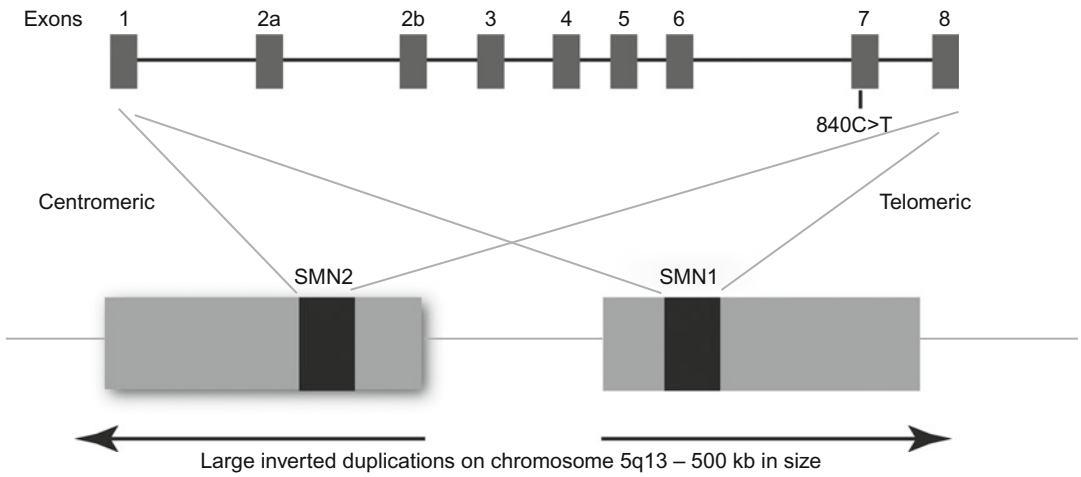


Fig. 20.3 Schematic of SMA genomic locus showing the inverted duplications and the positions of the *SMN1* and *SMN2* genes. The only coding sequence change between the two genes is shown in exon 7–840C > T *SMN1:SMN2*

  - Onset of weakness varies from before birth to adolescence and adulthood but is always progressive
    - Type I SMA (Werdnig–Hoffman disease): onset of weakness and hypotonia in first few months with fatal respiratory failure before 2 years of age (most common [60%])
    - Type II SMA: onset of weakness by 18 months and survival beyond 4 years of age
    - Types IIIa and IIIb have age of onset before and after 3 years, respectively
- Inheritance and prevalence
  - AR disorder
  - Disease frequency of 1 in 10,000
  - Carrier frequency ranges from 1 in 35 (Caucasian) to 1 in 117 (African–American) depending on ethnicity
- Survival motor neuron (*SMN*) genes
  - *SMN* genes (*SMN1* and *SMN2*) on 5q13 (Fig. 20.3)
    - Located within low copy repeat with neuronal apoptosis inhibitory protein (*NAIP*) and *p44* genes (500 Kb)
      - *SMN1* telomeric to *SMN2*
      - High sequence similarity between genes, no amino acid differences
      - Nine exons – 1, 2a, 2b, 3–8, and span 20 kb
    - Single-base differences between *SMN1* and *SMN2* in exons 7 and 8 are exploited in diagnosis
    - Single coding sequence difference in exon 7 (840C > T) important for splicing differences between *SMN1* and *SMN2* (synonymous change)
    - SMN1 encodes protein of 38 kDa, whereas *SMN2* encodes a protein that is lacking exon 7
    - Required for pre-mRNA splicing role in snRNP biogenesis and function
    - Loss of *SMN1* function is responsible for SMA phenotype
    - Mutation analysis
      - 95% of patients have homozygous loss of *SMN1* by deletion or gene conversion
      - Sequence differences of exons 7 and 8 are exploited in diagnosis (Fig. 20.3)
      - 5% have intragenic mutations
      - No genotype/phenotype correlations with *SMN1* mutations, but additional copies of *SMN2* are associated with less severe phenotype
- Diagnosis
  - Carrier testing detects the SMA deletion by dosage-sensitive techniques, such

as quantitative PCR or multiplex ligation-dependent probe amplification (MLPA)
- Can be problematic because some individuals have duplications on one chromosome and deletion of the other (silent 2 + 0 carriers)
  - 2% of mutations are de novo
  - 4–40% of individuals carry a duplication of *SMN1* (depending on ethnicity)
  - Not possible to detect the duplication carriers/deletion carriers, referred to as 2 + 0 carriers, from normal individuals using dosage-sensitive PCR techniques

20.2.5 Medium-Chain Acyl-CoA Dehydrogenase Deficiency

- Medium-chain acyl-CoA dehydrogenase deficiency (MCAD) is the most common disorder of fatty acid oxidation. During periods of fasting or prolonged aerobic exercise when glycogen stores are depleted, fatty acids become a main energy source by β-oxidation in the liver, skeletal, and cardiac muscles
- Clinical
  - Symptoms appear after periods of prolonged fasting or intercurrent infections
  - Hypoketotic hypoglycemia
  - Lethargy, seizures, coma, and death without treatment
  - Complications include hepatomegaly, acute liver disease, and brain damage
  - Disease typically presents before 2 years of age but after the newborn period
  - Variable onset in some patients – first few days of life to adults
  - Screening for MCAD included in tandem mass spectrometry of newborn screening program
- Inheritance and prevalence
  - AR disorder
  - Prevalent in individuals of Northwestern Europe
    - Highest overall frequency of 1 in 4,900 in Northern Germany
    - Incidence in the United States is estimated at 1 in 15,700
- Medium-chain acyl-CoA dehydrogenase
  - Enzyme that is intramitochondrial but is encoded by a nuclear gene
    - Normal function is the initial dehydrogenation of acyl CoAs with carbon chain lengths 4–12
    - Defect leads to the accumulation of metabolites of the medium-chain fatty acids, mainly dicarboxylic acids, acylglycine in urine, and acylcarnitine in plasma
      - Metabolites are at their highest concentration in the bloodstream in the first few days of life, so newborn period is the ideal time for detection
      - Specificity of this testing is 100% as no false-negatives have been reported
      - MCAD enzymatic activity can also be assayed in several different cell types
- Medium-chain acyl-CoA dehydrogenase gene (*ACADM*)
  - Located on chromosome 1p31 and spans 44 kb
  - Contains 12 exons
  - Encodes a protein of 421 amino acids
  - Mutations
    - Founder mutation in exon 11, 985A → G (K304E), represents 90% of all alleles in the Northern European population
    - Studies of the US population indicate that this mutation accounts for 79% of the total mutant alleles (greater ethnic diversity)
    - Additional mutations have been spread throughout the gene with no obvious mutation hot spot
      - Affect overall stability of the protein, improper folding (mostly missense mutations located away from the active center)
- Genetic testing
  - DNA testing for MCAD mutations offered as confirmatory testing after the initial diagnosis by biochemical testing

    - K304E allele is performed initially by PCR amplification followed by restriction enzyme digestion
    - Other methods that can discriminate between single-nucleotide changes, such as allele-specific oligonucleotide hybridization or ligation chain reaction amplification
    - Affected individuals heterozygous or negative for the K304E mutation have whole gene sequencing performed on all 12 exons
    - The majority of mutations identified in ACADM are thought to affect the folding of the protein and located away from the active center
- Treatment
  - With prompt postnatal diagnosis, MCAD can be treated
  - Precautions such as avoidance of fasting and saturated fats and ingestion of carbohydrates prior to bedtime can eliminate the symptoms and related complications of the disease

20.3 Autosomal Dominant Disorders

20.3.1 Nucleotide Repeat Expansion Disorders

20.3.1.1 Myotonic Dystrophy Type 1

- Myotonic dystrophy type 1 (DM1) is an AD multisystem disorder involving progressive muscle weakness, myotonia, cataracts, electrocardiogram (ECG) abnormalities, hypersomnia, and endocrine dysfunction
- Clinical
  - Categorized into overlapping phenotypes
    - Mild: cataract and mild myotonia, normal life span
    - Classic: muscle weakness, myotonia, cataract, cardiac arrhythmias, and balding; adults may have shortened life span
    - Congenital: neonatal hypotonia, motor and mental retardation, respiratory deficits, and early death
  - Exhibits disease anticipation
  - Caused by expansion of a CTG trinucleotide repeat in the 3′ UTR of the dystrophia myotonica protein kinase (*DMPK*) gene, located at 19q13.3 (Fig. 20.4 and Table 20.4)
- Prevalence
  - Estimated worldwide prevalence is 1/20,000
- Diagnosis
  - Based on a positive family history, characteristic clinical findings, and positive molecular testing
  - Molecular testing is clinically useful for diagnosis, prenatal diagnosis, and predictive testing
- Dystrophia myotonica protein kinase
  - Normal alleles are polymorphic and contain 5–35 repeats
  - Phenotypically normal individuals can have intermediate alleles that contain 35–49 repeats; can result in pathologic expansion in subsequent generations
  - Affected individuals have at least one allele with 50 or more CTG repeats, and severity generally correlates directly with repeat size
  - Mutant *DMPK* alleles display considerable somatic instability, resulting in mosaicism for the repeat in affected individuals
  - The largest alleles (>2,000 repeats) that produce the most severe form of the disorder, congenital DM1, are almost always maternally transmitted
  - DMPK protein shares homology to a cyclic AMP-dependent serine–threonine protein, kinase but its in vivo substrate remains unknown; expressed in specialized cells of the heart and skeletal muscle
  - As the CTG repeat occurs in the 3′ UTR of the DMPK transcript, allelic expansion does not alter DMPK protein structure
  - Exact mechanism as to how CTG expansion results in decreased DMPK protein production is not fully resolved; however, evidence exists implicating abnormalities in DMPK pre-mRNA processing and transport

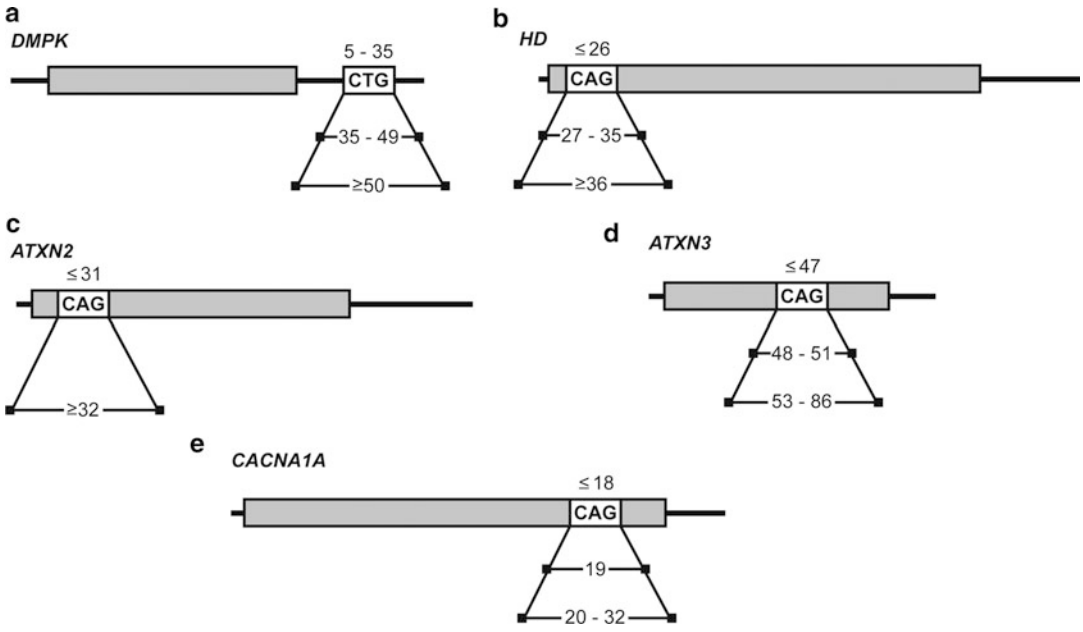


Fig. 20.4 Schematic of nucleotide expansion disorder genes. Depicted are the mRNA sequences for the genes involved in myotonic dystrophy (**a**), Huntington disease (**b**), and spinocerebellar ataxia types 2 (**c**), 3 (MJD) (**d**), and 6 (**e**). Sequences highlight the location of the trinucleotide repeats involved in expansion. Coding regions are displayed by horizontal *gray bars*, and the normal number of repeats is noted above the mRNA. Intermediate-range alleles, if they occur, and disease-causing alleles are displayed below the specific trinucleotide sequences

Table 20.4 Representative AD nucleotide repeat expansion disorders

				Number of repeats		
Disease	Nucleotide repeat	Gene	Repeat location	Normal	Unstable intermediate	Affected
DM1	CTG	*DMPK*	3′ UTR	5–35	35–49	≥50
HD	CAG	*HTT*	Coding region	≤26	27–35	≥36
SCA2	CAG	*ATXN2*	Coding region	≤31	–	≥32
SCA3/MJD	CAG	*ATXN3*	Coding region	≤47	48–51	53–86
SCA6	CAG	*CACNA1A*	Coding region	≤18	19	20–32

DM1 myotonic dystrophy type 1, *HD* Huntington disease, *SCA* spinocerebellar ataxia, *MJD* Machado–Joseph disease

- Molecular genetic testing
  - PCR amplification of the *DMPK* trinucleotide repeat region followed by gel electrophoresis
  - Alleles with >100 repeats may not be detectable by PCR
  - Individuals with a single allele size as detected by PCR should also be analyzed by Southern blot
  - Combination of PCR and Southern blot detects nearly 100% of DM1 cases

20.3.1.2 Huntington Disease

- Huntington disease (HD) is an AD neurodegenerative disorder characterized by involuntary movements, cognitive impairment, and emotional disturbance
- Clinical
  - Age of onset is typically in the third to fifth decade; however, symptoms may begin in childhood or after age 60
  - Juvenile form is defined by onset before 21 years of age

  - Early degenerative changes are marked in the striatum; advanced stage neuronal loss is widespread involving the cortex and cerebellum
  - Exhibits disease anticipation
  - Is caused by expansion of a CAG trinucleotide repeat in the coding region of the *HTT* gene, located at 4p16.3 (Fig. 20.4 and Table 20.4)
- Prevalence
  - More common in populations of Western European descent (prevalence of ~1/20,000) and less common in Asian and African populations
- Diagnosis
  - Based on a positive family history, characteristic clinical findings, and positive molecular testing
  - Molecular testing is clinically useful for diagnosis, prenatal diagnosis, and predictive testing
- *HTT* gene
  - Normal alleles have up to 26 repeats
  - Phenotypically normal individuals can have intermediate alleles that contain 27–35 repeats
  - Intermediate alleles can pathologically expand in subsequent generations, occurring almost exclusively by paternal transmission
  - Affected individuals have at least one allele with 36 or more CAG repeats
    - Full penetrance alleles: 40 or more repeats
    - Reduced penetrance alleles: 36–39 repeats
  - An inverse correlation exists between the number of CAG repeats and the age of onset; individuals with juvenile onset usually have >60 repeats
  - Wild-type HTT is widely expressed as two different-sized transcripts and is required for normal development
  - Normal function of the HTT protein is not well understood
  - Abnormal CAG expansion results in polyglutamine expansion in the HTT protein and likely confers an RNA or protein gain of function
  - Polyglutamine expansion may affect HTT function by influencing protein–protein interaction or may result in abnormal accumulation of toxic substance through the activity of transglutaminases
  - Polyglutamine expansion also increases affinity for Huntington-associated protein 1, which may play a role in aberrant neuronal cell death
- Molecular genetic testing
  - PCR amplification of the HTT trinucleotide repeat region followed by gel electrophoresis (Fig. 20.5)
  - Alleles with >100 repeats may not be detectable by PCR
  - Individuals with a single allele size as detected by PCR should also be analyzed by Southern blot
  - Combination of PCR and Southern blot detects nearly 100% of HD cases

20.3.1.3 Spinocerebellar Ataxia

- Spinocerebellar ataxia (SCA) is a group of AD neurodegenerative disorders, which share many clinical and neuropathologic features, most notably progressive cerebellar ataxia
- Characterized by poor coordination of movement and a wide-based, uncoordinated, unsteady gait; poor coordination of the limbs and speech is often present
- Categorized by causative gene or chromosomal locus
- The underlying genetic abnormalities have been identified in many subtypes and include trinucleotide repeat expansions (CAG and CTG) in coding and noncoding regions, pentanucleotide repeat expansions (ATTCT), and point mutations
- Pathogenic mutations are unidentified in 20–40% of SCA families
- More than 20 forms of SCA exist; however, molecular testing is only available for well-characterized subtypes. Causative genes and location
  - SCA1: *ATXN1*, 6p23
  - SCA2: *ATXN2*, 12q24.1
  - SCA3/Machado–Joseph disease (MJD): ATXN3, 14q24.3-q32.2

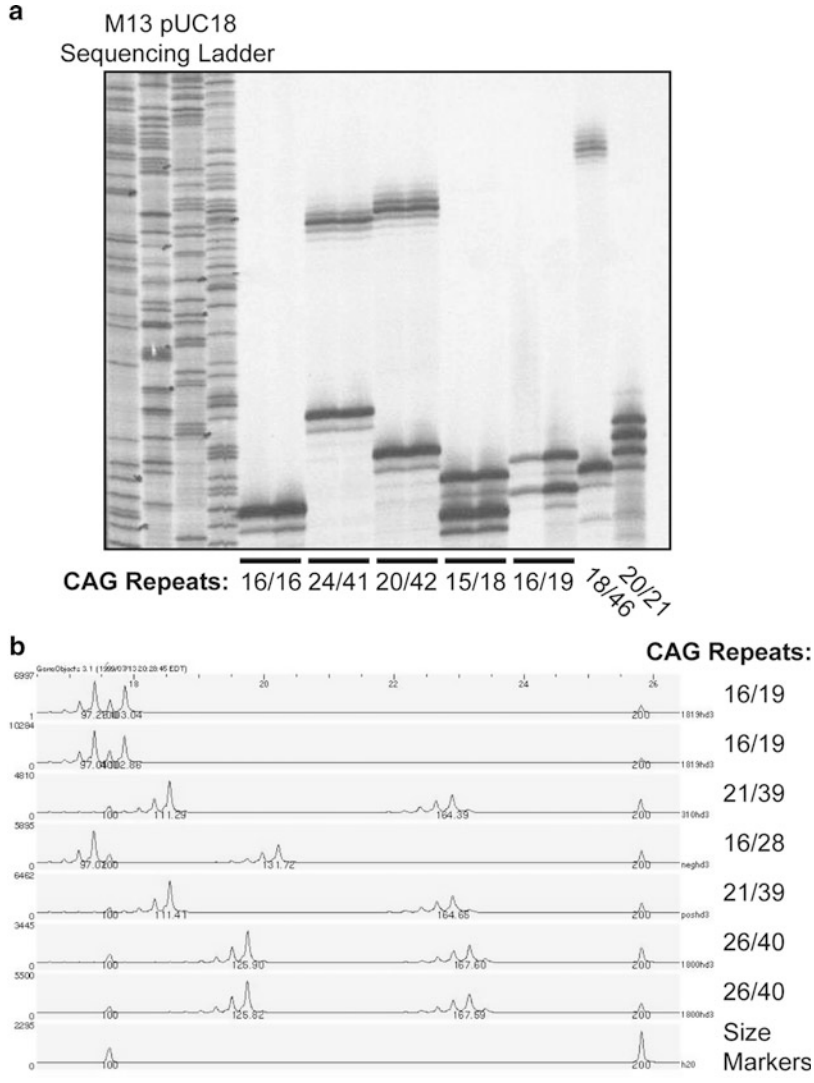


Fig. 20.5 Molecular genetic testing of Huntington disease. (**a**) An autoradiograph of radiolabeled PCR products that encompass the *HTT* CAG repeat following acrylamide gel electrophoresis. The number of individual CAG repeats is calculated by comparison to an M13 sequencing ladder and is noted below the gel. (**b**) Acrylamide electrophoresis of fluorescently labeled PCR products that encompass the *HTT* CAG repeat detected by an automated sequencer. The number of individual CAG repeats is calculated by comparison with standard size markers and is noted to the right of the graphical output (Images courtesy of M Galvez, P Scott, and D Rosenblatt, McGill University Health Centre, Division of Medical Genetics, Montreal, Canada)

  - SCA6: *CACNA1A*, 19p13.2-p13.1
  - SCA7: *ATXN7*, 3p21.1-p12
  - SCA8: *KLHL1AS*, 13q21
  - SCA10: *ATXN10*, 22q13.31
  - SCA12: *PPP2R2B*, 5q31-q33
  - SCA14: *PRKCG*, 19q13.4
  - SCA17: *TBP*, 6q27
- Majority of CAG trinucleotide repeat ataxias exhibit disease anticipation
- Molecular testing is clinically useful for diagnosis, prenatal diagnosis, and predictive testing
- Prevalence
  - Estimated worldwide prevalence is 1/100,000
  - Prevalence of individual subtypes varies by geographical area
  - SCA2, SCA3/MJD, and SCA6 are the most common forms of AD ataxia

SCA2

- Clinical
  - Age of onset is typically in the fourth decade
  - Variable findings include nystagmus, slow saccadic eye movements, and occasionally ophthalmoparesis and dementia

- Prevalence
  - Accounts for approximately 15% of all SCA
- Diagnosis
  - Difficult to distinguish from other hereditary ataxias; requires molecular genetic testing to detect an abnormal CAG trinucleotide repeat expansion in the coding region of the ataxin 2 (*ATXN2*) gene (Fig. 20.4 and Table 20.4)
- *ATXN2G* gene
  - Normal alleles have up to 31 repeats
  - Affected individuals have at least one allele with 32 or more CAG repeats
  - Most common disease-causing alleles contain 37–39 repeats
  - The CAG repeat is normally interrupted by CAA trinucleotides, which may enhance the meiotic stability of the repeat; expanded alleles that lack the CAA interruption have increased risk of disease anticipation in subsequent generations
  - CAG trinucleotide expansion results in polyglutamine expansion in the ATXN2 protein; however, the normal function of the ATXN2 protein is unknown
- Molecular genetic testing
  - PCR amplification of the *ATXN2* trinucleotide repeat region followed by gel electrophoresis
  - Alleles with >100 repeats may not be detectable by PCR
  - Individuals with a single allele size as detected by PCR should also be analyzed by Southern blot
  - Combination of PCR and Southern blot detects nearly 100% of SCA2 cases

SCA3/MJD

- Clinical
  - Age of onset varies but is typically in the second to fourth decade
  - Variable findings include pyramidal and extrapyramidal signs, nystagmus, amyotrophy fasciculations, and sensory loss
- Prevalence
  - Accounts for approximately 20% of all SCA
- Diagnosis
  - Difficult to distinguish from other hereditary ataxias; requires molecular genetic testing to detect an abnormal CAG trinucleotide repeat expansion in the coding region of the ataxin 3 (*ATXN3*) gene (Fig. 20.4 and Table 20.4)
- *ATXN3* gene
- Normal alleles have up to 47 repeats
  - Unlike SCA2, phenotypically normal individuals can have intermediate alleles that contain 48–51 repeats; can result in pathologic expansion in subsequent generations
  - Affected individuals have at least one allele with 53–86 CAG repeats
  - Increase in disease severity has been observed in individuals homozygous for expanded *ATXN3*
  - CAG trinucleotide expansion results in polyglutamine expansion in the ATXN3 protein; however, the normal function of the ATXN3 protein is unknown
- Molecular genetic testing
  - PCR amplification of the *ATXN3* trinucleotide repeat region followed by gel electrophoresis
  - Detects nearly 100% of SCA3 cases

SCA6

- Clinical
  - Characterized by adult onset, typically in the fifth to sixth decade
  - Variable findings include very slow progression, dysarthria, nystagmus, and occasionally diplopia
  - Can present with episodic ataxia
- Prevalence
  - Accounts for approximately 15% of all SCA
- Diagnosis
  - Difficult to distinguish from other hereditary ataxias; requires molecular genetic testing to detect an abnormal CAG trinucleotide repeat expansion in the coding region of the calcium channel, voltage-dependent, α-1A subunit (*CACNA1A*) gene (Fig. 20.4 and Table 20.4)
- *CACNA1A* gene
  - Has multiple transcript variants
  - Short form variants: The CAG repeat is located within the 3′ UTR and is not associated with any disease

  - Long form variant
    - Normal alleles have up to 18 repeats
    - Intermediate alleles with 19 repeats have unclear clinical significance
    - Affected individuals have at least one allele with 20–32 CAG repeats
    - Unlike many other AD ataxias, anticipation of SCA6 is not observed, as expansions of *CACNA1A* from parent to child rarely occur
  - *CACNA1A* encodes for the α-1A subunit of voltage-dependent calcium ion channels, which are involved in muscle contraction and hormone/neurotransmitter release; expressed predominantly in neuronal tissue
  - CAG trinucleotide expansion results in polyglutamine expansion in the CACNA1A protein
- Molecular genetic testing
  - PCR amplification of the *CACNA1A* trinucleotide repeat region followed by gel electrophoresis
  - As missense mutations in the *CACNA1A* gene cause disorders with phenotypic overlap to SCA6, mutation scanning of coding region is available
  - The combination of PCR and mutation scanning detects nearly 100% of SCA6 cases
  - *CACNA1A* mutation disorders (AD)
    - p.G293R causes a disorder similar to SCA6 but with a more severe clinical presentation
    - Episodic ataxia type 2
    - Familial hemiplegic migraine

20.3.1.4 Friedreich Ataxia: Autosomal Recessive

- Friedreich Ataxia (FRDA) is unique among the nucleotide expansion ataxias as it is most commonly caused by an unstable expansion of a GAA trinucleotide repeat, inherited in an AR fashion
- Is not associated with disease anticipation
- Clinical
  - Characterized by onset in the first to second decade; however, atypical cases presenting beyond 25 years of age have been observed
  - Associated with depressed tendon reflexes, dysarthria, Babinski responses, and loss of position and vibration senses
- Prevalence
  - It is the most common hereditary ataxia with an estimated prevalence of 1–2/50,000
  - Carrier frequency is 1/60–1/100
- Diagnosis
  - Requires molecular genetic testing to detect an abnormal GAA trinucleotide repeat expansion in the first intron of the frataxin *(FXN)* gene, located at 9q13-q21.1
  - >96% of FRDA individuals have *FXN* GAA expansion; approximately 4% of FRDA individuals are compound heterozygous for the GAA expansion and a deleterious *FXN* gene mutation
- *FXN* gene
  - Normal alleles have 5–33 repeats
  - Phenotypically normal individuals can have intermediate alleles that contain 34–65 repeats; can result in pathologic expansion in subsequent generations
  - Full penetrance alleles contain 66–1,700 GAA repeats
  - Other inactivating mutations of *FXN* include nonsense and missense mutations.
  - Encodes a mitochondrial protein and regulates mitochondrial iron transport and respiration
- Molecular genetic testing
  - PCR amplification of the *FXN* trinucleotide repeat region followed by gel electrophoresis
  - Alleles with >100 repeats may not be detectable by PCR
  - Individuals with a single allele size as detected by PCR should also be analyzed by Southern blot
  - Individuals who fulfill the clinical diagnostic criteria of FRDA but who are heterozygous for a pathogenic expanded allele should be tested for inactivating mutations by sequencing of the *FXN* coding region
  - Combination of PCR, Southern blot, and mutation scanning detects nearly 100% of FRDA cases

20.3.2 Skeletal and Connective Tissue Disorders

20.3.2.1 Achondroplasia

- An AD disorder characterized by short-limb dwarfism
- Clinical
  - Affected individuals exhibit short stature with disproportionate arms and legs, characteristic faces with frontal bossing and midface hypoplasia, exaggerated lumbar lordosis, limitation of elbow extension, genu varum, and trident hand
  - Intelligence and life span are usually normal; however, spinal cord and upper airway abnormalities increase risk of infant death
  - Caused by the p.G380R mutation in the fibroblast growth factor receptor 3 (*FGFR3*) gene, located at 4p16.3
  - Rare homozygous achondroplasia (ACH) has distinct radiologic findings and results in neurologic abnormalities and early death
- Prevalence
  - Is the most common form of inherited disproportionate short stature, occurring in approximately 1/26,000
- Diagnosis
  - Based on characteristic clinical and radiographic findings; molecular testing is available for atypical cases and those too young to diagnose with certainty
  - Molecular testing is clinically useful for prenatal diagnosis and confirmatory diagnostic testing
- *FGFR3* gene
  - Penetrance of mutated *FGFR3* is 100%
  - Majority of affected individuals have one of two point mutations resulting in the same amino acid substitution (p.G380R)
    - Major mutation (~98% of affected individuals): 1138 G > A
    - Minor mutation (~1% of affected individuals): 1138G > C
  - Affected individuals with p.G375C and p.G346E mutations have been reported
  - Mature FGFR3 protein is a receptor tyrosine kinase

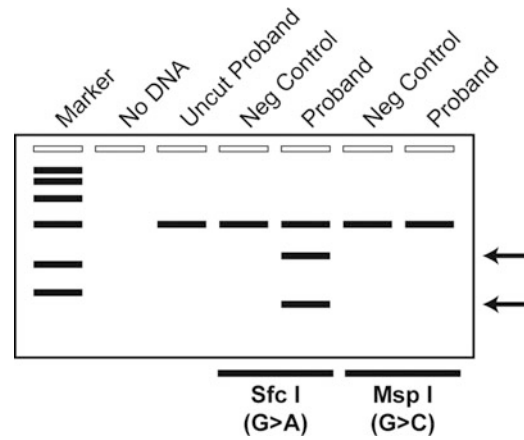


Fig. 20.6 Molecular genetic testing of ACH. The p.G380R *FGFR3* mutation results from either the 1138 G > A or the 1138 G > C nucleotide substitution, which creates *Sfc*I or *Msp*I restriction sites, respectively. The image depicts gel electrophoresis of PCR products that encompass the p.G380R mutation site and which are digested with either *Sfc*I or *Msp*I. In this illustration, the proband is heterozygous for the more common 1138 G > A mutation given the presence of the *Sfc*I restriction fragments (*arrows*)

  - FGFRs have highly conserved amino acid sequences, differing from one another in their ligand affinities and tissue distribution
  - Interacts with fibroblast growth hormone resulting in receptor dimerization, autophosphorylation, and signal transduction, ultimately modulating bone development and maintenance
  - p.G380R has been shown to result in constitutively activated *FGFR3*
  - As the *FGFR3* pathway normally exerts a negative growth control, *FGFR3* mutations result in gain of function
- Molecular genetic testing
  - Targeted p.G380R mutation analysis (Fig. 20.6) and/or DNA sequencing of select exons
  - Sequence analysis of remaining exons is recommended when the two common p.G380R mutations are not found and ACH is suspected based on clinical and radiographic findings
  - Combination of p.G380R mutation analysis and *FGFR3* exon sequencing detects nearly 100% of ACH cases

  - Other phenotypes associated with *FGFR3* mutations
    - Hypochondroplasia
    - Thanatophoric dysplasia
    - Severe ACH with developmental delay and acanthosis nigricans dysplasia
    - FGFR-related craniosynostosis

20.3.2.2 Fibroblast Growth Factor Receptor–Related Craniosynostosis Syndromes

- An AD spectrum of disorders comprised of Pfeiffer syndrome, Apert syndrome, Crouzon syndrome, Beare–Stevenson syndrome, *FGFR2*-related isolated coronal synostosis, Jackson–Weiss syndrome, Crouzon syndrome with acanthosis nigricans, and Muenke syndrome
- Clinical
  - Majority of syndromes are characterized by bicoronal craniosynostosis or cloverleaf skull, distinctive facial features, and variable hand and foot findings
  - Muenke syndrome and *FGFR2*-related isolated coronal synostosis are characterized only by uni- or bicoronal craniosynostosis
  - Abnormal skull may be detected by ultrasound prenatally or not until later infancy
  - Each syndrome has specific clinical features; however, most share common characteristics: hypertelorism, midfacial hypoplasia with proptosis, down-slanting palpebral fissures, high-arched palate, developmental delay/mental retardation, hydrocephalus, hearing loss, and visual impairment
  - Caused by mutations in the *FGFR1* (8p11.2-p11.1), *FGFR2* (10q26), and *FGFR3* (4p16.3) genes
- Prevalence
  - Overall incidence for all forms of craniosynostosis is 1/2,000–1/2,500 live births
- Diagnosis
  - Typically diagnosed based on clinical findings
  - Molecular testing of *FGFR1, FGFR2*, and *FGFR3* assists the diagnosis of suspected craniosynostosis
  - Molecular testing is available for prenatal diagnosis
- *FGFR* genes
  - A family of four tyrosine kinase receptors that nonspecifically bind FGFs – a family of signaling molecules that regulate cell proliferation, differentiation, and migration
  - *FGFR* sequence differences affect ligand-binding specificity
  - *FGFR4* is not involved in craniosynostosis syndromes
  - Normal function is to restrain limb growth
  - Mutations cause excessive activity
  - FGFR amino acids 252–253 are located within an extracellular "linker region" and are common mutation sites believed to alter ligand binding
  - *FGFR1*
    - Approximately 5% of individuals with Pfeiffer syndrome type 1 (mild form) have a p.P252R mutation
  - *FGFR2*
    - Most mutations are missense; however, deletions, insertions, and splice site mutations have been reported
    - Mutations identified in Pfeiffer, Apert, Crouzon, Beare–Stevenson, and Jackson–Weiss syndromes
    - p.S252W and p.P253R account for 71% and 26% of Apert syndrome individuals, respectively
    - Identical *FGFR2* mutations have been reported in Pfeiffer, Crouzon, and Jackson–Weiss syndromes
    - Cysteine residues 278 and 342 are common mutation sights for Pfeiffer and Crouzon syndromes
  - *FGFR3*
    - p.P252R is diagnostic for Muenke syndrome
    - p.A391E: majority of individuals with Crouzon syndrome with acanthosis nigricans
    - *FGFR3* mutations also cause ACH, hypochondroplasia, thanatophoric dysplasia, and severe ACH with developmental delay and acanthosis nigricans
- Molecular genetic testing

  - Involves initial analysis of recurrent mutations followed by selective gene sequencing
  - *FGFR1*- and *FGFR3*-targeted mutation analysis and sequencing of select exons
  - *FGFR2*-targeted mutation analysis, sequencing of select exons, and mutation scanning of entire coding region

20.3.2.3 Marfan Syndrome

- Marfan syndrome (MFS) is an AD systemic disorder characterized by ocular, skeletal, and cardiovascular abnormalities
- Clinical
  - Affected individuals show a very broad phenotypic spectrum
  - Symptoms may be present at birth or appear in childhood or adulthood
  - The four major diagnostic findings include dilation or dissection of the aorta at the level of the sinuses of Valsalva, ectopia lentis, dural ectasia, and four of eight typical skeletal features
  - Typical skeletal features: bone overgrowth, joint laxity, long extremities, pectus excavatum or carinatum, scoliosis, high-arched palate, positive wrist and thumb signs, reduced upper to lower segment, arm span to height ratio >1.05, and flat feet
  - Primarily caused by a mutation in the fibrillin 1 (*FBN1*) gene, located at 15q21.1
  - Approximately 75% of affected individuals have an affected parent; the remaining have a de novo *FBN1* mutation
- Prevalence
  - One of the most common connective tissue disorders, occurring in $\sim$1–2/10,000
- Diagnosis
  - Diagnosis of MFS is based on family history and characteristic clinical findings in multiple organ systems
- Requires "major" manifestations in at least two body systems, with "minor" involvement of a third body system; if a positive family history is established, diagnosis requires "major" manifestation in one body system with "minor" involvement of a second
- Molecular testing is clinically useful for prenatal diagnosis, predictive testing, and confirmatory diagnostic testing
- *FBN1* gene
  - Penetrance of mutated *FBN1* is 100% with variable expressivity
  - >500 *FBN1* mutations have been reported in MFS individuals
  - No definitive genotype–phenotype correlations have been observed; however, mutations associated with severe and rapidly progressive MFS intermittently cluster between exons 24 and 32
  - Is found in elastic and nonelastic connective tissues of the body
  - Wild-type FBN1 protein is an important component of extracellular microfibrils
  - Participates in the formation and homeostasis of elastic matrix and matrix cell attachments
  - Mutant *FBN1* is believed to be dominant negative as MFS individuals typically have reduced FBN1 protein expression below that expected from the remaining wild-type allele
- Molecular genetic testing
  - Mutation scanning of the *FBN1* gene is available
  - Given the large size of *FBN1*, cDNA sequencing is more efficient than genomic DNA sequencing
  - If a specific mutation is known within a family, targeted mutation analysis by bidirectional DNA sequencing is recommended
  - Linkage analysis may be used to determine if an individual has inherited an *FBN1* allele associated with MFS in multiple family members
    - Markers are highly informative and are within the *FBN1* gene
    - Not independently conclusive, as locus heterogeneity has not been definitively excluded in MFS
  - Mutations are detected in 70–93% of MFS individuals
  - Other phenotypes associated with *FBN1* mutations

- Mitral valve prolapse, aortic root diameter at upper limits, stretch marks of the skin, skeletal conditions similar to MFS (MASS) phenotype
- Mitral valve prolapse syndrome
- Familial ectopia lentis
- Shprintzen–Goldberg syndrome

20.3.2.4 Osteogenesis Imperfecta

- A primarily AD group of bone formation disorders characterized by low bone mass and propensity to fracture
- Clinical
  - Affected individuals may also exhibit blue sclera, dentinogenesis imperfecta, skin hyperlaxity, joint hypermobility, and hearing loss
  - Fractures are most common in extremities but can occur in any bone
  - Is a broad clinical entity but is artificially classified into seven types (I–VII) based on clinical presentation, radiographic findings, mode of inheritance, and molecular genetics
  - Severity
    - Type I: mild
    - Type II: perinatal lethal
    - Type III: severe
    - Type IV: moderate to mild
    - Types V–VII: moderate
  - In the majority of individuals is caused by a mutation in the collagen, type I, α-1 (*COL1A1*) or *COL1A2* genes, located at 17q21.33 and 7q22.1, respectively
  - *COL1A1* and *COL1A2* mutations are found in types I–IV; loci for types V–VII have not been accurately mapped
- Prevalence
  - Overall incidence for all forms of osteogenesis imperfecta (OI) is 6–7/100,000; types I and IV account for over half of all OI
- Diagnosis
  - Based on characteristic clinical and radiographic findings, family history, and biochemical and molecular testing
  - Molecular testing is clinically useful for prenatal diagnosis and confirmatory diagnostic testing
- *COL1A1* and *COL1A2 genes*
  - >200 structural mutations have been identified
  - Encode type I procollagen chains containing repeating sequences of uninterrupted Gly-X-Y that are essential for proper chain folding
  - Two pro-α-1(I) chains and one pro-α-2(I) chain form a triple helix from carboxy to amino terminus
  - Procollagen is secreted and terminal peptides removed forming type I collagen molecules, which are then assembled into collagen fibrils
  - Major protein in bone, connective tissue, and cartilage
  - OI type I is associated with heterozygous truncating *COL1A1* mutations resulting in haploinsufficiency
  - OI types II–IV are associated with heterozygous triple helix domain *COL1A1* and *COL1A2* mutations, resulting in glycine substitution and abnormal procollagen folding
  - Biochemical analysis is typically performed in vitro on cultured dermal fibroblasts by assaying the structure and quantity of synthesized type I collagen
- Molecular genetic testing
  - Mutation scanning of the *COL1A1* and *COL1A2* genes is available using both genomic DNA and cDNA for sequence analysis
  - If a specific mutation is known within a family, targeted mutation analysis by bidirectional DNA sequencing is recommended
  - Combination of targeted mutation and *COL1A1* and *COL1A2* coding region sequence analysis detects nearly 100% of OI types I and II and approximately 60–80% types III and IV
  - Other phenotypes associated with *COL1A1* and *COL1A2* mutations
    - Ehlers–Danlos syndrome (classic and arthrochalasia types)
    - Osteoporosis
    - Arterial dissection

20.3.2.5 Ehlers–Danlos Syndromes

- Ehlers–Danlos syndromes (EDS) are a group of connective tissue disorders characterized by joint hypermobility, skin hyperextensibility, and abnormal wound healing
- Clinical
  - Affected individuals may also exhibit tissue fragility, widened atrophic scars, smooth velvety skin, molluscoid pseudotumors, subcutaneous spheroids, joint sprains/dislocations/subluxations, hypotonia, easy bruising, hernia, chronic pain, and aortic root dilation
  - Elastic or fragile connective tissues due to an underlying defect of collagen structure
  - Classified into different types based on clinical presentation, mode of inheritance, and molecular genetics; the most common are listed below
    - Types I and II: classic type (AD)
    - Type III: hypermobility type, less severe (AD)
    - Type IV: vascular type (AD)
    - Type VI: kyphoscoliotic type (AR)
  - Types I and II are caused by mutations in the collagen, type 5, α-1 (*COL5A1*) or *COL5A2* genes, located at 9q34.3 and 2q31, respectively
  - *TNXB, COL3A1*, and *PLOD1* mutations are found in types III, IV and VI, respectively; however, many patients have an unknown molecular defect
- Prevalence
  - Overall incidence for all forms of EDS is estimated at ∼1/10,000
  - Hypermobility type III EDS is most common (although often not ascertained), followed by types I and II (1/20,000); other subtypes are very rare
- Diagnosis
  - Based on characteristic clinical findings, family history, and biochemical and molecular testing
  - Molecular testing is clinically useful for confirmatory diagnostic testing
- *EDS* genes
  - COL5A1 and COL5A2 form heterotrimers in skin, bone, and tendon; form heterotypic fibrils with type I collagen
  - Types I and II (classic) EDS are associated with heterozygous *COL5A1* and *COL5A2* mutations
    - Truncating mutations lead to haploinsufficiency
    - Structural mutations exert a dominant negative effect
  - TNXB is an extracellular matrix glycoprotein of uncertain specific function produced primarily by dermal and skeletal muscle fibroblasts
  - Type III (hypermobility) EDS is associated with heterozygous *TNXB* mutations in a small subset of patients
    - Truncating mutations lead to haploinsufficiency
    - Most type III EDS patients have unknown molecular defect
  - COL3A1 forms a homotrimer that is a major structural component of skin, blood vessels, and hollow organs
  - Type IV (vascular) EDS is associated with heterozygous *COL3A1* mutations
    - Majority of mutations disrupt the Gly-X-Y repeat of the triple helical region
    - Mutations typically lead to intracellular storage and impaired secretion of collagen chains
- Molecular genetic testing
  - Classic types I and II: Mutation scanning of the *COL5A1* and *COL5A2* genes is available using both genomic DNA and cDNA for sequence analysis
    - 50% of classic EDS have an identifiable *COL5A1* or *COL5A2* mutation
  - Hypermobility type III: Mutation scanning of *TNXB* is available
  - Vascular type IV: Mutation scanning by sequencing and deletion/duplication analysis of the *COL3A1* gene is available using genomic DNA
    - ∼97% of vascular type IV EDS have an identifiable *COL3A1* mutation
  - Other disorders with similar phenotypes
    - Marfan syndrome
    - Occipital horn syndrome
    - Loeys–Dietz syndrome
    - Familial aortic aneurysm

20.4 X-Linked Inheritance

20.4.1 Fragile X Syndrome

- Fragile X syndrome (FGLX) is the leading cause of X-linked mental retardation among males. The disorder was named for the cytogenetically visible fragile site (FRAXA) at band Xq27.3 that in some cases was heritable
- Clinical
  - Mental impairment, ranging from learning disabilities to mental retardation with delay of milestones in infancy
  - Attention deficit and hyperactivity
  - Anxiety and unstable mood
  - Autistic-like behaviors
  - Large head, long face, large ears, and flat feet
  - Macro-orchidism
  - Hyperextensible joints, especially fingers
  - Seizures (epilepsy) affect about 25% of people with fragile X
  - Sex-specific differences
    - Boys are typically more severely affected than girls
    - While most boys have mental retardation, only one-third to one-half of girls have significant intellectual impairment; the rest have either normal IQ or learning disabilities
- Prevalence
  - Prevalence of 1 in 4,000 affected males with 1/2 as many affected females
  - Prevalence of female premutation carriers was estimated at 1 in 259 in one study (possibly higher in specific populations)
- Inheritance
  - Premutation carrier females, but not males, are at risk for transmitting full mutation alleles to both male and female offspring
  - Many families transmit premutation fragile X mental retardation 1 (*FMR1*) alleles for generations with little or no presentation of clinical symptoms until a full mutation is produced, resulting in an affected individual
- *FMR1* gene
  - The *FMR1* gene is located at Xq27.3, contains 17 exons, and spans 38 kb
  - *FMR1* encodes an mRNA-binding protein of 632 amino acids – fragile X mental retardation protein (FMRP)
  - FMRP is thought to shuttle select mRNAs between the cytosol and nucleus and play a role in synaptic maturation and function
  - *FMR1* is highly expressed in the brain, testes, ovaries, esophageal epithelium, thymus, eye, and spleen
  - Mechanism of expansion (>99% of cases)
    - Expansion of CGG repeat located in 5′ UTR region of *FMR1* gene (Fig. 20.7)
    - Expansion of CGG leads to methylation of promoter CpG of *FMR1* gene, silencing the gene and resulting in lack of protein product – FMRP
    - In full mutation females, methylation of *FMR1* full mutation is independent of X-inactivation
  - Characterization of repeat size
    - Normal alleles: 6–44 repeats – with 29, 30, and 31 repeats most common
    - Normal "gray zone" alleles: 45–54 repeats – these alleles can exhibit instability, but have never been observed to expand to full mutation
    - Premutation alleles: 55–200 repeats – these alleles exhibit instability and are at risk for expansion to full mutation
    - 58 repeats is the smallest premutation allele observed to expand to a full mutation. The ACMG recommends using 55 repeats as smallest premutation to account for interlaboratory differences in size standards
      - Premutation carriers are at risk for additional adult-onset disorders
        - Premature ovarian failure in 20% of premature carriers
        - Fragile X tremor ataxia syndrome—higher penetrance in males
    - Full mutation alleles: >200–several thousand repeats – these alleles cause FGLX

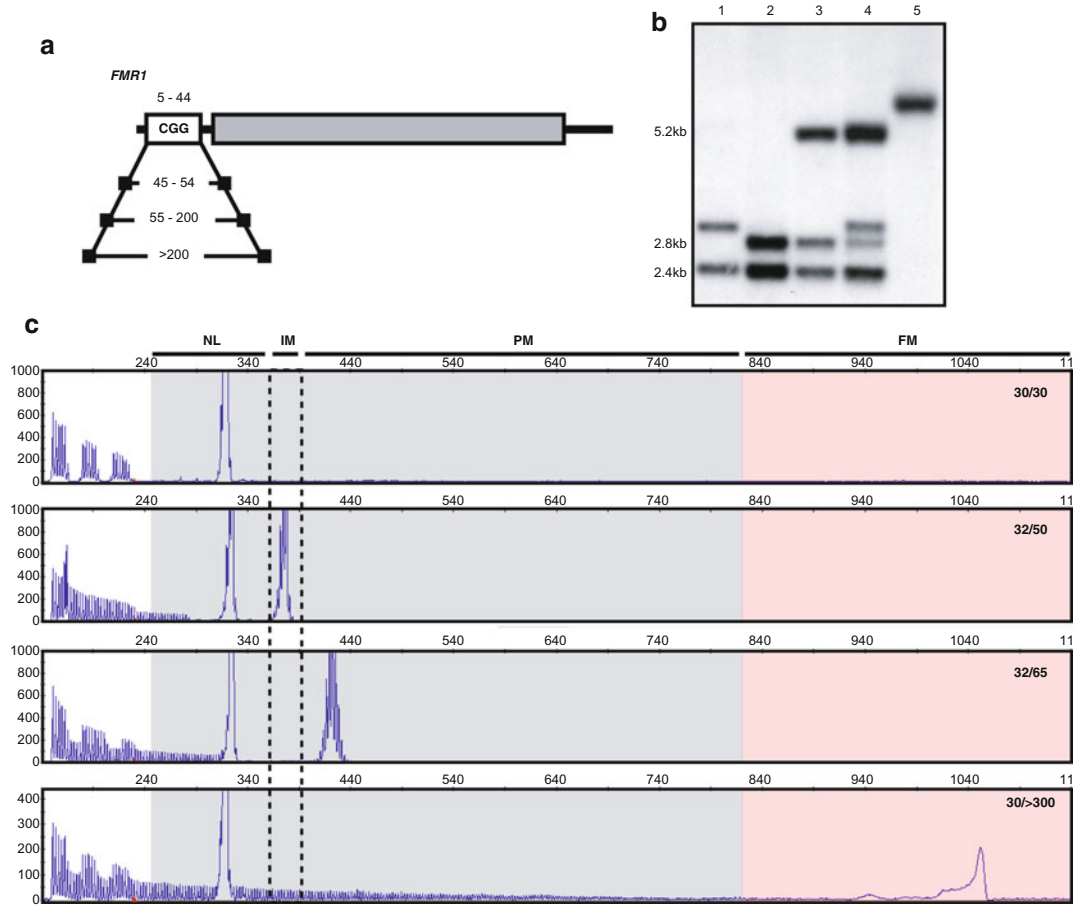


Fig. 20.7 (**a**) Schematic of *FMR1* gene trinucleotide repeat in the 5′ UTR. Ranges for *normal*, *gray zone*, premutation, and full mutations are shown. Fifty four repeats is the ACMG cutoff for gray zone alleles, and 55 repeats is now considered a premutation (*see* text). (**b**) Example of a fragile X Southern blot with genomic DNA that was digested with both *Eco*RI and *Xho*I (methylation sensitive). The 5.2-kb band represents the methylated alleles, unable to digest with *Xho*I, whereas the 2.8-kb and 2.4-kb bands represent unmethylated alleles that were cut with *Xho*I. (*Lane 1*) premutation male, (*lane 2*) normal male, (*lane 3*) normal female, (*lane 4*) premutation female, and (*lane 5*) full mutation male. (**c**) Example of electrophoresis of fluorescently labeled PCR products (gene-specific (GS) *FMR1* and CGG repeat primed (RP) *FMR1*) on an ABI genetic analyzer. From top to bottom, normal (*NL*), intermediate (*IM*), premutation (*PM*), and full mutation (*FM*) alleles were detected, respectively, in the samples. CGG repeat numbers of each sample were provided on the *right side* of each lane

    - Mosaicism for full mutation can complicate the analysis and may not be detectable if low level
- Diagnosis
  - Cytogenetic analysis of the fragile site (FRAXA) is not an acceptable method of diagnosis
  - PCR of CGG repeat is performed to determine allele sizes with use of appropriate control samples or size standards
    - PCR by traditional methods cannot distinguish a homozygous female from one with a nonamplifiable second allele and

can only detect alleles with less than 100–150 repeats
    - PCR by traditional methods is not adequate for the detection of mosaic individuals with both premutations and full mutations
- CGG repeat primed PCR – one primer pair flanks the repeat region and a third CGG repeat primer allows for hybridization of multiple targets producing a ladder of products (Fig. 20.7c)
    - Premutations and full mutations can be detected
    - Southern blot may still necessary for sizing of full mutations and accessing methylation status
- Southern blot used with double enzyme digestion with a second enzyme that is internal to the first and is methylation sensitive (Fig. 20.7b)
    - Females will have an undigested methylated (inactive X or full mutation) allele and digested unmethylated allele (active X)
    - Males will have digested allele only unless full mutation is present
    - Southern blot can detect premutation/full mutation mosaics
    - Prenatal analysis of chorionic villus sampling can be problematic
        - Methylation is absent or incomplete in this tissue at time of procedure
        - Full mutations, which are unstable and may have mosaic repeat size, are difficult to interpret on Southern blot
        - Followup amniocentesis may be necessary

20.4.2 X-Linked Muscular Dystrophy (Duchenne Muscular Dystrophy and Becker Muscular Dystrophy)

- Dystrophinopathies are a spectrum of X-linked muscle diseases that include Duchenne muscular dystrophy (DMD), Becker muscular dystrophy (BMD), and DMD-related dilated cardiomyopathy (DCM)
- Clinical
    - Characterized by a spectrum of muscle disease that ranges from mild to severe
    - DMD is most severe and rapidly progressive and presents in early childhood
        - Presents with delayed milestones (18 months–8 years)
        - Progressive muscle weakness that is symmetrical (proximal > distal)
        - Generalized motor delay, delay in sitting, standing, and walking
        - Gait problems including flat-footedness
        - Wheelchair bound by 13 years old
        - Cardiac involvement in 90% of patients
        - Approximately 80% of females show no signs or symptoms, but when present usually milder than males, later-onset muscle weakness and cramps
    - BMD is characterized by later onset of skeletal muscle weakness
        - Progressive muscle weakness (proximal > distal)
        - If wheelchair bound, after 18 years old
        - Activity-induced cramping
        - Cardiac involvement in 90% of patients
        - Approximately 80% of females show no signs or symptoms, but when present usually milder than males, later-onset muscle weakness and cramps, dilated cardiomyopathy in some
    - DCM shows no evidence of skeletal muscle disease
        - Dilated cardiomyopathy with congestive heart failure
        - Males present at ages 20–40 with rapid progression to death
        - Females present later in life with slower disease progression
- Prevalence and inheritance
    - 1 in 5,600 live male births
    - 2/3 are inherited by mother and 1/3 are new mutations
    - 75% of female carriers have no signs or symptoms

  - Risk for female carrier to have an affected male child is 25% with each pregnancy
  - Germline mosaicism is present in approximately 15% of carrier mothers, complicating risks to siblings
- *DMD* gene
  - The *DMD* gene spans 2.4 Mb of DNA on Xp21.2 and contains 79 exons (largest human gene known)
  - Encodes a membrane-associated protein, dystrophin, present in muscle cells and some neurons
    - Part of a complex that links the cytoskeleton with cell membrane and bridges the cytoskeleton with the extracellular matrix
  - Full-length protein is 427 kDa, but many different isoforms identified
  - Mutations
    - 6–10% of males with DMD or BMD have duplication of one or more exons
    - 25% of males with DMD and 5–10% of males with BMD have small insertions/deletions, point mutations, or splicing mutations
    - Mutations that obliterate or severely disrupt dystrophin function cause DMD, whereas mutations that affect the quantity of dystrophin or truncate the protein in frame result in BMD
    - DCM results from mutations that affect the expression or function of dystrophin in cardiac muscle (exon 1 and muscle-specific promoter mutations)
- Diagnosis
  - Serum creatine phosphokinase levels are elevated to >10X in all DMD males and >5X in all BMD males
  - Creatine phosphokinase levels are also elevated from 2X to 10X in approximately 50% of female carriers
  - Muscle biopsy with immunohistochemistry for dystrophin is informative in males and some females
  - Molecular analysis of DMD mutation status
    - Mutations found in 100% of DMD patients and 85% of BMD patients
    - Multiplex PCR, Southern blotting, and MLPA are used to detect deletions
    - Quantitative PCR and MLPA can be used to detect duplications
    - Mutation scanning methods, such as DHPLC and sequencing, are used to screen for small insertions/deletions and point mutations

20.4.3 X-linked Adrenoleukodystrophy

- X-linked adrenoleukodystrophy (X-ALD) is the most common of the peroxisomal disorders. It is a severe, often fatal, disease that manifests in a progressive demyelination of the central nervous system, dysfunction of the adrenal cortex, and testicular dysfunction in hemizygous males
- Clinical
  - Most common form has early onset that appears at 4–8 years of age resulting in a progressive irreversible dementia and death
  - Less severe presentations include adrenomyeloneuropathy with a later age of onset, adrenal insufficiency, and neurologic complications limited to spinal cord and peripheral nerves
- Prevalence
  - Incidence of all variant forms is 1 in 15,000 – most common genetic determinant of peroxisomal disease
- Inheritance
  - X-linked with males affected and up to 20% of carrier females with late-onset neurologic symptoms similar to adrenomyeloneuropathy
  - >93% of X-ALD patients inherit mutations from their mothers with remaining 7% carrying de novo mutations
- *X-ALD* gene
  - ATP-binding cassette, subfamily D, member1 (*ABCD1*) is located on Xq28, contains 10 exons, and spans 21 kb

  - Encodes a protein of 745 amino acids—ALD protein (ALDP)
  - Mutation analysis
    - Whole gene sequencing of exons and exon/intron boundaries performed as well as other mutation scanning techniques, such as DHPLC
    - Potential complications with PCR amplification because of paralogous gene segments exons 7–10 on chromosomes 2p11, 10p11, 16p11, and 22q11
    - Over 250 different lesions have been found in the *ABCD1* gene
    - Mutations in all 10 exons have been reported
    - Vast majority are point mutations (58.4%), although frameshifts and nonsense, and exon deletions have also been identified (http://www.x-ald.nl/)
    - Two-base pair AG deletion in exon 5 found in 10.3% of families with X-ALD (most common mutation identified)
    - No genotype–phenotype correlations are apparent, and wide phenotypic variation has been reported within families
    - 70% of missense mutations result in absent or reduced ALDP, indicating that most mutations in ABCD1 result in complete loss of protein function
- Diagnosis
  - Primary biochemical defect – impaired peroxisomal β-oxidation with accumulation of very long chain fatty acids – mostly C26 in plasma and tissues
  - In hemizygous males (99%) and 85% of carrier females, plasma concentration of very long chain fatty acids are elevated—used as a diagnostic marker for the disease

20.4.4 Rett Syndrome

- Rett syndrome (RTT) is an X-linked neurodevelopmental disorder involving regression of language and motor skills
- Clinical
  - Classic RTT—normal development during first 6–18 months of life, followed by a short period of developmental stagnation and rapid language and motor skill regression
  - Clinical features include repetitive, stereotypic hand movement, screaming and inconsolable crying, and autistic features
  - >99% are sporadic cases without family history
  - Caused by mutations in *MECP2* gene located at Xq28
  - Atypical RTT with confirmed *MECP2* mutations can be either milder or more severe than classic RTT
  - Primarily affects females
  - Males with *MECP2* mutations that would normally cause typical RTT in females are much more severely affected and usually do not survive infancy
- Prevalence
  - Estimated prevalence is 1/10,000 females
- Diagnosis
  - Clinical diagnosis uses three types of criteria: main, supportive, and exclusion
  - The presence of any of the exclusion criteria eliminates a diagnosis of classic RTT
  - Molecular testing is clinically useful for diagnosis, prenatal diagnosis, and predictive testing
- *MECP2* gene
  - Penetrance of mutated *MECP2* is almost 100%
  - Females with *MECP2* mutations may exhibit clinical variability due to highly skewed X-chromosome inactivation
  - >600 *MECP2* mutations have been reported in RTT individuals, including missense, nonsense, frameshift mutations, and large deletions
  - A range of *MECP2* mutations is also associated with milder forms of learning disabilities and autism
  - Genotype–phenotype correlations have not been consistently reported
  - >95% RTT *MECP2* mutations are de novo mutations
  - *MECP2* encodes MeCP2 protein (methyl CpG-binding protein 2), which contains methyl CpG-binding domain (MBD),

nuclear localization domain (NLS), transcription repression domain (TRD), and C-terminal domain (CTD)
  - A common mutation has not been reported in any population; however, clusters of mutations have been identified
- Molecular genetic testing
  - Denaturing high-performance liquid chromatography (DHPLC) and direct sequencing of all four exons identify ~80% of typical RTT patients and ~40% of atypical RTT patients
  - Deletion/duplication analysis identifies 8% of typical RTT patients and 3% of atypical RTT
  - Other phenotypes associated with *MECP2* mutations
    - Females: learning disability
    - Males: severe neonatal encephalopathy, manic-depressive psychosis, pyramidal signs, and parkinsonian and macro-orchidism (PPM-X) syndrome

20.5 Mitochondrial Disorders

- A clinically heterogeneous group of disorders that arise from mitochondrial respiratory chain dysfunction. Caused by mutations of mitochondrial (mtDNA) or nuclear DNA (nDNA) Clinical symptoms are first seen in tissues with high energy demands or low thresholds for energy deficiency; central nervous system and muscles often involved
- Prevalence
  - Including both mtDNA and nDNA mutations in children and adults, prevalence is approximately 1/5,000

20.5.1 The Mitochondrion

- An essential cytoplasmic organelle present in all eukaryotic cells that provides majority of cell energy
- Typical human cells have several hundred mitochondria; 1,000–2,000 in a single liver cell
- Energy-generating apparatus is the oxidative phosphorylation pathway (OXPHOS), composed of electron transport chain and ATPase (Fig. 20.8)
  - Located in the inner membrane and employs five multipolypeptide enzyme complexes and two electron carriers
  - Main function is coordinated transport of electrons and protons and production of ATP
  - Majority of OXPHOS proteins are nDNA encoded and imported from cytosol
- Has its own 16.5-kb double-stranded circular genome that contains two rRNA genes, 22 tRNA genes, and 13 structural genes, which encode OXPHOS subunits (for illustration, *see* http://www.mitomap.org/)
- mtDNA genome
  - Both strands are transcribed from a single promoter
  - Does not contain introns
  - Some differences in genetic code between mtDNA and nDNA
  - 10–17 times faster mutation rate than nDNA
  - Is maternally inherited
  - Each cell has 10^3–10^4 mtDNA molecules
- *Homoplasmy*—the state in which all mtDNA molecules are identical
- *Heteroplasmy*—the presence of more than one type of mtDNA molecule within a cell
- New mtDNA mutations are multiplied by replication and randomly divided into daughter mitochondria during cell division; leads to differences in level of heteroplasmy between tissues
- Penetrance of pathogenic mutation is increased with the degree of mutant heteroplasmy
- Tissue phenotype is normal until threshold level of mutant heteroplasmy is exceeded
- Organs with greatest ATP requirements are most sensitive to mtDNA mutations
- Mitochondrial disorders grouped into two major categories, those due to defects of mtDNA and those due to defects in nDNA
- All inheritance models are possible in connection with mitochondrial disorders
- Mitochondrial dysfunction is also observed in late-onset neurodegenerative disorders and aging

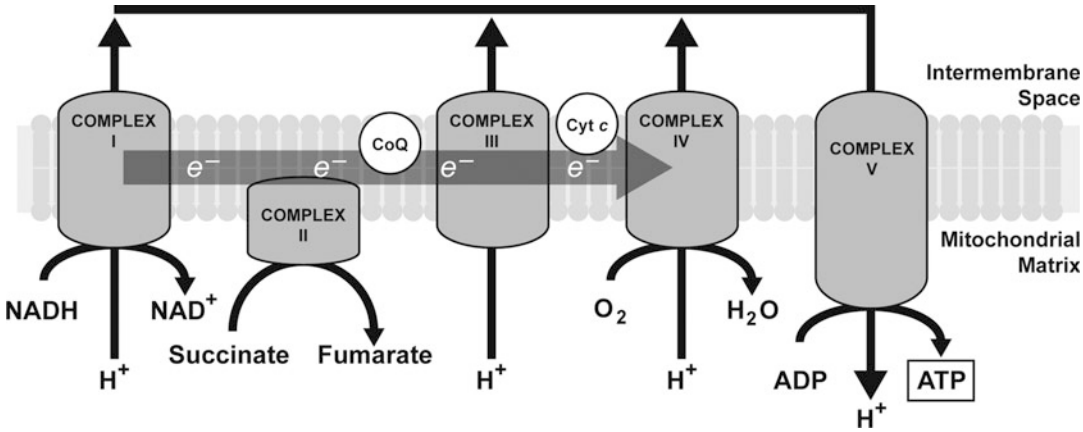


Fig. 20.8 The respiratory chain system in the inner membrane of the mammalian mitochondria. Electrons (e^-) are transferred from complexes I and II to coenzyme Q (*CoQ*). There they are transferred via complex III and cytochrome c (*Cyt c*) to complex IV, where oxygen is reduced to water. The movement of H^+ from the matrix to the intermembrane space is coupled with energy release from the electrons. The proton gradient is used for the production of ATP by complex V. Mutations in nDNA – which encodes for OXPHOS subunits or proteins involved in respiratory chain homeostasis – and in mtDNA lead to mitochondrial disease

- Majority of patients with enzymatically verified mitochondrial deficiency have an unidentified mutation

20.5.2 Defects of mtDNA

- Phenotypes of diseases vary between mtDNA mutations and between individuals with same mutation
- Probability of disease increases with age and diseases are often progressive

20.5.2.1 mtDNA Rearrangements

- Deletions vary in size and location, but a 5-kb common deletion has been observed in some sporadic disorders
- *Pearson syndrome (PS)*—typically early onset; sideroblastic anemia and exocrine pancreatic failure; often fatal
- *Kearns–Sayre syndrome (KSS)*—onset <20 years of age; progressive external ophthalmoplegia (PEO), pigmentary retinopathy, plus one of heart block, elevated cerebrospinal fluid (CSF) protein, or cerebellar ataxia
- *Chronic PEO*—onset >20 years of age; similar to KSS, bilateral ptosis
- Duplications of mtDNA have been observed in patients with KSS and diabetes mellitus with deafness
- Deletions/duplications usually encompass several essential coding and/or tRNA genes, which impairs mitochondrial protein synthesis

20.5.2.2 mtDNA Point Mutations

- Most are transition mutations occurring in tRNA/rRNA genes or respiratory chain subunit genes
- Are maternally inherited
- tRNA mutation disorders
  - *Mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS)*—characterized by recurrent vomiting, headache, and stroke-like episodes causing cortical blindness, hemiparesis, or hemianopia
    - Usually present in children or young adults after normal early development
    - Most common mutation: A3243G in $tRNA^{Leu(UUR)}$
    - Other less common mutations have been reported
  - *Myoclonus epilepsy with ragged red fibers (MERRF)*—characterized by myoclonus,

seizures, mitochondrial myopathy, and cerebellar ataxia
  - Less common signs include hearing loss and dementia
  - Most common mutations: A8344G, T8356C, and G8363A in $tRNA^{Lys}$
- *Nonsyndromic sensorineural deafness*—most common mutation: A7445G in $tRNA^{Ser(UCN)}$

- rRNA mutation disorders
  - *Aminoglycoside-induced nonsyndromic deafness*—most common mutation: A1555G in mitochondrial 12S rRNA gene
- Protein-encoding mutation disorders
  - *Neuropathy, ataxia, and retinitis pigmentosa (NARP)*—characterized by late childhood or adult onset, ataxia, pigmentary retinopathy, and dementia
    - Less common signs include sensorimotor neuropathy
    - Most common mutation: T8993G/C (heteroplasmy ~70%)
  - *Maternally inherited Leigh syndrome (MILS)*—characterized by early onset, devastating encephalopathy, hypotonia, cerebellar, and brain stem signs
    - Less common signs include ophthalmoplegia and respiratory depression
    - Most common mutation: T8993G/C (heteroplasmy ~90%)
  - *Leber hereditary optic neuropathy (LHON)*—characterized by young adult visual loss due to bilateral optic atrophy with a bias toward males
    - Less common signs include cardiac dysrhythmia and dystonia
    - Most common mutation: G3460A, G11778A, T14484C, and A14495G
    - Other less common mutations have been reported

20.5.3 Defects of Nuclear DNA

- Most of the respiratory chain subunits are encoded by the nDNA
- Structure and function of the respiratory chain require many steps, which are largely encoded by the nDNA
- nDNA mutation disorders can follow AR, AD, and X-linked recessive (XLR)-inheritance patterns

20.5.3.1 Structural Respiratory Chain Defects

- Complex I deficiency
  - *Leigh and Leigh-like syndrome* – (AR) – *NDUFS4* (5q11.1), *NDUFS7* (19p13.3), and *NDUFS8* (11q13)
  - *Hypertrophic cardiomyopathy and encephalomyopathy* – (AR) – *NDUFS2* (1q23)
  - *Macrocephaly, leukodystrophy, and myoclonic epilepsy* – (AR) – *NDUFV1* (11q13)
- Complex II deficiency
  - *Leigh and Leigh-like syndrome* – (AR) – *SDHA* (5p15)

20.5.3.2 Nonstructural Respiratory Chain Defects

- Intergenomic communication defects
  - *Mitochondrial neurogastrointestinal encephalomyopathy*–(AR)–*TP* (22q13.33)
  - *Dominant PEO*–(AD)–*ANT1* (4q35)
- Complex I assembly defects
  - *Early-onset progressive encephalopathy* – (AR) – *B17.2 L* (5q12.1)
- Complex III assembly defects
  - *Metabolic acidosis, tubulopathy, encephalopathy, and liver failure* – (AR) – *BCS1L* (2q33)
- Complex IV assembly defects
  - *Leigh syndrome* – (AR) – *SURF1* (9q34.2)
  - *Cardioencephalomyopathy* – (AR) – *SCO2* (22q13.33)
  - *Neonatal-onset hepatic failure and encephalopathy* – (AR) – *SCO1* (17p12–p13)
  - *Leigh and de Toni–Fanconi–Debre syndrome* – (AR) – *COX10* (17p12–17p11.2)
  - *Early-onset hypertrophic cardiomyopathy* – (AR) – *COX15* (10q24)
  - *French–Canadian Leigh syndrome* – (AR) – *LRPPRC* (2p21)

- Complex V assembly defects
  - *Early-onset encephalopathy, lactic acidosis* – (AR) – *ATPAF2* (17p11.2)
- Homeostasis and import
  - *Friedreich ataxia* – (AR) – *FXN* (9q13–q21.1)
  - *Hereditary spastic paraplegia* – (AR) – *SPG7* (16q24.3)
  - *mtDNA depletion myopathy* – (AR) – *TK2* (16q22–q23.1)
  - *Hepatocerebral mtDNA depletion* – (AR) – *DGUOK* (2p13)
  - *Wilson disease* – (AR) – *ATP7B* (16q24.3)
  - *PEO* – (AD or AR) – *ANT1* (4q35), *POLG* (15q25)
  - *Dominant optic atrophy* – (AD) – *OPA1* (3q28–q29)
  - *Deafness–dystonia syndrome* – (XLR) – *TIMM8A* (Xq22.1)
  - *Anemia, sideroblastic, and SCA* – (XLR) – *ABCB7* (Xq12–q13)
  - *Barth syndrome* – (XLR) – *TAZ* (Xq28)

20.5.4 Diagnostic Evaluation

- Some individuals have a clear characteristic phenotype of a specific disorder; can be confirmed by biochemical and molecular genetic testing
- Metabolic testing and muscle biopsy (respiratory chain activity) are useful for diagnosis
- Family history and inheritance evaluation is essential in directing molecular genetic testing
- Prenatal diagnosis is available for AR nDNA mutations
- Genetic counseling is complex based on the dual contribution of mtDNA and nDNA to the respiratory chain and the general characteristics of mitochondrial genetics
- Molecular genetic testing
  - Performed on DNA from blood (suspected nDNA mutations) or skeletal muscle (suspected mtDNA mutations)
  - Targeted mutation analysis of a panel of genes
  - Southern blot analysis may detect mtDNA rearrangements
  - If no recognized point mutation is identified, entire mtDNA sequencing and/or mutation scanning is available

Further Reading

Beutler E. Hemochromatosis: genetics and pathophysiology. Annu Rev Med. 2006;57:331–47.

Boehm CD, Cutting GR, Lachtermacher MB, Moser HW, Chong SS. Accurate DNA-based diagnostic and carrier testing for X-linked adrenoleukodystrophy. Mol Genet Metab. 1999;66:128–36.

Boileau C, Jondeau G, Mizuguchi T, Matsumoto N. Molecular genetics of Marfan syndrome. Curr Opin Cardiol. 2005;20:94–200.

Brandon MC, Lott MT, Nguyen KC, Spolim S, Navathe SB, Baldi P, Wallace DC. MITOMAP: a human mitochondrial genome database—2004 update. Nucleic Acids Res. 2005;33(Database Issue):D611–3. URL: http://www.mitomap.org.

Breighton P, De Paepe A, Steinmann B, Tsipouras P, Wenstrup RJ. Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). Am J Med Genet. 1998;77:31–3.

Chace DH, Kalas TA, Naylor EW. The application of tandem mass spectrometry to neonatal screening for inherited disorders of intermediary metabolism. Annu Rev Genomics Hum Genet. 2002;3:17–45.

Cohen Jr MM. Some chondrodysplasias with short limbs: molecular perspectives. Am J Med Genet. 2002;112:304–13.

DiMauro S. Mitochondrial diseases. Biochim Biophys Acta. 2004;1658:80–8.

Graff C, Bui TH, Larsson NG. Mitochondrial diseases. Best Pract Res Clin Obstet Gynaecol. 2002;16:715–28.

Gregersen N, Andresen BS, Corydon MJ, Corydon TJ, Olsen RKJ, Bolund L, Bross P. Mutation analysis in mitochondrial fatty acid oxidation defects: exemplified by acyl-CoA dehydrogenase deficiencies, with special focus on genotype phenotype relationship. Hum Mutat. 2001;18:169–89.

Grody WW, Cutting GR, Klinger KW, Richards CS, Watson MS, Desnick RJ. Subcommittee on cystic fibrosis screening, Accreditation of Genetic Services Committee. ACMG. Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. Genet Med. 2001;3:149–54.

Grosse SD, Kalman L, Khoury MJ. Evaluation of the validity and utility of genetic testing for rare diseases. Adv Exp Med Biol. 2010;686:115–31.

Hendrickson BC, Donohoe C, Akmaev VR, Sugarman EA, Labrousse P, Boguslavskiy L, et al. Differences in SMN1 allele frequencies among ethnic groups within North America. J Med Genet. 2009;46:641–4.

Kemp S, Pujol A, Waterham HR, van Geel BM, Boehm CD, Raymond GV, et al. ABCD1 mutations and the X-linked adrenoleukodystrophy mutation database: role in diagnosis and clinical correlations. Hum Mutat. 2001;18:499–515.

Maddalena A, Richards CS, McGinniss MJ, Brothman A, Desnick RJ, Grier RE, et al. Technical standards and guidelines for fragile X: the first of a series of disease-specific supplements to the Standards and Guidelines for Clinical Genetics Laboratories of the American College of Medical Genetics. Quality Assurance Subcommittee of the Laboratory Practice Committee. Genet Med. 2001;3:200–5.

McMurray CT. Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet. 2010;11:786–99.

Muntoni F, Torelli S, Ferlini A. Dystrophin and mutations: one gene, several proteins, multiple phenotypes. Lancet Neurol. 2003;2:731–40.

Ogino S, Wilson RB. Spinal muscular atrophy: molecular genetics and diagnostics. Expert Rev Mol Diagn. 2004;4:15–29.

Paulson HL, Fischbeck KH. Trinucleotide repeats in neurogenetic disorders. Annu Rev Neurosci. 1996;19:79–107.

Qaseem A, Aronson M, Fitterman N, Snow V, Weiss KB, Owens DK. Screening for hereditary hemochromatosis: a clinical practice guideline from the American College of Physicians. Ann Intern Med. 2005;143:517–21.

Scott SA, Edelmann L, Liu L, Luo M, Desnick RJ, Kornreich R. Experience with carrier screening and prenatal diagnosis for sixteen Ashkenazi Jewish Genetic Diseases. Hum Mutat. 2010;31:1–11.

Segawa M, Nomura Y. Rett syndrome. Curr Opin Neurol. 2005;18:97–104.

Sherman S, Pletcher BA, Driscoll DA. Fragile X syndrome: diagnostic and carrier testing. Genet Med. 2005;7:584–7.

Smeitink J, van den Heuvel L, DiMauro S. The genetics and pathology of oxidative phosphorylation. Nat Rev Genet. 2001;2:342–52.

Timchenko LT, Caskey CT. Trinucleotide repeat disorders in humans: discussions of mechanisms and medical issues. FASEB J. 1996;10:1589–97.

Watson MS, Cutting GR, Desnick RJ, Driscoll DA, Klinger K, Mennuti M, et al. Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel. Genet Med. 2004;6:387–91.

Biochemical Genetics and Inborn Errors of Metabolism

21

Chunli Yu, Melissa P. Wasserstein, and George A. Diaz

Contents

21.1 Introduction 562

21.2 Newborn Screening of Metabolic Diseases and Other Genetic Disorders 562
21.2.1 Overview 562
21.2.2 Newborn Screening Disorders 562
21.2.3 Newborn Screening Methodologies 564

21.3 Biochemical Genetic Laboratory Techniques and Methodologies 565
21.3.1 Amino Acid Analysis 565
21.3.2 Organic Acid Analysis 570
21.3.3 Acylcarnitine Analysis 572
21.3.4 Carnitine Analysis 573
21.3.5 Lysosomal Enzyme Assays 575

21.4 Amino Acid Disorders 576
21.4.1 Phenylketonuria 576
21.4.2 Tyrosinemia 577
21.4.3 Maple Syrup Urine Disease (MSUD) 578
21.4.4 Homocystinuria 579
21.4.5 Urea Cycle Disorders 579

21.5 Organic Acidemias 581
21.5.1 Propionic Acidemia 581
21.5.2 Methylmalonic Acidemia (MUT Deficiency) 582
21.5.3 3-Methylcrotonyl-CoA Carboxylase Deficiency (3-MCC) 582
21.5.4 Glutaric Acidemia Type I (GA I) 583

21.6 Fatty Acid Disorders 584

21.7 Lysosomal Storage Diseases 585
21.7.1 Hurler Syndrome 585
21.7.2 Gaucher Disease 586
21.7.3 Niemann-Pick Disease Types A and B 587
21.7.4 Fabry Disease 588
21.7.5 Krabbe Disease 589
21.7.6 Pompe Disease 589

Further Reading 590

C. Yu, MD (✉)
Department of Genetics & Genomic Sciences and Pediatrics, Mount Sinai School of Medicine, New York, NY, USA

M.P. Wasserstein, MD
Mount Sinai School of Medicine, New York, NY, USA

G.A. Diaz, MD, PhD
Department of Genetics & Genomic Sciences, Mount Sinai School of Medicine, New York, NY, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_21,

21.1 Introduction

- Inborn errors of metabolism (IEM) include a broad spectrum of defects of various gene products that affect intermediary metabolism in the body, resulting in clinical disease. Studying the molecular and biochemical mechanisms of these inherited disorders, systematically summarizing the disease phenotype and natural history, and providing diagnostic rationale and methodology and treatment strategy comprise the context of human biochemical genetics
- This chapter focuses on newborn screening of IEM and laboratory diagnosis of these conditions and summarizes the clinical features, genetic basis, and clinical managements of some common IEM

21.2 Newborn Screening of Metabolic Diseases and Other Genetic Disorders

21.2.1 Overview

- Population newborn screening (NBS) is a public health program that aims for early detection and early intervention of clinically important conditions to prevent mortality, morbidity, and disability
- NBS is a multitiered system consisting of screening, diagnosis, treatment, long-term followup, education, and continuous system evaluation
- NBS is usually performed in the first week of life on dried blood spot (DBS) collected from a heel stick (with exceptions of screening for hearing loss and screening for critical congenital heart disease)
- NBS tests for biochemical (and molecular) markers from DBS
  - Many biochemical markers detected by liquid chromatography tandem mass spectrometry (LC-MS/MS) require differential diagnosis between multiple conditions
- Many of the disorders in the NBS panel are treatable inborn errors of metabolism or metabolic diseases
  - Individually rare, but collectively have a frequency of 1:2,000–3,000
  - Readily identified by abnormal amino acid and/or acylcarnitine profiles from DBS
- NBS is becoming more uniform nationally and worldwide with standardization of disease panels, analytical and postanalytical performance, and system evaluation processes

21.2.2 Newborn Screening Disorders

21.2.2.1 Recommended Uniform Screening Panels

- Recommended uniform screening panel of the Secretary's Advisory Committee on Heritable Disorders in Newborns and Children (SACHDNC) (Table 21.1)
  - 31 core conditions are recommended for all screening programs
  - 25 secondary conditions revealed as part of the differential diagnosis of screening results for the core conditions do not meet the criteria to be included in the core panel, due to lack of effective treatment or with uncertain clinical significance; such information should be made available to the healthcare professional and or family members by the screening laboratories

21.2.2.2 Nomination of Disorders for NBS

- SACHDNC is responsible for systematically evaluating conditions nominated for addition to the uniform NBS panel and making recommendations to the secretary of the US Department of Health and Human Services (HHS)
- Review process is an evidence-based evaluation on the magnitude of benefit, adequacy of evidence, and certainty of net benefit/harm
- Severe combined immunodeficiency (SCID) was added to the recommended uniform screening panel in October 2010
- Critical congenital heart disease (CCHD) was added to the recommended uniform screening panel in September 2011

Table 21.1 Recommended uniform screening panel of the Secretary's Advisory Committee on Heritable Disorders in Newborns and Children: core conditions and secondary conditions. Classified by disease category, key analytes, and testing methodologies

Category	Condition	Key analytes	Methodology
	Core conditions (31)		
Amino acids (6)	PKU	Phe	LC-MS/MS
	MSUD	Leu/Ile, Val	
	HCY	Met	
	TYR I	Succinylacetone	
	CIT, ASA	Cit, ASA	
Organic acids (9)	IVA	C5	LC-MS/MS
	GA I	C5DC	
	MUT, Cbl A and B, PROP, MCD	C3	
	3-MCC, BKT, HMG	C5OH	
Fatty acid oxidation disorders (5)	MCAD	C8	LC-MS/MS
	VLCAD	C14:1	
	LCHAD, TFP	C16-OH	
	CUD	C0 (L)	
Hb pathies (3)	HB SS, HB S/β Th, Hb S/C	Hb variants	HPLC, electrophoresis
Others (8)	CH	T4, TSH	Immunoassay
	CAH	17-OHP	
	BIOT	Biotinidase activity	Colorimetric assay
	GALT	GALT activity, total galactose	Fluorometric and colorimetric assays
	HEAR	Hearing response	OAE and AABR
	CF	IRT	Immunoassay
	SCID	TREP	Molecular
	CCHD	Oxygenation of Hb	Pulse oximetry
	Secondary conditions (25)		
Amino acids (8)	H-PHE, BIOPT (BS), BIOPT (REG)	Phe	LC-MS/MS
	MET	Met	
	ARG	Arg	
	TYR II, TYR III	Tyr	
	CIT II	Cit	
Organic acids (6)	Cbl C and D	C3	LC-MS/MS
	MAL	C3DC	
	IBG	C4	
	2MBG, 2M3HBA, 3MGA	C5OH	
Fatty acid oxidation disorders (8)	SCAD	C4	LC-MS/MS
	GA 2	C4, C5	
	M/SCHAD, MCAT	C10-OH	
	CPT II, CACT	C16, C18:1	
	CPT IA	C0 (H), C0/16 + 18	
	DE RED	C10:2	
Hb pathies (1)	Var Hb	Hb variant	HPLC, electrophoresis
Others (2)	GALK, GALE	GALT activity, total galactose	Fluorometric and colorimetric assays

21.2.3 Newborn Screening Methodologies

21.2.3.1 Bacterial Inhibition Assays (BIA)

- BIA for screening phenylketonuria (PKU) was first developed by Dr. Robert Guthrie in the late 1950s
- BIA are no longer used and have been replaced with multiplexed LC-MS/MS method

21.2.3.2 Fluorometric and Colorimetric Assays

- Emerged in the 1990s with automated procedures
- Disorders that are screened with fluorometric and colorimetric methods
  - Galactosemia (galactose metabolites and GALT activity)
  - Biotinidase (biotinidase activity)
- Testing principles
  - These assays use coupled enzymatic reactions that lead to release of color or fluorescence-generating compounds to measure analyte concentrations and enzymatic activities
- Limitations
  - Unable to multiplex
  - High false-positive rate
  - Enzymes on DBS are vulnerable to heat and humidity

21.2.3.3 Immunoassays

- Disorders that are screened with immunoassay methods
  - Congenital adrenal hyperplasia (CAH)
  - Congenital hypothyroidism (CH)
  - Cystic fibrosis (CF)
- Testing principles
  - Use antibodies toward a specific antigen, such as thyroid-stimulating hormone (TSH) for CH, 17-hydroxypragestorone (17-OHP) for CAH, and immunoreactive trypsinogen (IRT) for CF
  - Antibodies or antigens are prelabeled with fluorochrome (europium)
  - Formation of antibody-antigen complex to a solid phase on the microplate
  - Use enhancing reagents to dissociate the fluorochrome and release fluorescence
- Limitations
  - Cross-reacting substance in the blood will interfere with the antibodies and result in false-positive results
  - For example, placental and fetal steroid can cross-react with 17-OHP antibody and cause falsely elevated 17-OHP

21.2.3.4 LC-MS/MS Method

- It is widely used to screen for amino acid, organic acid, and fatty acid disorders
- Method uses flow injection, positive electrospray ionization, and tandem mass spectrometry
- 40–50 amino acids and acylcarnitines are multiplexed and quantified in one test in approximately 2 min or less
- Sample preparation
  - One 1/8″ blood spot punch is required for the analysis
  - Amino acids and acylcarnitines are extracted from DBS with methanol containing stable isotope-labeled amino acid and acylcarnitine internal standards
    - Extracted amino acids and acylcarnitines are then derivatized with 3N butanolic HCL to form butyl esters
  - There are also underivatized methods used in some NBS laboratories
- MS/MS analysis (Fig. 21.1)
  - Precursor ion scan: acylcarnitine profile is obtained via precursor ion scan of m/z 85 over a typical mass range of 200–500
  - Neutral loss scan: neutral and acidic amino acids are detected by neutral loss scan of 102 Da, an uncharged molecule butyl formate that is lost after fragmentation in the collision cell
    - Neutral loss of 119 Da is used for detection of dibasic amino acids for additional loss of ammonia
  - Selective reaction monitoring (SRM): acylcarnitines and amino acids of interests can be quantified via SRM
  - The concentrations of acylcarnitine and amino acids are calculated by comparing the ion intensity to their respective or

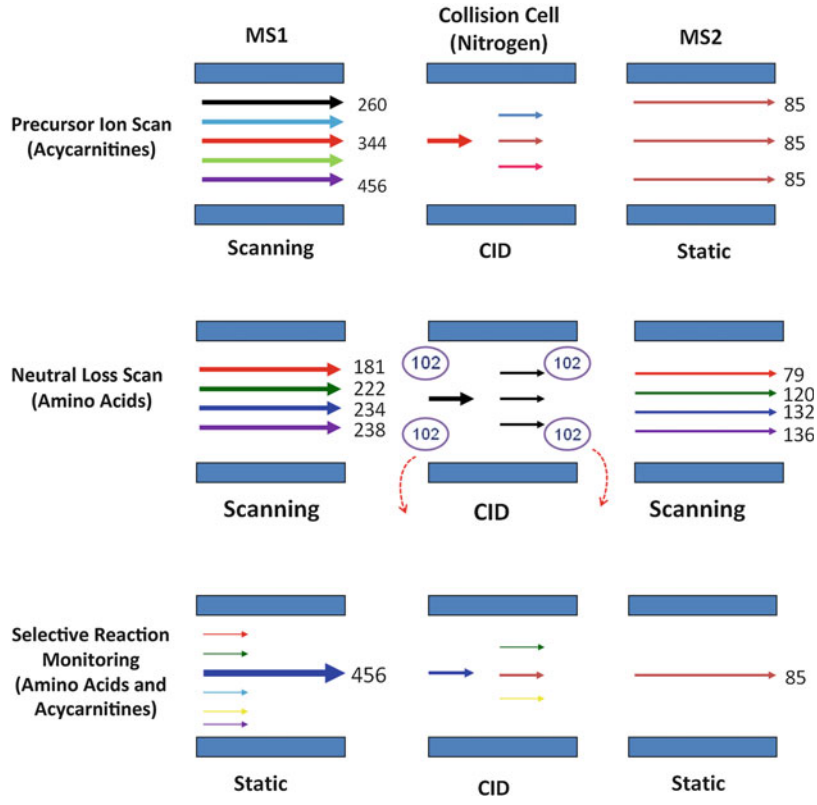


Fig. 21.1 MS/MS functions used in newborn screening (NBS). *Top panel*: precursor ion scan of *m*/*z* of 85 for free carnitine and acylcarnitines. *Middle panel*: neutral loss scan of *m*/*z* of 102 for neutral and acidic amino acids. *Bottom panel*: selective reaction monitoring (SRM) for both amino acids and acylcarnitines

closest *m*/*z* internal standards of known concentration

- Important points
  - LC-MS/MS detects levels of amino acids and acylcarnitines
  - Abnormal result may not be specific for a particular disorder and needs to be evaluated by confirmatory testing
  - Confirmation of abnormal NBS results usually requires combining biochemical findings of plasma amino acids and acylcarnitine, urine organic acids, and in some cases enzyme and mutation studies

21.2.3.5 DNA-Based Screening

- Primary screen uses DNA-based method
  - Quantitation of T cell receptor gene excision circles (TRECs) by quantitative PCR assay for screening SCID
- NBS of CF is a two-tier process
  - CFTR mutation is performed on samples with elevated IRT
- DNA-based testing is not feasible to screen metabolic disorders in general but can be used as second-tier testing to reduce false-positives in NBS laboratories (i.e., Q188R and Duarte variant for positive galactosemia screens, common mutations for positive biotinidase screens)

21.3 Biochemical Genetic Laboratory Techniques and Methodologies

21.3.1 Amino Acid Analysis

- Free amino acids in physiological fluids are widely used for clinical diagnosis and monitoring
  - Plasma is the preferred specimen for diagnosis and monitoring most of the amino acid disorders and some organic acidemias
  - Urine is useful for evaluation of specific disorders of renal transport, such as

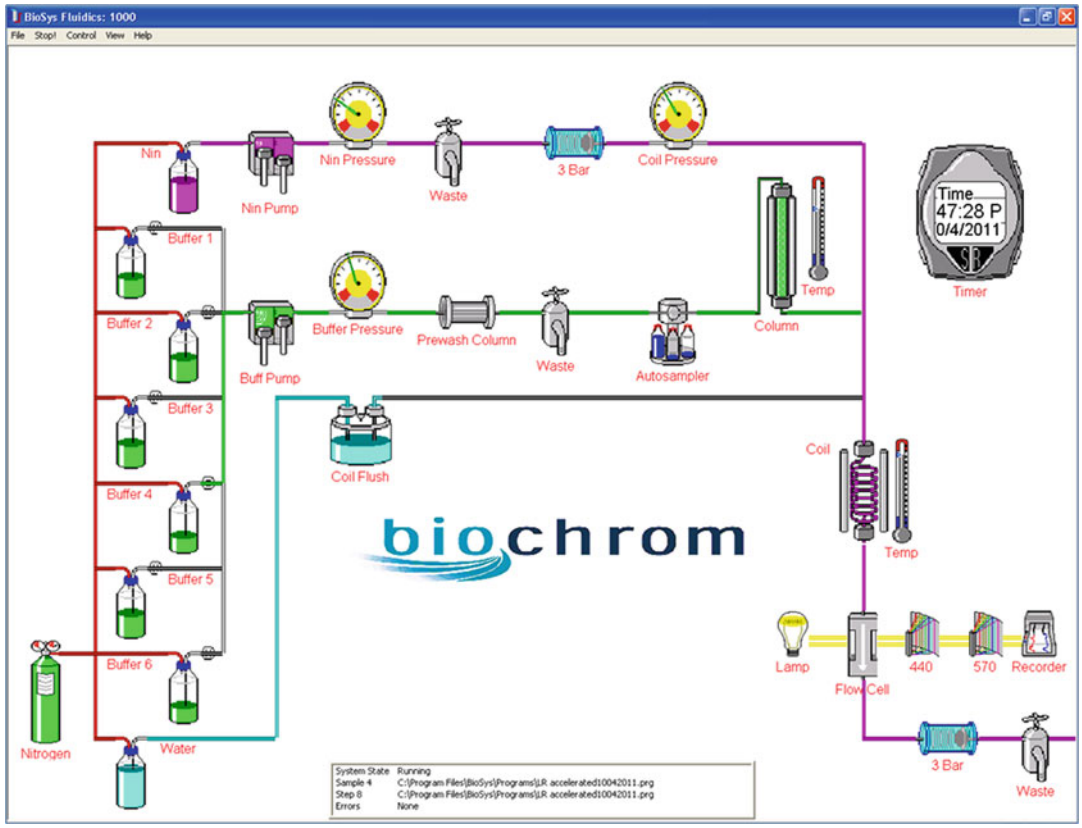


Fig. 21.2 Representative ion exchange chromatography (IEC) amino acid analyzer (Permission by Biochrom, Inc.)

cystinuria and lysinuric protein intolerance, as well as generalized renal tubular dysfunction
  - CSF is needed with a concurrent plasma sample for the evaluation of glycine encephalopathy or serine deficiency disorders
- Preanalytical requirements
  - Plasma: morning fasting or preprandial blood specimen is required
    - For infants, the blood should be collected prior to the next scheduled feeding
  - Plasma should be separated from red cells as soon as possible and stored frozen until analysis
    - Inappropriate sample handling can affect concentrations of some amino acids (Table 21.3 for sample-handling artifacts)
  - Urine: first-morning-void or random urine collected in a clean container without any preservatives is required; sample should be kept frozen until analysis
  - CSF: blood contamination may result in a generalized amino acid elevation in CSF
- Analytical methods
  - Ion exchange chromatography (IEC)
    - Gold standard method to analyze free amino acids in various physiological fluids; provides a comprehensive profile of 30–40 amino acids
    - Performed on a dedicated analyzer (Fig. 21.2, representative IEC system) that uses lithium citrate buffer system, cation exchange column separation, and postcolumn ninhydrin derivatization
      - A typical IEC chromatogram of amino acid calibrators shown in Fig. 21.3

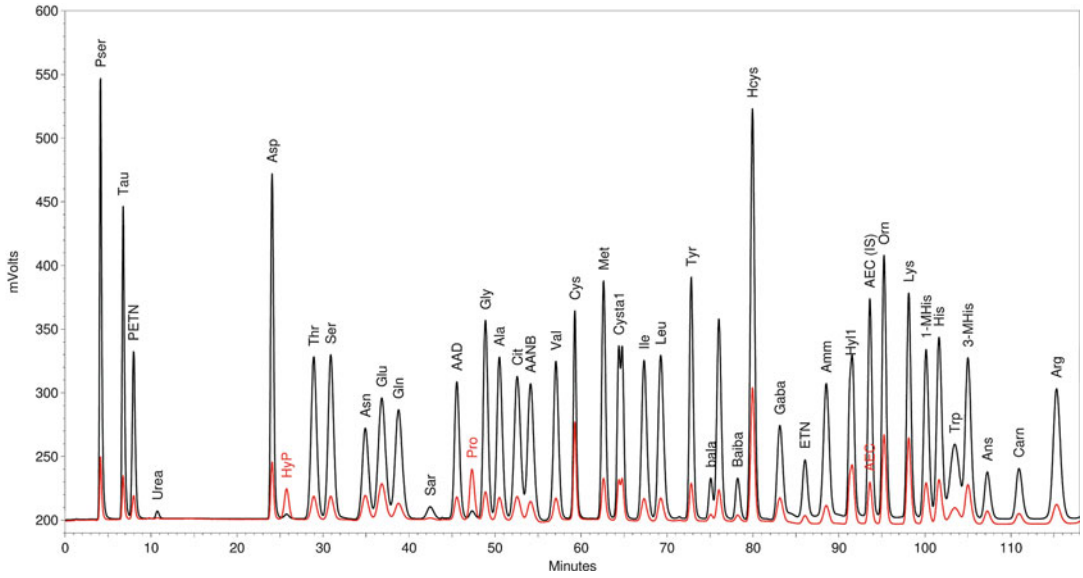


Fig. 21.3 Chromatogram of amino acids calibration mixture by ion exchange chromatography method (IEC). *Black trace*: amino acids and internal standard (AEC) on channel 570 nm. *Red trace*: amino acid and internal standard (AEC) on channel 440 nm. Proline (Pro) and hydroxyproline (Hyp) are quantitated on channel 440 nm

- Minimal sample preparation by deproteinization with sulfosalicylic acid and addition of a single internal standard
- One-point calibration using commercially available calibration mixtures is used for quantification
- Long analysis time (~2 h or longer per sample) is not suitable for high-throughput testing laboratories with expected fast turnaround time

– Ultraperformance liquid chromatography (UPLC) method
  - Derivatization with AccQ-Tag reagent
  - Derivatized amino acids are separated by UPLC and detected by UV detector
  - Enhanced sensitivity and dynamic range
  - Reduced analysis time (approximately 45 min)
  - Alternative to IEC but requires dedicated instrument
– LC-MS/MS method
  - Both nonderivatization and derivatization methods exist
  - HPLC (or UPLC) coupled with tandem MS in positive ESI mode
  - Amino acid detection is based on the retention time and its specific *m/z* (SRM transition)
  - Enhanced sensitivity (limit of quantification of approximately 1 μmol/L) and specificity
  - Fast analysis time (approximately 25 min)
  - Alternative to IEC for laboratory experienced in quantitative LC-MS/MS analysis

• Interpretation of results
  – Reportable 30–40 amino acids against age-matched reference ranges
  – Specific amino acid patterns for certain disorders (Table 21.2)
  – Clinically relevant atypical amino acids may be reported when present; these include, for example
    - Alloisoleucine (MSUD)
    - Argininosuccinic acid (ASA lyase deficiency)
    - Mixed disulfide peak (homocystinuria or CblC disease)
  – Amino acid results should be interpreted in the context of dietary history, nutritional status, and concurrent medication history, as well as sample condition

Table 21.2 Biochemical patterns of inborn errors of amino acid, organic acid, and fatty acid metabolism

Disease	Amino acid patterns	Organic acid patterns	Acylcarnitine patterns
Amino acid metabolism			
Phenylketonuria (PKU)	Phenylalanine (P)↑ Tyrosine (P)↓ or normal	Phenylpyruvic acid ↑ Phenyllactic acid ↑ Phenylacetic acid ↑ 2-Hydroxyacetic acid ↑ (normal when patient is under good metabolic control)	
Maple syrup urine disease (MSUD)	Alloisoleucine, leucine, isoleucine, valine (P)↑ Alanine (P)↓ when decompensated	2-Ketoisocaproic acid ↑ 2-Keto-3-methylvaleric acid ↑ 2-Ketoisovaleric acid ↑ 2-Hydroxyisovaleric acid ↑ (2-Keto acids may be absent when patient is under good metabolic control)	
Tyrosinemia type I (TYR I)	Tyrosine (P) ↑	Succinylacetone↑ 4-Hydroxyphenyllactic acid ↑ 4-Hydrophenylpyruvic acid ↑ 4-Hydrophenylacetic acid ↑	
Tyrosinemia II and III (TYR II and III)	Tyrosine (P) ↑↑	4-Hydroxyphenyllactic acid ↑ 4-Hydrophenylpyruvic acid ↑ 4-Hydrophenylacetic acid ↑	
Homocystinuria (HCY) or CBS deficiency	Methionine (P) ↑ Homocysteine (P) ↑		
Glycine encephalopathy	Glycine (CSF) ↑↑ Glycine (P) ↑ CSF/plasma glycine >0.08		
Ornithine transcarbamylase deficiency (OTC)	Glutamine (P) ↑ Citrulline (P) ↓ Arginine (P) ↓	Orotic acid ↑	
Argininosuccinic acid synthase deficiency (CIT)	Citrulline (P) ↑↑	Orotic acid ↑ or normal	
Argininosuccinic acid lyase deficiency (ASL)	Citrulline (P) ↑ Argininosuccinic acid (P) ↑	Orotic acid ↑ or normal	
Arginase deficiency (ARG)	Arginine (P) ↑	Orotic acid ↑ or normal	
Hyperammonemia hyperornithinemia homocitrullinuria syndrome (HHH)	Ornithine (P) ↑ Homocitrulline (P, U) ↑	Orotic acid ↑ or normal	
Organic acid metabolism			
Methylmalonic acidemia (MUT)	Glycine (P) ↑	Methylmalonic acid ↑↑ 3-Hydroxypropionic acid ↑ Methylcitric acid ↑	C3 ↑ C4DC ↑ or normal
Cbl A and B		Methylmalonic acid ↑ 3-Hydroxypropionic acid ↑ Methylcitric acid ↑	C3 ↑ C4DC ↑ or normal
CblC disease	Homocysteine (P) ↑ Methionine (P) ↓ or normal	Methylmalonic acid ↑ 3-Hydroxypropionic acid ↑ Methylcitric acid ↑	C3 ↑

(*continued*)

Table 21.2 (continued)

Disease	Amino acid patterns	Organic acid patterns	Acylcarnitine patterns
Propionic acidemia (PROP)	Glycine (P) ↑	3-Hydroxypropionic acid ↑ Propionylglycine ↑ Tiglylglycine ↑ Methylcitric acid ↑	C3 ↑
Isovaleric acidemia (IVA)	Glycine (P) ↑ or normal	3-Hydroxyisovaleric acid ↑ Isovalerylglycine ↑	C5 ↑
3-Methylcrotonyl-CoA carboxylase deficiency (3-MCC)		3-Hydroxyisovaleric acid ↑ 3-Methylcrotonylglycine ↑	C5-OH ↑
3-Hydroxy-3-methylglutaryl-CoA lyase deficiency (HMG)		3-Hydroxyisovaleric acid ↑ 3-Methylglutaric acid ↑ 3-Methylglutaconic acid ↑ 3-Hydroxy-3-methylglutaric acid ↑	C5-OH ↑ C6DC ↑
2-Methylbutyryl-CoA dehydrogenase (2MBG)		2-Methylbutyrylglycine ↑	C5 ↑
3-Ketothiolase deficiency (BKT)		2-Methyl-3-hydroxybutyric acid ↑ 2-Methylacetoacetic acid ↑ Tiglylglycine ↑	C5:1 ↑ C5-OH ↑ or normal
Isobutyryl-CoA dehydrogenase deficiency (IBG)		Isobutyrylglycine ↑ or normal	C4 ↑
Glutaric acidemia type I (GA I)		Glutaric acid ↑ 3-Hydroxyglutaric acid ↑	C5DC ↑
Fatty acid metabolism			
Carnitine uptake defect (CUD)			C0 ↓↓
Carnitine palmitoyltransferase I deficiency (CPT I)			C0 ↑ or normal C0/C16 + 18 ↑
Carnitine palmitoyltransferase II deficiency/carnitine-acylcarnitine translocase deficiency (CPT II/CACT)			C16, C18:1 ↑
Short-chain acyl-CoA dehydrogenase deficiency (SCAD)		Ethylmalonic acid ↑ Methylsuccinic acid ↑	C4 ↑
Medium-chain acyl-CoA dehydrogenase deficiency (MCAD)		Hexanoylglycine ↑ Suberylglycine ↑ Phenylproprionylglycine ↑ C6–C10 dicarboxylic acids ↑ or normal	C8 ↑ C6, C10 and C10:1 ↑
Very long-chain acyl-CoA dehydrogenase deficiency (VLCAD)		C6–C14 dicarboxylic acids ↑ or normal	C14:1 ↑ C14, C14:2 and C16 ↑ or normal
Long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency/trifunctional protein deficiency (LCHAD/TFP)		C6–C14 3-hydroxy dicarboxylic acids ↑ or normal	C14-OH, C16-OH, C18:1-OH and C18-OH ↑
Multiple acyl-CoA dehydrogenase deficiency (GA 2)		Glutaric acid ↑ 2-Hydroxyglutaric acid ↑ Ethylmalonic acid ↑ Hexanoylglycine ↑ Isovalerylglycine ↑ C6–C10 dicarboxylic acids ↑ or normal	C4, C5, C5DC ↑ C6–C18 ↑ or normal

Table 21.3 Common interferences in amino acid, organic acid, and acylcarnitine analyses from various sources

	Amino acids	Organic acids	Acylcarnitines
Sample-handling artifacts			
Hemolysis	Taurine, asparagine, glycine, and glutamic acid (P) ↑		
Prolonged storage at room temperature	Ornithine, glutamic acid (P) ↑, arginine, glutamine (P)↓ Homocysteine (P)↓		Free carnitine ↑ Short-chain species ↓
Bacterial contamination		3-Hydroxypropionic acid ↑ 4-Hydrophenylacetic acid ↑ D-Lactic acid ↑ 2-Ketoglutaric acid ↑ Succinic acid ↑ Phenylpropionylglycine ↑	
Dermatological soap/ ointment contamination		Palmitic, stearic ↑ (soap) Glycerol ↑ (ointment)	
Dietary/nutritional			
Dietary	Taurine (P)↑ (shellfish) Homocitrulline (P,U)↑ (canned food, cow's milk) Anserine, 1-MHis, carnosine (U) ↑ (white meat from fowl)	Adipic acid ↑ (jello) Furoic acid ↑ (chocolate) Benzoic acid, hippuric acid ↑ (preservatives in food) Adipic < suberic < sebacic acid ↑ with 7-hydroxyoctanoic acid ↑ (MCT oil)	C6 < C8 < C10 ↑ (MCT oil)
TPN	Nonspecific ↑	N-acetyltyrosine ↑	C18:2, C18:1 ↑ (intra-lipid)
Starvation	Leucine, isoleucine, valine (P)↑, alanine (P)↓		
B12 deficiency	Homocysteine (P) ↑	Methylmalonic acid ↑	C3 ↑
Illness/diseases			
Prematurity	Tyrosine (P) ↑ Generalized amino acid (U) ↑		Free carnitine ↓
Liver diseases	Tyrosine, Methionine (P) ↑	4-Hydrophenylacetic acid ↑ 4-Hydrophenyllactic acid ↑ 4-Hydrophenylpyruvic acid ↑	
Ketosis	Leucine, isoleucine, valine (P)↑, alanine (P)↓		C2, C4-OH ↑
Medications			
Valproic acid	Glycine (P, U) ↑	Various valproic acid metabolites	C8 ↑, Free carnitine ↓
Pivalic acid containing antibiotics		Pivalic acid	C5 ↑
Paracetamol	Interfering peak around phenylalanine (P, U)	5-Oxoproline ↑	
L-DOPA		Vanillin lactic acid ↑	

    - See Table 21.3 for the most common interferences to amino acid concentrations from various sources
- A clinical report of amino acid analysis should include explanation of result, possible diagnoses, and recommendation for further testing if required

21.3.2 Organic Acid Analysis

- Organic acids are water-soluble compounds that contain one or more carboxyl groups and other nonamino functional groups
    - Urine organic acid analysis identifies primary disorders of intermediary metabolism

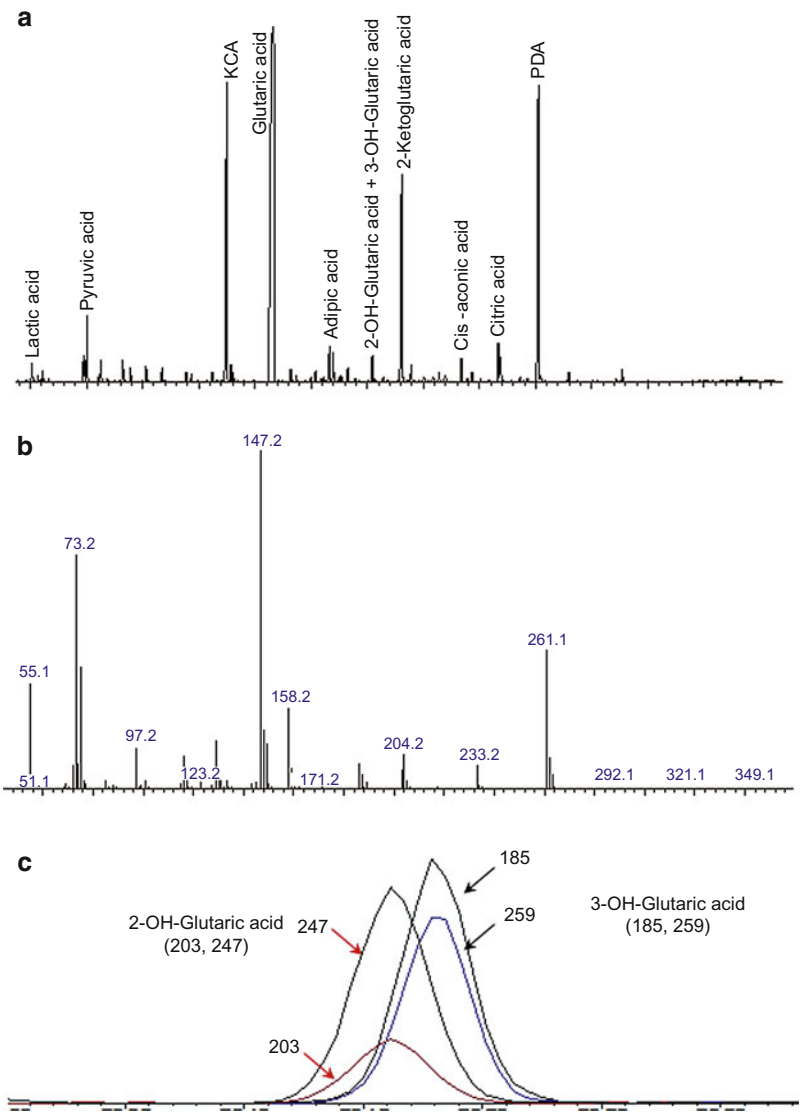


Fig. 21.4 Urine organic acid profile. (**a**) Total ion chromatogram (TIC) of an organic acid profile of glutaric acidemia type I (GA I). Markedly elevated glutaric acid and presence of 3-hydroxyglutaric acid are typical for classic GA I. (**b**) Electron impact (EI) mass spectrum of di-TMS derivatives of glutaric acid. (**c**). Extracted ion chromatogram (EIC) of 185 and 259 (characteristic for 3-hydroxyglutaric acid) and 203 and 247 (specific for 2-hydroxyglutaric acid) displayed in overlay

(amino acids, fatty acids, carbohydrates, nucleic acids, etc.) and provides a metabolic profile
  - See Fig. 21.4 for an organic acids profile, an EI mass spectrum, and an extracted ion chromatogram
- Preanalytical requirements
  - First-morning-void or random urine collected in a clean container without any preservatives is required for the analysis
    - Keep sample frozen until analysis
  - It is best to collect urine for organic acid analysis when patient is acutely ill, as abnormal organic acids often decrease, sometimes to near-normal concentration when patient is well
- Analytical methods
  - Sample preparation
    - Extraction of organic acids from urine by organic solvents and derivatization with trimethylsilane (TMS)
    - α-keto acids (pyruvate, 2-keto acids from BCAA, succinylacetone, and α-ketoglutaric acid) are oximated with hydroxylamine in order to form TMS derivatives

- GC/MS analysis
  - TMS-derivatized organic acids are separated on a capillary column (i.e., HP-5 column, 25 m, 0.20 mm, 0.33 μm) by an automated oven temperature program using helium as carrying gas
  - Eluted organic acids are ionized in positive electron impact (EI) mode via removal of electron and subsequent fragmentation of the molecules
  - The basis of compound identification is retention time and EI mass spectrum – the fingerprint of a specific molecule
- Special considerations
  - Some coeluting organic acids are often critical for diagnosis of certain disorders; ion extraction is necessary so that these organic acids are not missed during analysis
    - 4-Hydroxybutyric acid (204, 233) vs urea peak II (171, 189)
    - Hexanoylglycine (261, 274, 302) vs 4-hydroxyphenylacetic acid (251, 281, 296)
    - 3-Hydroxyglutaric acid (185, 259) vs 2-hydroxyglutaric acid (157, 203)
    - Orotic acid (254, 357) vs cis-aconitic acid (211, 229)
- Specific organic acid quantitation (methylmalonic acid, glutaric acid, orotic acid, and succinylacetone) is measured by stable isotope dilution GC/MS or LC-MS/MS for monitoring known patients
- Interpretation of results
  - Recognizable organic acid patterns for specific metabolic disorders (Table 21.2)
  - Organic acid results should be interpreted in the context of dietary history, nutritional status, and medical information including medication, as well as other metabolic studies
  - See Table 21.3 for the most common interferences from dietary and iatrogenic sources

21.3.3 Acylcarnitine Analysis

- Carnitine (β-hydroxy-γ-trimethylaminobutyric acid) is required for oxidation of long-chain fatty acids and is conjugated with fatty acids or organic acids to form acylcarnitines
  - Acylcarnitine analysis measures concentrations of acylcarnitine species from chain length C2 to C18 including some dicarboxylic and hydroxylated species
  - It is clinically useful in diagnosing various fatty acid oxidation disorders and organic acidemias
  - Definitions of clinically relevant acylcarnitine species and their associated diseases and conditions are summarized in Table 21.4
- Preanalytical requirements
  - Plasma or serum; sample should be kept frozen until analysis
  - DBS, 1–2 circles, fully saturated
  - Acylcarnitines are stable in −80 °C indefinitely
    - Prolonged storage at room temperature will result in degradation of acylcarnitine to free carnitine, particularly short-chain species
- Analytical methods
  - Sample preparation (see Sect. 21.2, newborn screening methodologies)
  - LC-MS/MS analysis (see Sect. 21.2, newborn screening methodologies)
  - Precursor ion scan if preferred in clinical biochemical genetics laboratories as it provides comprehensive profile in addition to numerical values (Fig. 21.5 for acylcarnitine profile of normal, SCAD, MCAD, and VLCAD via precursor ion scan of 85)
- Interpretation of results
  - Acylcarnitine abnormalities of specific fatty acid oxidation disorders and organic acidemias are summarized in Table 21.2 in conjunction of their respective organic acid pattern and amino acid abnormalities
    - Acylcarnitine result is best interpreted in the context of urine organic acid and plasma amino acid findings
  - In some diseases, acylcarnitine profile can be normal when patients are stable under good metabolic control (i.e., VLCAD and LCHAD)

Table 21.4 Clinically relevant list of acylcarnitine butyl esters and their associated diseases and conditions

Abbreviations	*m/z*	Acylcarnitines	Associated diseases and conditions
C0 (L)	218	Free carnitine	CUD Secondary carnitine deficiencies
C2	260	Acetyl-	Ketosis (with C4-OH) L-carnitine supplementation
C3	274	Propionyl-	PROP MUT (with C4DC?) SUCLA2 deficiency (with C4DC)
C4	288	Butyryl-	SCAD
		Isobutyryl-	IBG
C5:1	300	Tiglyl-	BKT (with C5-OH)
C5	302	Isovaleryl-	IVA
		2-Methylbutyryl-	2MBG
		Pivalyl-	Pivalic acid containing antibiotic use
C4-OH	304	3-Hydroxybutyryl-	Ketosis (with C2) SCHAD
C5-OH	318	3-Hydroxyisovaleryl-	3-MCC BIOT MCD HMG (with C6DC) 3 MGA
		2-Methyl-3-hydroxybutyryl-	BKT (with C5:1) 2M3HBA
C8	344	Octanoyl-	MCAD (with C6, C10:1, and C10)
C3-DC	360	Malonyl-	MAL
C4-DC	374	Methylmalonyl- Succinyl-	MUT (with C3) SUCLA2 deficiency (with C3)
C5-DC C10-OH	388	Glutaryl- 3-Hydroxydecanoyl-	GA I M/SCHAD, MCAT
C14:1	426	Tetradecenoyl-	VLCAD (with C14:2 and C14)
C16	456	Palmitoyl-	CPT II (with C18:1, C18:2, and C18)
C16-OH	472	3-Hydroxypalmitoyl-	LCHAD/TFP (with C16:1-OH, C18:1-OH, and C18-OH)

  - Acylcarnitine profile can be influenced by nutritional and medical status
    - For instance, B12 deficiency can result in mild C3 elevation
      - Some medications can cause artifacts (e.g., pivalic acid-based antibiotic leads to C5 elevation) (Table 21.3)

21.3.4 Carnitine Analysis

- Carnitine analysis involves the direct measurement of free carnitine and the total carnitine following alkaline hydrolysis of acylcarnitines
  - The fraction of acylcarnitines or esterified carnitines is obtained by subtraction
  - Plasma carnitine analysis is used to monitor carnitine homeostasis
  - Urine carnitine is evaluated when primary carnitine uptake defect is clinically suspected
- Preanalytical requirements
  - Plasma or serum; store frozen until analysis
  - Random urine collected in a clean container without preservatives; store frozen until analysis
- Analytical methods
  - Hydrolysis of acylcarnitines by potassium hydroxide

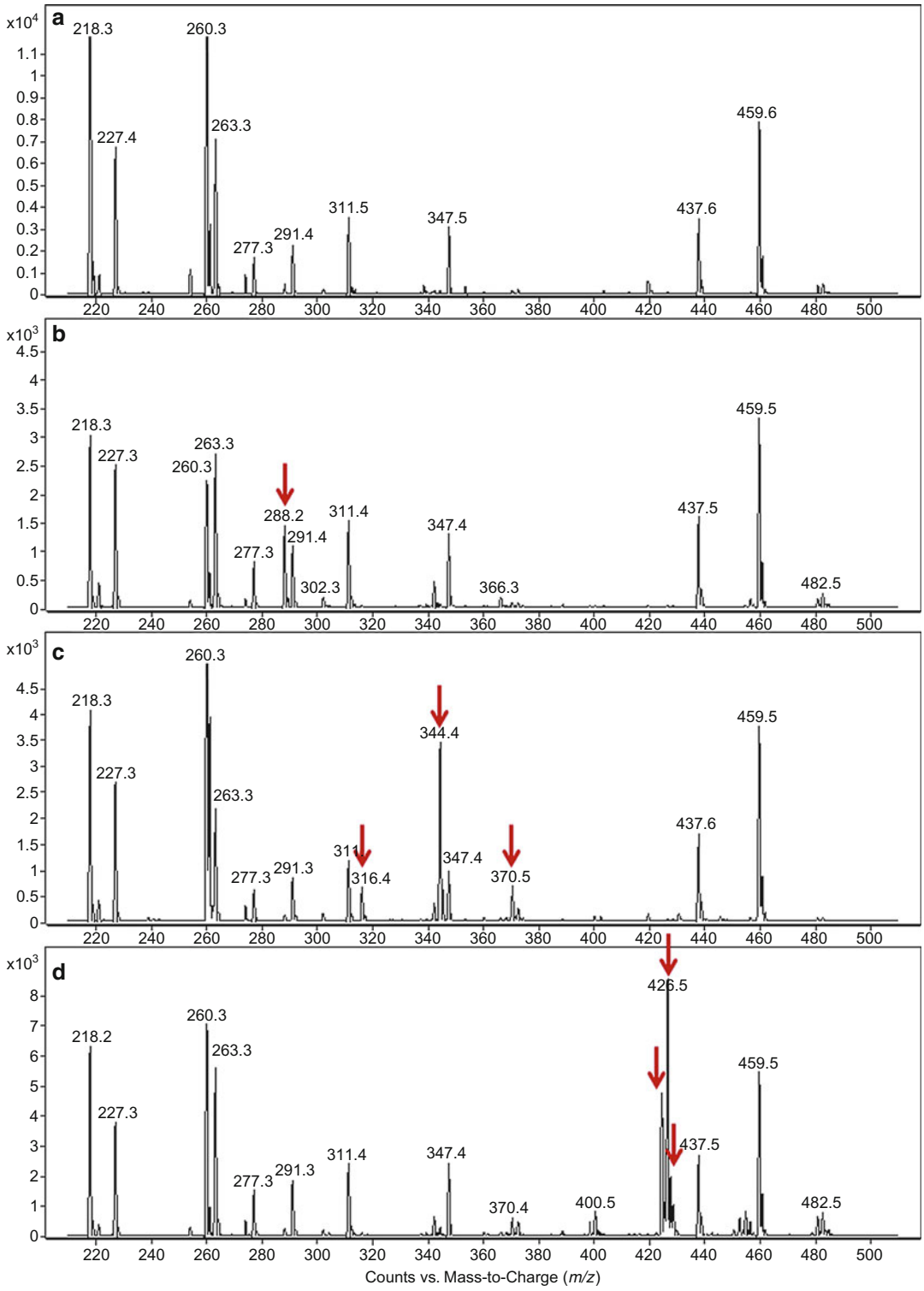


Fig. 21.5 (continued)

  - Measurement of underivatized free and total carnitine on LC-MS/MS via SRM transition of *m/z* 165/85
    - C2 (acetyl-) is often monitored to access hydrolysis efficacy in total carnitine
  - Other methods include radioenzymatic assays and spectrophotometric methods on automated chemistry analyzers
- Interpretation of results
  - Significantly decreased plasma carnitine (<5 μmol/L) with markedly increased urine carnitine is suggestive of a primary carnitine uptake defect
    - Diagnosis can be confirmed by mutation analysis or decreased carnitine transport activity in cultured cells
  - Secondary carnitine deficiencies can be seen in fatty acid (particularly long chain) oxidation disorders and organic acidemia
    - The acyl/free carnitine ratio is usually elevated in these cases
  - Low carnitine can also be seen in premature infants and patients taking valproic acid and benzoic acid

21.3.5 Lysosomal Enzyme Assays

- Lysosomal storage diseases (LSD) include a group of over 40 inherited disorders in the categories of sphingolipidoses, mucopolysaccharidoses, oligosaccharidoses, and neuronal ceroid lipofuscinosis
  - Lysosomal enzyme analysis screens and confirms the specific LSD
  - This chapter focuses on LSD enzymes that are already performed in NBS laboratories or piloted for NBS (see Sect. 21.7)
  - Most LSD enzyme assays are performed in leukocytes isolated from whole blood
    - LSD enzyme assay can also be performed in cultured skin fibroblasts
  - LSD enzymes are stable on DBS, and this makes DBS samples useful for LSD enzyme testing
- Preanalytical requirements
  - Peripheral blood collected in various anticoagulant-containing tubes (heparin, EDTA, or ACD); refrigerate blood and keep cool during shipment
    - The blood should be received and processed by the testing laboratory within 24–48 h after collection to ensure the quality of extracted leukocytes
  - Extracted leukocyte pellets; ship frozen on dry ice to the testing laboratory
  - Cultured skin fibroblasts; ship at room temperature
  - DBS, 1–2 circles, fully saturated; ship at room temperature
  - DBS specimen may vary in protein amount which cannot be normalized as in leukocytes, and presence of hemoglobin may interfere with the assay
- Analytical methods
  - Prior to enzymatic reaction
    - Isolation of leukocytes using standard dextran sedimentation method
    - Making cell lysates by sonication at low power using microprobe
    - Determination of protein concentration of lysates using Lowry assay
    - Enzymatic reaction
  - Enzyme reaction system (Table 21.5)
    - Citrate-phosphate or acetate buffer to acidic pH
    - Inhibitors of interfering isoenzymes
    - Detergents to solubilize the membrane-associated enzyme
    - Substrates

Fig. 21.5 Plasma acylcarnitine profile. Acylcarnitine (butyl esters) of various chain lengths are shown via precursor ion scan of *m/z* 85 from 200 to 550. Internal standards include D9-C0 (227.3), D3-C2 (263.3), D3-C3 (277.4), D3-C4 (288), D9-C5 (311.4), D3-C8 (347.4), D9-C14 (437.5), and D3-C16 (459.5). (**a**) Normal profile. (**b**) Short-chain acyl-CoA dehydrogenase deficiency (SCAD). C4 is elevated. (**c**) Medium-chain acyl-CoA dehydrogenase deficiency (MCAD). C8 is elevated with C6 and C10:1. (**d**) Very long-chain acyl-CoA dehydrogenase deficiency (VLCAD). C14:1 is elevated with C14:2, C14:0, and other long chains

Table 21.5 Fluorometric and radioactive LSD enzyme assay conditions

Diseases	Enzymes (abbreviation)	Substrates	Inhibitors or other additives	Detergent/others
Hurler syndrome	α-L-iduronidase (IDUA)	1.2 mM 4-MU-α-iduroside in H_2O		0.1 M sodium formate, pH 2.8
Gaucher disease	β-Glucocerebrosidase (GBA)	5.0 mM 4-MU-β-Glc in 0.2 M C-P buffer, pH 5.4		1% TC and 0.5% Triton-X
Fabry disease	α-Galactosidase A (GLA)	10.0 mM 4-MU-α-Gal in 0.2 M C-P buffer, pH 4.4	117 mM GalNAc	
Niemann-Pick type A/B	Acid sphingomyelinase (ASM)	0.25 mM ^{14}C-SM in 1.0 M Ac buffer, pH 5.0		1% TC and 0.5% Triton-X
Krabbe disease	β-Galactocerebrosidase (GALC)	0.25 mM ^{3} H-Gal-cer in 0.2 M C-P buffer, pH 4.2		1% TC and 0.5% Triton-X
Pompe disease	α-Glucosidase (GAA)	1.4 mM 4-MUG in 40 mM Ac, pH 3.8	8 μM acarbose	

C-P citrate-phosphate, 4-MU-β-Glc 4-MU-β-glucopyranoside, TC taurocholate, 4-MU-α-Gal 4-MU-α-galactoside, GalNAc N-acetylgalactosamine, ^{14}C-SM ^{14}C-sphingomyelin, Ac sodium acetate, ^{3}H-Gal-cer ^{3}H-galactosylceramide, 4-MUG 4-methylumbelliferyl- α-D-glucopyranoside

  - Different substrates used for selective LSD enzymes
    - 4-Methylumbelliferone (4-MU) artificial substrates (IDUA, GBA, GLA, and GAA)
    - Radiolabeled natural lipid substrates (ASM and GALC)
    - Proprietary synthetic structural analogs to the natural substrates and internal standards for tandem MS analysis (all six)
  - Fluorometric enzyme assays
    - Pros: simple, end-point assay
    - Cons: unable to multiplex as different enzyme assays release the same end product free 4-MU
  - Tandem mass spectrometry methods
    - Multiplex assay
    - Specific and accurate quantitation of products and substrates by LC-MS/MS from complex mixtures
  - Enzyme activity
    - Leukocytes or cultured cells: nmol/h/mg protein
    - DBS sample: nmol/h/punch
- Interpretation of results
  - Absent or significantly reduced (<10%) enzyme activity in most cases establishes the clinical diagnosis
    - Exception 1: >10% residual GBA activity in some Gaucher patients when 4-MU substrates are used
    - Exception 2: enzyme pseudodeficiency with 4-MU artificial substrates should be followed by molecular confirmation and clinical correlations
  - Testing laboratory should establish reference ranges for normal and affected
  - Carrier ranges vary and overlap with normal ranges

21.4 Amino Acid Disorders

21.4.1 Phenylketonuria

- PKU is one of the most common inborn errors of amino acid metabolism
  - It is an inherited deficiency of phenylalanine hydroxylase, which results in an inability to convert phenylalanine to tyrosine
  - Ingestion of dietary phenylalanine (Phe) results in elevated phenylalanine concentrations in the blood and brain, which are neurotoxic
- Clinical
  - Different metabolic phenotypes of hyperphenylalaninemia depend on

mutations and varying degrees of residual PAH activity which inversely correlates with Phe tolerance
  - Classic PKU: untreated Phe levels >1,000 μmol/L
  - Variant PKU: untreated Phe levels 600–1,000 μmol/L
  - Non-PKU hyperphenylalaninemia (non-PKU HPA): untreated Phe levels 120–600 μmol/L
- Classic PKU
  - Untreated PKU patients present with severe, irreversible mental retardation, epilepsy, "mousy" or musty odor, eczema, and hypopigmentation
  - Treated PKU patients have essentially normal intellect
    - Some may have lower IQs and abnormal neurocognitive function
    - Anxiety and depression are common despite adequate treatment
- Non-PKU HPA typically remains asymptomatic with no treatment or dietary restriction
- Maternal PKU syndrome (gestational hyperphenylalaninemia)
  - Microcephaly
  - Congenital heart disease
  - Intrauterine growth retardation
  - Craniofacial abnormalities

- Genetics
  - Classic PKU, variant PKU, and non-PKU HPA are allelic autosomal recessive disorders
  - The human *PAH* gene is located on 12q22–24.1 and contains 13 exons
  - More than 500 mutations have been identified in the *PAH* gene
  - An up-to-date summary of all known *PAH* mutations is kept at PAHdb (http://www.pahdb.mcgill.ca/)
- Diagnosis
  - Elevated plasma Phe and Phe/Tyr ratio
  - Evidence of normal biopterin metabolism
    - Normal BH4 loading test
    - Normal urinary pterins profile
    - Normal blood DHPR activity
- Management
  - Low-phenylalanine diet using phenylalanine-free, tyrosine-supplemented medical foods
  - Some patients respond to sapropterin dihydrochloride (Kuvan)
  - Phenylalanine ammonium lyase is in clinical trial

21.4.2 Tyrosinemia

- The tyrosinemias are a group of disorders of tyrosine metabolism
  - Type I (hepatorenal) is the most common and severe form and is caused by deficient activity of fumarylacetoacetate hydrolase (FAH)
  - Type II (oculocutaneous) is caused by cytosolic tyrosine aminotransferase (TAT) deficiency
  - Type III is caused by 4-hydroxyphenylpyruvate dioxygenase deficiency (4HPPD)
    - 4HPPD dysfunction also causes hawkinsinuria
  - Transient tyrosinemia of the newborn is caused by immaturity of 4HPPD
- Clinical
  - Type I: acute liver failure, cirrhosis, hepatocellular carcinoma, porphyria-like neurologic crises, renal Fanconi syndrome, and "rotten cabbage" odor
  - Type II: painful corneal erosions, palmoplantar keratosis, and mental retardation (~50%)
  - Type III: neurological abnormalities
  - Hawkinsinuria: metabolic acidosis, failure to thrive, and excretion of hawkinsin
- Genetics
  - With the exception of hawkinsinuria (autosomal dominant), the tyrosinemias are inherited in an autosomal recessive manner
  - Hepatorenal tyrosinemia (type I) due to mutation of *FAH* is more prevalent in the French-Canadian (common mutation: c.1062+5G>A [IVS12+5G>A]) and Finnish (common mutation: W262X) populations

  - Type II is caused by mutation of the *TAT* gene located on 16q22.1–q22.3
  - Type III is caused by mutation of the *HPD* gene located on 12q24–qter
- Diagnosis
  - Elevated plasma Tyr in types I, II, and III
  - Elevated 4-hydroxyphenyl derivatives in types I, II, and III
  - Type I
    - Elevated succinylacetone levels in plasma and urine
    - Elevated serum alpha-fetoprotein (AFP)
  - Hawkinsinuria
    - Presence of 4-hydroxycyclohexylacetic acid in urine organic acid profile in adult patients
    - 5-Oxoprolinuria in infants during metabolic acidosis
  - Molecular analysis
    - Sequence analysis for types II and III
- Management
  - Type I is managed using NTBC, a 4HPPD inhibitor, plus a phenylalanine- and tyrosine-restricted diet
  - Liver imaging for hepatocellular carcinoma surveillance in type I patients
  - Type II is managed with a phenylalanine- and tyrosine-restricted diet
  - Hawkinsinuria is treated with ascorbic acid and dietary protein restriction
  - Transient tyrosinemia of the newborn often improves spontaneously; some patients are treated with ascorbate and protein restriction

21.4.3 Maple Syrup Urine Disease (MSUD)

- MSUD, also known as branched-chain keto acid dehydrogenase complex (BCKAD) deficiency, is an inborn error of leucine, isoleucine, and valine catabolism
  - MSUD has five clinical and biochemical phenotypes
    - Classic form (most common)
    - Intermediate form
    - Intermittent form
    - Thiamine-responsive form
    - Dihydrolipoyl dehydrogenase (E3)-deficient form
- Clinical
  - Classic MSUD
    - Neonatal onset, mostly between 4 and 7 days of age, poor feeding, lethargy, metabolic acidosis, and hyperleucinemic encephalopathy, which can progress to seizures, coma, and death
    - The odor of burnt sugar/maple syrup is typically present in the urine, cerumen, and sweat and is associated with hyperisoleucinemia/hyperisoleucinuria
    - Recurrent metabolic decompensations are common
  - Intermediate MSUD patients have persistent elevations of the BCAA, but usually do not have severe metabolic decompensation
  - Intermittent MSUD patients can have acute metabolic decompensation, with normal BCAA and neurological status between episodes
  - Thiamine-responsive MSUD is clinically similar to intermediate MSUD and responds to pharmacological doses of thiamine
  - Dihydrolipoyl dehydrogenase (E3)-deficient MSUD manifests after neonatal period with hypoglycemia, lactic acidosis, developmental delay, and progressive neurological deterioration
    - E3 is a subunit in BCKAD, pyruvate dehydrogenase, and alpha-ketoglutarate dehydrogenase complex
- Genetics
  - MSUD is an autosomal recessive panethnic disorder
  - Mutations in any of four genes can cause MSUD
    - *BCKDHA* (19q13.1–q13.2) encoding the E1α subunit
    - *BCKDHB* (6q14.1) encoding the E1β subunit
    - *DBT* (1p31) encoding the dihydrolipoyl transacylase (E2) subunit
    - *DLD* (7q31–q32) encoding the dihydrolipoyl dehydrogenase (E3) subunit

  - The worldwide frequency of classic MSUD is estimated to be 1 in 185,000
  - The incidence of classic MSUD is higher in Old Order Mennonite population of Lancaster and Lebanon Counties, Pennsylvania (1 in 176), and in Ashkenazi Jews (~1 in 28,000)
  - Dihydrolipoyl dehydrogenase (E3) deficiency is also slightly more prevalent in the Ashkenazi Jewish population
- Diagnosis
  - Elevation of plasma leucine, isoleucine, alloisoleucine, and valine
    - Alloisoleucine *is a reliable diagnostic marker for MSUD*
  - Elevation of urinary excretions of 2-ketoisocaproic acid, 2-keto-3-methylvaleric acid, and 2-ketoisovaleric acid along with 2-hydroxyisovaleric acids
  - Deficient BCKAD activity in cultured fibroblast
  - Identification of mutations in related genes by sequence analysis for variant forms
- Management
  - Treatment is a branched-chain amino acid-restricted diet with prevention of isoleucine and valine deficiencies
  - Liver transplantation is an alternative treatment option

21.4.4 Homocystinuria

- Cystathionine β-synthase (CBS) deficiency is the most frequently encountered cause of isolated homocystinuria
  - CBS catalyzes the first step of the transsulfuration of homocysteine to cystathionine
  - Defects of cobalamin metabolism can also lead to isolated homocystinuria (cobalamin E and G diseases)
  - The CBS enzyme requires pyridoxine (vitamin B6) and heme as cofactors
- Clinical
  - Characteristic features of untreated CBS deficiency include mental retardation, thromboembolism, myopia, ectopia lentis, and a marfanoid habitus (genu valgum, scoliosis, pectus excavatum/carinatum)
  - Other less consistent features include risk for seizures, malar flush, dystonia, hypopigmentation, and psychiatric disease
  - There is significant clinical variability for all of the clinical features
  - Pyridoxine responsiveness generally indicates milder disease
- Genetics
  - CBS deficiency is inherited as an autosomal recessive trait
  - The *CBS* gene is located on chromosome 21q22 and contains 23 exons
  - Of >300 described mutants, two are common
    - c.833T>C (p.I278T) – pyridoxine responsive
    - c.919G>A (p.G307S) – pyridoxine nonresponsive
- Diagnosis
  - Elevated plasma total homocysteine and methionine
  - Pyridoxine-responsive variants may be missed by newborn screening
- Management
  - Determination of pyridoxine responsiveness by pyridoxine challenge
  - Dietary methionine restriction for pyridoxine nonresponders using low-methionine, cystine-supplemented medical formulas
  - Control of homocysteine levels with use of betaine
  - Folate and B_{12} supplementation to maximize conversion of homocysteine to methionine

21.4.5 Urea Cycle Disorders

- Primary urea cycle disorders result in accumulation of ammonia secondary to a defect in one of six enzymes or two transporter molecules involved in urea cycle function (Fig. 21.6)
  - Based on the nature of the defect, characteristic deficiencies or accumulation of urea cycle intermediates permits diagnosis and

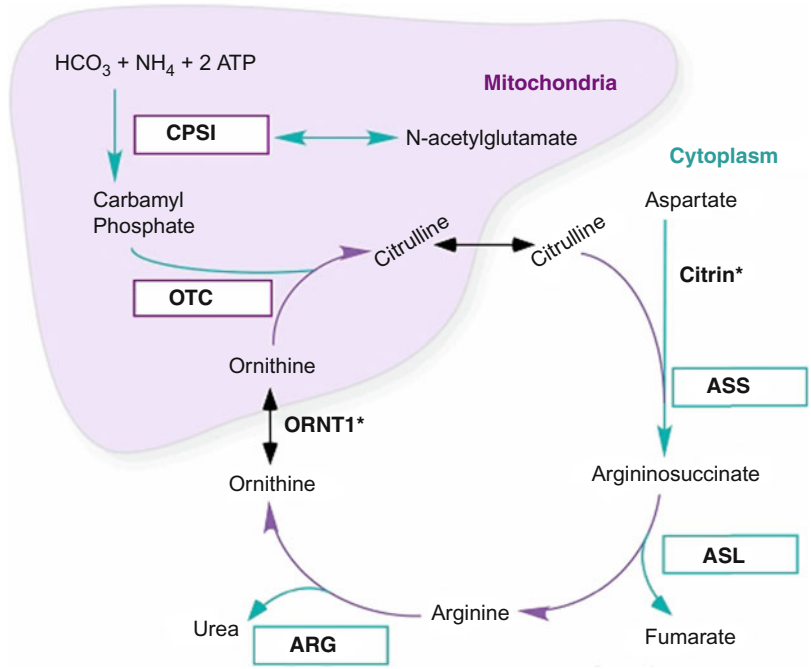


Fig. 21.6 The urea cycle. *CPS I* carbamoyl phosphate synthase I, *OTC* ornithine transcarbamylase, *citrin* mitochondrial citrulline transporter, *ASS* argininosuccinate synthase, *ASL* argininosuccinate lyase, *ARG* arginase, *ORNT1* mitochondrial ornithine transporter (Reprinted from Fig. 1, Urea Cycle Disorders Overview, www.genetests.org, Copyright University of Washington, Seattle)

treatment of most of the disorders prior to molecular confirmation

- Clinical
  - Hyperammonemia – more severe in "proximal" disorders (NAGS, CPS1, OTC deficiencies) vs "distal" disorders
  - Manifestations of acute hyperammonemic crises may include vomiting, neurologic compromise, liver dysfunction, and stroke
  - Female OTC carriers may manifest symptoms early in life or later during catabolic stress (i.e., postpartum)
  - See Table 21.6 for disease-specific features of the various UCD
  - Higher prevalence of SLC25A13 mutation (citrin deficiency) in Asian populations presenting with two distinct clinical phenotypes
    - Adult onset: hyperammonemia with neuropsychiatric symptoms, fatty liver, or mild fibrosis
    - Neonatal onset: intrahepatic cholestasis and liver fibrosis/dysfunction
- Genetics
  - All UCD except OTC (X-linked) are inherited as autosomal recessive traits
  - Cumulative incidence is estimated at approximately 1:30,000
  - HHH syndrome principally reported in French-Canadian population
- Diagnosis
  - Elevated blood ammonia
  - Accumulation of characteristic amino acids (see Table 21.2)
  - Elevated urinary orotic acid in OTC and other UCD and negative in NAGS and CPS1 deficiency
  - Molecular analysis – necessary to distinguish NAGS from CPS1 deficiency and useful in female OTC
- Management
  - Dietary restriction of natural protein and essential amino acid formula to decrease nitrogen burden
  - Supplementation of arginine (ASS, ASL deficiency) or citrulline (NAGS, CPS1, OTC)

Table 21.6 Summary of urea cycle defects

Disorder	Inheritance	Gene	Disease-specific features
N-Acetylglutamate synthase deficiency (NAGS)	Autosomal recessive	*NAGS* 17q21, 7 exons	Activator defect, urea cycle enzyme activities normal
Carbamoyl phosphate synthase I deficiency (CPS I)	Autosomal recessive	*CPS1* 2q34, 38 exons	
Ornithine transcarbamylase deficiency (OTC)	X-linked recessive	*OTC* Xp11, 10 exons	Female carriers may be symptomatic
Citrullinemia (CIT)	Autosomal recessive	*ASS1* 9q34, 16 exons	
Argininosuccinic aciduria (ASL)	Autosomal recessive	*ASL* 7q11, 16 exons	Predisposition to hepatic cirrhosis, renal tubular disease
Arginase deficiency (ARG)	Autosomal recessive	*ARG1* 6q23, 8 exons	Spastic diplegia of lower extremities
Hyperammonemia hyperornithinemia homocitrullinemia syndrome (HHH)	Autosomal recessive	*SLC25A15* 13q14, 7 exons	
Citrullinemia 2 (CIT II) or citrin deficiency	Autosomal recessive	*SLC25A13* 7q21, 18 exons	Distinct neonatal and adult forms

  - Ammonia scavenger therapy (sodium phenylbutyrate, sodium benzoate)
  - For NAGS deficiency, a synthetic form of N-acetylglutamate, carbamyl glutamate (Carbaglu), activates downstream urea cycle function
  - Protein-rich, low-carbohydrate diet for citrin deficiency
  - Liver transplantation

21.5 Organic Acidemias

21.5.1 Propionic Acidemia

- Propionyl coenzyme A, which is produced by the catabolism of specific amino acids (isoleucine, valine, methionine, threonine), odd-chain fatty acids, and cholesterol, accumulates in cases of deficiency of propionyl-CoA carboxylase, a biotin-dependent enzyme
- Clinical
  - Originally referred to as "ketotic hyperglycinemia"
  - Crises are characterized by severe metabolic ketoacidosis, with elevated lactate and ammonia as secondary perturbations
  - Other manifestations of acute crises include vomiting, neurologic compromise, bone marrow suppression leading to pancytopenia, and risk for developing pancreatitis
  - Increased risk for strokes, particularly of the basal ganglia, and cerebellar hemorrhage
  - Cardiomyopathy can occur
  - There is evidence that cardiac rhythm disturbances may occur as patients become older
  - Despite early therapy, patients with more severe disease will have significant developmental disabilities
- Genetics
  - Propionic acidemia is inherited as an autosomal recessive trait
  - Propionyl-CoA carboxylase is composed of six α and six β subunits (α6β6)
  - The α subunit gene, *PCCA*, contains 23 exons and is located on chromosome 13q32
  - The β subunit gene, *PCCB*, contains 15 exons and is on chromosome 3q22
- Diagnosis
  - Characteristic plasma amino acids, acylcarnitine, and urine organic acid pattern (Table 21.2)
  - Enzyme assay and molecular diagnosis might be needed for variant form

- Management
  - Dietary restriction of propiogenic amino acids and supplementation with medical formula
  - Oral antibiotic therapy to reduce gut propionate production
  - Carnitine supplementation
  - Bicarbonate replacement as needed to maintain acid–base balance
  - Transplantation (liver) has been tried in cases refractory to medical management

21.5.2 Methylmalonic Acidemia (MUT Deficiency)

- Methylmalonyl coenzyme A, which is produced by the catabolism of specific amino acids (isoleucine, valine, methionine, threonine), odd-chain fatty acids, and cholesterol, accumulates in cases of deficiency of methylmalonyl-CoA mutase, an adenosylcobalamin (vitamin B12)-dependent enzyme
  - Methylmalonyl-CoA is the product of the reaction catalyzed by propionyl-CoA carboxylase
  - Isolated methylmalonic acidemia can also be seen in defects of cobalamin metabolism that impair the production of adenosylcobalamin (cbl A, B, and D2 variant)
  - Cobalamin metabolic defects that impair both adenosylcobalamin and methylcobalamin production will lead to combined methylmalonic acidemia and homocystinuria (cbl C, D, F)
- Clinical
  - Clinical features of overlap those seen with propionic acidemia
  - Patients are characterized on basis of no (MUT^0) or some (MUT^-) residual functional mutase activity
  - Renal failure in older patients due to progressive interstitial nephritis
- Genetics
  - Methylmalonic acidemia is inherited as an autosomal recessive trait
  - Methylmalonyl-CoA mutase is a dimer of identical subunits (α2)
  - The gene encoding methylmalonyl-CoA mutase, *MUT*, contains 13 exons and is located on chromosome 6p21
- Diagnosis
  - Characteristic plasma amino acids, acylcarnitine, and urine organic acid pattern (Table 21.2)
  - Differential diagnosis from cobalamin disorders may require complementation study or enzyme analysis or mutational analysis
- Management
  - Cobalamin supplementation for vitamin B_{12}-responsive forms
  - Dietary restriction of propiogenic amino acids and supplementation with medical formula
  - Oral antibiotic therapy to reduce gut propionate production
  - Carnitine supplementation
  - Bicarbonate replacement as needed to maintain acid–base balance
  - Transplantation (liver and combined kidney/liver) has been tried in cases refractory to medical management

21.5.3 3-Methylcrotonyl-CoA Carboxylase Deficiency (3-MCC)

- 3-Methylcrotonyl-CoA carboxylase catalyzes the conversion of 3-methylcrotonyl-CoA to 3-methylglutaconyl-CoA in the leucine metabolic pathway
  - Most cases described prior to the advent of expanded newborn screening were detected after evaluation for mental retardation or for Reye-like illness, hypoglycemia, or ketoacidosis
  - Currently, most infants detected by newborn screening have been asymptomatic
- Clinical
  - Broad clinical spectrum – asymptomatic to severe Reye-like crises, severe ketoacidosis, and elevated lactate and ammonia
  - Other manifestations – weakness/fatigue, neurologic damage, myopathy (including cardiomyopathy), and hepatic fatty infiltration

  - Crises may be triggered by catabolic stress
  - Most patients detected by newborn screening have been asymptomatic but cases described of asymptomatic infants later manifesting symptoms
  - Secondary carnitine deficiency is common
  - Maternal 3-MCC is identified in infants with positive C5-OH screens
- Genetics
  - 3-MCC is inherited as an autosomal recessive trait
  - 3-Methylcrotonyl-CoA carboxylase is a heterodimer of α and β subunits
  - The gene encoding *MCCC1* is located on chromosome 3q27 and contains 19 exons
  - The gene encoding *MCCC2* is located on chromosome 5q13 and contains 17 exons
  - Incidence on newborn screening is approximately 1:36,000
- Diagnosis
  - Urine organic acids to exclude other disorders with similar acylcarnitine profile (beta-ketothiolase, HMG-CoA lyase, biotinidase/holocarboxylase deficiencies; 2-methyl, 3-hydroxybutyric, or 3-methylglutaconic aciduria)
  - Molecular analysis possible if diagnosis uncertain
  - Enzyme assay for definitive diagnosis possible with skin fibroblasts
  - Test mothers of infants detected by newborn screening to exclude maternal 3-MCC
- Management
  - Surveillance in apparently well children detected by newborn screening with caloric support during catabolic stress and illness-induced fasting
  - Dietary protein restriction if symptomatic
  - Leucine restriction with medical formula if necessary
  - Carnitine supplementation as needed

21.5.4 Glutaric Acidemia Type I (GA I)

- GA I is a disorder caused by deficiency of glutaryl-CoA dehydrogenase resulting in the accumulation and excretion of glutaric and 3-hydroxyglutaric acids produced by impaired metabolism of the glutarigenic amino acids lysine and tryptophan
- Clinical
  - Characterized as a "cerebral" organic aciduria
  - Crises can cause neurological injury – dystonia and dyskinesia
  - Pronounced macrocephaly
  - Striatal degeneration of the caudate and putamen
  - Frontotemporal atrophy, arachnoid cysts, or both before the onset of symptoms
  - Striatal damage and neurologic phenotype do not develop in all patients
  - Presentations with chronic subdural effusions may be mistaken for shaken baby syndrome
- Genetics
  - GA I is inherited as an autosomal recessive trait
  - The glutaryl-CoA dehydrogenase gene (*GCDH*) is located on chromosome 19p13.2 and contains 11 exons
  - In the Old Order Amish in Lancaster County, Pennsylvania, a single founder allele encoding the mutant protein A421V is present at high prevalence
  - In outbred populations, most GA I patients are heterozygous for two different mutant alleles
  - Worldwide incidence is currently estimated at 1:100,000
- Diagnosis
  - Characteristic plasma acylcarnitine and urine organic acid pattern (Table 21.2)
    - 3-Hydroxyglutaric acid is a reliable diagnostic marker for GA I
    - Low-excretor biochemical phenotype exists
  - Enzymatic and molecular testing for the low-excretor variants
  - Prenatal diagnosis is possible and is based on increased concentrations of glutaric acid in amniotic fluid
- Management
  - Supplementation with L-carnitine
  - Dietary restriction of lysine and tryptophan

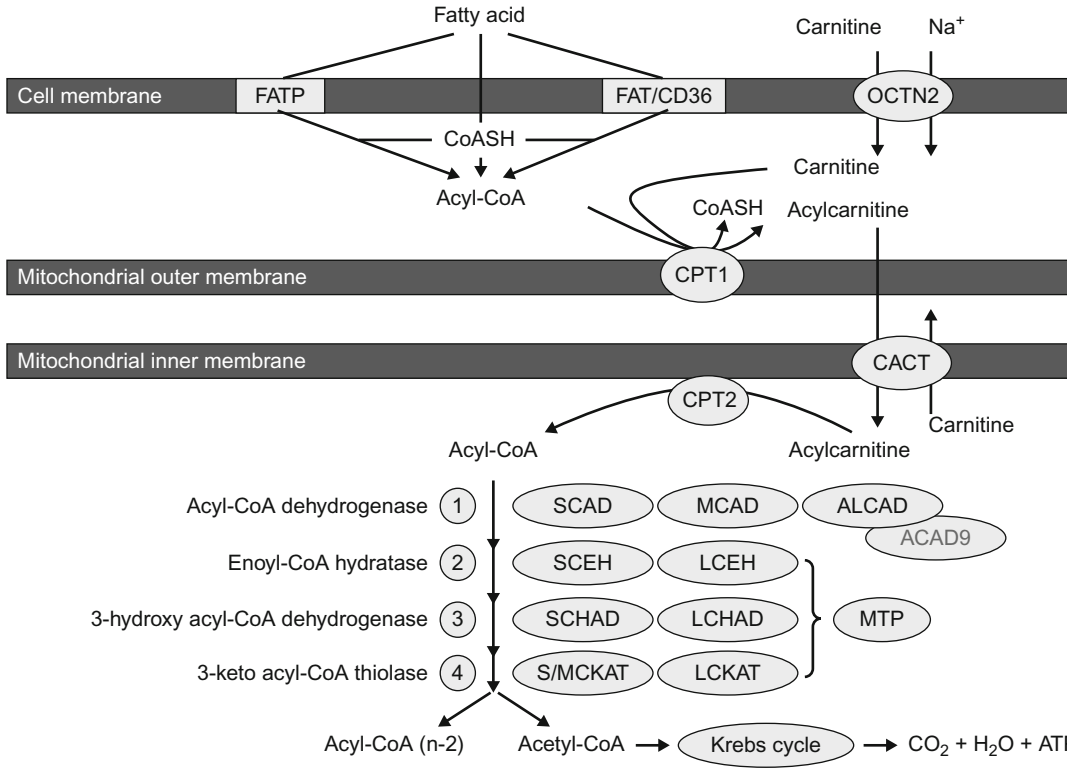


Fig. 21.7 Schematic representation of the mitochondrial fatty acid beta-oxidation pathway. The pathway starts with the uptake of FAs and carnitine from the plasma compartment into the cell, followed by the transport of acyl-CoA esters into the mitochondria via the carnitine cycle and the actual stepwise degradation of acyl-CoAs via the beta-oxidation spiral with acetyl-CoA units as end product, which can either be converted into ketone bodies or combusted in the citric acid cycle to CO_2 and H_2O (Reprinted with permission from Fig. 1, Wanders et al. 2010)

- Aggressive treatment of intercurrent infections with intravenous fluids, glucose and insulin, and dietary restriction
- There is evidence that following recommended treatment measures (dietary restriction, carnitine supplementation, sick day anabolic support) decreases the risk of neurological crises by up to 90%

21.6 Fatty Acid Disorders

- Primary fatty acid oxidation disorders result in accumulation of short-, medium-, or long-chain triglycerides secondary to a defect in one of carnitine transport enzymes involved in carnitine-facilitated fatty acid or one of the inner mitochondrial membrane-associated or mitochondrial matrix enzymes involved in fatty acid oxidation (Fig. 21.7)
  - Representative fatty acid oxidation disorders are summarized in Table 21.7
  - Based on the nature of the defect, characteristic accumulations of fatty acid metabolites or carnitine species permit diagnosis and treatment of most of the disorders prior to molecular confirmation
- Clinical
  - Manifestations principally seen after prolonged fasting and during illness or metabolic stress
  - May present with hypoketotic hypoglycemia and/or dysfunction of liver (Reye-like presentation), muscle (rhabdomyolysis, cardiomyopathy, arrhythmia), or nervous system (encephalopathy, seizures)

Table 21.7 Summary of common fatty acid oxidation disorders

Disorder	Inheritance	Gene	Disease-specific features
Short-chain acyl-CoA dehydrogenase deficiency (SCAD)	Autosomal recessive	*ACADS* 12q12, 10 exons	Most apparently asymptomatic
Medium-chain acyl-CoA dehydrogenase deficiency (MCAD)	Autosomal recessive	*ACADM* 1p31, 12 exons	Significant clinical variability even within families
Very long-chain acyl-CoA dehydrogenase deficiency (VLCAD)	Autosomal recessive	*ACADVL* 17p31, 20 exons	Rhabdomyolysis frequent in adult-onset form
Long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency (LCHAD)	Autosomal recessive	*HADHA* 2p23, 20 exons *HADHB* 2p23, 16 exons	Pigmentary retinopathy Maternal HELLP syndrome
Carnitine uptake defect (CUD)	Autosomal recessive	*SLC22A5* 5q31, 10 exons	Maternal cases may present on newborn screening
Carnitine palmitoyltransferase II deficiency (CPT II)	Autosomal recessive	*CPT2* 1p32, 5 exons	Lethal neonatal or adult-onset forms

  - Nearly 20% of undiagnosed MCAD may initially manifest with sudden death
  - LCHAD is more likely to have cardiac and hepatic involvement
  - Maternal cases of CUD have been identified on evaluation of abnormal newborn screens
  - Other disease-specific features of the various FAO/carnitine metabolic defects include
    - Maternal pregnancy complications, particularly LCHAD, when carrying affected fetus – HELLP syndrome (hemolysis, elevated liver enzymes, low platelets) and acute fatty liver of pregnancy (AFLP)
    - LCHAD is associated with pigmentary retinopathy
    - CPT II deficiency can present as a lethal neonatal form with dysmorphia and renal cysts, a severe late infantile form with peripheral and cardiomyopathy and hepatic failure, or an adult-onset form with myopathy and rhabdomyolysis
- Genetics
  - Summarized in Table 21.7
  - Common mutations are observed in
    - MCAD – *ACADM* c.985A>G (p.K329E)
    - LCHAD – *HADHA* c.1528G>C (p.E510Q)
    - CPT II – *CPT2* c.338C>T (p.S113L)
  - MCAD is the most common FAO with an estimated incidence of 1:10,000 Caucasians and 1:17,000 in the USA
- Diagnosis
  - Characteristic urine organic acids (including acylglycine) and acylcarnitine profiles (Table 21.2)
  - Molecular testing
  - Fibroblast fatty acid oxidation probe or enzyme assay
- Management
  - Avoidance of prolonged fasting and low-fat diet
  - Medium-chain triglyceride supplementation for VLCAD and LCHAD
  - Carnitine supplementation in CUD, as needed in MCAD

21.7 Lysosomal Storage Diseases

21.7.1 Hurler Syndrome

- Hurler syndrome, also known as mucopolysaccharidosis (MPS) type I, is a progressive, multisystem disorder of glycosaminoglycan degradation
  - It is caused by deficient activity of alpha-L-iduronidase

  - There are three clinical types that range in severity
    - Type 1H (Hurler – most severe)
    - Type 1S (Scheie – milder)
    - Type 1H/S (intermediate)
- Clinical
  - Type IH
    - Inguinal or umbilical hernias
    - Organomegaly – liver and spleen
    - Dysmorphia – coarse facial features and macroglossia
    - Skeletal involvement – dysostosis multiplex and short stature
    - Coronary artery disease and progressive cardiac failure
    - Mental retardation and communicating hydrocephalus
    - Corneal clouding
    - Hearing loss
    - Frequent respiratory tract and nasal infections
    - The average life expectancy 10 years due to obstructive airway disease, heart failure, and infection
  - Type IS
    - Normal intelligence and stature
    - Joint stiffness
    - Corneal clouding
    - Valvular thickening
- Genetics
  - MPS IH, IS, and IH/S are allelic autosomal recessive disorders
  - Incidence of Hurler syndrome 1/76,000–1/200,000
  - Hurler-Scheie and Scheie are less common
  - The alpha-L-iduronidase gene, *IDUA*, is located on 4p16.3
  - Three mutations, p.W402X, p.Q70X, and p.P533R, account for about one-half of the mutant alleles in the Caucasian MPS IH population
- Diagnosis
  - Deficient activity of alpha-L-iduronidase in WBC or fibroblasts
  - Urinary excretion of heparan and dermatan sulfate
- Management
  - Enzyme replacement therapy (ERT) with recombinant human alpha-L-iduronidase is available for patients with MPS IH and IH/S
  - Bone marrow transplantation is another treatment option
  - Supportive treatment of hydrocephalus, joint stiffness, airway obstruction, and valvular heart disease

21.7.2 Gaucher Disease

- Gaucher disease is caused by deficient activity of glucocerebrosidase, which results in the intralysosomal accumulation of glycosylceramide in cells of the monocyte/macrophage system
  - There are three subtypes of Gaucher disease
    - Type 1 is nonneuronopathic and the most common form
    - Type 2 is a fatal neurodegenerative disorder of infancy
    - Type 3 has a severity intermediate between type 1 and type 2
- Clinical
  - Type 1
    - Hepatosplenomegaly
    - Hypersplenism can cause clinically significant thrombocytopenia and mild anemia
    - Skeletal disease (bony lesions, episodic painful bone crises, Erlenmeyer flask deformity of the distal femur)
    - Pulmonary infiltration
    - Absence of primary central nervous system disease
  - Type 2
    - Massive hepatosplenomegaly
    - Oculomotor abnormalities
    - Progressive neurologic abnormalities
    - Most infants with type 2 disease die by 2 years of age
  - Type 3
    - Hepatosplenomegaly
    - Neurologic manifestations at a later age of onset compared to type 2

- Genetics
  - Type 1 Gaucher disease is panethnic but is more common among Ashkenazi Jews
  - There are four common Ashkenazi mutations in the *GBA* gene which account for 92–95% mutations in Ashkenazi Jewish patients with Gaucher disease
    - p.N409S (previously known as N370S)
    - c.84dupG (previously known as 84GG)
    - c.115+1G>A (previously known as IVS2+1G>A)
    - p.L483P (previously known as L444P)
  - Type 3 Gaucher disease is more common among Norrbottnian Swedes
  - Presence of *GBA* mutation is associated with increased risk for Parkinson disease
- Diagnosis
  - Deficient β-glucosidase activity in WBC
  - Elevated serum/plasma chitotriosidase activity which is recommended for monitoring response to therapy
    - 40% of Caucasians are homozygotes or heterozygotes of a null CHIT1 allele (dup24)
  - Elevated serum/plasma PARC/CCL18
  - Mutation analysis (see Chap. 22)
- Management
  - Intravenous ERT is available for patients with type 1 Gaucher disease
  - Substrate deprivation therapy with an oral glycosidase is another treatment option
  - Supportive treatment of osteopenia with bisphosphonates
  - Avoidance of contact sports to minimize risk of splenic rupture

21.7.3 Niemann-Pick Disease Types A and B

- Types A and B Niemann-Pick disease are allelic disorders caused by deficient activity of acid sphingomyelinase, which normally converts sphingomyelin to ceramide
  - Niemann-Pick cells are lipid-laden macrophages found in the bone marrow, liver, lungs, and spleen of all patients, as well as the brains of type A patients
  - Type A Niemann-Pick disease (NPD-A) is a progressive neurodegenerative disorder of infancy
  - Type B disease (NPD-B) is a nonneuronopathic disorder with a broad range of phenotypic severity
  - An intermediate phenotype has the features of type B with a slowly progressive neurodegenerative course
- Clinical
  - Type A
    - Massive hepatosplenomegaly
    - Neurodegeneration: development plateaus and then regresses at approximately 9 months of age
    - Retinal cherry-red spots
    - Gastroesophageal reflux
    - Interstitial lung disease
    - Progressive liver dysfunction
    - Failure to thrive
    - Death usually by age 3 years, most commonly of respiratory failure
  - Type B
    - Hepatosplenomegaly during childhood
    - Restrictive lung disease due to progressive infiltration of the lungs with sphingomyelin
    - Elevated total cholesterol, LDL cholesterol, and triglycerides and low HDL cholesterol
    - Liver dysfunction
    - Children have restricted growth and delayed skeletal age and onset of puberty, but most reach normal adult height
    - Cherry-red spots in up to 20% of NPD-B patients despite absence of neurological symptoms
- Genetics
  - The gene encoding acid sphingomyelinase, *SMPD1*, is located on chromosome 11p15
  - NPD-A is a panethnic disorder, but it is more common among Ashkenazi Jews
    - Common Ashkenazi mutations include
      - p.L304P (previously known as L302P)
      - p.R498L (previously known as R496L)

  - c.996delC (previously known as fsp330)
  - NPD-B is panethnic
    - The most common NPD-B mutations are p.R610del (previously known as p.delR608) which has a neuroprotective effect
  - The p.Q294K (previously known as Q292K) mutation is associated with the intermediate phenotype of NPD
- Diagnosis
  - Enzyme assay: fluorometric assay (see Sect. 21.3, Lysosomal Enzyme Assays)
  - Elevated chitotriosidase
  - Mutation analysis (see Chap. 22)
- Management
  - At present, there is no specific treatment for either type A or type B NPD
  - Treatment for NPD-A is supportive
  - Clinical trial of ERT with human recombinant acid sphingomyelinase for NPD-B
  - Supportive care for NPD-B
    - Pharmaceutical management of dyslipidemia
    - Nutritional support for severely affected children
    - Avoidance of contact sports to minimize risk of splenic rupture

21.7.4 Fabry Disease

- Fabry disease is an X-linked disorder of glycosphingolipid catabolism caused by deficient activity of alpha-galactosidase A
  - Glycosphingolipids with terminal alpha-galactosyl moieties, primarily globotriaosylceramide (GL-3), accumulate in the lysosomes of endothelial, perithelial, and smooth muscle cells, as well as cells in the kidneys, heart, eyes, and ganglion cells
- Clinical
  - Typically presents in males during childhood or adolescence
    - Burning pain in the extremities (acroparesthesias)
    - Angiokeratomas in skin and mucous membranes
    - Hypohidrosis
    - Lenticular changes and corneal dystrophy
    - Renal – proteinuria and renal insufficiency to overt renal failure
    - Cardiac – left ventricular enlargement, valvular involvement, conduction abnormalities, angina pectoris, congestive heart failure, and myocardial infarction
    - Cerebrovascular – cerebral thromboses and hemorrhage
    - Gastrointestinal – episodic diarrhea and abdominal pain
  - Fabry "cardiac" variant
    - No classical systemic manifestations
    - Disease is limited to the heart and, to a lesser extent, kidney
  - Female heterozygotes can be asymptomatic to as severely affected as males
- Genetics
  - The *GLA* gene which encodes alpha-galactosidase A is located at Xq22.1
  - The prevalence of Fabry disease is about 1 in 40,000–60,000 males
  - The c.639+919G>A (previously known as IVS4+919G>A) mutation is associated with the cardiac variant in the Taiwan Chinese population
- Diagnosis
  - Deficient α-galactosidase A (α-Gal A) enzyme activity in plasma, WBC, and/or cultured cells
  - Only 60–70% of female Fabry patients demonstrate enzyme deficiency
  - Gb3 and lyso-Gb3 can be used as biomarker
  - Mutational analysis for females and variant forms
- Management
  - ERT – agalsidase alpha or agalsidase beta
  - Clinical trials of the safety and efficacy of pharmacologic chaperone therapy are underway
  - Supportive therapy for cardiac and renal disease
  - Pain management

21.7.5 Krabbe Disease

- Krabbe disease, also known as globoid cell leukodystrophy, is an autosomal recessive disorder neuronopathic lysosomal storage disease caused by deficient activity of galactocerebroside β-galactosidase (GALC)
- Clinical
  - Infantile Krabbe disease
    - Irritability and hypersensitivity before 6 months of age
    - Rapid progression to spasticity, dysphagia, and mental deterioration
    - In later stages of disease – flaccid, hypotonic, blind, and deaf
    - Death usually occurs by age 2 years
  - Later onset forms
    - The clinical manifestations and age of presentation are variable
      - Late infantile (age of onset 6 months to 3 years)
        - Irritability, developmental regression, ataxia, and visual loss
        - Rapidly progressive neurodegenerative course
      - Juvenile (age of onset 3–8 years of age)
        - Slow progression ataxia, visual loss, and psychomotor regression
        - Progressive neurodegenerative course
      - Adult onset (age of onset variable, up to sixth decade of life)
        - Abnormal gait, tremors, and visual loss
        - Slow, progressive neurologic deterioration
- Genetics
  - The *GALC* gene is located at chromosome 14q31 and contains 17 exons
  - Common mutations
    - 30 kb deletion starting near the middle of intron 10
    - c.1586C>T, c.1700A>C: 10–15% of mutations in infantile patients of European ancestry
  - Common polymorphisms in the *GALC* gene which affect enzyme activity
    - c.1685T>C (p.I562T): allele frequency 40–45%
    - c.550C>T (p.R184C): allele frequency 4–5%
    - c.742G>A (p.D248N): allele frequency 8–10%
- Diagnosis
  - Deficient GALC activity in DBS, WBC, or cultured fibroblasts
  - Mutation analysis
- Management
  - Hematopoietic stem cell transplantation (HSCT) with umbilical cord blood is an option for newborns with infantile Krabbe disease
  - Transplantation ideally before 6 weeks of age when asymptomatic
  - Transplanted symptomatic infants have had progressive deterioration

21.7.6 Pompe Disease

- Pompe disease is an autosomal recessive disorder of glycogen metabolism caused by deficient activity of lysosomal acid maltase (acid α-glucosidase)
  - There are several clinical phenotypes of Pompe disease with a broad range of severity (classic infantile Pompe disease; nonclassic infantile Pompe presents without cardiomegaly, juvenile onset, and adult onset)
- Clinical
  - Classic infantile disease: presents in the first few months of life with hypotonia, cardiomegaly, severe muscle weakness, macroglossia, hepatomegaly, and death before age 1 year of cardiorespiratory failure
  - Nonclassic infantile Pompe: hypotonia and weakness and no cardiomegaly
  - Juvenile onset disease: later age of presentation, slower progression of muscle weakness, less frequent cardiomegaly, and macroglossia
  - Adult onset disease: often presents as a proximal myopathy, cardiomegaly not typical, and progressive respiratory muscle failure

- Genetics
  - The α-glycosidase gene, *GAA*, resides at 17q25.2–q25.3
  - The estimated frequency of infantile onset disease ranges from 1/43,000–1/500,000
    - Infantile onset disease is panethnic
    - There is a higher frequency in African-Americans
      - The common African-American mutations are R854X and M318T
    - The common Chinese mutation is p.Asp645Glu
    - The estimated frequency of adult onset disease is approximately 1/57,000
      - The common Caucasian adult onset mutation is a splice site mutation: C. –32–13T>G (previously known as IVS 1–13T>G)
- Diagnosis
  - Deficient acid alpha-glucosidase (GAA) activity in WBC or DBS or cultured cells
  - Elevated urinary Hex4, an oligosaccharide that is used as a diagnostic biomarker
- Management
  - ERT – recombinant human alglucosidase alfa in infantile, juvenile, and adult onset disease
  - Respiratory support is often required
  - Physical therapy to maintain muscle strength is an adjunct to treatment

Further Reading

Calonge N, Green NS, Rinaldo P, Lloyd-Puryear M, Dougherty D, Boyle C, et al. Committee report: method for evaluating conditions nominated for population-based screening of newborns and children. Genet Med. 2010;12:153–9.

Dietzen DJ, Rinaldo P, Whitley RJ, Rhead WJ, Hannon WH, Garg UC, et al. National academy of clinical biochemistry laboratory medicine practice guidelines: follow-up testing for metabolic disease identified by expanded newborn screening using tandem mass spectrometry; executive summary. Clin Chem. 2009;55:1615–26.

Duffey TA, Bellamy G, Elliott S, Fox AC, Glass M, Turecek F, et al. A tandem mass spectrometry triplex assay for the detection of Fabry, Pompe, and mucopolysaccharidosis-I (Hurler). Clin Chem. 2010a;56:1854–61.

Duffey TA, Sadilek M, Scott CR, Turecek F, Gelb MH. Tandem mass spectrometry for the direct assay of lysosomal enzymes in dried blood spots: application to screening newborns for mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). Anal Chem. 2010b;82:9587–91.

Enns GM, Steiner RD, Cowan TM. Lysosomal disorders. In: Sarafoglou K, Hoffmann GF, Roth KS, editors. Pediatric endocrinology and inborn errors of metabolism. New York: McGraw-Hill; 2009. p. 721–55.

Garg U, Dasouki M. Expanded newborn screening of inherited metabolic disorders by tandem mass spectrometry: clinical and laboratory aspects. Clin Biochem. 2006;39:315–32.

Held PK, White L, Pasquali M. Quantitative urine amino acid analysis using liquid chromatography tandem mass spectrometry and aTRAQ reagents. J Chromatogr B Analyt Technol Biomed Life Sci. 2011;879:2695–703.

Howell RR, Lloyd-Puryear MA. From developing guidelines to implementing legislation: actions of the US Advisory Committee on Heritable Disorders in Newborns and Children toward advancing and improving newborn screening. Semin Perinatol. 2010;34:121–4.

Lanpher BC, Gropman A, Chapman KA, Lichter-Konecki U, Urea Cycle Disorders Consortium, Summar ML. Urea cycle disorders overview. In: Pagon RA, Bird TD, Dolan CR, Stephens K, editors. Gene reviews. Seattle: Published online by University of Washington; 1993. http://www.ncbi.nlm.nih.gov/books/NBK1217/

Magera MJ, Gunawardena ND, Hahn SH, Tortorelli S, Mitchell GA, Goodman SI, et al. Quantitative determination of succinylacetone in dried blood spots for newborn screening of tyrosinemia type I. Mol Genet Metab. 2006;88:16–21.

Narayan SB, Ditewig-Meyers G, Graham KS, Scott R, Bennett MJ. Measurement of plasma amino acids by Ultraperformance(R) Liquid Chromatography. Clin Chem Lab Med. 2011;49:1177–85.

Puck JM. Population-based newborn screening for severe combined immunodeficiency: steps toward implementation. J Allergy Clin Immunol. 2007;120:760–8.

Rinaldo P, Zafari S, Tortorelli S, Matern D. Making the case for objective performance metrics in newborn screening by tandem mass spectrometry. Ment Retard Dev Disabil Res Rev. 2006;12:255–61.

Rinaldo P, Cowan TM, Matern D. Acylcarnitine profile analysis. Genet Med. 2008;10:151–6.

Wanders RJ, Ruiter JP, et al. The enzymology of mitochondrial fatty acid beta-oxidation and its application to follow-up analysis of positive neonatal screening results. J Inherit Metab Dis. 2010;33(5):479–94.

Watson MS, Mann MY, Lloyd-Puryear MA, Rinaldo P. Newborn screening: toward a uniform screening panel and system. Genet Med. 2006;8(Suppl 5):1S–252.

Zhang XK, Elbin CS, Turecek F, Scott R, Chuang WL, Keutzer JM, et al. Multiplex lysosomal enzyme activity assay on dried blood spots using tandem mass spectrometry. Methods Mol Biol. 2010;603:339–50.

Molecular Genetics of Congenital Heart Disease

22

Abigail S. Carey and Bruce D. Gelb

Contents

22.1 Introduction 592

22.2 Overview of Normal Cardiac Development 592

22.3 Prevalence of Congenital Heart Disease 595

22.4 Etiology of Congenital Heart Disease ... 595

22.5 Genetics Techniques for Evaluating CHD 595
22.5.1 Standard Chromosome Analysis 595
22.5.2 FISH Technology 596
22.5.3 DNA Mutational Analysis 596

22.6 Overview of the Genetics of Normal Cardiac Development 597
22.6.1 GATA4 597
22.6.2 NKX2.5 599
22.6.3 TBX5 600

22.7 Syndromes Involving Congenital Heart Defects Secondary to Aneuploidy or Microdeletions 600
22.7.1 Trisomy 21, 18, and 13 600
22.7.2 Turner Syndrome 601
22.7.3 DiGeorge and Velocardiofacial Syndromes 601
22.7.4 Williams–Beuren Syndrome 602

22.8 Syndromes Associated with CHD Resulting from Single Gene Defects 602
22.8.1 Noonan Syndrome and Related Disorders 602
22.8.2 Alagille Syndrome 603
22.8.3 Char Syndrome 603
22.8.4 CHARGE Syndrome 603
22.8.5 Kabuki Syndrome 604

22.9 Conclusion: Impact on Patient and Families 604

Further Reading 605

A.S. Carey, MD
Department of Pediatrics, Mount Sinai School of Medicine, New York, NY, USA

B.D. Gelb, MD (✉)
Mount Sinai School of Medicine, New York, NY, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_22,

22.1 Introduction

- The goal of this chapter is to review the genetic basis of congenital heart disease (CHD)
  - While the etiology of CHD is thought to be multifactorial, this chapter primarily highlights specific mutations that have been found to cause CHD
- As techniques for detecting genomic changes in children with CHD continue to improve, and such technology becomes available on a clinical basis, researchers and clinicians alike stand to gain a better understanding of the genetic changes that underlie CHD
- In order to best comprehend its molecular basis and the larger complexities of cardiovascular genetics, it is fundamental to first review normal human cardiac development and the particular genes that drive healthy human cardiac development
  - Following this section, currently available genetic testing modalities will be reviewed
  - Finally, specific genes/loci implicated in syndromic and isolated CHD as well as genetic syndromes caused by single gene defects will be reviewed. This will include Alagille, Holt–Oram, Noonan, Char, and CHARGE syndromes

22.2 Overview of Normal Cardiac Development

- Developmental timeline
  - Cardiac tissue is first recognizable at approximately day 15 of gestation
  - Beating occurs approximately 1 week later at day 22–23 of gestation
  - Systemic blood flow occurs during the fourth week of gestation
- The cardiovascular system is the first system in the developing embryo to become functional (Fig. 22.1)
  - Cardiac development begins with the formation of the cardiac crescent, which arises from the anterior lateral plate mesoderm
  - The cardiac crescent contains cardiac precursor cells, now referred to as the first heart field, which ultimately organize into a ventral midline heart tube
  - The linear heart tube is composed of an inner layer of endothelial cells, which are separated from an outer layer of myocardial precursor cells by cardiac jelly
  - Once the heart tube has formed, rhythmic beating begins so that this organ is able to assist the embryo with meeting its nutritional needs
  - As the heart tube matures, it undergoes D-looping during the end of the third week of gestation. This process results in the convergence of the inflow and outflow limbs, with superior positioning of the atria above the ventricles. During D-looping, the heart tube simultaneously constricts and dilates, creating a heart primordium composed of four chambers
    - Bulbus cordis (comprised of the truncus arteriosus (TA), conus arteriosus, and conus cordis), ventricle, atrium, and sinus venosum
  - From the fourth to the eighth week, the atrioventricular canal, primordial atrium, and ventricle undergo partitioning
    - During this process, the common outflow track becomes septated into the aorta and pulmonary artery (PA)
      - While in utero, the open ductus arteriosus does not allow for the separation of pulmonary and corporal circulations
  - Heart valves develop in the embryo from precursor structures called endocardial cushions
    - Development of the semilunar valves results from swelling of the subendocardial tissue surrounding the orifices of the aorta and pulmonary trunk. The atrioventricular valves (mitral and tricuspid) develop from a similar process of subendocardial tissue swelling. Valvular remodeling continues to occur as morphogenesis progresses

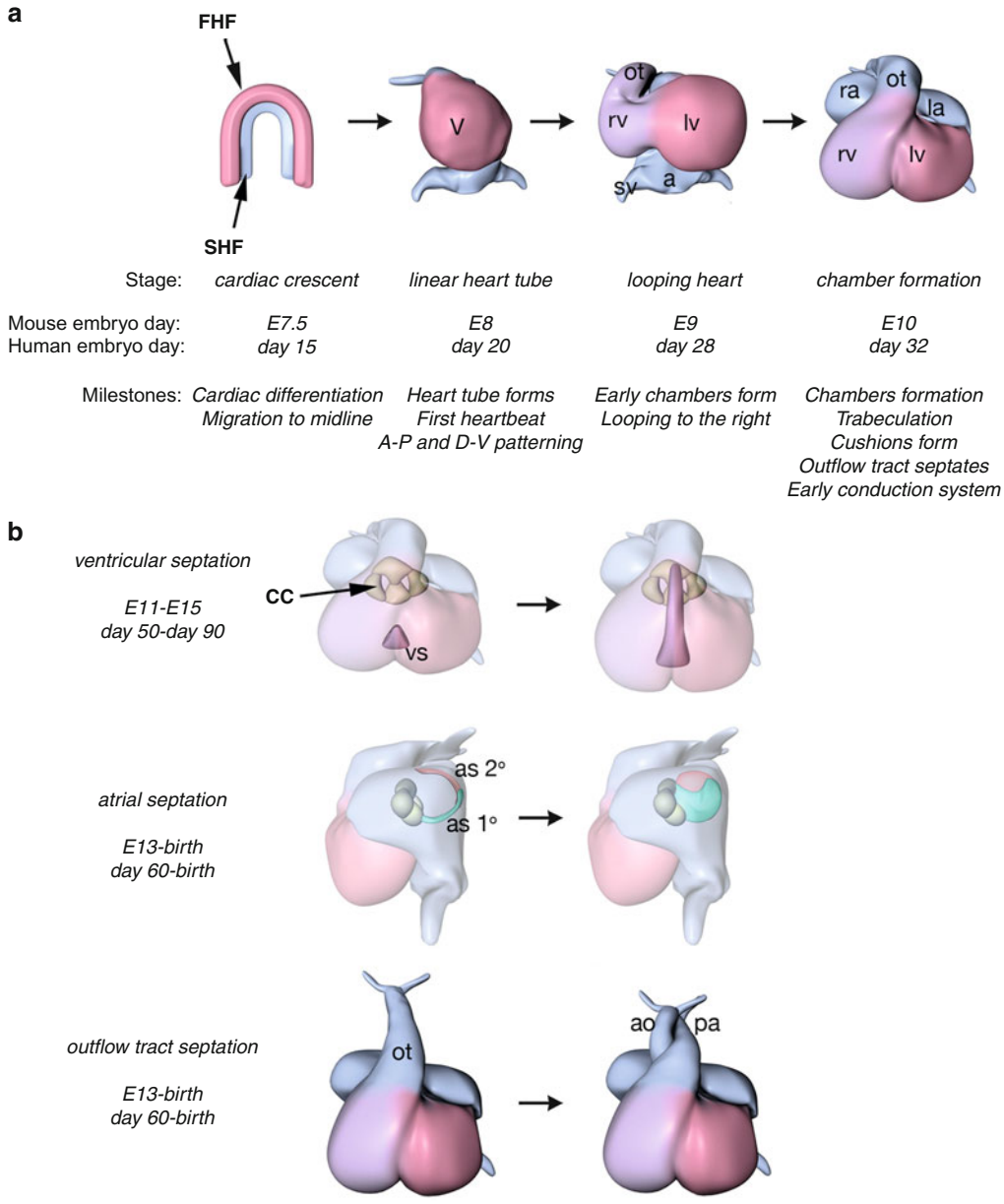


Fig. 22.1 (**a**) Early steps in heart development. Diagrams of heart development are shown in ventral views. At the earliest stages of heart formation (cardiac crescent), two pools of cardiac precursors exist. The first heart field (*FHF*) contributes to the left ventricle (*LV*), and the second heart field (*SHF*) contributes to the right ventricle (*RV*) and later to the outflow tract (*OT*), sinus venosus (*SV*), and left and right atria (*LA* and *RA*, respectively). *V*, ventricle. (**b**) Maturation of the heart. The cardiac cushions (*CC*) will give rise to the atrioventricular valves. The ventricular septum (*VS*) arises from myocardium from the left and right ventricles. Atrial septation (*AS*) occurs by the growth of two septa: the primary septum (*green*) and the secondary septum (*pink*). Outflow tract septation separates the common outflow tract (*OT*) into the aorta (*AO*, connected to the left ventricle) and the pulmonary artery (*PA*, connected to the right ventricle) (Images courtesy of F. Yeung, University of Toronto, Canada. Reprinted with permission from Bruneau B. The developmental genetics of congenital heart disease. *Nature* 2008;451:943–8)

- Myocytes originate from four primary heart fields
  - Cells arising from the first heart field are located in the anterior lateral plate mesoderm and are signaled by the surrounding endoderm to become cardiac cells
  - A smaller subset of contributing cells arises from the second heart field (SHF). These cells are located in pharyngeal mesoderm, caudal to the outflow tract of the heart
    - During looping, these SHF-derived cells contribute to the developing atria, right ventricle, and common outflow tract
  - Some consider contributing cardiac neural crest cells to comprise a third heart field. Cardiac neural crest cells invade the mesodermal core of the posterior pharyngeal arch to contribute to the smooth muscle of the pharyngeal arch arteries and derivatives of the arterial pole of the heart
    - Neural crest cells also migrate through the SHF into the outflow tract to form the endocardial cushions
    - Neural crest cell migration participates in outflow tract septation and arterial pole maturation
  - Recently, a fourth field has been proposed, derived from the proepicardial organ that substantially contributes myocytes to the formation of the ventricular septum, as well as to the developing atrial and ventricular walls
- The cardiac conduction system develops during the fifth week of gestation, beginning with the origination of the SA node
  - Originally located in the right wall of the sinus venosus (SV), the node (along with the SV wall) is ultimately incorporated into the right atrium
  - The atrioventricular (AV) node and bundles develop from the cells of the left wall of the SV and are located superior to the endocardial cushions
  - The bundle of His results from the extension of the AV bundles into the ventricles. They split into left and right bundle branches and penetrate into the ventricular myocardium, thereby allowing the conduction impulses to reach the ventricles and trigger rhythmic contractions
- The aortic arches (or pharyngeal arch arteries) are a set of six paired embryologic vascular structures that give rise to the primary arteries of the developing fetus
  - First pair of arch arteries: These pharyngeal arches largely disappear, but remnants contribute to the maxillary arteries, which supply the ears, teeth, and muscles of the eye/face. Arteries from this arch also contribute to the external carotid arteries
  - Second pair of arch arteries: Dorsal parts of these arteries persist and contribute to the stapedial arteries, which run through the stapes bone in the inner ear
  - Third pair of arch arteries: Proximal parts of these arteries form the common carotid arteries and thereby supply the structures of the head. Distal parts of the third arch arteries join the dorsal aortas to form the internal carotid arteries
  - Fourth pair of arch arteries
    - The left fourth arch artery ultimately forms part of the arch of the aorta
    - The right fourth arch artery forms the proximal part of the right subclavian artery
  - Fifth pair of arch arteries: These arch arteries are generally regarded as rudimentary vessels that do not develop into any vascular structures
  - Sixth pair of arch arteries
    - The proximal portion of the left sixth arch artery remains as part of the left pulmonary artery. The distal portion of this artery forms the ductus arteriosus, a prenatal shunt from the left pulmonary artery to the dorsal aorta
    - The proximal portion of the right sixth arch artery contributes to the right pulmonary artery. The distal portion of the sixth arch artery deteriorates
- Epicardial-derived mesenchymal cells serve as the progenitor cells for the coronary

endothelium and smooth muscle cells of the developing heart
- Coronary vasculature formation begins with the migration of proepicardial cells to form the epicardium
- Through epithelial–mesenchymal transformation, coronary endothelial cells, smooth muscle, and fibroblast cells differentiate and delaminate from the epicardium
- Following formation of a tubular network of endothelial cells, an aortic ring of endothelial cells penetrates the aorta to form the two ostia
  - Such penetration initiates rapid smooth muscle recruitment, and the coronary artery network begins forming as blood flow beings

22.3 Prevalence of Congenital Heart Disease

- CHD affects approximately 1% of newborns and, despite recent advances in surgical and catheterization-based techniques, remains the most deadly birth defect
  - CHD has an estimated prevalence of 4–50 per 1,000 live births. This value most likely underestimates the number of children with CHD; however, as many estimations exclude the presence of the most common form of CHD, bicuspid aortic valve
    - Bicuspid aortic valve is particularly associated with late morbidity and mortality, often occurs in isolation, and affects approximately 10–20 per 1,000 people
  - In total, CHD comprises approximately 25% of all congenital abnormalities. With the increased rate of surgical and medical innovations in CHD therapy, infants with CHD are now surviving past the neonatal period
    - It is critical for all clinicians to have a firm understanding of CHD inheritance so as to provide comprehensive long-term care for these patients

22.4 Etiology of Congenital Heart Disease

- The etiology of CHD is thought to be multifactorial, with both genetic and environmental factors contributing to disease development
  - Environmental factors include viral infections (such as maternal rubella), toxic exposures (including retinoic acid, Dilantin, ethanol, and halogenated hydrocarbons), and maternal diseases (such as diabetes mellitus)
  - The genetic component to CHD development is thought to be the predominant cause. CHD is associated with particular chromosomal abnormalities (e.g., trisomy 21, 22q11 microdeletion), and there is an increased CHD recurrence risk for future pregnancies for couples with one CHD+ child
    - Abnormal chromosomal structure, genetic mutations, abnormal RNA involvement, epigenetic changes, and copy number alterations have all been implicated in CHD development
    - CHD more often does not follow a Mendelian inheritance pattern, consistent with complex genetics

22.5 Genetics Techniques for Evaluating CHD

- Several methods are currently available to diagnose CHD, including karyotyping, fluorescence in situ hybridization (FISH), array comparative genomic hybridization, and DNA sequence analysis

22.5.1 Standard Chromosome Analysis

- Karyotyping has long been used as a screening tool to identify genetic abnormalities and is well suited to detect aneuploidy and extremely large aberrations

  - Standard chromosome analysis can detect abnormalities in 5–8% of neonates with CHD
  - Using molecular cytogenetic detection techniques, however, the prevalence of detected chromosomal changes in CHD is considerably higher
- High-resolution metaphase karyotyping (550–850 bands) can be performed using a variety of sample sources: peripheral blood lymphocytes, umbilical cord blood, amniocytes, or chorionic villous samples
  - Most postnatal karyotypes are performed using peripheral blood lymphocytes. Prenatal karyotyping is performed using chorionic villus samples or amniocytes at 10–12 or 15–18 weeks of gestation, respectively
  - High-resolution karyotyping, which is currently used in cytogenetic analysis, has the ability to detect inversions, translocations, duplications, or terminal deletions
  - The resolution of high-resolution karyotyping limits the ability to detect microdeletion or microduplications smaller than 3–5 Mb
    - For example, deletions occurring at the region 22q11, which results in DiGeorge and velocardiofacial syndromes, would most likely be undetected using standard karyotyping

22.5.2 FISH Technology

- This method works by hybridizing DNA probes to metaphase chromosomes to detect whether a deletion or duplication is present
  - FISH allows for the additional detection of rearrangements and microdeletion and microduplication syndromes by examining the telomeres of each chromosome (subtelomere FISH)
  - Given its specificity, FISH requires targeting of clinically suspected loci. For example, Alagille, Williams, and the 22q11 deletion syndromes have all been associated with consistent microdeletions that can be detected easily via FISH
- As reviewed by Pierpont and colleagues, cardiac malformations reported in children with subtelomeric chromosomal rearrangements include aortic arch abnormalities, ventricular septal defects (VSD), atrial septal defects (ASD), mitral valve insufficiency, and concomitant pulmonary stenosis with VSD
- A limitation of FISH technology is that it does not scan the entire genome, making it most useful for detecting clinically suspected genomic lesions

22.5.3 DNA Mutational Analysis

- DNA mutational analysis targets microscopic or submicroscopic alterations in the sequence of a gene
  - Sanger sequencing and next generation sequencing
    - Frederick Sanger developed a rapid DNA sequencing method in the 1970s using dideoxynucleotide triphosphates (ddNTPs) as DNA chain terminators
      - Using gel electrophoresis, banding indicates a DNA fragment as the result of chain termination after incorporation of ddNTPs. The relative difference of band position among lanes on the gel enables the DNA sequence to be read
      - Building on this technology, automated sequencing uses similar principles but allows more DNA to be sequenced in a shorter amount of time. This process involves labeling ddNTPs with fluorescent dye and conducting capillary electrophoresis to detect fluorescent peaks, corresponding to specific DNA bases. This automated process is considered by some to be "first generation sequencing"
    - Next generation sequencing, in contrast to the time intensive process of nucleotide sequencing with the Sanger method, enables massive parallel

sequencing of millions of sequences at once
  - This technology can be applied to whole-genome sequencing, targeted resequencing, and noncoding RNA expression profiles
  - In the future, scientists hope to use this technology to isolate sequence data from the genome of a single cell, which is exceedingly important to better understand diseases such as cancer
- Single nucleotide polymorphism (SNP) arrays
  - Alterations such as small indels and point mutations can be detected using SNP arrays
  - In addition to detecting copy number variation, SNP arrays are useful for genome-wide association analysis and are thus often used in research settings
- Microarray-based comparative genomic hybridization (array CGH) (Fig. 22.2)
  - aCGH enables whole-genome scanning at high resolution capable of detecting large genetic aberrations
  - aCGH is often regarded as the best method for detecting changes in copy number and is particularly useful when standard karyotyping fails to detect a chromosomal abnormality, but a phenotypic defect is evident
  - A limitation of aCGH is that it does not detect balanced translocations or inversions and that an array's resolution is limited by the number of represented elements on the microarray chip

- When considering submicroscopic aberrations, it is important to not only focus on changes in gene coding regions but to also detect aberrations in the regions flanking gene-containing areas and in noncoding regions
  - Such genetic changes may play a role in gene dosage effects and are often implicated in gene activity regulation
  - Most detection methods employ polymerase chain reaction-based assays, but more expensive exome sequencing techniques have recently emerged
- Given the nature of the aberrations detected, it is critical to consider whether the variation is disease related
- One problem with large-scale DNA mutational analysis is that scientists are still trying to determine which gene products play a role in disease phenotype
  - For example, the involved genes for Alagille syndrome were not known to be expressed in the heart prior to mutational identification
  - Thus, the challenge of using such sophisticated technology becomes separating pathogenic aberrations from normal population variation

22.6 Overview of the Genetics of Normal Cardiac Development

- Early inductive signals from the underlying endoderm trigger the development of primitive cardiac tissue
  - Specifically, bone morphogenic protein (BMPs), basic fibroblast growth factors, and Wnt proteins are critical to this process
- While the mechanism of early cardiac development is not completely understood, four highly conserved transcription factors (TFs) are known to play a fundamental role in cardiac development: *GATA4, NKX2.5*, and *TBX5* (Fig. 22.3)

22.6.1 GATA4

- GATA4, a member of the zinc finger superfamily of TFs, is one of the key factors necessary for expression of cardiac-specific genes and is considered an essential regulator of the earliest stages of cardiogenesis
  - In concert with SMAD proteins, GATA4 has been shown to regulate TGF/BMP signaling, which leads to the activation of another critical cardiac TF, *NKX2.5*. In addition to enabling the heart to respond

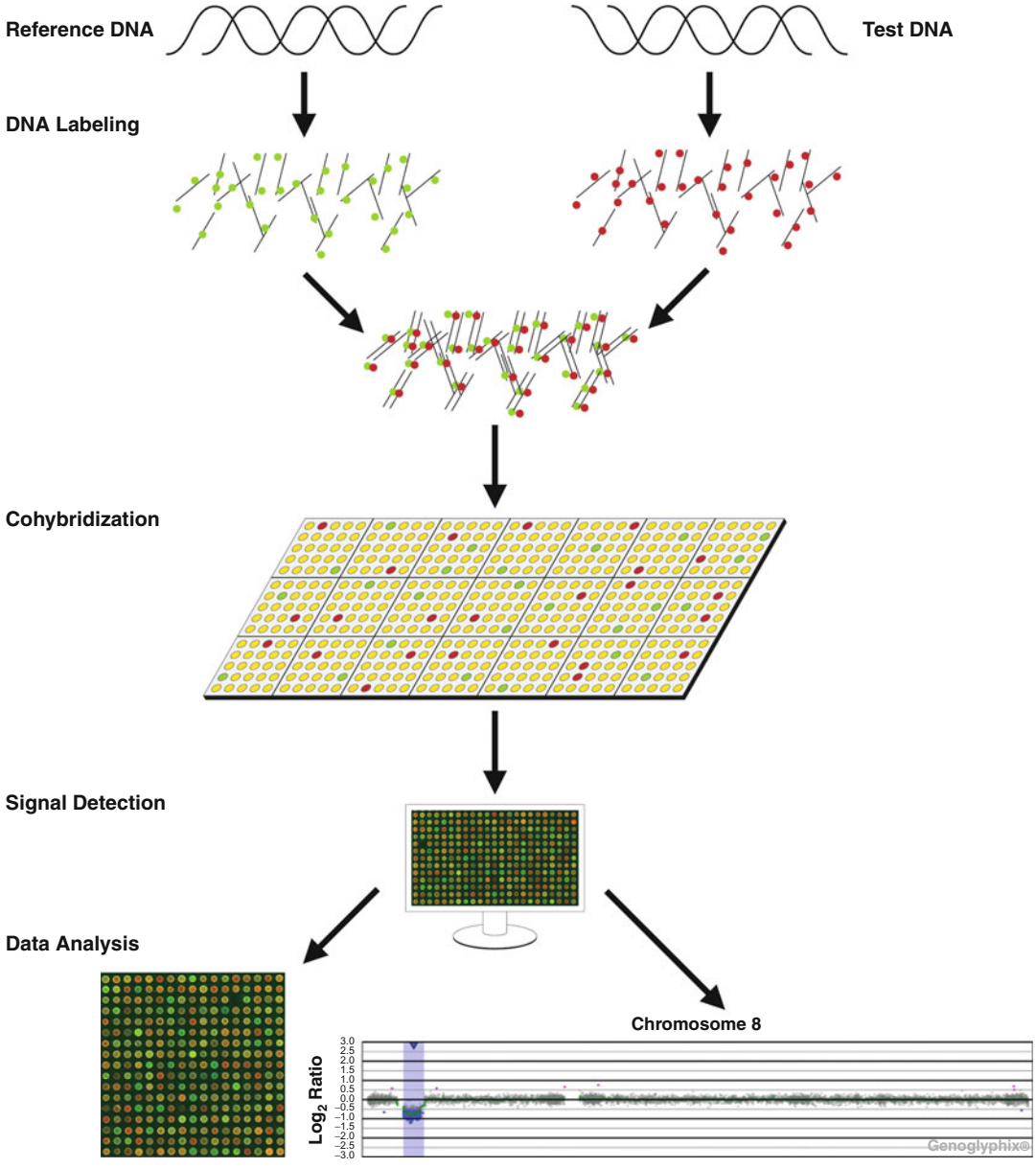


Fig. 22.2 Schematic representation of CGH microarray technology. Whole genomic DNA from a control or reference (*left*) and genomic DNA from a test or patient (*right*) are differentially labeled with two different fluorophores. The two genomic DNA samples are competitively cohybridized with DNA targets that have been robotically printed onto the microarray (*middle*). A computer imaging program assesses the relative fluorescence levels of each DNA for each target on the array (*lower left*). The ratio between control and test DNA for each target can be linearly plotted using data analysis software to visualize dosage variations (*lower right*), indicated by a deviation from the normal $\log_2$ ratio of zero. The lower right image is a representative plot from an oligonucleotide-based CGH microarray showing the whole chromosome 8 with an 8p23.1 deletion (*shaded region*), visualized using Genoglyphix (Signature Genomic Laboratories, Spokane, WA) (Reprinted with permission from American Society for Investigative Pathology, Bejjani BA et al. Application of array-based comparative genomic hybridization to clinical diagnostics. *J Mol Diagn*. 2006; 8:528–33)

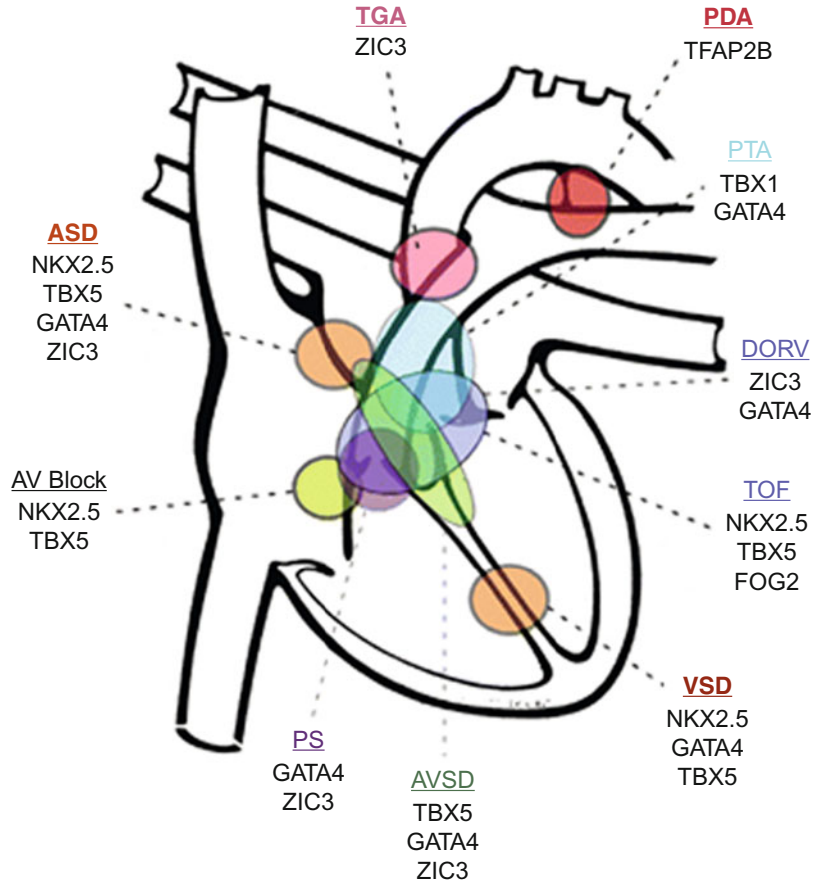


Fig. 22.3 Sites of some of the major structural CHDs. Note how the same transcription factor is linked to multiple anomalies and how each anomaly can be caused by more than one gene. *ASD* atrial septal defect, *AV block* atrioventricular block, *AVSD* atrioventricular septal defect, *DORV* double outlet right ventricle, *PS* pulmonary stenosis, *PTA* tricuspid atresia, *TOF* tetralogy of Fallot, *VSD* ventricular septal defect (Reproduced with permission from Elsevier from Nemer M. Genetic insights into normal and abnormal heart development. *Cardiovasc Pathol*. 2008;17:48–54)

to cardioinductive signals, when coexpressed with NKX2.5, *GATA4* activates BMP-4, as well as other TFs necessary for heart development

- *GATA4* mutations have been specifically implicated in familial ASDs and VSDs (Fig. 22.3), inherited in an autosomal dominant fashion
  - Rajagopal and colleagues created a murine *Gata4* knockout mouse model. Gata4 knockout mouse embryos were found to possess atrial and ventricular septation defects, endocardial cushion defects (ECD), RV hypoplasia, and cardiomyopathies. In humans, *GATA4* missense mutations cause the same CHD phenotypes, but not cardiomyopathy
  - *GATA4* mutations appear to account for substantially less than 5% of CHD including ASD cases specifically
- *GATA4* deletions, duplications, and frame shift mutations have all been associated with septal defects. Interestingly, GATA4 interacts with additional key cardiac TFs, NKX2.5 and TBX5
  - In studying *Gata4* mutations in mice models, Garg et al. observed that *Gata4* mutation disrupts the gene's ability to interact with *Tbx5*. Given that the same group also proved a similar disruption between *Gata4* and *Nkx2.5*, the authors postulated that TBX5, NKX2.5, and GATA4 function in complex to regulate the genes required for cardiac septation

22.6.2 NKX2.5

- As the heart develops, proper chamber compartmentalization and orientation becomes

increasingly important for successful morphogenesis. Work done with knockout mice has demonstrated that *Nkx2.5* is critical for chamber specification
  - *Nkx2.5* is uniformly expressed throughout the heart
- *Nkx2.5* mutations have been associated with septal and conduction defects in humans (Fig. 22.3). Studies examining *Nkx2.5* knockout mice have demonstrated progressive loss of the AV nodal conduction tissue, ultimately resulting in complete AV conduction block
  - Gene dosage appears to correlate with the number of functional cells in the cardiac conduction pathway
  - *Nkx2.5* mutations have also been associated with other forms of CHD including tetralogy of Fallot and VSDs
- In addition to interacting with GATA4 to regulate septation, NKX2.5 has also been shown to interact with MEF2C to initiate ventricular formation
  - While mice with individual mutants of *Nkx2.5* and *Mef2c* had identifiable ventricles, double mutants displayed ventricular hypoplasia and a more severe cardiac phenotype
- *Nkx2.5* mutations affecting the homeodomain result in a reduction or loss of DNA binding that translates into decreased protein–protein interactions, particularly involving GATA4 and TBX5

22.6.3 TBX5

- TBX5, a member of the T-box transcription factor family, is critical for cardiac and limb development
  - *TBX5* is mutated in approximately 70% of those with the autosomal dominant disorder, Holt–Oram syndrome (HOS)
  - The prevalence of HOS is approximately 1 of 100,000 live births
    - All patients with HOS have preaxial radial malformations, and approximately 90% of patients have associated cardiac defects (most often secundum ASD and/or VSD, but rarely complex heart diseases) and/or conduction abnormalities (Fig. 22.3)
- TBX5 regulates several genes, including downstream targets: atrial natriuretic factor (ANF) and connexin 40 (CX40), and functions in synergy with NKX2.5
- Basson et al. established genotype–phenotype correlations between HOS and *TBX5*
  - Truncation mutations result in severe cardiac and skeletal malformations
- A single missense mutation was shown to cause significant cardiac abnormalities with only minor skeletal defects, while two missense mutations result in severe skeletal malformations with minor cardiac phenotype
  - Since all three of these missense mutations are believed to affect DNA binding, *TBX5* is believed to bind different targets in the heart and limb

22.7 Syndromes Involving Congenital Heart Defects Secondary to Aneuploidy or Microdeletions

- While a majority of children have apparently isolated CHD, a proportion (between 20% and 40% of cases) have congenital cardiac anomalies associated with a larger syndrome

22.7.1 Trisomy 21, 18, and 13

- Down syndrome, caused by a complete or partial trisomy of chromosome 21, is the most common chromosomal defect among live births and the most frequent cause of intellectual disability
- Approximately 50% of children with Down syndrome have CHD, a majority of which are ASDs, VSDs, or atrioventricular septal defects
  - While the overexpression of type IV collagen, whose gene resides on chromosome 21, has been suggested to play a role in the pathogenesis of CHD in Down

syndrome, no chromosome 21 gene's increased dosage has been confirmed to cause CHD
- While much rarer diagnoses, almost 100% of patients with trisomy 18 have septal defects, whereas those with trisomy 13 also commonly have heterotaxy in addition to septal defects

22.7.2 Turner Syndrome

- Turner syndrome (TS), due most commonly to monosomy of chromosome X but also resulting from X-mosaicism and X structural defects, affects approximately 1 in 4,000 female live births
  - The four characteristics of TS are growth failure, cardiovascular disease, learning difficulties, and gonadal failure
  - CHDs affect 75% of fetuses and 25–45% of live-born females with TS
- The most common cardiac abnormalities in TS are bicuspid aortic valve and coarctation of the aorta
  - The genetic etiology of CHD in TS remains unknown

22.7.3 DiGeorge and Velocardiofacial Syndromes

- Deletion at chromosome 22q11 is associated with greater than 80% of patients with DiGeorge syndrome (DGS) and velocardiofacial syndrome (VCFS)
  - Approximately 75% of those with VCFS have cardiac abnormalities
  - In addition to cardiac defects, 22q11 deletion causes palatal defects, impaired immune response secondary to defective thymic development, hypocalcemia due to defective parathyroid development, ophthalmologic defects, and learning difficulties
  - Approximately 10% of those with 22q11 deletion have DGS, which consists of having at least two of the following: conotruncal cardiac defects, hypoparathyroidism, hypocalcemia, thymic aplasia, or immune deficiency
- Conotruncal defects are most commonly associated with 22q11 deletions. These include truncus arteriosus, aortopulmonary window, interrupted aortic arch type B, tetralogy of Fallot, and conal septal VSDs
  - Given that the phenotypes of DGS and VCFS are thought to arise from defective migration of neural crest cells and/or the secondary heart field cells and the developing embryonic outflow tract of the heart tube is derived from a subpopulation of those cells, it is logical that these patients often have conotruncal heart defects
- Studies are underway to identify the gene or genes residing at 22q11 relevant to the disease phenotype
  - The 22q11.2 region has been mapped, cloned, and sequenced. More than 35 genes are present within the commonly deleted region
  - Three candidate genes (*TBX1, CRKL*, and *ERK2*), whose haploinsufficiency causes neural crest dysfunction and anomalies associated with 22q11.2 phenotype, have been identified
    - *TBX1* and *CRKL* are critical for the function of fibroblast growth factor 8 (FGF8)
      - In mice, loss of *FGF8* phenocopies DGS
    - Missense *TBX1* aberrations result in proteins with reduced transcriptional activity, and *TBX1* haploinsufficiency results in the underdevelopment of the pharyngeal pouches (Fig. 22.3)
    - Recently, haploinsufficiency of *ERK2* was also shown to also have an effect on neural crest development
    - Newbern and colleagues demonstrated that patients with small (approximately 1 Mb) microdeletions distal to 22q11.2 exhibit haploinsufficient *ERK2* expression
    - Deletion of *Erk2* in mouse neural crest cells specifically resulted in craniofacial and cardiac abnormalities observed in human patients with DGS

  - Taken together, it appears that deletion of multiple genes within and proximate to the 22q11.2 region underlies the DGS and VCFS phenotypes
    - Bernice Morrow's group investigated the molecular mechanism underlying the recurrent deletions at 22q11.2
      - They determined that the two most common deletions at 22q11 causing DGS/VCFS result from nonhomologous recombination events

22.7.4 Williams–Beuren Syndrome

- Williams–Beuren syndrome (also known as Williams syndrome (WS)) is characterized by distinctive facial dysmorphism, growth retardation, intermittent infantile hypercalcemia, learning difficulties, a distinctive personality, connective tissue abnormalities, vasculopathies, and heart defects
  - Cardiac involvement most typically includes supravalvular aortic stenosis (SVAS) and peripheral pulmonic stenosis. Additional associated cardiac defects include pulmonary arterial stenosis, multiple arterial stenoses, aortic/mitral valve defects, and systemic hypertension
- WS is caused by a 1.5-Mb deletion at 7q11.23, which causes haploinsufficiency of the elastin gene (*ELN*)
  - In addition to *ELN* involvement, WS affects at least 15 other contiguous genes, including *FZD3, BCL7B, STX1A, LIMK1*, and *CYLN2*
  - This deletion is detected via FISH in 99% of individuals with WS
- A related disorder, familial SVAS, is an autosomal dominant trait with the vasculopathy of WS without the extracardiovascular issues
  - Familial SVAS is caused by point mutations in *ELN*
  - Based on this, it is clear that the vasculopathy in WS is attributable to *ELN* haploinsufficiency

22.8 Syndromes Associated with CHD Resulting from Single Gene Defects

22.8.1 Noonan Syndrome and Related Disorders

- Noonan syndrome (NS) is an autosomal dominant trait characterized by short statue, distinct facial features, skeletal abnormalities, ectodermal and hematologic abnormalities, neurocognitive disability, and cardiac disease
  - Approximately 85% of NS patients have some form of CHD. Most commonly, patients have pulmonary stenosis with or without dysplastic pulmonary valve and hypertrophic cardiomyopathy (HCM)
  - Although validated scoring systems are available to clinically diagnose NS, genetic testing is now available for diagnosing suspected patients
- In addition to NS, related disorders include Costello, LEOPARD, and cardiofaciocutaneous syndromes
  - These related disorders have overlapping phenotypes with NS and have similar underlying genetic defects
- NS and related disorders primarily result from defects in the RAS/mitogen-activated protein kinase (MAPK) signal transduction pathway
  - Approximately 75% of patients with NS harbor a specific genetic defect in *PTPN11, SOS1, KRAS, NRAS, RAF1, BRAF*, or *MEK1*
  - Most mutations of the RAS/MAPK pathway result in gain of function mutations; however, some mutations, such as *PTPN11* mutations causing LEOPARD syndrome, result from loss of function mutations
  - While specific phenotypes have been related to some genotypic abnormalities (e.g., *RAF1* defects have been associated with HCM), the position of the defect along the RAS/MAPK pathway does not appear to confer a specific phenotype overall

22.8.2 Alagille Syndrome

- Alagille syndrome is an autosomal dominant disorder characterized by a paucity of bile ducts, as well as cardiac, skeletal, renal, facial, and ocular defects
  - There is significant variability in the phenotypic expression ranging from patients with subclinical disease to those with severe manifestations requiring heart and/or liver transplantation
    - Common cardiac defects include pulmonary stenosis and/or atresia and tetralogy of Fallot
  - The incidence is approximated to be 1 in 70,000 live births
- Mutations in *JAG1*, a cell surface protein that functions as a ligand for the Notch transmembrane receptor, cause Alagille syndrome
  - *JAG1* is located at chromosome 20p12
  - Point mutations and small deletions in *JAG1* can be identified in 70–95% of patients with clinical Alagille syndrome, and they are inherited in 30–50% of case
  - Aberrations involving *JAG1* are thought to cause Alagille syndrome either through haploinsufficiency of the Jagged1 protein or by causing a dominant negative effect
    - Specifically, deletions at *JAG1* involving the whole gene, as well as mutations, including frameshift, missense, and nonsense aberrations, can underlie Alagille syndrome
  - Through investigating *JAG1* within the context of Alagille syndrome, scientists and clinicians have come to understand that the Notch signaling pathway is critically important for organogenesis, particularly involving the ocular, cardiac, and hepatic systems
- *NOTCH2* mutations account for a small percentage of Alagille syndrome
  - *NOTCH2* is a transmembrane receptor that is essential for hepatic, renal, and vascular development
  - Located on chromosome 1p12
  - Although heterozygote *Jag1* knockout mice do not exhibit Alagille syndrome phenotype, *Jag1*/*Notch2* double heterozygote mice possess ocular, cardiac, liver, and renal manifestations similar to those demonstrated in Alagille syndrome
    - This finding suggests that *Notch2* modifies *Jag1* expression

22.8.3 Char Syndrome

- Char syndrome is an autosomal dominant trait characterized by a patent ductus arteriosus (PDA), facial defects, and deformities of the fifth digit
- *TFAP2B* encodes a transcription factor expressed in neural crest cells that maps to the Char syndrome critical region at chromosome 6p12–21.1
  - Cardiac neural crest cells are known to contribute to the development of the sixth branchial arch, which ultimately becomes the medial layer of the ductus arteriosus
  - TFAP2B proteins bind DNA as homodimers or as heterodimers formed with other TFAP2B family members
- *TFAP2B* sequencing, which is only available on a research basis, detects mutations in only approximately 50% of patients (Fig. 22.3)
- Studies of mutated *TFAP2B* reveal a dominant negative mechanism of action, suggesting inhibition of other TFAP2B dimerization members

22.8.4 CHARGE Syndrome

- CHARGE syndrome is a constellation of multiple congenital abnormalities including coloboma of the eye, heart defects (primarily conotruncal defects, AV canal defects, and aortic arch malformations), choanal atresia, growth retardation, genital malformations, and ear abnormalities
  - CHARGE syndrome occurs in an autosomal dominant inheritance pattern and affects approximately 1 in 10,000 live births

  - Almost all mutations occur de novo, but parent to child transmission has occasionally been reported
- Through conducting whole-genome analysis on individuals with clinical CHARGE syndrome, a 2.3 Mb microdeletion was detected on chromosome at 8p21 in one patient
  - This observation led to the genetic sequencing of the presumed CHARGE syndrome critical region and revealed a mutation in *CHD7*
  - The *CHD7* gene consists of 37 coding exons and one noncoding exon
    - CHD7 analysis detects mutations in more than 90% of patients fulfilling clinical criteria for CHARGE syndrome
- *CHD7* encodes a chromodomain helicase DNA-binding protein expressed during embryogenesis
  - Haploinsufficiency of CHD7 is believed to lead to CHARGE syndrome in greater than 50% of patients with CHARGE syndrome
  - Although most patients with CHD7 mutations have a truncating mutation, missense and frameshift mutations also commonly occur
- In mouse embryos, *Chd7* is highly expressed in the cardiac outflow tract and truncus arteriosus

22.8.5 Kabuki Syndrome

- Kabuki syndrome (also known as Niikawa–Kuroki syndrome) is a rare genetic syndrome characterized by multiple congenital abnormities, including deformed facies, skeletal and growth abnormalities, and dermatologic and visceral defects
  - Kabuki syndrome is reported as having a prevalence of 1 in 32,000 infants; however, this number is most likely an underestimation
  - Kabuki syndrome often occurs as a sporadic occurrence, but familial occurrences, inherited in an autosomal dominant fashion, have been reported
- In one study, CHD was found in approximately 58% of patients with Kabuki syndrome
  - The most frequent cardiac defects include aortic coarctation, ASD, and VSD
  - Additionally, multiple studies have reported patients with hypoplastic left heart syndrome
- Recently, advances in exome sequencing have demonstrated that nonsense or frameshift mutations altering *MLL2* result in Kabuki syndrome in approximately 72% of cases
  - MLL2 is part of the trithorax family of proteins that confer histone methyltransferase activity and are important in the epigenetic control of active chromatin states

22.9 Conclusion: Impact on Patient and Families

- With the completion of the Human Genome Project, the identification of genes for monogenetic traits has increased exponentially
- As more sophisticated detection methods are developed, additional genetic mechanisms underlying CHD are being elucidated and are improving genetic counseling and patient care
- The ability to offer more precise genetic counseling stands to vastly improve CHD patient care
- With a better understanding of the genetic changes underlying phenotypic CHD, not only will clinicians be better able to forecast the likelihood of CHD in future pregnancies and assist with family planning, but they will also be better equipped to identify individuals at risk for developing cardiac and noncardiac complications later in life
  - For example, individuals with ASD and *NKX2*.5 mutations are at risk for late-onset complete heart block, a potentially lethal event, which can be anticipated and intervened on through placement of a pacemaker

Further Reading

Garg V. Insights into the genetic basis of congenital heart disease. Cell Mol Life Sci. 2006;63:1141–8.

Gelb BD. Genetic basis of syndromes associated with congenital heart disease. Curr Opin Cardiol. 2001; 16:188–94.

Huang JB, Liu YL, Sun PW, et al. Molecular mechanisms of congenital heart disease. Cardiovasc Pathol. 2010; 19:183–93.

Nemer M. Genetic insights into normal and abnormal heart development. Cardiovasc Pathol. 2008;17:48–54.

Pierpont ME, Basson CT, Benson DW, et al. Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. Circulation. 2007;115:3015–38.

Richards AA, Garg V. Genetics of congenital heart disease. Curr Cardiol Rev. 2010;6:91–7.

Schott JJ, Benson DW, Basson CT, et al. Congenital heart disease caused by mutations in the transcription factor NKX2-5. Science. 1998;281:108–11.

Tomita-Mitchell A, Maslen CL, Morris CD, et al. GATA4 sequence variants in patients with congenital heart disease. J Med Genet. 2007;44:779–83.

Vissers LE, van Ravenswaaij CM, Admiraal R, et al. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nat Genet. 2004;36:955–7.

Familial Cancer Syndrome

23

Beth T. Harrison, Rodolfo Montironi, Antonio Lopez-Beltran, Michelle P. Elieff, and Liang Cheng

Contents

23.1 Overview 608

23.2 Retinoblastoma 610

23.3 Li–Fraumeni Syndrome 616

23.4 Hereditary Breast and Ovarian Cancers 617
23.4.1 Overview 617
23.4.2 BRCA1 617
23.4.3 BRCA2 618

23.5 Fanconi Anemia 620

23.6 Hereditary Gastrointestinal Cancers 621
23.6.1 Overview 621
23.6.2 Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer) 621
23.6.3 Familial Adenomatous Polyposis 623
23.6.4 MYH-associated Polyposis Syndrome 625
23.6.5 Peutz–Jeghers Syndrome 625
23.6.6 Juvenile Polyposis Syndrome 626
23.6.7 Hereditary Diffuse Gastric Cancer 627

23.7 Hereditary Endocrine Tumor Syndromes 627
23.7.1 Multiple Endocrine Neoplasia, Type 1 (Wermer Syndrome) 627
23.7.2 Multiple Endocrine Neoplasia, Type 2 MEN2A (Sipple Syndrome), MEN2B, and Familial Medullary Thyroid Cancer 628

23.8 Hereditary Kidney Tumor Syndromes 629
23.8.1 Familial Wilms Tumor 629
23.8.2 Hereditary Leiomyomatosis and Renal Cell Carcinoma 630
23.8.3 Hereditary Papillary Renal Cell Carcinoma 631
23.8.4 Hereditary Nonpapillary Renal Cell Carcinoma 632
23.8.5 von Hippel–Lindau Syndrome 632
23.8.6 Birt–Hogg–Dubé Syndrome 633
23.8.7 Tuberous Sclerosis Complex 633

23.9 Genodermatoses 633
23.9.1 Hereditary Melanoma 633
23.9.2 Nevoid Basal Cell Carcinoma Syndrome (Gorlin Syndrome) 634
23.9.3 Xeroderma Pigmentosum 635
23.9.4 Carney Complex 636
23.9.5 PTEN Hamartoma Tumor Syndrome (Cowden Syndrome) 637
23.9.6 Birt–Hogg–Dubé Syndrome 638

23.10 Tuberous Sclerosis Complex 639

23.11 Neurofibromatosis Type 1 (von Recklinghausen Syndrome) 640

23.12 Neurofibromatosis Type 2 641

23.13 Ataxia-Telangiectasia 642

Further Reading 642

B.T. Harrison, MD (✉)
Department of Pathology, NYU Langone Medical Center, New York, NY, USA

R. Montironi, MD
Institute of Pathological Anatomy and Histopathology, Polytechnic University of the Marche Region (Ancona), School of Medicine, United Hospitals, Ancona, Italy

A. Lopez-Beltran, MD, PhD
Unit of Anatomic Pathology, Department of Surgery, Cordoba University School of Medicine, Cordoba, Spain

M.P. Elieff, MD
Sparrow Forensic Pathology Services, Lansing, MI, USA

L. Cheng, MD
Indiana University School of Medicine, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_23,

23.1 Overview

In 2010, 1.5 million individuals were newly diagnosed with cancer and 500,000 died from cancer-related causes in the USA. An estimated 5–10% of all cancers are considered hereditary. Well-defined familial cancer syndromes arise from pathogenic germline mutations that confer a significantly increased lifetime risk of cancer to the mutation carrier. Genetically, these syndromes are characterized by Mendelian inheritance of rare, highly penetrant, single gene variants. Most of them display autosomal dominant inheritance with a 50% risk of transmission to offspring and a pedigree of successive affected generations. A few autosomal recessive and X-linked syndromes do exist.

In the late 1980s and early 1990s, studies employing linkage analysis and positional cloning identified the susceptibility genes for many of the common cancer syndromes. Tumor suppressor genes underlie the majority of syndromes, suggesting that mutation carriers develop tumors later in life after somatic inactivation of the second copy of a gene, in accordance with Knudson's "two-hit" hypothesis. Some familial cancer syndromes represent exceptions, notably multiple endocrine neoplasia (MEN) type 2 due to mutations in the *RET* proto-oncogene and Lynch syndrome (LS) due to mutations in mismatch repair genes.

Over the past decade, the fields of cancer genetics and epidemiology have shifted attention to a polygenic model of predisposition. Much of the inherited susceptibility to cancer cannot be explained by highly penetrant variants and rather reflects complex inheritance of genetic variants with rare to common frequency and low to moderate penetrance. Genome-wide association studies (GWAS) have become a new tool to identify common low-penetrance genes that are individually associated with modest increases in risk. Unfortunately, the findings from these studies have fueled direct to consumer genetic tests with unclear interpretation and clinical implications.

Appropriate diagnosis and management of familial cancer syndromes still depends on clinical recognition of suggestive personal or family history and referral to genetics professionals. Genetic testing should occur in the setting of pretest and posttest counseling and should emphasize screening for rare disease-causing mutations rather than common risk-increasing polymorphisms for greatest clinical utility.

- Personal and family history suggestive of a familial cancer syndrome
  - Multifocal or bilateral primary tumors of the same origin
  - Multiple primary tumors in different organs
  - Younger age at diagnosis
  - Characteristic pathology (i.e., triple-negative, basal-like breast cancer in *BRCA1* mutation carriers)
  - Associated genetic traits, congenital anomalies, or rare diseases
  - Member of certain ethnic groups or genetically isolated populations
  - Two or more first- or second-degree relatives with the same tumor or tumors characteristic of a syndrome
- Genetic counseling
  - Careful construction of a pedigree with verification of cancer diagnoses by clinical and pathologic records. It is important because the accuracy of family history can be quite variable, depending on the age of the respondent, relatedness of the affected relative, recency of the diagnosis, site of cancer, etc.
  - Risk assessment employing clinical criteria, available models, and professional interpretation of such tools
  - Education about the benefits, risks, and limitations of genetic testing, early detection, and preventive strategies
  - Discussion of psychosocial, financial, ethical, and legal concerns
  - Acquisition of informed consent
- Test characteristics
  - Genetic tests should only be offered if they provide reliable and accurate results (*analytical and clinical validity*) and aid in diagnosis or management of the disorder (*clinical utility*)

Table 23.1 Applications of molecular genetic testing

Application	Description
Diagnostic testing	Confirms or rules out a suspected genetic disorder in a symptomatic individual
Predictive testing	Offered to asymptomatic individuals with a positive family history in order to identify a mutation that increases risk or leads to eventual manifestation of genetic disease. Referred to as *predispositional* and *presymptomatic* testing, respectively
	Ideally, follows identification of the specific gene mutation in an affected family member in order to obtain the most meaningful and cost-effective results
Carrier testing	Offered to individuals with an ethnic background or family history concerning for autosomal recessive or X-linked diseases
	Carriers, who have a mutation in one copy of the relevant gene, are not at risk of disease. However, testing of both parents provides information on risks to offspring
Prenatal diagnosis (PND)	Offered to parents at increased risk of having a child with a genetic disorder due to known pathogenic mutation in the family
	Performed on genetic material from amniocentesis or chorionic villus sampling. Noninvasive prenatal diagnosis using cell-free fetal DNA in the maternal serum may be clinically available in the near future
	Available for most familial cancer syndromes listed in this chapter, but a laboratory with the test of interest must be found
Preimplantation genetic diagnosis (PGD)	Performed on early embryos of infertility patients who have a high risk of a child with a serious disorder due to a known pathogenic mutation in the family
	Currently, it is expensive, limited to selected disorders, available at only a few centers, and controversial due to ability for reproductive selection

- Applications
  - Genetic testing is performed for various purposes. It includes diagnostic testing, predictive testing, carrier testing, prenatal diagnosis, and preimplantation genetic diagnosis
    - Please see Table 23.1 for a detailed description of these applications
- Methodologies
  - Genetic testing is performed by commercial laboratories and select academic centers certified in accordance with the Clinical Laboratory Improvement Amendments (CLIA) and found in a directory on www.genetests.org
  - At present, germline testing usually involves direct testing of DNA from peripheral blood lymphocytes by a combination of *sequence analysis* and *deletion/duplication analysis*
    - Full gene sequencing is often the first-line approach in screening for unknown mutations given the significant improvements in efficiency, accuracy, and cost of next-generation platforms. One major caveat is that it does not detect large deletions, duplications, or rearrangements
    - Multiplex ligand-dependent probe amplification (MLPA) is a multiplex PCR-based method that determines the relative copy number of all exons simultaneously. It is currently the preferred method for deletion/duplication analysis due to its high sensitivity and efficiency
    - *Family-specific mutation analysis* with sequence analysis limited to the region of interest provides more cost-effective and meaningful results in individuals who have a known mutation in the family
    - *Targeted mutation analysis* for a specific mutation or panel of mutations is performed prior to sequencing if the individual belongs to a population with founder mutations or the syndrome features highly recurrent mutations
  - In general, there are variations in the technologies and strategies employed depending on the syndrome of interest and its particular genetics

  - See Table 23.2 for a more comprehensive summary of the methodologies available for genetic testing of familial cancer syndromes
  - Please see Chap. 8, Diagnostic Methodology and Technology for a detailed description of technologies mentioned in this table
- Interpretation
  - Accurate interpretation of test results is at the crux of diagnostic medicine
  - As sequencing technologies continue to improve and to have more widespread clinical application, interpretation of sequence analyses in the context of both sporadic and hereditary cancer may fall increasingly within the realm of the clinical laboratorian/molecular genetic pathologist
  - Genetic tests have four possible outcomes
    - True positive
      - Detected mutation prevent normal gene function and confers an increased risk of cancer
    - True negative
      - Known family-specific mutation is not detected. The individual is not at increased risk for cancers associated with the syndrome affecting his/her family
    - Uninformative
      - Mutation is not detected in an individual who has a strong family history, and a mutation has not been previously identified in the family
      - Negative results do not exclude the possibility that there is an undetectable mutation in the gene of interest or a mutation in another susceptibility gene
      - Individual and family members may still be at risk of cancer and should be managed based on family history
    - Variant of uncertain significance (VUS)
    - Mutation is detected, but its effect on protein function and its clinical significance are not known. It is typically a missense, intronic, or distal mutation
    - Individual and family members may still be at increased risk of cancer and should be managed based on family history
    - An attempt to determine its significance may involve approaches such as segregation analysis (genetic testing of parents or affected relatives to see if the variant cosegregates with cancer), protein function assays, clinicopathologic and family history criteria, and computational modeling.
    - It is a significant issue in BRCA testing for hereditary breast and ovarian cancer syndrome
- Overview of familial cancer syndromes, see Table 23.3

23.2 Retinoblastoma

- Overview
  - Retinoblastoma is a malignant tumor of the developing retina that occurs in children under 5 years of age
  - Incidence is 1 in 18,000 to 1 in 30,000 live births
  - Survival is 99% in the developed world, but 50% in developing nations
  - It is the model for Knudson's "two-hit" hypothesis of carcinogenesis due to loss of a tumor suppressor gene
    - Retinoblastoma occurs in cells that have disease-causing mutations in both copies of the gene
    - Hereditary retinoblastoma arises when a germline mutation is followed by somatic inactivation of the second allele. It is frequently multifocal and bilateral
    - Sporadic retinoblastoma is due to somatic mutations in both alleles. It is unifocal and unilateral
- Clinical features
  - Most commonly presents as leukocoria (white pupillary reflex) or strabismus
  - Unilateral disease is found in 60% and is diagnosed at a mean age of 24 months
  - Bilateral disease (40%) presents earlier at a mean age of 15 months

Table 23.2 Methodologies for genetic testing of familial cancer syndromes

Type of genetic test	Definition	Methodologies	Strategy	Examples
Biochemical genetic	Analysis of protein structure or function	Protein expression assays Enzyme activity assays Analyte assays	For certain syndromes, it provides a preliminary diagnosis or testing strategy prior to direct DNA testing	For HLRCC, measurement of enzyme activity of fumarate hydratase or the level of its analyte provides an initial diagnosis For Lynch syndrome, IHC demonstrates loss of expression of one of four mismatch repair proteins, identifying the gene for sequence analysis
Cytogenetic	Examination of whole chromosomes for structural abnormalities	Karyotype analysis FISH	Useful for multiple syndromes that occasionally arise from recurrent cytogenetically visible abnormalities involving the locus of interest	For hereditary retinoblastoma, gross deletions can be detected by FISH using commercially available probes for 13q14 For Fanconi anemia, chromosome breakage studies provide the initial diagnosis
Molecular genetic	Direct DNA testing			
Mutation scanning	Screening the entire coding region or select exons for variant regions prior to determination of the specific sequence alteration by sequencing	HRM DHLPC DGGE CSGE SSCP	Many clinical laboratories skip mutation scanning and proceed directly to sequencing	
Sequence analysis	Direct sequencing of the full gene or select exons in order to detect missense, nonsense, or splice site mutations and small intragenic insertions or deletions	"Next-" or second-generation high-throughput sequencing (emerging third-generation technologies are highly anticipated)	Often first-line approach due to significant improvement in the efficiency, accuracy, and cost-effectiveness of technologies It does not detect large deletions, duplications, or rearrangements	For certain syndromes, i.e., tuberous sclerosis, sequence analysis is complicated by the size of the gene, number of distinct mutations, and lack of hotspots
Deletion/duplication analysis	Analysis of submicroscopic deletions and duplications by molecular or molecular cytogenetic methods	MLPA PCR-based methods (i.e., qPCR, long-range PCR) FISH Array CGH	Used in conjunction with sequence analysis to increase the diagnostic yield MLPA has become the preferred method due to its high sensitivity and efficiency	For Peutz–Jeghers syndrome, deletion analysis by MLPA greatly increases the mutation detection rate, as 15–40% of mutations are large deletions
Targeted mutation analysis	Testing for the presence of a specific mutation, a specific type of mutation, or a panel of mutations	Sequence analysis limited to the region of interest Real-time PCR-based methods ASO hybridization	Cost-effective and efficient first-line approach if the individual comes from a population with founder mutations or the suspected syndrome has highly recurrent mutations	For HBOC, mutation analysis for three different founder mutations in *BRCA1* and *BRCA2* is the first step in the diagnosis of Ashkenazi Jewish women

(*continued*)

Table 23.2 (continued)

Type of genetic test	Definition	Methodologies	Strategy	Examples
Family-specific mutation analysis	Testing for the specific disease-causing mutation previously identified in a family member	Sequence analysis limited to the region of interest	More cost-effective and meaningful first-line approach if there is a known mutation in the family Important role in diagnostic, predictive, carrier, and prenatal testing	For HBOC, an effort should be made to perform genetic testing on an affected family member prior to predictive testing in an individual with a strong family history
Linkage analysis	Indirect DNA testing Markers (polymorphisms) cosegregate with a gene of interest and track the gene within the family without actually knowing the mutation		Employed when a mutation is not known or is not detected on direct DNA testing and at least two family members are affected	For FAP, linkage analysis is clinically available at multiple laboratories worldwide and displays excellent accuracy of 98%

Methodologies: *FISH* fluorescence in situ hybridization, *HRM* high-resolution melting, *DHLPC* denaturing high-performance liquid chromatography, *DGGE* denaturing gradient gel electrophoresis, *CSGE* conformation-sensitive gel electrophoresis, *SSCP* single-strand conformation polymorphism, *MLPA* multiplex ligand-dependent probe amplification, *PCR* polymerase chain reaction, *CGH* comparative genomic hybridization, *ASO* allele-specific oligonucleotide, *IHC* immunohistochemistry. Cancer syndromes: *HLRCC* hereditary leiomyomatosis and renal cell carcinoma, *HBOC* hereditary breast and ovarian cancer, *FAP* familial adenomatous polyposis

  - Trilateral disease refers to unilateral or bilateral retinoblastoma plus intracranial neuroblastoma and/or pineoblastoma
  - Second primary tumors occur in 25% (up to 50% after external beam radiation)
    - Brain tumors
    - Osteosarcoma
    - Soft tissue sarcomas
    - Melanoma
    - Leukemias and lymphomas
- Molecular genetics
  - *RB1* is a large gene with 27 exons at chromosome 13q14
  - Over 1,000 distinct mutations have been identified
    - Majority is single base substitutions, frameshift mutations, or splice mutations that result in a premature termination codon and protein truncation
    - Recurrent mutations at 14 methylated CpG dinucleotides account for 25% of mutations
  - Chromosome deletion of band 13q14 is found in 5–8% of index cases and is associated with developmental delay and birth defects
  - Wild-type pRB is a 928 amino acid protein that regulates cell cycle progression and transcription
    - Binds and inhibits the transcription factor, E2F, during G0 and G1
    - Phosphorylated by cyclin D-cdk 4/6 and releases E2F
    - E2F induces genes that mediate S-phase entry
- Molecular genetic diagnosis
  - In patients with family history or bilateral retinoblastoma
    - Sequence analysis or mutation scanning of *RB1* full gene is performed on peripheral blood DNA to identify small deletions, insertions, and base substitutions of germline origin. The mutation detection rate is 70%

Table 23.3 Genes associated with common syndromes

Syndrome	Gene	Chromosome	Key clinical features
Retinoblastoma	*RB1*	13q14	Retinoblastoma
			Intracranial neuroblastoma or pineoblastoma
Li–Fraumeni syndrome	*TP53*	17p13.1	Adrenocortical carcinomas Breast cancer Brain tumors Soft tissue and bone sarcomas
Hereditary breast and ovarian cancer	*BRCA1*	17q11–21	Breast cancer Ovarian cancer Prostate cancer Pancreatic cancer Breast cancer in men and women Ovarian cancer Prostate cancer
	BRCA2	13q12–13	
Fanconi anemia	*FANC A–P*		Physical abnormalities Bone marrow failure Acute myelogenous leukemia Squamous cell carcinoma of the head and neck Vulvar and cervical carcinoma Hepatocellular carcinoma Breast cancer
Lynch syndrome	*MLH1*	3p21–23	Colorectal cancer with microsatellite instability Gastric cancer Small intestinal cancer Ovarian and endometrial cancers Upper urinary tract transitional cell carcinoma
	MSH2	2p21	
	MHS6	2p16	
	PMS1	2q	
	PMS2	7p22	
Familial adenomatous polyposis	*APC*	5q21–22	Multiple polyps with high risk of colorectal cancer Duodenal or ampullary adenomas and carcinomas Desmoid tumors Congenital hypertrophy of retinal pigment epithelium Hepatoblastomas
Peutz–Jeghers syndrome	*STK11(LKB1)*	19p13.3	Multiple gastrointestinal hamartomas Pigmented oral/labial macules Breast, colon, gastric, and ovarian cancer Sex cord tumors with annular tubules
Juvenile polyposis syndrome	*SMAD4*	18q21	Gastrointestinal hamartomatous polyps
	BMPR1A	10q22	Colon and other gastrointestinal cancers
Hereditary diffuse gastric carcinoma	*CDH1*	16q22.1	Diffuse signet ring gastric cancer Lobular breast cancer in women
Wermer syndrome	*MEN1*	11q13	Pituitary adenoma Parathyroid adenoma Pancreatic neuroendocrine tumors
Multiple endocrine neoplasia, MEN2A	*RET*	10q11.2	Medullary thyroid cancer Pheochromocytoma Hyperparathyroidism
Multiple endocrine neoplasia, MEN2B	*RET*	10q11.2	Medullary thyroid cancer Pheochromocytoma Marfanoid habitus Gangliomas and neuromas

(*continued*)

Table 23.3 (continued)

Syndrome	Gene	Chromosome	Key clinical features
Familial Wilms tumor	*WT1*	11p13	Wilms tumor Syndromic features associated with *WT1* mutations
	FWT1	17q12–21	
	FWT2	19q	
	WT3	16q	
	WT5	7p11	
Hereditary leiomyomatosis and renal cell carcinoma (RCC)	*FH*	1q42.1	Type II papillary RCC Multiple cutaneous leiomyomas Uterine fibroids
Hereditary papillary RCC	*MET*	7q31	Type I papillary RCC
			Renal papillary adenomas
Hereditary nonpapillary RCC	Chromosome 3p translocations	3p14.2	Nonpapillary clear cell RCC
von Hippel–Lindau	*VHL*	3p25	Central nervous system hemangioblastomas Endolymphatic sac papillary adenocarcinoma Renal cell carcinoma Renal and pancreatic cysts Pancreatic endocrine tumors Pheochromocytomas
Hereditary prostate cancer	*HPC 1*	1q24–25	Early onset prostate cancer
	HPCX	Xq27–28	
	BRCA1	17p12	
	BRCA2	13q12–13	
Hereditary melanoma	*CDKN2A*	9p21	Melanoma Multiple dysplastic nevi Pancreatic cancer Breast cancer
	CDK4	12q14	
Gorlin syndrome	*PTCH*	9q22.3	Multiple basal cell carcinomas Basal cell nevi Palmar or plantar pits Odontogenic keratocyst Ectopic calcifications Skeletal abnormalities including bifid, fused, or absent ribs or vertebrae Macrocephaly (>97 percentile) Cardiac or ovarian fibroma Medulloblastoma Cleft lip or palate
Xeroderma pigmentosum	*XP A–G*		Skin and eye cancer Acute UV sensitivity Corneal opacities Neurocognitive deficits
	ERCC1		
Carney complex	*PRKAR1A*	7q23–24	Cutaneous pigmented lesions Cutaneous, breast, and cardiac myxomas Psammomatous melanotic schwannomas Large cell calcifying Sertoli cell tumors of the testis Primary pigmented nodule adrenocortical disease Thyroid carcinoma Breast adenomas

(*continued*)

Table 23.3 (continued)

Syndrome	Gene	Chromosome	Key clinical features
Cowden syndrome	*PTEN*	10q22–23.3	Hyperkeratotic oral papules, facial trichilemmomas, acral keratoses Cerebellar gangliocytic tumors Macrocephaly Breast fibroadenomas and carcinomas Thyroid follicular tumors Endometrial cancer Mental retardation Multiple gastrointestinal hamartomas
Birt–Hogg–Dubé syndrome	*FLCN*	17p11.2	Fibrofolliculomas, trichodiscomas, and acrochordons Lung cysts and spontaneous pneumothorax Oncocytic hybrid tumors, chromophobe RCCs, and benign oncocytomas of the kidney
Tuberous sclerosis	*TSC1*	9q34	Cortical tubers Subependymal glial nodules Subependymal giant cell astrocytomas Cardiac rhabdomyomas Renal angiomyolipomas Pulmonary lymphangiomatosis Retinal hamartomas Facial angiofibromas, periungual fibromas, hypopigmented macules, and Shagreen patches
	TSC2	16p13.3	
von Recklinghausen syndrome	*NF1*	17q11.2	Plexiform neurofibroma Multiple neurofibromas Café au lait macules Melanocytic iris hamartomas (Lisch nodules) Optic nerve glioma Axillary or inguinal freckling Specific bone abnormalities
Neurofibromatosis type 2	*NF2*	22q12.2	Acoustic schwannomas Meningiomas Ependymomas Astrocytomas Lens opacities
Ataxia-telangiectasia	ATM	11q22.3	Progressive cerebellar ataxia Oculomotor apraxia Choreoathetosis Telangiectasias Leukemia and lymphoma Breast cancer

- Analysis for submicroscopic whole-exon and multiexon deletions, insertions, and rearrangements by MLPA or other methods detects 15% of mutations
- Gross deletion/duplication analysis by FISH using commercially available probes for 13q14
- Targeted mutation analysis for a specific panel of the recurrent point mutations

– In patients with bilateral retinoblastoma and no family history
  - If no mutation is identified in peripheral blood DNA, tumor DNA can be tested by DNA sequencing analysis, loss of heterozygosity, and methylation analysis
  - *RB1* mutations identified in the tumor DNA should be verified by mutation analysis of peripheral blood DNA.

If these mutations are not detected, mosaicism is assumed
  - In patients with unilateral disease and no family history
    - Molecular genetic testing for *RB1* gene mutation or hypermethylation is initially performed on tumor DNA, followed by targeted testing of peripheral blood DNA
    - 15% of patients with unilateral retinoblastoma have *RB1* mutations in peripheral blood DNA indicating either inherited germline mutation or mosaic state
  - Predictive testing is standard management for at-risk family members
- Management
  - Goal is preservation of life and sight through early diagnosis and treatment
  - In patients with germline mutations, eye examination is performed every 4 weeks for the first year of life and less frequently thereafter
  - Photocoagulation, cryotherapy, chemotherapy, and enucleation are potential treatments
  - Limit exposure to radiation, UV light, and tobacco

23.3 Li–Fraumeni Syndrome

- Overview
  - Rare, autosomal dominant syndrome with predisposition to multiple primary cancers at a young age
- Clinical and pathologic features
  - Cancer penetrance is 50% by age 30 and 90% by age 60
  - Characteristic "SBLA" (sarcoma, breast, leukemia, and adrenal gland) tumors
    - Soft tissue and bone sarcomas
    - Premenopausal breast cancer
    - Brain tumors
    - Leukemia
    - Adrenocortical carcinomas
  - Individuals with childhood choroid plexus tumors or adrenocortical carcinomas have a high likelihood of germline *TP53* mutations
  - Estimated risk of second and third primary cancers is nearly 60% and 40%, respectively
- Genetics
  - *TP53* is a large tumor suppressor gene with 11 exons (exon 1 is noncoding) on chromosome 17p13.1
  - Almost 300 distinct *TP53* germline mutations
    - Mostly missense mutations, generating a truncated protein
    - Exons 5–8, which encode the DNA-binding region, account for 70% of mutations; 90% arise in exons 4–9
  - Wild-type p53 is a transcription factor considered the "guardian of the genome"
    - In unstressed cells, it is inactivated and targeted for degradation by *MDM2* ubiquitin ligase
    - In stressed cells, it is activated by phosphorylation, accumulates, and induces gene expression for
      - Cell cycle arrest to allow for DNA repair
      - Apoptosis to eliminate irreparably damaged cells
    - Mutant p53 does not effectively bind DNA and activate transcription, leading to loss of its tumor-suppressive functions
  - Of note, somatic mutations in p53 are present in 50% of all human cancers; however, the spectrum of sporadic and inherited tumors is quite different
  - CHEK2 mutations were originally reported in a subset of families who met Li–Fraumeni syndrome (LFS) criteria but who did not have detectable *TP53* mutations
    - Recent studies have determined that it is not a major cause of LFS or Li–Fraumeni-like (LFL) syndrome
- Clinical diagnosis
  - Classic LFS criteria
    - Proband diagnosed with sarcoma before age 45
    - First-degree relative with an LFS tumor (breast, brain, sarcoma, leukemia, adrenal) diagnosed before age 45

    - Another first- or second-degree relative with any cancer before age 45 or a sarcoma at any age
  - 2009 Chompret LFS criteria
    - Proband who has an LFS tumor before age 46 and at least one first- or second-degree relative with an LFS tumor before age 56 or multiple tumors
    - Proband with multiple tumors (two LFS tumors), the first diagnosed before age 46
    - Proband diagnosed with adrenocortical carcinoma or choroid plexus tumor, irrespective of family history
  - LFL syndrome shares the same spectrum of tumor predisposition as classic LFS but is defined by two sets of less stringent criteria, Birch criteria and Eeles definition
- Molecular genetic diagnosis
  - 70% of LFS and 10–20% of LFL syndrome can be attributed to detectable *TP53* mutations
  - Sequence analysis of the full gene detects 95% of *TP53* mutations
  - Limited sequence analysis of exons 4–9 may be preferred by some laboratories for its efficiency, since it detects 90–95% of *TP53* mutations
  - Deletion/duplication analysis of the coding region, exon 1, and promoter by MLPA detects an additional 1%
  - Predictive testing can be performed in at-risk relatives who have a known mutation in the family
- Management
  - Earlier mammography and colonoscopy
  - Risk-reducing mastectomy considered on a case by case basis
  - Avoidance of radiation to reduce risk of secondary malignancy
  - Families should be cautioned prior to genetic testing that there is no evidence demonstrating benefit of increased surveillance or early intervention
  - Advexin, an adenoviral-based *TP53* gene replacement, may have potential in the targeted treatment of LFS tumors

23.4 Hereditary Breast and Ovarian Cancers

23.4.1 Overview

- Breast cancer
  - Most common cancer and second most common cancer-related cause of death in women
    - Lifetime risk is 12%
  - Hereditary breast cancer accounts for 5–10% of all breast cancer cases
  - Hereditary breast and ovarian cancer (HBOC) syndrome due to *BRCA1* or *BRCA2* gene mutations represents 60% of hereditary breast cancer
    - Carrier frequency for both *BRCA* mutations
      - 1/300–1/800 in the US
      - 1/40 in Ashkenazi Jews
  - Other hereditary syndromes associated with increased risk of breast cancer
    - LS with *MLH1* germline mutations
    - LFS (*TP53*)
    - Cowden syndrome (*PTEN*)
    - Peutz–Jeghers syndrome (*LKB1*/*STK11*)
    - Hereditary diffuse gastric cancer syndrome (*CDH-1*)
    - Hereditary melanoma (*CDKN2A*)
    - Ataxia-telangiectasia (*ATM*)
    - *CHEK2* variant c.1100delC (twofold to threefold in females, tenfold in males)
- Ovarian cancer
  - Occurs in approximately 1 in 50 women
  - Hereditary predisposition accounts for 5–15% of all ovarian cancer cases
    - HBOC syndrome in 65–85%
    - Lynch syndrome in 10–15%

23.4.2 BRCA1

- Overview
  - Autosomal dominant transmission with a carrier frequency from 1/400 to 1/2,000
  - *BRCA1* mutations account for about 50% of hereditary breast cancers and 80% of hereditary ovarian cancers

  - Carriers exhibit 60–90% lifetime risk of breast cancer and 40% lifetime risk of ovarian cancer
  - Breast cancer penetrance is 20% by age 40 and 60% by age 50
  - Twofold to threefold increased risk of prostate cancer, but no increased risk of male breast cancer
  - Increased risk of pancreatic cancer in both sexes
- Clinical and pathologic features
  - *BRCA1*-asscociated breast cancers
    - High-grade invasive ductal carcinoma with lymphoplasmacytic inflammation
    - Medullary carcinoma
    - Lack of an in situ component
    - Triple-negative tumors (unlike *BRCA2*)
    - Basal cell phenotype with expression of high molecular weight cytokeratins CK5/6, CK14, and CK17
    - High frequency of *TP53* mutations
    - Prognosis similar to sporadic breast cancer when controlled for grade and stage
  - BRCA1-associated ovarian carcinomas
    - Serous carcinoma (80%) of the ovaries, fallopian tubes, or peritoneum
    - Aggressive, high-grade "type II" histopathology
    - Overexpression of p53, MIB1 (high proliferation), and cyclin E
    - Serous tubal intraepithelial carcinoma is a likely precursor lesion (found in 5–10% of prophylactic bilateral salpingo-oophorectomy [PBSO] specimen)
    - Survival advantage possibly due to increased susceptibility to platinum-based chemotherapy
- Molecular genetics
  - *BRCA1* is a tumor suppressor gene on chromosome 17q11–21
  - Large gene with 24 exons, 22 of which are coding
    - Exon 11 is the largest region, encodes 60% of the protein, and accounts for over 500 mutations
    - Amino-terminal RING finger domain is important to its function
  - Over 1,600 distinct mutations, polymorphisms, and variants have been identified
    - Majority are frameshift mutations causing protein truncation
    - Most are unique to kindred
  - Founder mutations
    - Ashkenazi Jewish: 185delAG and 5385insC, 1.1% and 0.1% carrier frequency, respectively
    - Dutch: deletion in exon 13, mutation in exon 22
    - European: duplication in exon 13, deletion in exons 8–9 and 14–20
  - *BRCA1* gene product is involved in repair of DNA damage and regulation of transcription and cell cycle progression and maintains genomic integrity as a "caretaker gene"
    - Phosphorylated by ATM and CHK2 in response to DNA damage
    - Complexes with BARD1 to form heterodimer with ubiquitin ligase activity
    - Heterodimer interacts with BRCA2–RAD51 complex to repair double-strand breaks (DSBs) by homologous recombination (HR)
    - Also promotes growth arrest and DNA repair by binding to several cell cycle proteins and by activating p53-mediated transcription
  - BRCA1 is expressed in a wide range of tissues, especially testis and thymus
- Management (see below)

23.4.3 BRCA2

- Overview
  - Autosomal dominant with a carrier frequency of 8/100–1/1,000
  - Accounts for about 25% of hereditary breast cancers
  - Carriers have 40–70% lifetime risk of breast cancer and 20% lifetime risk of ovarian carcinoma
  - Men have increased risk of breast cancer, 100-fold greater than general population (not seen in *BRCA1* carriers)

  - Increased risk of prostate cancer (seen in both *BRCA1* and *BRCA2* carriers)
- Clinical and pathologic features
  - Histology and prognosis similar to sporadic breast cancer
- Molecular genetics
  - *BRCA2* is a large tumor suppressor gene with 27 exons on chromosome 13q12–13
  - Mutations usually truncate the protein and lead to loss of function
  - Common founder mutations
    - Ashkenazi Jewish: 6174delT, 8/100 carrier frequency
    - Icelandic: 999del5
  - BRCA2 gene product is involved in DNA repair and transcription regulation
    - Complexes with RAD51, localizes with BRCA1 at DSB through interaction with PALB2, and repairs DSB by HR
- Clinical diagnosis of BRCA1 and BRCA2 mutations
  - According to National Comprehensive Cancer Network 2010 guidelines, HBOC is suspected and risk evaluation is warranted if one of the following is present
    - Early onset breast cancer (<50 years)
    - Ovarian/fallopian tube/peritoneal cancer at any age
    - Two primary breast and/or ovarian/fallopian tube/peritoneal cancers in a single individual or in close relatives on the same side of the family
    - Male breast cancer in the family
    - At-risk populations
    - Member of the family with a known *BRCA1* or *BRCA2* mutation
  - Multiple breast cancer risk assessment models are available to estimate the likelihood that an individual or family harbors a mutation (BRCAPRO, Myriad II, BOADICEA, Gail, Claus, IBIS, etc.)
  - American Society of Clinical Oncologists (ASCO) cautions against reliance on a numerical threshold for referral for genetic risk assessment and testing. An experienced professional should interpret the pedigree and determine the appropriateness of testing
  - Of note, 30–50% of women with germline BRCA mutations have no significant family history
- Molecular genetic diagnosis of BRCA1 and BRCA2 mutations
  - Sequence analysis of full genes (or mutation scanning) is the first step if an individual comes from a family with unknown mutation
    - Variants of uncertain significance (VUS) are found in 10–15% of individuals who undergo full sequence analysis of *BRCA1* and *BRCA2*. Further segregation analysis or functional assays may be warranted
  - If no detectable mutation, deletion/duplication analysis is performed
  - Targeted mutation analysis is the initial test for an individual from a population with founder mutations
    - For those of Ashkenazi Jewish ancestry, analysis for the three founder mutations detects 90% of mutations
  - Family-specific mutation analysis is the primary approach if the individual comes from a family with a known pathogenic mutation
  - Ideally, predictive testing should occur after the family-specific mutation is identified in an affected relative for the most meaningful interpretation of genetic test results
- Management
  - Earlier screening mammography as well as MRI of the breasts
  - Annual or semiannual transvaginal ultrasound and CA-125
  - Prostate and breast cancer screening in males with *BRCA2* mutation
  - Prophylactic bilateral mastectomy reduces risk of breast cancer by 90%
  - PBSO reduces risk of breast cancer by 50% and ovarian cancer by 80–95%. The risk of primary peritoneal cancer is 2% after surgery
  - PARP inhibitors are promising targeted therapeutics
    - Prevent the repair of single-strand breaks (SSBs), which allows SSBs to progress to DSBs

- Selectively kill cancer cells with defective DNA repair and spare normal cells, i.e., exhibit "synthetic lethality" to cells with mutant BRCA1 or BRCA2
- Numerous clinical trials for treatment of hereditary and sporadic breast and ovarian cancer
- In the future, it could expand the application of *BRCA* genotyping to a greater proportion of breast and ovarian cancer patients as a companion diagnostic test

23.5 Fanconi Anemia

- Overview
  - Fanconi anemia (FA) is an autosomal recessive syndrome (except for FANCB, which is X-linked recessive) associated with malformations, bone marrow failure, and increased risk of malignancies
  - Carrier frequency is 1 in 300. It increases to 1 in 90 in Ashkenazi Jewish, 1 in 80 in Afrikaners, and 1 in 100 in black Africans
- Clinical and pathologic features
  - Physical abnormalities (60–75%): short stature, skeletal deformities, abnormal skin pigmentation, and renal, neurological, cardiac, eye, and ear anomalies
  - Progressive bone marrow failure in children
  - Adult-onset aplastic anemia
  - Hematologic malignancies (10–30%)
    - Myelodysplastic syndrome (MDS)
    - Acute myelogenous leukemia (AML)
  - Solid tumors (25–30%)
    - Squamous cell carcinoma of head, neck, and esophagus
    - Vulvar and cervical carcinoma
    - Hepatocellular carcinoma
    - Breast cancer (*BRCA2, BRIP1*, and *PALB2*)
  - Diverse endocrine abnormalities
  - Mental retardation (20%)
- Molecular genetics
  - Fifteen genes responsible for known FA complementation groups, *FANC A–P* (approved symbols: *BRCA2* [D1], *BRIP1* [J], *PALB2* [N], *RAD51C* [O], and *SLX4* [P])
  - Significant proportion of FA is attributable to mutations in *FANCA* (60–70%) and *FANCC* (15%)
  - Biallelic mutations in *BRCA2* (*FANCD1*) result in a severe phenotype, in which 97% develop cancer by age 5
  - Many of the FA gene products form a nuclear complex involved in monoubiquitination of FANCD2. FANCD2 interacts with other proteins to promote genomic stability
  - Hematopoietic somatic mosaicism is present in 10–20% of FA patients
- Diagnosis
  - Chromosome breakage studies
    - Peripheral blood lymphocytes from FA homozygotes demonstrate cytogenetic aberrations (breakage, rearrangement, radial chromosomes) when cultured with clastogenic agents diepoxybutane (DEB) or mitomycin C (MMC)
    - FA heterozygotes exhibit normal results
  - Complementation analysis
    - Responsible complementation group is identified by expressing cDNAs of the 15 genes in the patient's cells and determining which one corrects the sensitivity to DEB/MMC
    - Available clinically
  - Sequence analysis is subsequently performed to identify the mutation in the responsible complementation group
  - Deletion/duplication analysis is clinically available for *FANCA* and *FANCB*
  - Targeted mutation analysis is possible for the common c.345+4A mutation of *FANCC* in Ashkenazi Jewish patients
  - Carrier testing in relatives is performed for the pathogenic mutation identified in the kindred. Carriers are not at risk of developing the autosomal recessive or X-linked disorders
- Management
  - Regular blood counts and annual bone marrow aspirate or biopsy

Table 23.4 Hereditary colorectal cancer syndromes

Type of polyposis	Syndrome
Nonpolyposis	Lynch syndrome (LS; hereditary nonpolyposis colorectal cancer, HNPCC)
Adenomatous polyposis	Familial adenomatous polyposis (FAP)
	MUTYH-associated polyposis (MAP)
Hamartomatous polyposis	Peutz–Jeghers syndrome (PJS)
	Juvenile polyposis syndrome (JPS)
	PTEN hamartoma tumor syndrome (PHTS); see Sect. 23.9, "Genodermatoses"
Hyperplastic polyposis	Hereditary predisposition to multiple hyperplastic polyps (including sessile serrated adenomas) and colorectal cancer, without a clear genetic basis

  - Avoidance of radiation due to increased sensitivity
  - Androgens and growth factors to stimulate hematopoiesis
  - Hematopoietic stem cell transfers prior to multiple blood transfusions and hematologic malignancy as a curative measure

23.6 Hereditary Gastrointestinal Cancers

23.6.1 Overview

- Colorectal cancer
  - Third most common malignancy and cancer-related cause of death in men and women in industrialized countries
  - Lifetime risk is about 5%
  - 20–30% of colorectal cancers have a familial basis
  - Approximately 5% of colorectal cancers are due to well-defined inherited syndromes associated with highly penetrant single gene mutations
- Hereditary colorectal cancer, see Table 23.4

23.6.2 Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer)

- Overview
  - Autosomal dominant syndrome caused by deleterious mutations in several of the mismatch repair genes
  - Responsible for 2–4% of colorectal cancers (CRCs) and 2% of endometrial cancers
  - Incidence is about 1 in 1,000 live births
- Clinical features
  - Increased risk of colorectal, upper GI (small bowel and gastric), biliary, gynecologic (ovary and endometrium), and upper urinary tract (ureter and renal pelvis transitional cell carcinoma) cancers
  - *MLH1* and *MLH2* mutation carriers have a 70–80% risk of CRC with a mean age of diagnosis in the mid-40s. Female carriers have a 30–60% risk of endometrial cancer
  - *MSH6* mutation carriers have a lower risk of CRC (25%) with later age at diagnosis (mid-50s), higher risk of endometrial cancer (70%), and lower risk of other LS-related cancers
  - *PMS2* mutation carriers have substantially lower risk of cancer
  - Muir–Torre syndrome
    - Colorectal carcinoma
    - Multiple sebaceous tumors
    - Gastric, small intestinal, gynecologic, and kidney cancers
    - *MSH2* and *MLH1* mutations
  - Turcot syndrome
    - Colorectal carcinoma
    - Glioblastoma
    - *MLH1* and *PMS1* mutations
  - Familial colorectal cancer type X
    - Fulfills Amsterdam I criteria but does not have a detectable mismatch repair gene defect
- Pathologic features

  - Two-thirds of cancers in proximal colon (right-sided)
  - Accelerated carcinogenesis leads to few, if any, adenomatous polyps associated with the tumor
  - MSI-high CRC phenotype
    - Mucinous, signet ring, or poor differentiation
    - Tumor-infiltrating lymphocytes
    - Crohn disease-like lymphocytic reaction
    - Medullary growth pattern
  - Risk of synchronous or metachronous tumors
  - Stage for stage, LS patients have better prognosis than patients with sporadic CRC but do not seem to benefit from adjuvant chemotherapy with 5-FU
- Molecular genetics
  - Mutations in one of several DNA mismatch repair genes; loss of these housekeeper genes leads to "mutator phenotype"
    - *MLH1* at 3p21–23 (30–40%)
    - *MSH2* at 2p21 (30–40%)
    - *MSH6* at 2p16 (10%)
    - *PMS1* at 2q
    - *PMS2* at 7p22
  - Small subset of families has deletions in the 3′ region of the *EPCAM* gene resulting in *EPCAM–MSH2* fusion transcripts
  - Mismatch repair system recognizes and corrects base-pair mismatches and insertion–deletion loops during DNA replication
  - Tumors with mismatch repair gene mutation display "microsatellite instability" (MSI), in which thousands of insertion or deletion mutations are found in the repetitive sequences of microsatellites distributed throughout the genome
- Clinical criteria
  - Original Amsterdam I criteria did not account for extracolonic cancers and were too stringent for clinical purposes
  - Amsterdam II criteria
    - Three relatives with histologically verified CRC or other LS-associated cancer, one of whom is a first-degree relative of the other two
    - Two successive generations affected
    - One person diagnosed before age 50
    - Exclusion of FAP
  - 2004 Revised Bethesda guidelines
    - CRC diagnosed before age 50
    - CRC with MSI-H phenotype on histology diagnosed before age 60
    - Synchronous or metachronous CRC or other LS-associated cancer
    - CRC or other LS cancer diagnosed before age 50 in one first-degree relative
    - CRC or other LS cancer diagnosed at any age in two first- or second-degree relatives
  - Revised Bethesda guidelines show improved but still rather low sensitivity, as 30% of LS cases are not identified by these criteria
  - Models utilizing personal and family history to estimate the probability of mismatch repair gene mutation include PREMM, MMRPro, and MMRpredict
- Molecular genetic diagnosis
  - If patient meets any of the Bethesda guidelines and tumor is available, the most cost-effective approach begins with MSI and/or immunohistochemical (IHC) analysis
    - Microsatellite instability (see Chap. 1)
      - PCR-based testing for differences in the length of sequences of specific microsatellite markers between tumor and normal patient DNA
      - National Cancer Institute consensus workshop recommended a panel of five microsatellite markers for the detection of MSI, including BAT25, BAT26, D2S123, D5S346, and D17S250 ("Bethesda panel," please see Table 23.5)
        - MSI-high (MSI-H): instability in two or more markers, which is found in 95% of LS
        - MSI-low (MSI-L): instability in only one of five markers, which is not generally found in LS
        - MSI-stable (MSS): no instability in any of the markers

Table 23.5 Bethesda panel

Microsatellite marker	*Type of repeats*
BAT25	Mononucleotide repeats
BAT26	Mononucleotide repeats
D2S123	Dinucleotide repeats
D5S346	Dinucleotide repeats
D17S250	Dinucleotide repeats

  - Sensitivity is 80–90% for *MLH1* and *MSH2* mutations and 55–80% for *MSH6* mutations, and specificity is 90%
  - Of note, MSI is found in 15% of all colon cancers. In sporadic cases, it is due to promoter hypermethylation of the *MLH1* gene
- IHC for mismatch repair proteins
  - Loss of expression of one of four major mismatch repair proteins (due to pathogenic mutation) nearly always predicts the tumor is MSI-H
  - Absence of MLH1 indicates either germline mutation in LS or promoter hypermethylation in sporadic CRC
  - Sensitivity is 85% and specificity is 90%, comparable to the performance of MSI testing
  - More widely available and reduces costs by identifying the gene to sequence
- If tumor exhibits MSI-H phenotype on MSI or IHC testing, or if tumor is not available, proceed to direct DNA testing
  - Sequence analysis
    - Start with *MLH1* and *MSH2*, then "minor" genes
    - Mutation detection rate is 90–95% for *MLH1* and 50–80% for *MSH2*
    - Routine sequencing does not detect large deletions with breakpoints outside of exons because the wild-type allele masks the mutant one
  - Conversion analysis
    - Method to separate paternal and maternal alleles to unmask potential mutations for analysis
    - Not in routine clinical use, because it is cumbersome to perform in comparison to newer techniques, like MLPA
  - Deletion/duplication analysis (by MLPA, etc.)
    - Detection rate is 5% for *MLH1* and 20% for *MSH2*
- Additional testing is necessary to distinguish sporadic CRC cases with MSI in patients with loss of MLH1 expression on IHC or in patients without detectable mismatch repair gene mutation
  - BRAF V600E mutation status
    - Mutation is found in 40–90% of sporadic tumors with MSI, but not found in LS tumors
  - Hypermethylation of the MLH1 gene promoter
- Recently, some authors recommend and centers perform screening for mismatch repair deficiency in all colorectal and/or endometrial cancers
- Management
  - Earlier and more intense surveillance by colonoscopy
  - Prophylactic colectomy

23.6.3 Familial Adenomatous Polyposis

- Overview
  - Highly penetrant autosomal dominant syndrome conferring increased risk of gastrointestinal and other carcinomas
  - Incidence of 1 in 5,000 to 1 in 12,000 live births
  - Responsible for less than 1% of all CRC cases
- Clinical and pathologic features
  - Colorectal manifestations
    - Hundreds to thousands of small adenomatous polyps (Fig. 23.1)
    - Lifetime risk of CRC is almost 100%, with an average age at diagnosis of 40 years
  - Upper gastrointestinal manifestations
    - Fundic gland polyps, with or without foveolar dysplasia (25–60%)

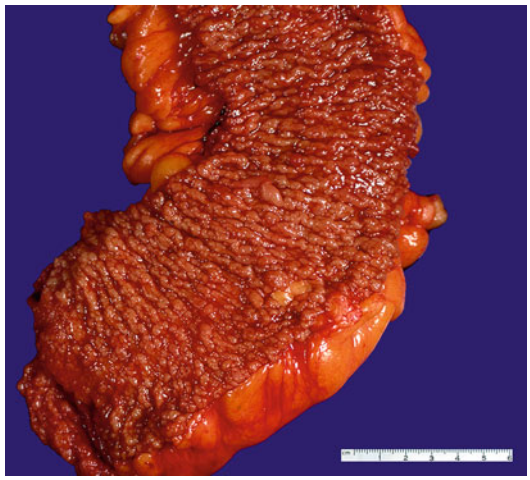


Fig. 23.1 Familial adenomatous polyposis (FAP). Numerous tubular adenomas are present in the colon resection specimen

- Gastric adenocarcinoma (<1%)
- Duodenal adenomas, especially papillary or periampullary (80%)
- Duodenal and ampullary adenocarcinoma (5–10%)

– Extraintestinal manifestations
  - Congenital hypertrophy of the retinal pigment epithelium (CHRPE) (70–80%)
    – Fundoscopic examination in at-risk infants is an early diagnostic test for FAP
  - Desmoid tumors (15%)
    – Locally invasive benign fibromatoses, which have a tendency to recur
    – Most commonly arise in the abdomen or abdominal wall, especially after surgery
    – Major cause of morbidity and mortality
  - Thyroid cancer (1–2%), three-fourths papillary
  - Hepatoblastomas
  - Supernumerary or missing teeth
  - Other benign tumors, including osteomas, adrenal cortical adenomas, and cutaneous lipomas

– Gardner syndrome
  - Colonic polyposis with osteomas of mandible and long bones, epidermoid cysts, and desmoid tumors
  - Still refers to patients with more prominent extraintestinal manifestations even though its distinction from classic FAP is historical

– Turcot syndrome
  - Colonic polyposis with CNS tumors (medulloblastoma and glioblastoma)

– Attenuated FAP
  - Colonic polyposis is less extensive (<100 polyps) and more often proximal
  - No extraintestinal features
  - Delayed age of CRC onset by 15 years
  - Often no significant family history

- Molecular genetics
  – *APC* tumor suppressor gene with 15 exons at 5q21–22
  – Over 700 different disease-causing APC mutations
    - Most introduce a premature stop codon leading to a truncated protein
    - Classic FAP involves mutations between codons 169 and 1,393
    - Mutation cluster region at 5′ end of exon 15, with 30% of all mutations in codons 1,061 and 1,309
  – Founder mutation, I1307K, in 6–8% of Ashkenazi Jews
    - T:A transversion creates an unstable adenine series that renders the gene susceptible to somatic mutations
  – 25% of cases due to de novo germline mutation
  – Genotype–phenotype correlations for APC gene mutations
    - Severe polyposis (>1,000 polyps) – codons 1,250–1,464
    - Attenuated FAP – extreme 3′ end and codons 1,596–2,644
    - Duodenal adenomas – codons 976–1,067
    - Desmoids – beyond codon 1,309, especially 1,445–1,580
  – APC plays a role in cell adhesion, migration, and proliferation
    - Part of the wingless (Wg)/Wnt signaling pathway

    - Complexes with axin and glycogen synthase kinase 3β
    - Complex phosphorylates β-catenin, leading to its degradation
    - Mutant APC allows for the accumulation of β-catenin
    - Excess nuclear β-catenin activates transcription of genes involved in proliferation and antiapoptosis, i.e., c-myc and cyclin D1
  - APC mutations present in 80% of sporadic colon cancers as well
- Clinical diagnosis
  - Greater than 100 polyps in the colon
  - Less than 100 polyps and a first-degree relative with FAP
- Molecular genetic diagnosis
  - Sequence analysis of the complete gene is the first-line screening tool and has a mutation detection rate of 90%
  - Deletion/duplication analysis by MLPA or conversion analysis increases the overall detection rate to 95%
  - Protein truncation testing has largely been replaced (detection rate is up to 80%)
  - Linkage analysis is very accurate (98%)
  - Predictive testing of at-risk individuals should begin around age 10 years
- Management
  - Screening for hepatoblastoma by liver ultrasound and serum alpha fetoprotein until age 5 years
  - Surveillance for other extraintestinal tumors
  - Early surveillance by colonoscopy and upper endoscopy
  - Prophylactic colectomy is recommended if multiple adenomas with advanced histology are present

23.6.4 MYH-associated Polyposis Syndrome

- Overview
  - Autosomal recessive disease with carrier frequency of about 1%
- Clinical features
  - Clinical and histopathologic features difficult to distinguish from attenuated FAP
  - Presents early in life with 20–100 adenomatous colorectal polyps
  - Lifetime risk of CRC is 8% in monoallelic carriers (threefold increase) and 80% in biallelic carriers
  - Increased risk of upper gastrointestinal and extraintestinal tumors seen in FAP
- Molecular genetics
  - Biallelic germline mutations of *MUTYH (MYH)* gene on chromosome 1p34.3
  - Y165C and G382D account for 80–85% of mutations in Caucasians
  - Gene product participates in base excision repair of oxidative DNA damage. Deficiency in this repair pathway allows G:C to T:A transversions, resulting in somatic mutation of the *APC* gene
- Diagnosis
  - Genetic testing is recommended for individuals with >10 colorectal adenomas and no detectable *APC* mutation. Biallelic *MUTYH* mutations are found in 20–25% of patients with negative *APC* testing
  - Targeted mutation analysis for Y165C and G382D or sequence analysis of the entire coding region may be performed, depending on the laboratory
  - Carrier testing is available
- Management
  - Similar to FAP

23.6.5 Peutz–Jeghers Syndrome

- Overview
  - Rare, autosomal dominant syndrome with hamartomatous polyps and increased risk of multiple cancers
  - Incidence of 1 in 8,000 to 1 in 280,000 live births
- Clinical and pathologic features
  - Early onset of multiple hamartomatous gastrointestinal polyps (usually <20) with characteristic pathologic features
    - Arborization of the muscularis mucosa

    - Involution of epithelial tissue and its infiltration into the muscularis mucosa appears as pseudocarcinomatous invasion
  - Characteristic pigmented macules of the lips, buccal mucosa, periorbital area, hands, and feet (95%)
  - Increased risk of cancer, lifetime risk is >90%
    - Breast (50%), colon (40%), pancreatic (35%), gastric (30%), and ovarian (20%)
    - Adenoma malignum of the cervix
  - Increased risk of benign ovarian sex cord tumors with annular tubules or testicular Sertoli cell tumors
- Molecular genetics
  - *STK11 (LKB1)* tumor suppressor gene at chromosome 19p13.3
  - Majority of mutations are truncating or missense; however, 15–40% may be large deletions
  - Up to 25% of cases are due to de novo mutations
  - Gene product is a serine threonine kinase that regulates cellular proliferation, polarity, and response to low energy levels
  - *STK11* mutation leads to dysregulation of the mTOR signaling pathway
- Diagnosis
  - Clinical criteria of two or more of the following
    - Characteristic mucocutaneous pigmentation
    - Two or more hamartomatous polyps
    - Family history of PJS
  - Molecular genetic diagnosis
    - Sequencing of *STK11* gene has a mutation detection rate of 70–80%
    - Deletion analysis by MLPA, etc. greatly increases the detection rate
    - Predictive testing can be performed in at-risk relatives for the kindred-specific mutation
    - Rare families do not have a detectable mutation on comprehensive testing and do not exhibit disease linkage to chromosome 19p13.3
- Management
  - Earlier screening schedules for gastrointestinal, breast, and testicular cancer

23.6.6 Juvenile Polyposis Syndrome

- Overview
  - Autosomal dominant syndrome with increased risk of gastrointestinal cancers
  - Incidence of 1 in 16,000 to 1 in 100,000 live births
  - Singular juvenile polyps (retention, hamartomatous, cystic) are found in 2% of children
  - True juvenile polyposis syndrome (JPS) involves multiple juvenile polyps
- Clinical and pathologic features
  - Multiple hamartomatous polyps of the GI tract
    - Dilated mucus-filled glands surrounded by edema and inflammation in lamina propria without smooth muscle proliferation
    - Indistinguishable from those arising sporadically and in Cowden syndrome
  - Increased risk of colon, stomach, small intestinal, and pancreatic cancers
    - 70% risk of colon cancer by age 60
    - *SMAD4* mutations linked to massive gastric polyposis and increased risk of stomach cancer
  - JPS–HHT
    - Symptoms of hereditary hemorrhagic telangiectasia (HHT) in patients with JPS (telangiectasias, epistaxis, arteriovenous malformations of the GI tract, and anemia)
    - *SMAD4* mutations at 5′ end
- Molecular genetics
  - Two gene mutations are well characterized
    - *SMAD4 (DPC4)* at chromosome 18q21 (15% of JPS cases)
    - *BMPR1A (ALK3)* at chromosome 10q22 (25% of JPS cases)
  - Deletions account for a significant proportion of pathogenic mutations

  - 25% of JPS cases are due to de novo mutations
  - Both genes are part of the TGF-β signaling pathway, which is involved in regulating the cell cycle, especially in colon cells
  - JPS cases originally attributed to *PTEN* mutations are most likely Cowden syndrome
- Diagnosis
  - Mainly based on clinical criteria, which require pathologic confirmation of all polyps. One or more of the following
    - At least 3–5 juvenile polyps of the colon
    - Multiple juvenile polyps throughout the gastrointestinal tract
    - Any number of polyps in a patient with a family history of JPS
  - Genetic testing
    - Sequence analysis of both genes has a detection rate of only about 20% for each gene
    - Deletion/duplication analysis of both genes has a detection rate of about 5–10% for each gene
    - If no mutation is found, genetic testing of PTEN is appropriate to determine if individual has PHTS rather than JPS
- Management
  - Screening by CBC, colonoscopy, and upper endoscopy beginning in midteens

23.6.7 Hereditary Diffuse Gastric Cancer

- Overview
  - Rare autosomal dominant disorder characterized by diffuse, poorly differentiated gastric carcinoma
- Clinical and pathologic features
  - Gastric cancer
    - Lifetime risk of 80%
    - Diffuse subtype ("linitis plastica") with signet ring cells
  - Breast cancer
    - Lifetime risk of 40% in women
    - Lobular subtype
- Molecular genetics
  - *CDH1* gene, 16 exons, at chromosome 16q22.1
  - Carriers develop cancer according to the "two-hit hypothesis," suggesting *CDH1* is a tumor suppressor gene
  - Over 100 distinct pathogenic mutations of various types have been identified
  - *CDH1* expresses E-cadherin, a calcium-dependent cell–cell adhesion molecule predominantly found in epithelial tissue
- Diagnosis
  - Clinical criteria
    - Two first- or second-degree relatives with diffuse gastric carcinoma, one diagnosed before age 50
    - Three first- or second-degree relatives, any age
  - Molecular genetic diagnosis
    - Sequence analysis of the entire coding region of *CDH1* gene detects mutations in 30% of individuals with a clinical diagnosis
- Management
  - Close surveillance with chromoendoscopy and endoscopic ultrasound
  - Prophylactic gastrectomy
  - Early mammography

23.7 Hereditary Endocrine Tumor Syndromes

23.7.1 Multiple Endocrine Neoplasia, Type 1 (Wermer Syndrome)

- Overview
  - Autosomal dominant syndrome with multiple endocrine tumors
  - Incidence is from 1 in 5,000 to 1 in 50,000 live births

  - Penetrance is 45% by age 30, 80% by age 50, and 95% by age 70
- Clinical features
  - Parathyroid adenomas (90%), hyperparathyroidism
  - Pancreaticoduodenal neuroendocrine tumors (30–80%), especially gastrinomas (Zollinger–Ellison syndrome) and insulinomas
  - Anterior pituitary adenomas (10–60%), especially prolactinomas
  - Adrenocortical tumors (40%)
  - Carcinoid of the thymus, bronchus, or stomach (10%)
  - Nonendocrine tumors such as facial angiofibromas, collagenomas, lipomas, and meningiomas
- Molecular genetics
  - *MEN1* tumor suppressor gene on chromosome 11q13
  - 70% of mutations lead to protein truncation
  - 10% of germline mutations are de novo
  - Gene product menin regulates transcription (e.g., inhibits transcription factor JunD), cell proliferation, apoptosis, and genome stability
  - Recently, germline mutations in *CDKN1B*/p27 gene have been associated with MEN1-like disease featuring familial pituitary tumors and hyperparathyroidism, termed "MEN4"
- Clinical criteria (either of the two)
  - Two major lesions (synchronous or metachronous) involving the parathyroid, endocrine pancreas, or anterior pituitary
  - One major lesion and first-degree relative with MEN1 syndrome
- Molecular genetic diagnosis
  - Direct DNA sequencing detects germline mutations in 80–90% of familial cases and 65% of simplex cases
  - Deletion/duplication analysis has a detection rate of 1–4%
  - Linkage analysis is available
  - Predictive testing can be performed for at-risk relatives who have a known mutation in the family
- Management
  - Screen with special clinical attention to endocrine system
  - Biochemical laboratory tests for hormones, glucose, and calcium
  - Imaging of head, chest, and abdomen

23.7.2 Multiple Endocrine Neoplasia, Type 2 MEN2A (Sipple Syndrome), MEN2B, and Familial Medullary Thyroid Cancer

- Overview
  - Autosomal dominant group of syndromes with medullary thyroid carcinoma (MTC) and other endocrine tumors
  - Incidence of 1 in 30,000 live births
  - Hereditary MTC accounts for 25% of patients with MTC
- Clinical and pathologic features
  - MEN2A
    - MTC (95%) – manifests before age 35
    - Pheochromocytoma (50%) – 10% are malignant in MEN2A and MEN2B
    - Hyperparathyroidism (10–30%)
    - Criteria: two or more tumors in an individual or close family member
  - MEN2B
    - MTC (100%) – diagnosed earlier than MTC in MEN2A, often before age 10
    - Pheochromocytoma (50%)
    - Ganglioneuromas of the GI tract
    - Mucosal neuromas of the lips and tongue
    - Distinctive facies with enlarged lips
    - Marfanoid habitus
  - Familial MTC (variant of MEN2A)
    - MTC (100%) – diagnosed later than MTC in MEN2
    - Criteria: four or more family members with MTC and no other MEN2 tumors
- Molecular genetics
  - *RET* proto-oncogene on chromosome 10q11.2

  - Large gene with 21 exons that encodes two major isoforms
  - MEN2A mutations
    - Exons 10 or 11 (98%)
    - Usually found in one of five cysteine residues at codons 609–634, especially codon 634 (almost 90%)
    - 5% of cases due to de novo mutations
  - MEN2B mutations
    - M918T substitution in tyrosine kinase region of exon 16 (95%)
    - A883F in exon 15
    - 50% of cases due to de novo mutations
  - Familial MTC mutations
    - Similar to MEN2A, one of five cysteine residues at codons 609–634
    - Codons 768 and 804
  - Gene product, RET, is a cell membrane receptor tyrosine kinase
    - Receptor ligands belong to the glial-derived neurotropic factor (GDNF) family
    - Binding of ligands leads to receptor dimerization, autophosphorylation, and signal transduction via the MAPK pathway
    - Point mutations in RET result in ligand-independent constitutive activity
  - Somatic *RET* mutations are found in 50% of sporadic MTC
  - Interestingly, 35–50% of cases of Hirschsprung disease have germline loss of function mutations in *RET* gene
- Molecular genetic diagnosis
  - Genetic testing should be offered to all patients diagnosed with MTC regardless of family history (as up to 7% of seemingly sporadic cases have germline mutations). It is especially important for patients with the following
    - MTC and family history for components of MEN2
    - MTC and pheochromocytoma, without family history
    - Bilateral adrenal tumors
  - Sequence analysis of select exons 10, 11, and 13–16 is first-line for MEN2A and familial MTC (detects 98% and 95% of mutations, respectively)
  - Targeted mutation analysis of M918T and A883F or sequence analysis of exons 15 and 16 is first-line for MEN2B (detects 98%)
  - Full gene sequence analysis is the next step, if *RET* mutation is not identified on the targeted analyses
  - Deletion/duplication analysis is not used because no large gene rearrangements have been reported
  - Linkage analysis is very accurate (95%)
- Management
  - American Thyroid Association Guidelines Task Force 2009 classified MEN2 genotypes into levels A–D based on the risk of aggressive MTC and offered management recommendations reflecting this classification
  - Prophylactic total thyroidectomy before age 5
  - Annual biochemical screening (calcitonin, urine/plasma catecholamines, PTH, calcium, etc.)

23.8 Hereditary Kidney Tumor Syndromes

23.8.1 Familial Wilms Tumor

- Overview
  - Various autosomal dominant syndromes confer an increased risk of embryonal tumors of the kidney (nephroblastomas; Wilms tumors)
  - Wilms tumor (WT) has an incidence of 1 in 10,000 live births
  - 10–15% of cases are attributed to heritable causes, which may or may not be associated with known syndromes
  - Nonsyndromic familial Wilms tumor (FWT) is genetically heterogeneous
    - Linked to various loci, including *FWT1* on 17q12–21, *FWT2* on 19q, *WT3* on 16q, and *WT5* on 7p11.2–15
    - Due to germline mutation in *WT1* gene in only a small number of families

  - Syndromic FWT, in which Wilms tumor occurs as part of a multiple congenital anomaly syndrome, arises from germline mutations in *WT1* gene on 11p13 (2–3% of WT cases)
    - WAGR syndrome: Wilms tumor, aniridia, genital anomalies, and mental retardation
    - Denys–Drash syndrome (DDS): Wilms tumor, nephropathy (mesangial sclerosis), and gonadal dysgenesis
    - Frasier syndrome (FS): nephropathy (focal segmental glomerulosclerosis), gonadoblastoma, gonadal dysgenesis, and rarely Wilms tumor
    - Genitourinary (GU) anomalies and Wilms tumor
  - Various overgrowth syndromes predispose to Wilms tumor. For example, Beckwith–Wiedemann syndrome (BWS) is linked to the locus referred to as *WT2* on 11p15.5
  - In 5% of apparently sporadic Wilms tumors, de novo germline *WT1* mutations are present
- Clinical and pathologic features
  - Clinically, it presents as abdominal mass in otherwise healthy child
  - Average age at presentation is 42 months for unilateral WT and 30 months for bilateral WT
  - Bilateral or multifocal Wilms tumors are present in 5–10% of patients with WT, but these are usually not hereditary
  - Pathologically, nephrogenic rests are precursor lesions
- Genetics of *WT1*-related Wilms tumor
  - *WT1* gene on 11p13
  - Syndromic genotype–phenotype correlations for *WT1*
    - WAGR: microdeletion of 11p13, encompassing *WT1* and *PAX6*
    - DDS: missense mutations in exons 8 and 9
    - FS: spice site point mutation in intron 9
    - GU anomalies: nonsense or frameshift mutations
  - Gene product is a zinc-finger transcription factor essential for normal development of the kidney and gonads
- Molecular genetic diagnosis
  - Individuals with WT in association with genitourinary anomalies or renal failure should undergo molecular genetic testing
    - Deletion analysis by cytogenetics, FISH, or MLPA for suspected WAGR syndrome
    - Sequence analysis of *WT1* for suspected DDS, FS, or GU anomalies syndrome
  - Individuals with nonsyndromic FWT, or bilateral WT without anomalies, may be considered for genetic testing on a case to case basis, but it is low yield
- Management
  - Screen individuals with syndromic FWT and siblings of affected individuals by abdominal ultrasound every 3–4 months until age 7 years

23.8.2 Hereditary Leiomyomatosis and Renal Cell Carcinoma

- Clinical and pathologic features
  - Rare, autosomal dominant syndrome
  - Renal cell carcinoma (RCC) (10–15%)
    - Most are type II papillary RCC, although other histologies have been reported
    - Median age at diagnosis is in mid-40s
    - Unilateral, solitary, and clinically aggressive
  - Multiple cutaneous leiomyomas (75%)
  - Uterine leiomyomas (fibroids) (almost 100%)
    - Early onset in 20s–30s
    - Large and numerous
  - Uterine leiomyosarcoma reported in a few cases
- Molecular genetics
  - *FH* tumor suppressor gene, 10 exons, at chromosome 1q42.1
  - Most mutations are missense ones (60%). Large deletions account for about 7% of mutations

  - Gene product, fumarate hydratase (FH), plays an important role in cellular energy metabolism
    - Catalyzes the conversion of fumarate to 1-malate in the mitochondrial tricarboxylic acid cycle
    - Inactivating mutation causes FH accumulation, decreased degradation of hypoxia-inducible factor (HIF), and overexpression of downstream genes involved in neoplasia
  - Homozygous mutations result in fumaric aciduria, a severe inborn error of metabolism characterized by rapidly progressive neurologic impairment
- Diagnosis
  - Clinical diagnosis is based on multiple cutaneous and uterine leiomyomas (MCULs) with at least one histologically confirmed leiomyoma or a single leiomyoma in the presence of a positive family history
  - Biochemical testing of fumarate hydratase enzyme activity in cultured skin fibroblasts or lymphoblastoid cells demonstrates reduced activity (<60%) in HLRCC patients
  - Molecular genetic testing
    - Sequence analysis/mutation scanning of the *FH* gene has a mutation detection rate of 80–100%
    - Deletion/duplication analysis by MLPA detects mutations in only a small fraction of individuals without variants on sequence analysis
  - Predictive testing can be performed in at-risk relatives when the family-specific mutation is known
- Management
  - Dermatologic and gynecologic examinations every 1–2 years
  - Abdominal/pelvic CT scan with contrast to evaluate for renal lesions every 2 years
  - Surgical excision, cryoablation, or laser removal of cutaneous lesions
  - Gonadotropin-releasing hormonal agonists or surgical resection (myomectomy and/or hysterectomy) for uterine fibroids

23.8.3 Hereditary Papillary Renal Cell Carcinoma

- Clinical and pathologic features
  - Rare, autosomal dominant syndrome that predisposes to multiple, bilateral kidney tumors with type I papillary RCC histology as well as papillary adenomas
- Molecular genetics
  - *MET* proto-oncogene on chromosome 7q31
  - Gene product is hepatocyte growth factor receptor (HGFR), which is a membrane receptor tyrosine kinase involved in an invasive growth program of stem cells and cancer cells
  - Cancer cells hijack the MET signaling pathway to activate key oncogenic pathways, angiogenesis, and metalloproteinase production leading to cellular dissociation and metastasis
- Diagnosis
  - Clinical diagnosis is based on bilateral or multifocal papillary RCC tumors in an individual without family history or a single tumor and a first- or second-degree relative
  - Differential diagnosis includes bilateral or multifocal tumors arising in the setting of acquired cystic kidney disease
  - Molecular genetic testing is available at a few laboratories worldwide and includes sequence analysis of the entire coding region and deletion/duplication analysis
- Management
  - Renal imaging, urinalysis, and urine cytology every 1–2 years beginning at age 25 years
  - Renal parenchymal-sparing surgery should be considered in patients with tumors larger than 3 cm
  - Foretinib, an oral dual-kinase small molecular inhibitor of the tyrosine kinase domains of MET and VEGFR2, is in clinical trials for papillary RCC

23.8.4 Hereditary Nonpapillary Renal Cell Carcinoma

- Overview
  - Autosomal dominant and recessive inheritance of nonpapillary clear cell RCC
  - Nonpapillary RCC accounts for 80% of all RCC
  - An estimated 2% of all RCC is inherited
- Clinical and pathologic features
  - Early onset cancer (<50 years of age) and bilateral disease suggest genetic predisposition
- Molecular genetics
  - Most families with hereditary RCC have normal chromosomes
  - A small number of families exhibit chromosome 3p translocations involving the 3p14.2 locus
- Diagnosis
  - von Hippel–Lindau, Birt–Hogg–Dubé, and tuberous sclerosis syndromes should be ruled out prior to testing; see below
  - Cytogenetic analysis of the peripheral blood of an affected family member may demonstrate the 3p translocation
- Management
  - Renal imaging every 1–2 years, beginning at age 35 years

23.8.5 von Hippel–Lindau Syndrome

- Overview
  - Autosomal dominant syndrome with nearly 100% penetrance characterized by multiple highly vascular tumors
  - Incidence is about 1 in 35,000 live births
  - Average life expectancy is 50 years, with RCC as the most common cause of death
- Clinical and pathologic features
  - CNS hemangioblastomas (cerebellum, spinal cord, and retina) (70%)
  - RCC, clear cell variant (40%)
  - Pheochromocytomas (10–20%)
  - Renal and pancreatic cysts
  - Pancreatic endocrine tumors
  - Endolymphatic sac papillary adenocarcinoma
  - Papillary cystadenomas of the epididymis or broad ligament of the uterus
  - VHL type 1
    - Low risk for pheochromocytoma
  - VHL type 2
    - High risk for pheochromocytoma associated with missense mutations
      - Type 2A: low risk for RCC
      - Type 2B: high risk for RCC
      - Type 2C: risk for pheochromocytoma alone
- Molecular genetics
  - *VHL* tumor suppressor gene at chromosome 3p25–26
  - Gene with 3 exons that encodes two functional isoforms by alternative splicing
  - Over 300 pathogenic germline mutations
    - Missense (70%), large deletions (30%)
    - Mutational hotspot at codon 167
  - 20% of germline mutations are de novo
  - Wild-type pVHL plays a major role in transcriptional regulation of hypoxia-inducible genes
    - pVHL forms a complex with elongin B and C to target hydroxylated hypoxia-induced factor-1α (HIF-1α) for degradation
    - Under hypoxic conditions, unhydroxylated HIF-1α accumulates and activates transcription of hypoxia-induced genes (VEGF, PDGF, TGF-α, etc.)
    - When VHL is mutated, HIF-1α accumulates and activates those genes as well, which induces cell proliferation and angiogenesis
  - Sporadic clear cell RCC occasionally features somatic mutation (50%) or hypermethylation (10–20%) of the VHL gene
- Clinical diagnosis
  - Clinical criteria for diagnosis of VHL in a proband

  - Two hemangioblastomas
  - One hemangioblastoma and at least one other characteristic tumor
  - At least one characteristic tumor and a family history of VHL syndrome
- Molecular genetic diagnosis
  - Genetic testing detects mutations in *VHL* gene in nearly 100% of individuals who meet clinical criteria
  - Sequence analysis of the exons has a mutation detection rate of 70%
  - Deletion/duplication analysis has a detection rate of 30%
  - Genetic counseling and predictive testing for all family members, starting with first-degree relatives, is recommended
- Management
  - Annual physical and eye examination and biochemical screening for urine or plasma catecholamines, beginning before age 5
  - Annual MRI of CNS, beginning in the early teens
  - Annual abdominal ultrasound, beginning in the midteens
  - Significant progress has been made in the development of targeted therapeutics for metastatic clear cell RCC. Multiple agents targeting the VEGF pathway are FDA approved and commonly used in clinical practice
    - Sunitinib and sorafenib are receptor tyrosine kinase inhibitors that target both VEGF and PDGF receptors
    - Bevacizumab is an anti-VEGF antibody
  - Everolimus, a small-molecule inhibitor of the mTOR pathway (mTOR regulates expression and stability of HIF-1α), may be considered in patients who fail anti-VEGF therapy

23.8.6 Birt–Hogg–Dubé Syndrome

- Increased risk of oncocytic hybrid tumors, chromophobe renal cell carcinomas, and oncocytomas; see Sect. 23.9, "Genodermatoses"

23.8.7 Tuberous Sclerosis Complex

- Increased risk of renal angiomyolipomas, renal cysts, and renal cell carcinoma; see below

23.9 Genodermatoses

23.9.1 Hereditary Melanoma

- Overview
  - About 10% of melanoma cases present in familial clusters, suggesting inherited mutation, shared environmental exposures, or both
  - Hereditary melanoma (HM) is an autosomal dominant syndrome attributable to high-risk, high-penetrance susceptibility alleles associated with early onset disease, multiple primary tumors, and multiple familial cases
    - *CDKN2A* mutations account for 20–40% of HM but only 0.2–2% of all melanoma cases
    - *CDK4* mutations have been reported in only a few families
  - Penetrance of hereditary melanoma due to *CDKN2A* varies by geographic region. By age 80, it is about 60% in Europe, 75% in the US, and 90% in Australia
  - *CDKN2A* mutation rate increases with the number of melanoma diagnoses in the family. Families with similar histories have a greater likelihood of mutation in lower-incidence countries
  - Low- to moderate-risk susceptibility alleles (primarily related to pigmentary traits, i.e., melanocortin-1 receptor gene, *MC1R*) are prevalent in the general population and serve as genetic modifiers
- Clinical and pathologic features
  - Multiple primary cutaneous malignant melanomas
  - Multiple dysplastic nevi
  - Increased risk of other cancers
    - Pancreas (38-fold increased risk, lifetime risk over 10%)
    - Breast

- Molecular genetics
  - *CDKN2A* gene on chromosome 9p21
    - Encodes two proteins, p14/Arf and p16/Ink4a, by alternative splicing
    - Most mutations affect the p16 transcript
    - Both gene products are potent tumor suppressors that regulate the cell cycle and apoptosis
      - p16 inhibits cyclin D1/cdk4 (see below), which leads to cell cycle arrest at G_1/S
      - p14 binds to MDM2 to prevent p53 destruction
  - *CDK4* gene on chromosome 12q14
    - *CDK4* regulates progression from G1 to S phase of the cell cycle
      - Phosphorylates pRB, pRB releases the transcription factor E2F, and E2F induces expression of S-phase genes
      - Mutant *CDK4* is resistant to inhibition from p16
  - Two additional high-risk susceptibility loci, 1p22 and 1p36, have been linked to hereditary melanoma
- Diagnosis
  - Mainly based on clinical grounds
  - Genetic testing has been a controversial issue due to the limited predictive value of the test
  - Criteria for genetic testing (2009)
    - Moderate- to high-melanoma-incidence areas
      - Individual with three primary melanomas
      - Family with one invasive melanoma and two other diagnoses of melanoma or pancreatic cancer in first- and second-degree relatives
    - Low-incidence areas
      - Two melanomas or pancreatic cancers in a family
  - Current commercial test for *CDKN2A* mutations focuses on sequence analysis of the coding regions of the p16/Ink4A transcript
    - Mutations are not detected in 50–60% of individuals with strong melanoma pedigrees, suggesting other high-risk susceptibility alleles remain to be determined
    - Negative results provide little reassurance and do little to inform management in this large proportion of suspicious familial cases
- Management
  - Increased surveillance
  - Avoid sun exposure

23.9.2 Nevoid Basal Cell Carcinoma Syndrome (Gorlin Syndrome)

- Overview
  - Autosomal dominant syndrome with almost 100% penetrance and variable expressivity
  - Incidence of about 1 in 50,000 live births
- Clinical and pathologic features
  - Diagnostic criteria: two major criteria or one major and two minor criteria
    - Major criteria
      - Multiple basal cell carcinomas (90% of Caucasians, 40% of African Americans)
      - Multiple basal cell nevi (70%)
      - Odontogenic keratocyst (90%)
      - Ectopic calcifications, especially in the falx cerebri (90%)
      - Palmar or plantar pits (85%)
      - Family history of the disease
    - Minor criteria
      - Skeletal abnormalities including bifid, fused, or absent ribs or vertebrae (30–60%)
      - Macrocephaly (>97th percentile), with frontal bossing (60%)
      - Cardiac and ovarian fibroma (2% and 20%, respectively)
      - Medulloblastoma (5%)
      - Congenital malformations (5%): cleft lip or palate, polydactyly, and eye anomaly
- Molecular genetics
  - *PTCH1* gene, 23 exons, on chromosome 9q22.3

  - Majority of mutations result in a premature stop codon and protein truncation
  - Rarely, this syndrome is attributable to cytogenetically visible deletions of chromosome 9q
  - 20–30% of cases represent de novo germline mutations
  - Gene product is a transmembrane receptor for the secreted factor sonic hedgehog (SHH) and mediates its signaling
    - In the absence of its ligand, PTCH1 inhibits expression of genes that control cell fate, patterning, and growth
    - Mutant PTCH1 no longer represses these genes
- Diagnosis
  - Based mainly on clinical criteria
  - Molecular genetic testing
    - Sequence analysis of coding region has mutation detection rate of 60–85%
      - Missense mutations are common (15%) and their significance may be difficult to interpret
      - Other types of mutations are considered pathogenic
    - Deletion/duplication analysis has a detection rate of about 5%
    - Tumor DNA analysis is performed for suspected mosaics
    - Predictive testing can be performed in at-risk relatives when there is a known mutation in the family
- Management
  - Monitoring of head circumference in infancy
  - Annual jaw radiograph, beginning at age 8
  - Dermatologic examination every 4–12 months, beginning in adolescence
  - Avoidance of sun exposure and radiation therapy
  - For basal cell carcinomas, the goal is to eradicate disease while preventing disfigurement. Options include surgery, cryotherapy, laser ablation, photodynamic therapy, and topical 5-fluorouracil
  - Topical sonic hedgehog antagonists have shown promise in clinical trials

23.9.3 Xeroderma Pigmentosum

- Overview
  - Xeroderma pigmentosum (XP), an autosomal recessive syndrome with increased risk of skin and eye cancer due to UV hypersensitivity
  - Incidence is 1 in 1,000,000 in the US and 1 in 40,000 in Japan
  - XP variant clinically resembles XP, but due to different molecular genetic origin, it demonstrates normal response to UV-damaged DNA on laboratory testing (see below)
- Clinical and pathologic features
  - Cutaneous
    - Acute sun sensitivity in the first year of life
    - Marked freckling of the face
    - Xerosis
    - Poikiloderma
  - Ocular
    - Severe keratitis and corneal opacification
    - Lid pigmentation and atrophy
    - Loss of lashes
  - Neurologic
    - Microcephaly
    - Diminished deep tendon reflexes
    - Progressive sensorineural hearing loss
    - Progressive cognitive impairment
  - Oncologic
    - 1,000-fold increased risk of cancer of UV-exposed sites of skin and eyes
    - Median age of onset of nonmelanoma skin cancer is under age 10 years
- Molecular genetics
  - Mutations in XP genes for complementation groups A–G that participate in NER of UV-damaged DNA
    - Specific genes include *XPA*, *ERCC3 (XPB)*, *XPC*, *ERCC2 (XPD)*, *DDB2 (XPE)*, *ERCC4 (XPF)*, *ERCC5 (XPG)*, and *ERCC1*
    - Majority of mutations in *XPA* (25%), *XPC* (25%), and *XPD* (15%)
  - Mutation in *POLH (XP-V)* results in XP variant (20%), in which an error-prone

DNA polymerase continues replication of damaged DNA by bypassing UV-induced thymidine dimers

- Diagnosis
  - Multiple assays available in research laboratories identify XP
    - Cellular UV sensitivity test
      - XP cells are hypersensitive to killing by UV
      - XP variant cells have normal post-UV survival but can be sensitized by caffeine
    - Unscheduled DNA synthesis (UDS)
      - XP cells with defective NER display reduced levels of UDS when exposed to UV light, as measured by the incorporation of radioactive thymidine into nondividing cells
    - Host cell reactivation
      - XP cells are unable to efficiently repair UV damage to transfected plasmid DNA, resulting in reduced expression of a reporter gene
    - Complementation group analysis
      - Responsible gene is determined when transfection of its cDNA rescues NER and restores normal results in one of the above tests
  - Molecular genetic testing
    - Once the defective complementation group is identified, the following can be performed
      - Sequence analysis is clinically available for *XPA* and *XPC*
      - Direct DNA testing by sequence analysis, mutation scanning, etc. is available on a research basis for the other genes
    - Targeted mutation analysis is available for a founder mutation responsible for 90% of *XPA*-related disease in Japan
    - Carrier testing is available
- Management
  - Aggressive avoidance of sun and other UV exposure
  - Regular skin, eye, and neurological examination
  - Oral isotretinoin to prevent new neoplasms in individuals with multiple skin cancers

23.9.4 Carney Complex

- Overview
  - Carney complex (CNC) is a rare, autosomal dominant syndrome previously referred to as NAME (nevi, atrial myxomas, ephelides) or LAMB (lentigines, atrial myxoma, blue nevi) syndrome
  - Very heterogeneous phenotype
  - Average life expectancy is 50 years. Cardiac complications related to atrial myxomas are the most common cause of death
- Clinical and pathologic features
  - Spotty skin pigmentation on the vermilion border of the lips, eyelids, conjunctiva, sclera, and vaginal and penile mucosa (60–70%)
    - Lentigines, café au lait macules, compound nevi, etc.
  - Blue nevi
  - Cutaneous, breast, and cardiac myxomas (30–60%)
  - Psammomatous melanotic schwannomas (10–20%)
    - Mostly gastrointestinal or paraspinal
    - Benign or malignant
  - Large cell calcifying Sertoli cell tumors of the testis (LCCSCT) (30–50%)
  - Endocrine disorders
    - Primary pigmented nodule adrenocortical disease (PPNAD), with or without Cushing syndrome (25–30%)
    - Pituitary tumors, especially somatotropinomas with acromegaly (10%)
    - Thyroid adenoma or carcinoma (10–25%)
  - Breast ductal adenomas (25%)
- Molecular genetics
  - Over half of CNC cases are due to mutation in *PRKAR1A* tumor suppressor gene on chromosome 7q23–24
    - Majority result in a truncated protein
    - Mutations are unique to kindred
    - 30% of germline mutations arise de novo
    - PRKAR1A is the R1α regulatory subunit of protein kinase A, the main mediator of cyclic AMP signaling

  - 20–30% of families with CNC have been linked to 2p15–16
- Diagnosis
  - Clinical diagnosis, either of the following
    - At least two major features (listed above)
    - One major feature and family history
  - Molecular genetic diagnosis
    - Sequence analysis of the *PRKAR1A* full gene has a detection rate of 60%
    - Deletion/duplication analysis has a detection rate of 2%
    - Linkage analysis is only available on a research basis
    - Predictive testing can be performed in at-risk relatives in families with known mutation
- Management
  - Annual echocardiogram, starting in infancy
  - Thyroid and testicular ultrasound
  - Biochemical screening for endocrine tumors; i.e., cortisol, growth hormone, and IGF-1

23.9.5 PTEN Hamartoma Tumor Syndrome (Cowden Syndrome)

- Overview
  - PHTS refers to a group of autosomal dominant syndromes characterized by multiple hamartomas, including Cowden (CS), Bannayan–Riley–Ruvalcaba (BRRS), Proteus (PS), and Proteus-like (PLS) syndromes. Cowden syndrome is the focus of this section
  - Autosomal dominant transmission with almost complete penetrance by the third decade of life and quite variable expressivity
  - Incidence of 1 in 200,000 live births
- Clinical and pathologic features
  - Pathognomonic features
    - Mucocutaneous lesions: trichilemmomas, acral keratoses, and papillomatous papules
    - Lhermitte–Duclos disease with cerebellar gangliocytoma
  - Major features
    - Macrocephaly (30%)
    - Breast carcinoma (30–50%)
    - Thyroid carcinoma, especially follicular (5–10%)
    - Endometrial carcinoma (5–10%)
  - Minor features
    - Mental retardation (10%)
    - Gastrointestinal hamartomas (see Sect. 23.6, Hereditary Gastrointestinal Cancers for differential diagnosis; note JPS associated with *PTEN* mutations is most likely Cowden syndrome)
    - Others include thyroid lesions, fibrocystic disease of the breast, genitourinary tumors, uterine fibroids, lipomas, and fibromas
- Genetics
  - *PTEN* tumor suppressor gene on chromosome 10q22–23.3
  - More than 150 unique mutations
    - Most are unique to kindred
    - Three-fourths lead to protein truncation
    - 40% are located in exon 5
  - Gene product, PTEN, is a dual-specificity protein and lipid phosphatase (dephosphorylates tyrosine and serine/threonine) that regulates important oncogenic signaling pathways
    - Cytoplasmic PTEN converts PIP3 to PIP2, which inhibits the PI3K/AKT pathway and elicits apoptosis
    - Nuclear PTEN predominantly signals down the MAPK pathway to mediate cell cycle arrest
    - Mutant PTEN results in uncontrolled cell survival and proliferation
  - Germline variants in *SDHB* and *SDHD* genes and elevated manganese superoxide dismutase have been discovered in patients with CS-like phenotype but no detectable *PTEN* mutation
  - Interestingly, somatic inactivation of PTEN is very frequently found in endometrioid endometrial carcinoma
- Diagnosis
  - Operational diagnostic criteria proposed by the International Cowden Syndrome

Consortium and updated annually by the National Comprehensive Cancer Network (NCCN) are based on combinations of the pathognomonic, major, and minor features listed above
  - Molecular genetic diagnosis
    - Sequence analysis detects *PTEN* mutations in 80% who meet clinical criteria for CS; it detects 60% in BRRS, 50% in PLS, and 20% in PS
    - Promoter analysis by direct sequencing of the promoter region is currently available on a research basis; it detects function-altering mutations in 10% of individuals with CS
    - Deletion/duplication analysis of *PTEN* is only useful in distinguishing BRRS, as over 10% of these individuals will have large deletions; large deletions are rarely identified in CS patients
    - Predictive testing in at-risk relatives when there is a known mutation in the family
- Management
  - Earlier mammography and breast MRI as well as endometrial biopsy or ultrasound in females
  - Baseline thyroid ultrasound and annual neck examination

23.9.6 Birt–Hogg–Dubé Syndrome

- Overview
  - Rare, autosomal dominant disorder characterized by hamartomatous skin lesions, pulmonary cysts, and renal carcinoma
  - Very heterogeneous phenotype
- Clinical and pathologic features
  - Cutaneous lesions present as multiple, small papules all over the face, neck, and trunk in the third and fourth decades of life. Triad
    - Fibrofolliculomas
    - Trichodiscomas (angiofibroma)
    - Acrochordons (skin tag)
  - Lung cysts (90%)
    - Spontaneous pneumothorax (30–40% of BHDS patients)
  - Renal tumors (30%)
    - Oncocytic hybrid tumors with features of both oncocytomas and chromophobe renal cell carcinoma (two-thirds of renal tumors)
    - Chromophobe RCCs (one-fourth of renal tumors)
    - Benign oncocytomas
    - Other RCC subtypes at a low frequency
- Molecular genetics
  - *FLCN* is a gene with 14 exons on chromosome 17p11.2
  - All mutations lead to protein truncation
  - Half of families have deletion or duplication of a single cytosine nucleotide in a polycytosine tract in exon 11
  - Gene product, folliculin, is involved in mTOR and AMPK signaling, but its exact function largely remains unknown
- Diagnosis
  - Clinical diagnosis is primarily based on dermatologic criteria, but molecular genetic testing is extended to those with the following
    - Five or more papules on the neck or trunk, with at least one histologically confirmed fibrofolliculoma
    - Histologically confirmed facial angiofibroma, without other clinical features suggestive of tuberous sclerosis or MEN1
    - Multiple, bilateral characteristic renal tumors
    - Single-characteristic renal tumor and family history of renal cancer
    - Family history of autosomal dominant spontaneous pneumothorax
  - Molecular genetic diagnosis
    - Sequence analysis of the entire coding region, exons 4–14, has a mutation detection rate of 88%
    - Sequence analysis of exon 11 only may be the preferred initial approach, because over half of mutations occur at the mutational hotspot

    - Deletion/duplication analysis by MLPA is available
    - Predictive testing in at-risk relatives from families with a known pathogenic mutation
- Management
  - Baseline chest X-ray, abdominal/pelvic CT/MRI, and renal ultrasound in adults with a positive skin biopsy
  - Continued surveillance for renal cancer by imaging every 2–3 years
  - Nephron-sparing surgery for renal tumors larger than 3 cm
  - Laser ablation of skin lesions results in substantial improvement, but relapse is common

Other familial cancer syndromes classified as genodermatoses include von Hippel–Lindau syndrome (please see Sect. 23.8, Hereditary Kidney Tumors) as well as tuberous sclerosis complex, neurofibromatosis types 1 and 2, and ataxia-telangiectasia (please see individual sections below).

23.10 Tuberous Sclerosis Complex

- Overview
  - Highly penetrant autosomal dominant complex characterized by hamartomas of multiple organs including the CNS, skin, kidney, and heart
  - Markedly variable expressivity
  - Incidence of up to 1 in 6,000 live births
  - CNS disease is the leading cause of death
- Clinical and pathologic features
  - Major features
    - Cortical tubers (glial hamartomas) (70%)
    - Subependymal glial nodules (90%) and subependymal giant cell astrocytomas (10%)
    - Cardiac rhabdomyomas (50–70%)
    - Renal angiomyolipomas (70%)
    - Pulmonary lymphangiomyomatosis (females only)
    - Multiple retinal hamartomas (50%)
    - Skin lesions (90%), including facial angiofibromas, periungual fibromas, hypopigmented macules, and Shagreen patches
  - Minor features
    - Cerebral white matter radial migration lines
    - Nonrenal hamartomas
    - Renal cysts
    - Bone cysts
    - Dental enamel pits
    - Gingival fibromas
  - Additionally, patients often have associated seizures (60–90%) and neurocognitive impairment (50%)
  - Increased risk of renal cell carcinoma
- Molecular genetics
  - *TSC1* on chromosome 9q34 expresses hamartin (20% of cases)
  - *TSC2* on chromosome 16p13.3 expresses tuberin (60% of cases)
  - Both *TSC1* and *TSC2* are large genes exhibiting a large number of pathogenic mutations
    - All *TSC1* and 70% of *TSC2* mutations produce a truncated protein
    - Only *TSC2* mutations include large deletions or rearrangements
    - Because *TSC2* is adjacent to *PKD1*, large deletions may lead to a contiguous gene syndrome of severe autosomal dominant polycystic kidney disease in infancy
  - De novo germline mutations account for two-thirds of cases
  - Somatic mosaicism is found in 10–25% of cases
  - Hamartin and tuberin form a dimer with a major role in the inhibition of the mammalian target of rapamycin (mTOR) signaling pathway
    - Dimer functions as a GTPase-activating protein (GAP)
    - Directly hydrolyzes and inactivates Rheb, the small GTPase that activates mTOR
- Diagnosis
  - Based on clinical criteria

    - Two major features or one major and two minor features
    - Isolated combination of lymphangiomyomatosis and renal angiomyolipomas is nondiagnostic
  - Molecular genetic diagnosis is complicated by the size of the genes, the number of mutations, and the frequent mosaicism. Overall mutation detection rate is only 85–90%, even when using newer diagnostic techniques like sequencing and MLPA
    - Full gene sequence analysis
      - Mutation detection rate is 15–30% for *TSC1* and 50–70% for *TSC2*
    - Deletion/duplication analysis
      - Mutation detection rate is 5% for *TSC2*, but the diagnostic yield is very low for *TSC1*
    - DNA analysis from tumor or other tissues is performed for suspected mosaics
    - Predictive testing is a helpful way to exclude risk in the relatives of those very common de novo cases
- Management
  - CNS imaging every 1–3 years
  - Renal ultrasound every 1–3 years or semiannually in the presence of angiomyolipomas
  - Sirolimus (rapamycin), inhibitor of mTOR, is under investigation in clinical trials for astrocytomas, lymphangiomyomas, and renal angiomyolipomas in TS patients

23.11 Neurofibromatosis Type 1 (von Recklinghausen Syndrome)

- Overview
  - Autosomal dominant inherited disorder with almost complete penetrance by age 20 years and highly variable expressivity
  - Incidence of 1 in 3,000 live births
- Clinical and pathologic features
  - Clinical features include the diagnostic criteria below in conjunction with learning disabilities (30–60%) or mental retardation (10%)
  - National Institutes of Health diagnostic criteria, at least two of the following
    - Six or more café au lait macules (CALMs) (100%)
    - Axillary or inguinal freckling (90%)
    - Two neurofibromas (99%) or one plexiform neurofibroma (25%)
    - Two or more Lisch nodules (melanocytic iris hamartomas) (95%)
    - Optic nerve glioma (15%)
    - Specific bone abnormalities
    - Family history of NF
  - Malignant peripheral nerve sheath tumors (neurofibrosarcomas), usually arise from transformation of plexiform neurofibromas (10%)
  - Increased risk for myeloid leukemia
- Molecular genetics
  - *NF1* is a tumor suppressor gene on chromosome 17q11.2
  - Very large gene with 60 exons that encodes three alternatively spliced transcripts
  - Most mutations result in severe protein truncation, often by altered splicing
  - One-half of cases are due to de novo germline mutation
  - Mosaicism may manifest as a segmental or milder generalized phenotype
  - Genotype–phenotype correlations
    - Severe variant: whole gene deletion
    - Mild variant: 3-bp inframe deletion of exon 17
  - Gene product, neurofibromin, is a GTP-activating protein for RAS
    - Normally accelerates inactivation of Ras by hydrolyzing RAS-GTP to RAS-GDP
    - Loss of function leads to accumulation of active RAS-GTP and uncontrolled cellular proliferation
- Diagnosis
  - Diagnosis is mainly based on clinical criteria
  - Molecular genetic testing is infrequently indicated clinically; it is useful when suspicious clinical features do not meet criteria

  - Multistep protocol addresses difficulties due to the large size of the gene and its lack of mutation hotspots; it detects 95% of *NF1* mutations
    - Sequence analysis of both genomic DNA and mRNA due to the high frequency of splicing mutations (mutation detection rate of 90%)
    - Small, intragenic deletion/duplication analysis (detection rate of 1%)
    - Large, whole gene deletion/duplication analysis (detection rate of 5%)
    - Comprehensive testing of lesional tissue for suspected mosaics
    - Linkage analysis
      - Historically, commercially available protein truncation testing was the first-line strategy
      - Prenatal diagnosis and PGD are available, but the phenotypic severity cannot be predicted
- Management
  - Children with six café au lait spots alone and no family history should be followed as if they have NF1 since most will develop it
  - Annual physical and eye examination
  - Screening for developmental delays
  - Monitoring for malignant transformation
  - Surgery when needed
  - Farnesyl protein transferase inhibitors, which inhibit the Ras signaling pathway, are in clinical trials for NF1

23.12 Neurofibromatosis Type 2

- Clinical features
  - Autosomal dominant syndrome with penetrance of 100%
  - Incidence of 1 in 40,000 live births
  - Characterized by acoustic schwannomas (100%), meningiomas (50%), ependymomas, astrocytomas, and lens opacities
- Molecular genetics
  - *NF2* tumor suppressor gene, 17 exons, on chromosome 22q12.2
  - Over 200 different mutations, mostly truncating point mutations
  - Cytogenetically visible deletions and ring chromosome 22 are responsible for NF2 phenotype as well
  - One-half of cases are due to de novo mutations
  - Mosaicism is found in 25% of individuals with de novo mutations
  - Somatic inactivation of both alleles commonly occurs in sporadic schwannomas (90%) and meningiomas (50%)
  - Gene product, merlin, appears to play a role in growth factor receptor signaling and cell adhesion, although its function requires further elucidation. It regulates Schwann cell and leptomeningeal cell proliferation
- Clinical diagnosis
  - Manchester (modified NIH) criteria, one of the following
    - Bilateral vestibular schwannomas
    - First-degree relative with NF2 and unilateral vestibular schwannoma or any two of the characteristic tumors
    - Multiple meningiomas and unilateral vestibular schwannoma or any two NF2 tumors
    - Unilateral vestibular schwannoma and two NF2 tumors
- Molecular genetic diagnosis
  - Sequence analysis of NF2 exons 1–15 detects 60% of mutations in simplex cases and 75% in cases with family history
  - Deletion/duplication analysis by MLPA detects an additional 10–15%
  - Linkage analysis is very accurate
  - Tumor DNA testing is usually necessary for suspected mosaics
  - Predictive testing is recommended for presymptomatic at-risk relatives, age 10 and older
- Management
  - Annual hearing, vision, and neurological examination
  - Annual MRI of the CNS
  - Minimally invasive treatment of vestibular schwannomas by endoscopic tumor resection or stereotactic radiosurgery

23.13 Ataxia-Telangiectasia

- Overview
  - Autosomal recessive syndrome with an incidence of 1 in 30,000–100,000 live births
  - Most patients survive into their 30s
- Clinical and pathologic features
  - Progressive cerebellar ataxia, appearing at age 1–4 years (100%)
  - Oculomotor apraxia
  - Choreoathetosis
  - Telangiectasias of skin and conjunctivae
  - Immunodeficiency and frequent infections (60–80%)
  - Increased risk of cancer, especially leukemia and lymphoma (30–40%)
  - Carriers have a fourfold increased risk of cancer, primary due to breast cancer
  - Characteristic laboratory abnormalities include elevated serum AFP (95%) and cytogenetically visible 7;14 translocation (5–15%)
- Genetics
  - *ATM* is a very large gene with 66 exons on 11q22.3
  - Four complementation groups, defined by their phenotype of cellular radiosensitivity, map to the same locus. The distribution is A (55%), C (28%), D (14%), and E (3%)
  - Over 500 distinct mutations of *ATM* reported, with 70–85% producing a truncated protein
  - Founder mutations described in many different ethnic populations
  - Gene product is serine protein kinase that senses double-stranded DNA breaks (DSBs), coordinates cell cycle checkpoints, and recruits repair proteins to site of damage
    - Phosphorylates over 700 substrates in response to DNA damage, including SMC1, p53, MDM2, Chk2, BRCA1, FANCD2, and c-abl
- Diagnosis
  - Biochemical testing
    - Immunoblotting for ATM protein
      - Most sensitive and specific diagnostic test
      - Ataxia-telangiectasia patients exhibit undetectable ATM in 95%, trace amounts in 5%, and a normal amount of "kinase-dead" protein in 1%
    - Radiosensitivity assay
    - Flow cytometry to measure phosphorylation of ATM substrates
    - ATM serine/threonine kinase activity
  - Molecular genetic testing
    - Sequence analysis of coding region using genomic DNA or cDNA
      - Mutation detection rate of 90% or 95%, respectively
      - Lower sensitivity and much more expensive than immunoblot assay
      - Difficult to distinguish pathogenic missense mutations from normal polymorphisms
    - Deletion/duplication analysis is low yield
    - Targeted mutation analysis for founder mutations
    - Linkage analysis/ethnic haplotype analysis
    - Carrier testing
      - Rapid flow cytometric assay measures phosphorylation of SMC1 protein to identify heterozygotes
      - Targeted mutation analysis for kindred-specific mutation
- Management
  - Increased surveillance for malignancy and careful use of radiotherapy

Further Reading

Adley BP, Smith ND, Nayer R, et al. Birt–Hogg–Dubé syndrome: clinicopathologic findings and genetic alterations. Arch Pathol Lab Med. 2006;130:1865–70.

Allain DC. Genetic counseling and testing for common hereditary breast cancer syndromes. A paper from the 2007 William Beaumont Hospital Symposium on Molecular Pathology. J Mol Diagn. 2008;10:383–95.

Arnold CN, Goel A, Blum HE, et al. Molecular pathogenesis of colorectal cancer. Implications for molecular diagnosis. Cancer. 2005;104:2035–47.

Coleman JA, Russo P. Hereditary and familial kidney cancer. Curr Opin Urol. 2009;19:478–85.

Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group. Recommendations from the EGAPP Working Group: genetic testing strategies in newly diagnosed individuals with colorectal cancer aimed at reducing morbidity and mortality from Lynch syndrome in relatives. Genet Med. 2009;11:35–41.

Evans DG. Neurofibromatosis type 2 (NF2): a clinical and molecular review. Orphanet J Rare Dis. 2009;4:16.

GeneTests—www.genetests.org.

Gerstenblith MR, Goldstein AM, Tucker MA. Hereditary genodermatoses with cancer predisposition. Hematol Oncol Clin North Am. 2010;24:885–906.

Hobert JA, Eng C. PTEN hamartoma tumor syndrome: an overview. Genet Med. 2009;11:687–94.

Honrado E, Benitez J, Palacios J. The molecular pathology of hereditary breast cancer: genetic testing and therapeutic implications. Mod Pathol. 2005;18:1305–20.

Jett K, Friedman JM. Clinical and genetic aspects of neurofibromatosis 1. Genet Med. 2010;12:1–11.

Lindor NM, McMaster ML, Lindor CJ, et al. Concise handbook of familial cancer susceptibility syndromes - second edition. J Natl Cancer Inst Monographs 2008;1–93.

Linehan WM, Bratslavsky G, Pinto PA, et al. Molecular diagnosis and therapy of kidney cancer. Annual Rev Med.. 2010;61:329–43.

Lo Muzio L. Nevoid basal cell carcinoma syndrome (Gorlin syndrome). Orphanet J Rare Dis. 2008;3:32.

Lynch HT, Casey MJ, Snyder CL, et al. Hereditary ovarian cancer: molecular genetics, pathology, management, and heterogeneity. Mol Oncol. 2009a;3:97–137.

Lynch HT, Lynch PM, Lanspa SJ, et al. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. Clin Genet. 2009b;76:1–18.

Nichols KE, Walther S, Chao E, et al. Recent advances in retinoblastoma genetic research. Curr Opin Ophthalmol. 2009;20:351–5.

Orlova KA, Crino PB. The tuberous sclerosis complex. Ann N Y Acad Sci. 2010;1184:87–105.

Palmero EI, Achatz MIW, Ashton-Prolla P, et al. Tumor protein 53 mutations and inherited cancer: beyond Li–Fraumeni syndrome. Curr Opin Oncol. 2010;22:64–9.

Robson ME, Storm CD, Weitzel J, et al. American society of clinical oncology policy statement update: genetic and genomic testing for cancer susceptibility. J Clin Oncol. 2010;28:893–901.

Rubinstein WS. Endocrine cancer predisposition syndromes: hereditary paraganglioma, multiple endocrine neoplasia type I, multiple endocrine neoplasia type II, and hereditary thyroid cancer. Hematol Oncol Clin North Am. 2010;24:907–37.

Rustgi AK. The genetics of hereditary colon cancer. Genes Dev. 2007;21:2525–38.

Udayakumar D, Tsao H. Melanoma genetics: an update on risk-associated genes. Hematol Oncol Clin North Am. 2009;23:415–29.

Van Lier MGF, Wagner A, van Leerdam ME, et al. A review on the molecular diagnostics of Lynch syndrome: a central role for the pathology laboratory. J Cell Mol Med. 2010;14:181–97.

Wilkes D, McDermott DA, Basson CT. Clinical phenotypes and molecular genetic mechanisms of Carney complex. Lancet Oncol. 2005;6:501–8.

Genetic Counseling

24

Kimberly A. Quaid and Lisa J. Cushman

Contents

24.1 Definition of Genetic Counseling 645

24.2 Role and Training of Genetic Counselors 645

24.3 History of Genetic Counseling 646

24.4 Models of Genetic Counseling 646
24.4.1 Eugenic Model 646
24.4.2 Medical/Preventive Model 646
24.4.3 Decision-Making Model 646
24.4.4 Psychotherapeutic Model 647

24.5 Common Genetic Counseling Terms 647

24.6 Common Genetic Counseling Problems 648

24.7 The Process of Clinical Genetics and Genetic Counseling 648
24.7.1 Prior to Clinic Visit 648
24.7.2 Clinic Visit 648
24.7.3 Followup Care 648

24.8 Determining Genetic Risks 648
24.8.1 Recurrence Risk Based on Known Genotype 648
24.8.2 Recurrence Risks Using Empiric Data 649
24.8.3 Risk Assessment in Cases with Consanguinity 649
24.8.4 Bayesian Analysis in Risk Estimation 649

Further Reading 652

K.A. Quaid, PhD (✉)
Professor of Medical and Molecular Genetics, Indiana University School of Medicine, Faculty Investigator, Indiana University Center for Bioethics, Indianapolis, IN, USA

L.J. Cushman, PhD, CGC
Hayes, Inc., Landsale, PA, USA

24.1 Definition of Genetic Counseling

- A communication process which deals with the human problems associated with the occurrence, or the risk of occurrence, of a genetic disorder in a family
- Involves an attempt by one or more appropriately trained persons to help the individual or family to
  - Comprehend the medical facts, including the diagnosis, probable course of the disorder, and available management
  - Appreciate the way heredity contributes to the disorder and the risk of recurrence in specified relatives
  - Understand the alternatives for dealing with the risk of recurrence
  - Choose the course of action which seems to them appropriate in view of their risk, their family goals, and their ethical and religious standards and to act in accordance with that decision
  - Make the best possible adjustment to the disorder in an affected family member and/or to the risk of recurrence of that disorder (ASHG 1975)

24.2 Role and Training of Genetic Counselors

- Major role in the investigation and management of genetic disorders

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_24,

- Graduates of 2-year master's level training programs in medical genetics and counseling
- Board certification by the National Society of Genetic Counselors
- Licensing required in some states

24.3 History of Genetic Counseling

- In 1865, Gregor Mendel finds that individual traits are determined by discrete factors, later called genes, that are inherited from parents
- In 1883, Francis Galton suggests that eugenics (Greek meaning wellborn) become the study of social policies that may improve or impair racial qualities of future generations either physically or mentally
- In 1906, Bateson suggests the term "genetics" for the biological and medical study of heredity
- By mid-1940s, heredity clinics were being established
- By the 1950s, medicine began to focus on prevention and clinics established to advise people about inherited traits
- In 1969, Sarah Lawrence College established first program in human genetics
- In 1971, first ten "genetic associates" graduated from Sarah Lawrence
- In 1975, term "genetic counseling" was coined by Sheldon Reed

24.4 Models of Genetic Counseling

24.4.1 Eugenic Model

- Goal of eugenicists to improve the human species by better breeding
- In 1904, Genetics Records Office opened at Cold Spring Harbor
- In 1907, the state of Indiana passed the first sterilization law
- When approved by a board of experts, the law mandated the sterilization of imbeciles, idiots, criminals, and others in state institutions
- By 1926, 23 of the 48 states in the USA had laws mandating sterilization of the "mentally defective," and over 6,000 people had been sterilized
- Concept of eugenic sterilization found a receptive audience among the leadership of Nazi Germany
- Sterilization of the unfit became a signature policy in the Nazi's quest for racial purity and superiority

24.4.2 Medical/Preventive Model

- Nazi excesses lead to a retreat from eugenics practices
- Structure of DNA discovered in 1953
- Few diagnostic tests available
- Information about risk based on empirical observations offered to families so families could avoid recurrence of a disorder that had already occurred
- Goal to prevent genetic disorders by offering families information and the option to avoid childbearing
- Presumption that "rational" families would want to prevent recurrence

24.4.3 Decision-Making Model

- Human diploid complement of 46 chromosomes reported in 1956
- In late the 1950s and early 1960s, findings led to an understanding of the cytogenetics of Down, Klinefelter, and Turner syndromes
- Trisomies 13 and 18 discovered
- Became possible to identify those heterozygous for B-thalassemia and Tay-Sachs disease
- Amniocentesis first used in 1956 for prenatal diagnosis initially for sex selection and later for karyotyping
- Provided families with new options for assessing genetic risks and avoiding a genetic disorder
- Legalization of abortion in 1972 meant that hard choices needed to be made regarding termination for genetic defects
- Educating families about choices was labor intensive and time consuming

- Because of the almost exclusive focus on reproduction, the ideal of nondirective counseling was embraced with emphasis on patient autonomy in decision making
- Goals of counseling shifted from providing information to a process in which individuals were not only educated but helped to make decisions that were consistent with their own values and needs

24.4.4 Psychotherapeutic Model

- Recognition that families cannot process or act on information they have been given unless they have dealt with the powerful emotions evoked by such information
- New emphasis on exploring client's experiences, emotional responses, goals, cultural expectations, religious beliefs, financial and social resources, family and interpersonal dynamics, and coping styles

24.5 Common Genetic Counseling Terms

- Allele – One of the alternative versions of a gene that may occupy a given locus
- Bayesian analysis – A mathematical method widely used in genetic counseling to calculate the risk of recurrence of a genetic disorder. This method combines information from a variety of sources including genetics, pedigree information, and test results to determine the probability that a specific individual may develop or transmit a specific disorder
- Conditional probability – In Bayesian analysis, this is the chance of an observed outcome given the prior probability of the consultand's genotype
- Consanguinity – Relationship by descent from a common ancestor
- Consultand – The individual requesting genetic counseling
- Dominant – A trait is dominant if it is phenotypically expressed in heterozygotes
- Empiric risk – The probability that a trait will occur or recur in a family based on past experience rather than on knowledge of the causative mechanism
- Expressivity – The extent to which a genetic defect is manifested
- Pedigree – A diagram of a family history indicating the family members, their relationship to the proband, and their status with regard to a particular genetic condition (affected vs. unaffected)
- Joint probability – The product of the prior and conditional probabilities
- Karyotype – The chromosomes of an individual. Also, a picture of the chromosomes of an individual arranged in a standard presentation
- Penetrance – The probability that a mutant genotype will have any phenotypic expression
- Phenotype – The observed biochemical, physiological, and morphological characteristics of an individual as determined by his or her genotype and the environment in which it is expressed
- Proband – The family member through whom the family is ascertained. If affected, may be called the index case
- Recessive – A trait of gene that expressed only in individuals who have inherited two copies of the gene
- Recurrence risk – The probability that a genetic disorder present in one or more members of a family will recur in another member of the family of the same or subsequent generation
- Teratogen – An environmental agent, medication, X-ray, or pathogen that produces or raises the incidence of congenital malformation
- Variable expressivity – When the manifestation of a phenotype differs in people who have the same genotype
- Variable penetrance – When a condition is expressed in less than 100% of individuals who carry the responsible allele
- X-linked – Genes on the X chromosome or traits determined by such genes
- Y-linked – Genes on the Y chromosome or traits determine by such genes

24.6 Common Genetic Counseling Problems

- Single-gene disorders known or suspected
- Multifactorial disorders known or suspected
- Chromosomal disorders diagnosed in the consultand or family member
- An abnormal trait or carrier state identified by genetic screening
- Prenatal diagnosis for advanced maternal age or other indications
- Consanguinity
- Teratogen exposure
- Repeated pregnancy loss or infertility

24.7 The Process of Clinical Genetics and Genetic Counseling

24.7.1 Prior to Clinic Visit

- Reason for referral
- Collection of family history information and construction of a pedigree
- Collection and review of medical records and laboratory test results on consultand and other family members

24.7.2 Clinic Visit

- Clinical examination
- Diagnosis or ordering of further tests in order to make a diagnosis
- Recurrence risk estimation
- Genetic counseling
  - Nature and consequences of disorder
  - Recurrence risk
  - Means of modification of consequences
  - Means of prevention of recurrence
  - Management plan

24.7.3 Followup Care

- Referral to appropriate clinical specialists
- Referral to appropriate health agencies
- Referral to appropriate support groups
- Further clinical assessment if warranted
- Continued contact and support by genetic counselor if needed

24.8 Determining Genetic Risks

24.8.1 Recurrence Risk Based on Known Genotype

24.8.1.1 Autosomal Recessive Disorders

- For example, Tay-Sachs disease
- Recurrence risk if both parents are known or obligate carriers is 25% for each future pregnancy [1/2 (chance that the mother passes on the mutation) $\times$ 1/2 (chance that the father passes on the mutation) = 1/4]

24.8.1.2 Autosomal Dominant Disorders

- For example, achondroplasia
- Recurrence risk if one parent is affected is 50% for each future pregnancy (chance that the parent passes on the mutation)
- Factors to consider when counseling about an autosomal dominant condition in the absence of a positive family history
  - It could be the result of a new mutation in the proband
  - There may be decreased penetrance
  - There may variable expressivity
  - Germline mosaicism may be present in a parent

24.8.1.3 X-Linked Recessive Disorders

- For example, Duchenne muscular dystrophy
- Recurrence risk if mother is a known or obligate carrier is 25% for each future pregnancy [1/2 (chance that she passes on the mutation) $\times$ 1/2 (chance that she has a son) = 1/4]

24.8.1.4 X-Linked Dominant Disorders

- For example, incontinentia pigmenti
- These conditions are often lethal in males
- One-third of the children of an affected female will be affected
- All of the live-born males will be unaffected
- Half of the females will be unaffected

- One may also see an increase in the number of spontaneous abortions (affected male fetuses)

24.8.1.5 Mitochondrial Disorders

- Cause by an mtDNA mutation, for example, Leber hereditary optic neuropathy
- Recurrence risk if mother is a known mtDNA mutation carrier in a homoplasmic form is 100% for each future pregnancy

24.8.2 Recurrence Risks Using Empiric Data

- Empiric recurrence risks are those in which the chance of having another affected individual in the family is based on observed data as opposed to mathematical calculations
- Multifactorial or polygenic conditions
  - Empiric recurrence risks are available for numerous conditions, such as congenital heart defects, cleft lip and palate, diabetes, psychiatric disorders, and cardiovascular disease
  - The risk is greatest among first degree relatives and decreases with the distance of the relationship
  - The recurrence risk depends on the incidence of the condition
  - An estimate of the recurrence risk, when specific risk figures are not available, is the square root of the incidence of the condition (i.e., the recurrence risk is approximately one in 100 for a condition with an incidence of 1/10,000)
  - If there is an unequal sex incidence, the recurrence risk is greater for relatives of a proband of the sex in which the disorder is less common
  - The recurrence risk may increase if there are multiple affected family members and/or if the condition is more severe

24.8.2.1 Structural Chromosome Rearrangements

- Empiric recurrence risks for those carrying common balanced chromosome rearrangements, such as a 14;21 Robertsonian translocation, are available
  - These risks may be dependent on sex of the transmitting parent (Table 24.1)

Table 24.1 The risk of having an abnormal liveborn may depend on the sex of the parent who carries the Robertsonian translocation

Translocation	Transmitting parent	Risk of abnormal live-born
Rob(13;14)	Mother	~0.5%
Rob(13;14)	Father	~0.5%
Rob(14;21)	Mother	10–15%
Rob(14;21)	Father	1%

24.8.2.2 Autosomal Dominant Conditions with Germline Mosaicism

- There are some disorders in which the risk of germline mosaicism has been determined
- For example, the chance that a parent of a child with osteogenesis imperfecta type 2 will have another affected child due to germline mosaicism is approximately 7%

24.8.3 Risk Assessment in Cases with Consanguinity

- Consanguinity refers to relationships involving persons who share a common ancestor (i.e., are blood relatives)
- The primary concern for children of these relationships is an increased risk of autosomal recessive disorders
- The risk depends on the degree of relationship
  - The offspring of siblings have a one in eight chance of having an autosomal recessive disorder, while the risk for offspring of first cousins is 1 in 32

24.8.4 Bayesian Analysis in Risk Estimation

- Bayesian analysis, which is based on Bayes theorem of probability, is a method for modifying one's "prior" risk (i.e., risk based on Mendelian inheritance pattern or general

population risk) using "conditional," or phenotypic, information

24.8.4.1 Autosomal Dominant Disorders

- Situations in which one may use Bayesian analysis include when the condition has reduced penetrance or variable expressivity or when the disorder has a late age of onset
- For example, a man, whose mother had Huntington disease (HD), is currently asymptomatic at age 60. Given that 3/4 of individuals with a HD gene mutation will have symptoms by age 60, what is the chance that this man inherited the gene mutation from his mother?
- Prior probability is the chance that this man inherited the mutation based on Mendelian risks (he had a one in two chance of inheriting it and a one in two chance of not inheriting it) (Table 24.2)
- Conditional probability is the chance that this man is asymptomatic at age 60 if he did inherit the mutation [1/2 (chance he received the mutation) × 1/4 (chance he is asymptomatic with the mutation)] or that he is asymptomatic at age 60 if he did not (there is essentially a 100% chance he would be unaffected if he did not inherit the mutation) (Table 24.2)
- The joint probabilities are the product of the prior and conditional probabilities (Table 24.2)
- The posterior probabilities are the joint probability for the particular hypothesis divided by the sum of both joint probabilities (Table 24.2)
- Therefore, this man has a one in nine (or approximately 11%) chance of having inherited the HD mutation from his mother given that he is asymptomatic at age 60 (compared to his prior risk of 50%) (Table 24.2)

Table 24.2 The posterior probability the man inherited the mutation given that he is asymptomatic at age 60

Hypothesis	Man inherited mutation	Man did not inherit mutation
Prior probability	1/2	1/2
Conditional probability – asymptomatic at 60	$1/2 \times 1/4 = 1/8$	1
Joint probability	$1/2 \times 1/8 = 1/16$	$1/2 \times 1 = 1/2$
Posterior probability	$(1/16)/(1/16 + 1/2) = 1/9$	$(1/2)/(1/16 + 1/2) = 8/9$

24.8.4.2 Autosomal Recessive Disorders

- One may use Bayesian analysis when the genotypes of one or both parents are not known
- For example, a Caucasian couple of Northern European ancestry is referred for a maternal family history of cystic fibrosis
  - Cystic fibrosis is an autosomal recessive condition with a carrier frequency of approximately one in 25 in individuals with a Northern European ethnic background
  - The woman is a known mutation carrier
  - The man has also had CFTR gene testing
  - However, after being tested for the 87 most common mutations, no mutation was identified
  - The detection rate for this panel in individuals of his ethnicity is approximately 90% (one in ten will have a mutation that is not identified)
  - What is the chance that this couple will have a child with cystic fibrosis given the man's gene test results?
- Prior probability is the chance that this man has a mutation based on the carrier frequency in the population (he had a one in 25 chance of being a carrier and a 24 in 25 chance of not being one) (Table 24.3)
- Conditional probability is the chance that this man tests negative if he does have a mutation [1/25 (chance he has mutation) × 1/10 (chance the test did not identify the mutation)] or that he tests negative if he is not a carrier (there is a 100% chance he would test negative if he does not have a mutation) (Table 24.3)
- The joint probabilities are the product of the prior and conditional probabilities (Table 24.3)
- The posterior probabilities are the joint probability for the particular hypothesis divided by the sum of both joint probabilities (Table 24.3)

Table 24.3 The probability that the father is a mutation carrier given that he had a negative gene test

Hypothesis	Man is a mutation carrier	Man is not a mutation carrier
Prior probability	1/25	24/25
Conditional probability – negative gene test	1/25 × 1/10 = 1/250	1
Joint probability	1/25 × 1/250 = 1/6,250	24/25 × 1 = 24/25
Posterior probability	(1/6,250)/(1/6,250 + 24/25) = 1/6,001	(24/25)/(1/6,250 + 24/25) = 6,000/6,001

- Therefore, this man has a one in 251 (or a less than 1%) chance of being a carrier given that he tested negative (compared to his prior risk of one in 25 or 4%) (Table 24.3)
- The chance that this couple will have a child with cystic fibrosis is one in 24,004 [1/6,001 (chance man is a carrier) × 1/4 (chance both parents pass on the mutation)]

24.8.4.3 X-Linked Recessive Disorders

- One may use Bayesian analysis when the possible carrier has had previously unaffected sons
- For example, a 25-year-old woman is referred for prenatal genetic counseling because of a family history of hemophilia A (her sister and mother are obligate carriers of the factor VIII gene mutation)
  - Given that the client has already had two unaffected sons, what is the chance that the male fetus she is currently carrying will have hemophilia?
- Prior probability is the chance that this woman is a carrier based on the Mendelian risks (she had a one in two chance of being a carrier and a one in two chance of not being one) (Table 24.4)
- Conditional probability is the chance that this woman has two unaffected sons if she is a carrier [1/2 (chance that first son is unaffected) × 1/2 (chance that the second son is unaffected)] or that she has two unaffected sons if she is not a carrier (there is essentially a 100% chance she would have unaffected sons if she is not a carrier) (Table 24.4)
- The joint probabilities are the product of the prior and conditional probabilities (Table 24.4)

Table 24.4 The probability the client is a carrier given she already had two unaffected sons

Hypothesis	Client is a carrier	Client is not a carrier
Prior probability	1/2	1/2
Conditional probability – has two unaffected sons	1/2 × 1/2 = 1/4	1
Joint probability	1/2 × 1/4 = 1/8	1/2 × 1 = 1/2
Posterior probability	(1/8)/(1/8 + 1/2) = 1/5	(1/2)/(1/8 + 1/2) = 4/5

- The posterior probabilities are the joint probability for the particular hypothesis divided by the sum of both joint probabilities (Table 24.4)
- Therefore, this woman has a one in five (or 20%) chance of being a carrier given that she has two unaffected sons (compared to her prior risk of one in two or 50%) (Table 24.4)
- The chance that this woman's unborn son is affected is one in ten [1/5 (chance that she is a carrier) × 1/2 (chance that she passes the mutation to him)]

24.8.4.4 Use of Molecular Genetics in Risk Assessment

- Molecular genetic testing is currently available for hundreds of genetic conditions (www.ncbi.nlm.nih.gov/sites/GeneTests/)
- Many techniques exist for the identification of a particular gene mutation (gene sequencing, PCR analysis, SSCP, other mutation scanning techniques)
- Establishing a diagnosis, or identifying a gene carrier, by determining the gene mutation an individual carries allows a genetic counselor to provide a precise recurrence risk

Further Reading

Ad hoc Committee on Genetic Counseling of the American Society of Human Genetics (ASHG). Genetic counseling. Am J Hum Genet. 1975; 27:240–1.

Bennet RL, Steinhaus KA, Uhrich SB, et al. Recommendations for standardized human pedigree nomenclature. Am J Hum Genet. 1995;56:745–52.

Gardner RJM, Sutherland GR, Shaffer LG. Chromosomal abnormalities and genetic counseling. 4th ed. Oxford: Oxford University Press; 2011.

Harper PS. Practical genetic counselling. 7th ed. New York: Hodder Arnold; 2010.

Uhlmann WR, Schuette JL, Yashar BM, editors. A guide to genetic counseling. 2nd ed. New York: Wiley-Blackwell; 2009.

Young ID. Introduction to risk calculation in genetic counseling. 3rd ed. Oxford: Oxford University Press; 2007.

Section 4

Molecular Microbiology

Molecular Virology

25

Fei Ye, Miao Cui, Rame H. Khasawneh, Robert Shibata, Josephine Wu, Mona Sharaan, and David Y. Zhang

Contents

25.1 General 656

25.2 Human Immunodeficiency Virus 658
25.2.1 General Characteristics 658
25.2.2 Clinical Presentation 660
25.2.3 Diagnostic Methods 660
25.2.4 Clinical Utility 671

25.3 Hepatitis C Virus 671
25.3.1 General Characteristics 671
25.3.2 Clinical Presentation 673
25.3.3 Diagnostic Methods 673
25.3.4 Clinical Utility 678

25.4 Hepatitis B Virus 678
25.4.1 General Characteristics 678
25.4.2 Clinical Presentation 680
25.4.3 Diagnostic Methods 680
25.4.4 Clinical Utility 683

25.5 Cytomegalovirus 684
25.5.1 General Characteristics 684
25.5.2 Clinical Presentation 684
25.5.3 Diagnostic Methods 685
25.5.4 Clinical Utility 692

25.6 Epstein–Barr Virus 692
25.6.1 General Characteristics 692
25.6.2 Clinical Presentation 694
25.6.3 Diagnostic Methods 694
25.6.4 Clinical Utility 697

25.7 Herpes Simplex Virus 697
25.7.1 General Characteristics 697
25.7.2 Clinical Presentation 698
25.7.3 Diagnostic Methods 698
25.7.4 Clinical Utility 700

25.8 Varicella Zoster Virus 701
25.8.1 General Characteristics 701
25.8.2 Clinical Presentation 701
25.8.3 Diagnostic Methods 701
25.8.4 Clinical Utility 703

25.9 Human Papillomavirus 703
25.9.1 General Characteristics 703
25.9.2 Clinical Presentation 703
25.9.3 Diagnostic Methods 705
25.9.4 Clinical Utility 709

25.10 Influenza A, B, and C Viruses 709
25.10.1 General Characteristics 709
25.10.2 Clinical Presentation 711
25.10.3 Diagnostic Methods 712
25.10.4 Clinical Utility 715

25.11 Avian Influenza (Bird) and Influenza (Flu) A Viruses 716
25.11.1 General Characteristics 716
25.11.2 Clinical Presentation 716
25.11.3 Diagnostic Methods 717

25.12 Adenovirus 717
25.12.1 General Characteristics 717
25.12.2 Clinical Presentation 717
25.12.3 Diagnostic Methods 718
25.12.4 Clinical Utility 720

25.13 Respiratory Syncytial Virus 720
25.13.1 General Characteristics 720
25.13.2 Clinical Presentation 720
25.13.3 Diagnostic Methods 721
25.13.4 Clinical Utility 722

F. Ye, PhD (✉) • M. Cui, MD • R.H. Khasawneh, MD • R. Shibata, MD • D.Y. Zhang, MD, PhD, MPH
Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

J. Wu, DDS, CLSP(MB), CLDIR
Mount Sinai School of Medicine, New York, NY, USA

M. Sharaan, MD
Molecular Pathology Laboratory, Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_25,

25.14 Severe Acute Respiratory Syndrome 722
25.14.1 General Characteristics 722
25.14.2 Clinical Presentation 723
25.14.3 Diagnostic Methods 723
25.14.4 Clinical Utility 724

25.15 Enterovirus 724
25.15.1 General Characteristics 724
25.15.2 Clinical Presentation 725
25.15.3 Diagnostic Methods 725
25.15.4 Clinical Utility 726

25.16 JC/BK Virus 726
25.16.1 General Characteristics 726
25.16.2 Clinical Presentation 727
25.16.3 Diagnostic Methods 727
25.16.4 Clinical Utility 728

25.17 HHV6 729
25.17.1 General Characteristics 729
25.17.2 Clinical Presentation 729
25.17.3 Diagnostic Methods 729

25.18 Clinical Utility 730

Further Reading 730

25.1 General

- Limitation and pitfalls
  - In the past decade, molecular methods for detection and quantification of virus infections have replaced many traditional viral culture and serological methods
  - The molecular tests significantly improved clinical turnaround time and reduced hands-on time in addition to increase the diagnostic sensitivity and specificity
  - Many commercial assay kits and automatic instruments are available which allow many clinical microbiology laboratories to offer molecular tests
  - However, the routine implementation of nucleic acid (both DNA and RNA) amplification and hybridization methodologies in clinical laboratories is still associated with a number of limitations
    - Increased cost/test due to expensive instrumentation and reagents
    - Amplification carryover contamination
    - Standardization of positive, negative, and quantitative controls
    - Integrated coamplified internal DNA control to demonstrate absence of polymerase chain reaction (PCR) inhibitors and amplification
    - Prevention of false-positive and false-negative reports due to antigenic and pathogen nucleic acid sequence drift and accurate interpretation of data and software analyses
- Specimens
  - Collection of adequate specimens is important for molecular diagnosis of virus. These are specimen types commonly used for molecular diagnosis
  - Whole blood: 3–5 mls collected in an EDTA (lavender top) tube. Store at 4–25 °C. Do not freeze
  - Plasma: Collect 7–10 mls of whole blood in EDTA, ACD solution A, or PPT sterile tube. Store whole blood at room temperature (18–30 °C) for no more than 4 h. Remove plasma from cells

within 4 h of collection by centrifugation at 1,000 × g for 10–15 min. Do not clarify by filtration or further centrifugation. Store plasma at –60 to –80 °C within 30 min of separation. Plasma may also be stored at –20 °C in nonfrost-free freezer for up to 72 h if colder freezer is not available. Ship on dry ice for overnight delivery. The minimum volume of specimen is 2 mls of plasma
  - Urine: first 10–20 mls of voided urine collected in a sterile urinalysis container (15 ml sterile screw cap tube preferred). Store at 4–25 °C for less than 24 h or store at –70 °C for long term
  - Bronchial lavage/tracheal aspirate: 1–4 mls, collected in a sterile tube. Store at 4–25 °C for less than 24 h
  - Bone marrow: 1–2 mls, collected in EDTA tube. Store at 4–25 °C. Do not freeze
  - Tissue: ~0.5-cm tissue block collected in a sterile screw-top container, add small amount of saline to keep it moist. Avoid the use of viral transport media to avoid potential inhibition of PCR. Fresh tissues should be stored at –72 °C immediately to preserve the nucleic acids
    - Paraffin-embedded tissue is acceptable. Usually 5–10 sections (5 μm thickness) are sufficient for PCR analysis. The tissue sections must be deparaffinized with xylene before DNA extraction
  - Fecal: sterile swab (plastic shaft only) or very small fecal sample placed in 1–2 mls sterile saline in a container with tight fitting lid. Do not use viral transport media to avoid potential inhibition of PCR
  - Swab: sterile swab (plastic shaft only) placed in 1–2 mls sterile saline. Do not use viral transport media to avoid potential inhibition of PCR
  - Cerebrospinal fluid (CSF): 1–1.5 mls fluid, submitted in a sterile, leakproof tube, store at 4–25 °C for less than 24 h or store at –70 °C for long term
- Assay performance analysis
  - Analytical performance
    - Analytical sensitivity: to determine the lowest number of targets that can be detected by the assay
    - Cross-reactivity (specificity): to determine if the assay can produce false-positive results in the presence of high concentration of other similar or unrelated pathogens (bacteria, yeast, and virus)
    - Linearity: to evaluate the log differences from the expected concentration; this difference should be within ±0.1 log (or a ratio of observed mean quantitation to expected concentration within 95%)
    - Quantitative range: the measured concentrations within the linear range with a good reproducibility
  - Clinical performance
    - Limits of detection: the lowest concentration of target nucleic acids that can be detected (at or above the detection cutoff in 95% of replicates, usually 10 replicates)
    - Detection cutoff: the point on the assay quantitation scale such that 95% of negative specimens produce results below this cutoff with 95% confidence
    - Limits of quantification: the lowest concentration of target nucleic acids that can be quantified in 95% of replicates
    - Reproducibility: The reproducibility of the test is usually established by testing three to six sample panels with known concentrations of target in triplicate or quadruplicate. A commercial panel should be used to establish this parameter, if available. Reproducibility is expressed as percent correlation coefficient. For quantitative assays, the CVs range from 10% to 50%
    - Precision: the reproducibility of a test result (e.g., inter- and intratechnologist and inter- and intra-assay)
    - Sensitivity: true positive samples, % of true positive samples above the limits of detection
    - Specificity: true negative samples, % of true negative samples below the limits of detection

- Quality controls: For quantitative assay, additional quality control procedures should be performed, including calibration and calibration verification
  - Calibration is the set of operations that establish, under specified conditions, the relationship between reagent system/instrument response and the corresponding concentration/activity values of an analyte. Calibration procedures are typically specified by a method manufacturer, but may also be established by the laboratory
  - Calibration verification denotes the process of confirming that the current calibration settings remain valid for a method
  - Recalibration or calibration verification and analytical measurement range validation must be performed at least once every 6 months
  - For each run, sensitivity controls should be included in addition to positive and negative controls. It is recommended that two levels of controls (high and low) should be included

25.2 Human Immunodeficiency Virus

25.2.1 General Characteristics

- Human immunodeficiency virus (HIV) is an RNA retrovirus belonging to the lentivirus family. HIV1 and HIV2 are genetically different; HIV2 shares 40% nucleotide homology with HIV1. HIV2 is more related to SIV than to HIV1. Both types appear to cause clinically indistinguishable AIDS. However, it seems that HIV2 is less easily transmitted, and the period between initial infection and illness is longer in the case of HIV2. Worldwide, the predominant virus is HIV1, and generally, when people refer to HIV without specifying the type of virus, they will be referring to HIV1. The relatively uncommon HIV2 is concentrated in West Africa and is rarely found elsewhere
- Structure of HIV virion (Fig. 25.1)
  - HIV virus consists of a spherical viral particle encased in a lipid bilayer derived from host cell covered by protruding peg-like structures composed of gp41 and gp120 glycoproteins
  - The virus core nucleocapsid contains the major capsid protein, p24; two copies of genomic RNA; and three viral enzymes (protease, reverse transcriptase, and integrase)
- Viral replication (Fig. 25.2)
  - The first step of infection is entry into the host cell, which requires binding of the gp120 molecule on the virus to CD4 molecules on the host cell's surface, and is mediated by the gp41 molecule. Two surface molecules CCR5 and CXCR4, chemokine receptors for beta-chemokines and alpha-chemokines are also required for entry
  - Once bound, the viral envelope fuses with the cell membrane and the virus' RNA and enzymes enter the cytoplasm
  - Reverse transcriptase catalyzes, first, the synthesis of a DNA copy of the viral RNA and, second, the synthesis of a second DNA strand complementary to the first one. Therefore, a double-stranded DNA (dsDNA) is generated
  - Integrase then facilitates the integration of viral DNA into the cellular chromosome when the cell divides and provides latency enabling the virus to effectively evade host responses
  - Transcription of the DNA results in the production of RNA. This RNA can serve as the genome for new viruses and can be translated to produce viral proteins. Viral proteins are facilitated by protease and assembled into viral particles using the host cell's protein-making machinery
  - Complete HIV particles are assembled. In macrophages, HIV buds out of the cell without rupturing the cell, and the cycle begins again. In T cells, HIV exits the cell by rupturing it, effectively killing the cell
- The gag, pol, and env genes encode for structural proteins for new virus particles.

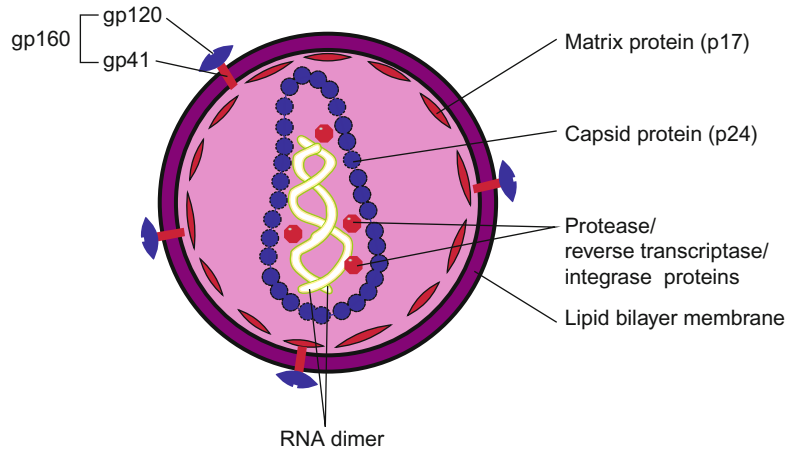


Fig. 25.1 HIV viral particle (www.mcld.co.uk/hiv)

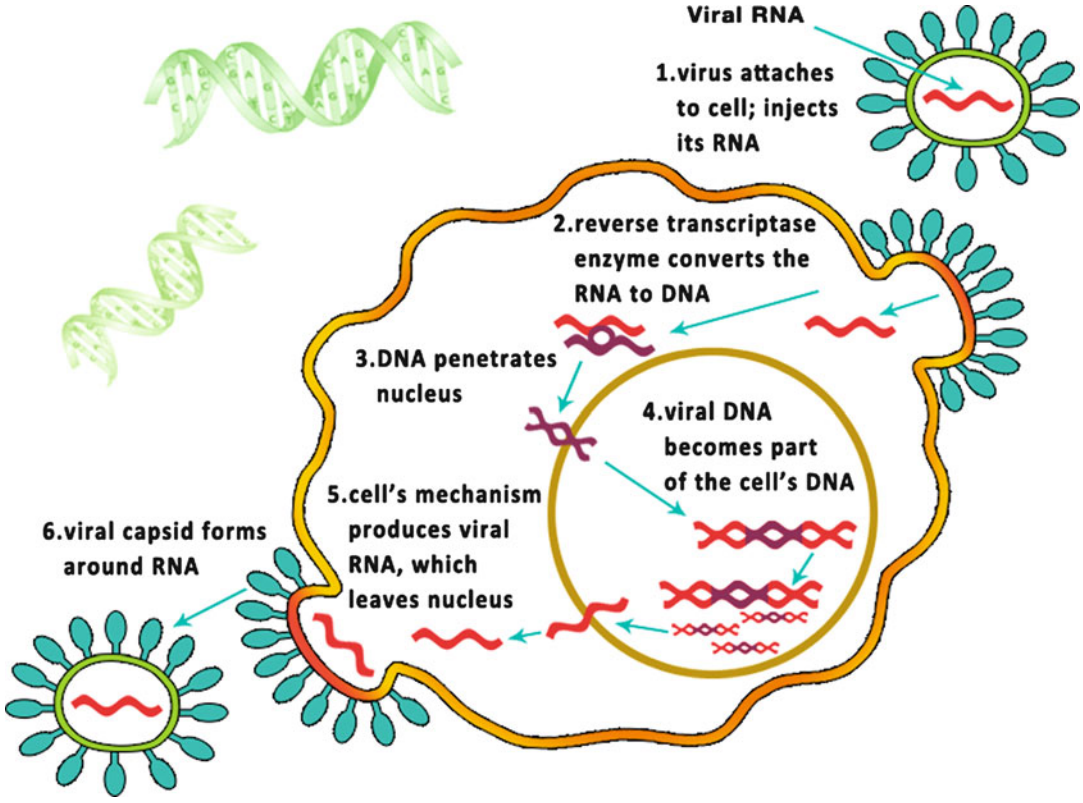


Fig. 25.2 The replication cycle of HIV

The other six genes, tat, rev, nef, vif, vpr, and vpu, regulate the synthesis and assembly of viral particles

- The phylogenic analysis of the nucleotide sequences of the env gene has enabled classification of HIV1 into three groups: M (major), N (non-M), and O (outlier). The group M of HIV1 infection has been classified into nine different genetic subtypes A–K. More than 90% of HIV1 infections belong to HIV1 group M. Subtype/class B is the most prevalent in the developed world

- HIV is transmitted via sexual contact, blood (via transfusion, blood products, or contaminated needles), or passage from mother to child (in utero, during birth, or ingestion of breast milk). Although saliva can contain small quantities of the virus, the virus cannot be spread by kissing. HIV is not spread by the fecal–oral route, aerosols, insects, or casual contact

25.2.2 Clinical Presentation

- HIV is the causative agent of acquired immunodeficiency syndrome (AIDS), the leading cause of death in humans between the ages of 25–44 years
- Two main targets of HIV: immune system and central nervous system. HIV targets CD4+ T cells, monocytes/macrophages, and Langerhans cells/dendritic cells causing severe immunosuppression and neuropathologic symptoms such as dementia, meningitis, and encephalopathy in the host
- Common opportunistic infections: *Pneumocystis carinii*, candidiasis, tuberculosis, *Cryptococcus*, cytomegaloretinitis
- Common malignancies: Kaposi sarcoma, lymphoma (non-Hodgkin and brain primary), and uterine carcinoma
- Individuals who have HIV face a long challenging road. The disease has a steady natural history, starting with an asymptomatic state and progressing toward AIDS. Natural history includes three phases
  - Early-stage HIV infection is defined as the presence of HIV with a CD4 count greater than 500. Early stage develops 3–6 weeks after initial exposure with self-limited flu-like symptoms resolving 2–4 weeks later in 50–60% of patients by high level of viral production, viremia, and widespread seeding of lymphoid tissues
  - Chronic phase HIV infection occurs when the CD4 count is between 200 and 500. Chronic phase is associated with a period of latency in which the immune system is intact, but there is continuous HIV replication that may last for years. Patients are either asymptomatic or develop persistent lymphadenopathy with minor opportunistic infections, such as candidiasis or herpes zoster
  - When the CD4 count drops below 200, the HIV infection has entered the crisis phase. This is when certain infections that are easily handled by an intact immune system take advantage of this immunocompromised state (opportunistic infections). Certain cancers may also appear for the same reason. When a patient has a CD4 count less than 200 and at least one opportunistic infection or cancer specifically seen in crisis phase HIV, he or she is officially designated as having AIDS

25.2.3 Diagnostic Methods

- Specimens – whole blood, serum, and plasma (Table 25.1)
- Conventional tests and problems
  - Lymphocyte count
    - D4 cells (also called T cells or T-helper cells) are the primary targets of the HIV virus. Quantitation of CD4 cells was the first effective predictor of HIV progression. The CD4 count is one of many factors (including clinical status, HIV viral load, and medication adherence) that should be assessed before starting or changing antiretroviral (ARV) treatment
    - The CD4 cell count (<200 cells/mm^3) is important in determining the staging of HIV disease and for indicating the need for prophylaxis against opportunistic pathogens
    - Most laboratories report the CD4 count as part of a list of several types of lymphocytes, as both an absolute count and a relative percentage. Measurement and trending of CD4 percentage in addition to absolute count must be performed prior to initiation or adjustment of ARV treatment management decisions
    - The CD4 percentage sometimes is used in coordination with the absolute value

Table 25.1 Specimens handling in different HIV assays

Assay	Collection	Transport	Storage	Comments
Antibody screening assay	Serum (including serum collected in serum separator tubes) or plasma containing heparin, EDTA, citrate, or CPDA1 anticoagulants		For long-term storage, specimens should be stored frozen. Specimens can be stored at 2–8 °C for a maximum of 14 days	Samples may be tested up to three freeze–thaw cycles
HIV monitoring assay	Plasma specimens anticoagulated with EDTA or ACD only. Specimens must **not** be anticoagulated with heparin	Whole blood should be stored at 2–25 °C for no longer than 6 h. Plasma must be separated within 6 h of collection by centrifugation at 800–1,600x g for 20 min at room temperature and transferred to a polypropylene tube to prevent viral degradation	Plasma may be stored at 2–8 °C for up to 5 days or frozen at –70 °C	Specimens should be stored in 600–700 ul aliquots in sterile, 2 ml polypropylene tubes. Freeze–thaw studies have shown that specimens may be tested for up to three freeze–thaw cycles without loss of viral RNA
HIV genotyping	Plasma specimens anticoagulated with EDTA. Specimens must not be anticoagulated with heparin	Whole blood should be stored at 2–25 °C for no longer than 2 h. Plasma should be separated within 30 mins, but no later than 120 min by centrifugation at 1,000–2,000x g for 15 min at 15–25 °C and transferred to a polypropylene tube	Plasma may be stored frozen at –65–80 °C for up to 6 months	Samples may be tested up to two freeze–thaw cycles. Plasma specimens containing the following have been shown to interfere with results: lipids up to 30 mg/ml bilirubin up to 0.6 mg/ml hemoglobin up to 5 mg/ml

to assess the significance of changes in the absolute CD4 count. The absolute CD4 count can fluctuate as overall lymphocyte counts vary, but the CD4 percentage often remains stable during insignificant CD4 fluctuations. CD8 cell (or cytotoxic T cell) counts do not appear to predict clinical outcomes

- For monitoring purposes, the CD4 count should be repeated approximately every 3–4 months both in stable untreated patients and in patients on stable ART. The CD4 count should be checked more frequently according to the clinical situation

– Viral culture
  - Although very specific, single positive culture must be confirmed with a second specimen
  - Rarely used due to high cost, labor-intensive, and less sensitivity than antibody testing
  - Negative culture may be caused by technical problems, a defective virus, or the inability of the virus to replicate in culture
– Serological studies
  - p24 antigen
    – Early developed assay to detect HIV infection and screen donated blood for HIV
    – Advantage is to detect HIV infection prior to development of antibodies
    – Disadvantage is limited utility due to the short window of time and should only be used when other tests are unavailable

- Antibody screening assays (qualitative)
  - Detection of antibodies to HIV is the most common way to diagnose HIV infection in adults and children >18 months old
  - These antibodies are usually detectable within 3–6 weeks after infection
  - Most individuals seroconvert by 12 weeks, although may not be detectable for months or years
  - The window period is the time between infection and the development of antibodies. When a person is infected with HIV, it takes a few weeks for the body to make antibodies to the virus. Most people develop antibodies within a month of infection, although some people can take up to 3 months. In very rare cases, it can take 6 months of the test to be positive. In general, repeat testing for HIV should occur at 3 months if there is a significant concern of recent HIV infection. Additionally, because the level of virus in blood is high during the window period, people can more easily transmit HIV
  - Serologic HIV antibody screening testing is highly sensitive (ELISA, rapid test, or home test), but requires followup of preliminary positive specimens with a highly specific HIV antibody confirmatory assay (Western blot) (Fig. 25.3)
  - ELISA method is most common and earliest developed antibody screening assay
  - Home Access HIV1 test system analyzes a dried-blood spot from finger stick collected on filter paper at home and sent to a testing facility
  - Rapid tests for HIV are assays that detect antibodies to HIV within minutes. The rapid test is highly specific: negative means negative except during window period; and the test is also highly sensitive: positive means most likely has HIV, but must be confirmed using Western blot for HIV diagnosis
- Confirmatory antibody assays: Western blot
  - Gold standard for HIV diagnostic testing
  - The virus is disrupted, and the individual proteins are separated by molecular weight via differential migration on a polyacrylamide gel and blotted onto a membrane support. HIV serum antibodies from the patient are allowed to bind to the proteins in the membrane support, and patterns of reactivity can be visibly read
  - Detects three major proteins/viral bands: p24 core protein and two envelope proteins, gp41 and gp120/160
  - Reactive WB demonstrates antibody to two of the three major bands; nonreactive WB will have no detectable viral bands (Fig. 25.4)
  - Repeated reactivity by ELISA and reactivity by the confirmatory assay are reported as positive for antibody to HIV1
  - Nonreactive specimens by ELISA or repeatedly reactive by ELISA and nonreactive by the confirmatory assay are negative for antibody to HIV1
  - WB in which serum antibodies bind to any other combination of viral bands is considered indeterminate; followup blood specimen should be obtained 1 month later for repeat HIV antibody testing
  - Individuals with repeat indeterminate results should undergo further testing using molecular assays, such as PCR
  - At least as sensitive as and more specific than screening assays, although they are not as sensitive in the detection of early seroconversion
  - Disadvantages: more labor-intensive, more prone to subjective interpretation, and more costly than screening assays

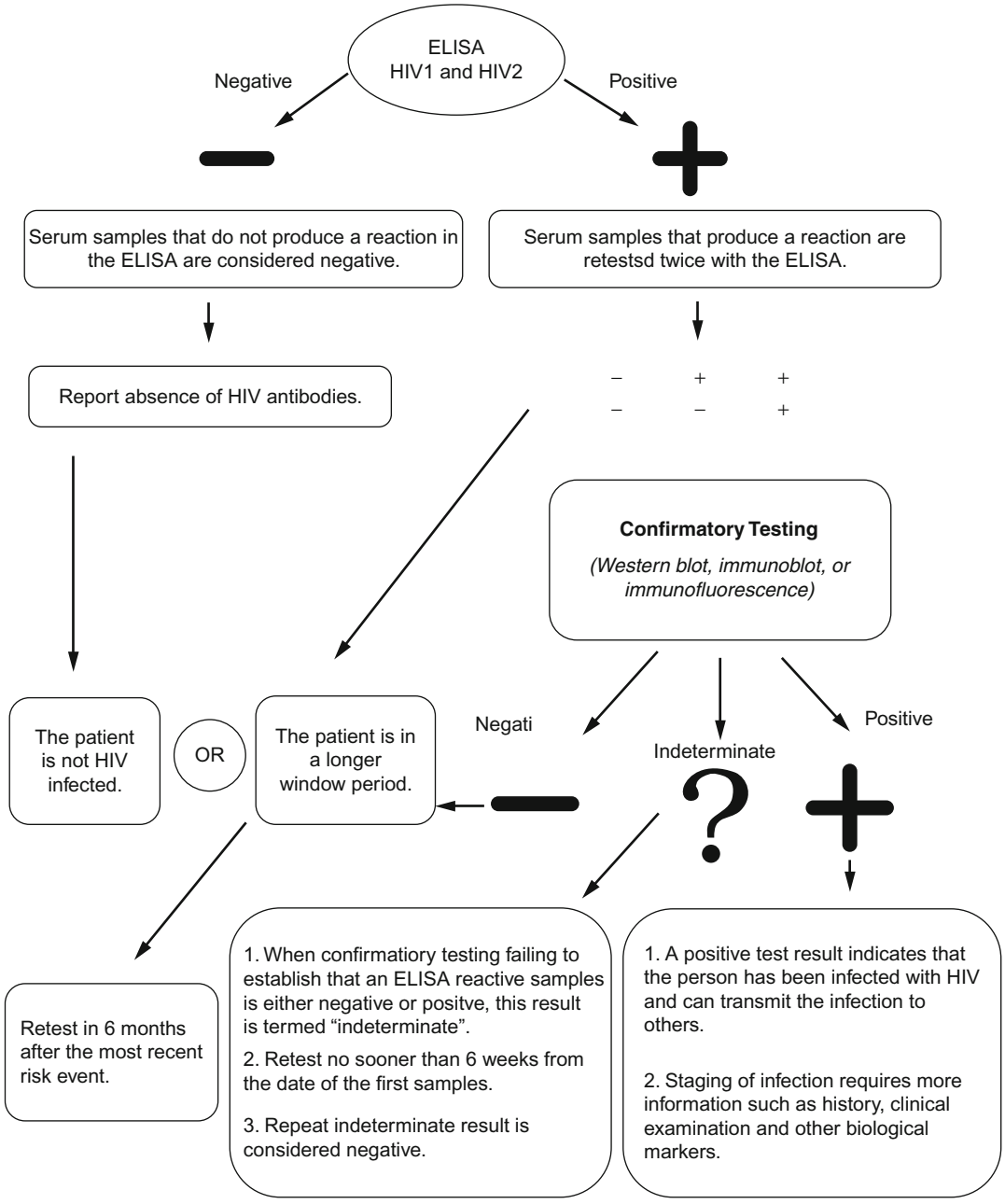
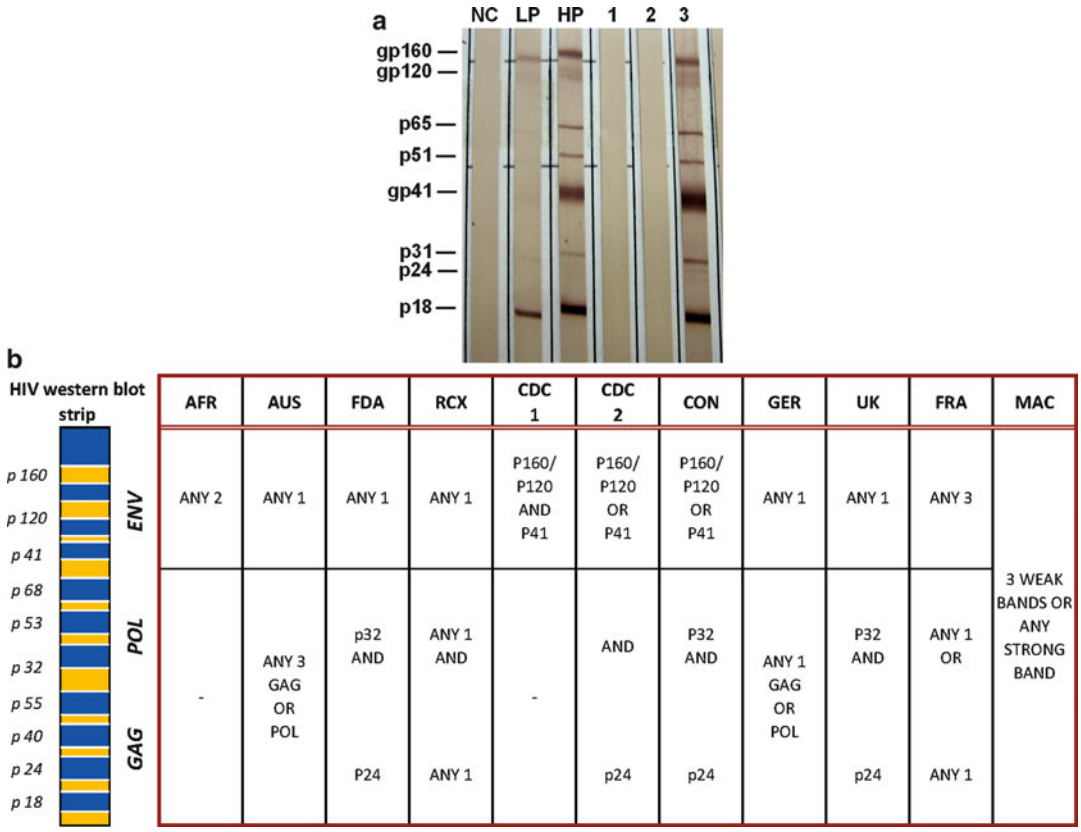


Fig. 25.3 HIV antibody screening algorithm

- Alternative antibody screening assay (qualitative)
  - US Food and Drug Administration (FDA) has approved assays that test body fluids other than blood to detect HIV1 antibodies, although sensitivity and specificity are less reliable
  - Utilizes same testing algorithm as serum (ELISA followed by WB)



	AFR	AUS	FDA	RCX	CDC 1	CDC 2	CON	GER	UK	FRA	MAC
ENV	ANY 2	ANY 1	ANY 1	ANY 1	P160/ P120 AND P41	P160/ P120 OR P41	P160/ P120 OR P41	ANY 1	ANY 1	ANY 3	3 WEAK BANDS OR ANY STRONG BAND
POL	-	ANY 3 GAG OR POL	p32 AND	ANY 1 AND	-	AND	P32 AND	ANY 1 GAG OR POL	P32 AND	ANY 1 OR	
GAG			P24	ANY 1		p24	p24		p24	ANY 1	

Fig. 25.4 (**a**) Western blot analysis and band pattern interpretation: *NR* negative control, *LR* low reactive control, *HR* high reactive control, *1–2* nonreactive, *3* reactive. (**b**) Criteria used to define a positive HIV Western blot. *AFR* Africa, 1 *AUS* Australia, 2 *FDA* US Food and Drug Administration, 3 *RCX* US Red Cross, 3 *CDC* US Center for Disease Control, 3 *CON* US Consortium for Retrovirus Serology Standardization, 3 *GER* Germany, *UK* United Kingdom, *FRA* France, *MACS* US Multicenter AIDS Cohort Study 1983–1992. *Bands not in electrophoretic order

  - Advantages are noninvasive sample collection, increased safety due to lack of needles, and disposal of infectious waste minimized
  - Oral fluid (oral mucosal transudate): antibodies detectable, but significantly lower (800– 1,000-fold) than those of serum
  - Urine: Interpretative criteria for a reactive WB requires only presence of visible band at gp160 region
- Molecular methods
  - Qualitative assay: viral identification assays – recommended for resolving indeterminate Western blot results
    - DNA PCR (Roche)
    - The HIV DNA PCR test is a new technology used for HIV early detection testing. The new HIV PCR test looks for the DNA copy of the HIV virus itself in peripheral blood. It is extremely accurate, sensitive, and the cutting edge of HIV early detection testing
    - The average window period with HIV1 antibody tests is 25 days. Antigen testing cuts the window period to approximately 16 days, and HIV DNA test further reduces this period to 12 days
    - Detection of HIV DNA in peripheral blood mononuclear cells by PCR is recommended for children <18 months old born to HIV1-infected mothers

Table 25.2 Comparison of commonly used HIV viral load assays

Characteristic	1) Amplicor HIV-1 Monitor; 2) COBAS® AmpliPrep/ COBAS® TaqMan® HIV-1 Test,v2.0 [*Roche*]	Branched Chain DNA (bDNA) [*Versant*]	Nucleic Acid Sequence-based Assay (NASBA) [bioMérieux]
Amplification method	Target amplification	Signal amplification	Target amplification
Specimen type	Plasma in ACD or EDTA tube	Plasma in EDTA tube	Plasma in ACD, EDTA, or heparin tube
Specimen volume	Standard 1.5:0.2 mL Ultrasensitive 1.5:0.5 mL V2.0:0.5 mL	1.0–2.0 mL	1.0 mL
Specimen transport	Prepare plasma within 6 h of collection; store specimens at –20°C or –70°C	Prepare plasma within 4 h of collection; store specimens at –20°C or –70°C	Prepare plasma within 4 h of collection; store specimen at –20°C or –70°C
Sensitivity (copies/mL)	Standard 1.5 (400) Ultrasensitive 1.5 (50) v2.0 (20)	Version 3.0 (75)	NucliSens QT (176)
Dynamic range (copies/mL)	Standard 1.5 (400–750,000) Ultrasensitive 1.5 (50–100,000) V2.0: (20–10,000,000)	Version 3.0 (75–500,000)	NucliSens QT (80–3,470,000)
Area of HIV genome selected for amplification	Gag or Gag/LTR	Pol	Gag

- False-positive reactions common due to small amounts of background "noise" or contamination
- All initial positive DNA PCR reactions must be confirmed with a second PCR test on a separate specimen
- Currently, recommended only for detection in infants born to mothers infected with HIV1. However, potential for false-positive result must still be recognized
- Plasma HIV RNA
- Surrogate marker of HIV disease progression
- During acute infection, viral load levels are very high (ranging from 100,000 to over 10 million copies/mL) and detectable before seroconversion
- Important to use both a plasma HIV RNA assay and antibody/Western blot testing to establish diagnosis in acute and primary infections
- Low levels of virus (<5,000 copies/mL) may be indicative of a false-positive result and should not be considered diagnostic of primary HIV infection. Standard antibody testing should be repeated

– Quantitative
  - Viral monitoring (Table 25.2 and Fig. 25.5)
  - Amplicor HIV1 Monitor and COBAS® AmpliPrep/COBAS® TaqMan® HIV1 Test,v2.0 (Roche) (FDA approved)
    – Quantitation of HIV1 RNA in plasma
    – Standard assay limit of detection for Amplicor HIV1 Monitor: >400 copies/ml and used for monitoring patient not on ARV therapy
    – Ultrasensitive assay and v2.0: as low as 50 or 20 copies/ml and used for monitoring viral loads <400 copies/ml. Viral load results of <50 or 20 copies/mL do not necessarily indicate absence of HIV1 viral replication. Inhibitory substances may be present in the plasma specimen, leading to negative or falsely low HIV1 RNA results. Improper specimen collection

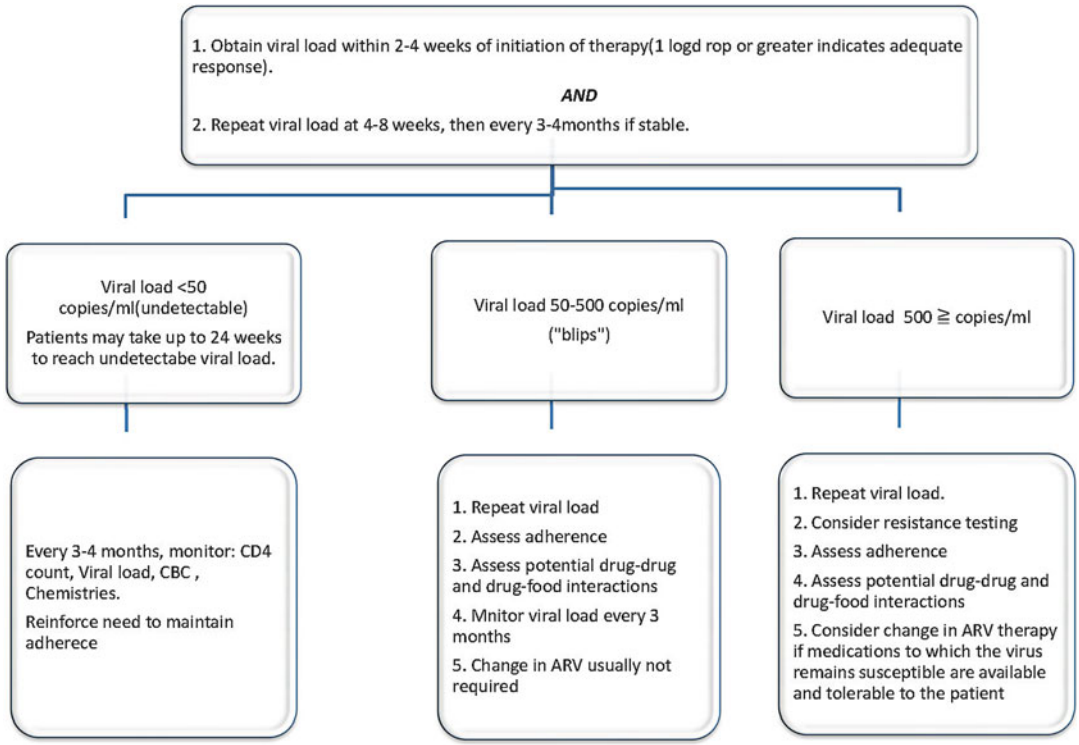


Fig. 25.5 HIV viral monitoring algorithm

or storage may falsely lower the plasma viral load results

- Plasma specimens collected using ACD anticoagulant show quantitative HIV1 RNA levels that are ~15% lower than those collected in tubes containing EDTA
- Acute concurrent illness and/or recent vaccination may cause transient rise in viral load
- Calculation of HIV viremia for Amplicor HIV1 Monitor and COBAS® AmpliPrep test is based on optical density reading. Input QS copies are lot specific and provided with each kit. Standard sample volume factor = 40; ultrasensitive sample volume factor = 4. OD_{450}, optical density at 450nm; QS, quantitation standard. The well with the lowest OD_{450} reading between 0.2 and 2.0 is selected for calculation. (Figs. 25.6 and 25.7)

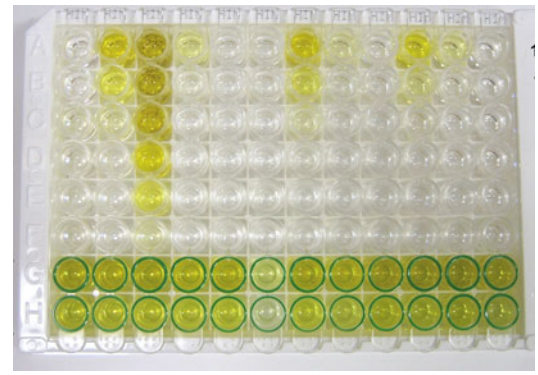


Fig. 25.6 HIV assay plate layout. *N* Negative, *L* low positive control, *H* high positive control, *S1–S9* samples, *DF* dilution factor, *QS* quantitation standard

- The HIV TaqMan test v2.0 uses three LTR primers with one FAM-labeled LTR probe in conjunction with four GAG primers and one FAM-labeled GAG probe. Therefore, the v2.0 offers primers and probes that are

$$\left[\frac{\text{HIV OD}_{450} \times \text{DF}}{\text{QS OD}_{450} \times \text{DF}}\right] \times \text{Input QS copies} \times \text{Sample volume factor} = \text{HIV -1 RNA copies/ml}$$

Calculation Example: Low Positive Control

$$\left[\frac{0.166 \times 25}{0.931 \times 5}\right] \times 51 \times 4 = \text{HIV -1 RNA copies/ml (ULTRA)} = 182 \text{ HIV RNA copies/ml}$$

	N	L	H	S1	S2	S3	S4	S5	S6	S7	S8	S9
A	0.015	2.808	-	**0.288**	0.009	0.011	2.301	**0.156**	0.013	1.343	**0.257**	0.017
B	0.010	0.703	-	0.055	0.008	0.009	**0.530**	0.031	0.006	**0.239**	0.043	0.008
C	0.085	**0.166**	-	0.022	0.010	0.013	0.117	0.011	0.008	0.056	0.014	0.011
D	0.060	0.046	2.076	0.015	0.013	0.009	0.031	0.008	0.007	0.018	0.008	0.012
E	0.044	0.018	**0.726**	0.012	0.011	0.031	0.008	0.004	0.004	0.005	0.005	0.006
F	0.046	0.013	0.120	0.009	0.010	0,009	0.003	0.003	0.002	0.004	0.003	0.003
G	2.833	2.729	-	2.628	2.700	**0.313**	2.618	2.183	2.775	2.579	2.865	2.228
H	**1.003**	**0.931**	**0.972**	**0.850**	**0.795**	0.112	**0.864**	**0.581**	**1.130**	**0.892**	**1.116**	**0.593**
HIV	**UND**	**182**	**19,046**	**< 50**	**UND**	**UND**	**125**	**<50**	**UND**	**55**	**<50**	**UND**

Fig. 25.7 HIV viral load (ultrasensitive) calculation

used to amplify the gag and LTR regions; it also provides diagnostic accuracy of test results even if mutations occur in one of the two regions and compensates for the possibility of mismatch occurring with a primer/probe region. It also ensures enhanced reliability of test results and more confidence in assessing viral loads (Fig. 25.8)

- A single HIV1 viral load test result should not be used as the only criteria to guide therapeutic decision and intervention in the clinical care of HIV1-infected patients. Viral load results should be correlated with patient symptoms, clinical presentation, and CD4 cell count
- Due to the inherent variability in the assay, physiologic variation, and concurrent illnesses in the infected patients, <100-fold (<2 log) change in plasma HIV1 viral load should not be considered as significant change

- Branched DNA (bDNA) (Bayer)
  - The VERSANT@ HIV1 RNA 3.0 Assay (bDNA) is a signal amplification nucleic acid probe assay for the direct quantitation of HIV1 RNA in plasma of HIV1-infected individuals using the Bayer@ System 340 bDNA Analyzer. The test can quantitate HIV1 RNA over the range of 75–500,000 HIV1 RNA copies/ml
  - bDNA is based on a series of hybridization procedure followed by an enzyme substrate reaction. The basic principles and procedures for bDNA 3.0 include viral lysis, overnight hybridization of viral nucleic acid, and target capture in a 96-well plate, followed by wash steps and the serial addition of probes that allow for signal amplification, detection, and quantification
  - bDNA methods have progressed from first generation assays, which were accurate and reproducible but

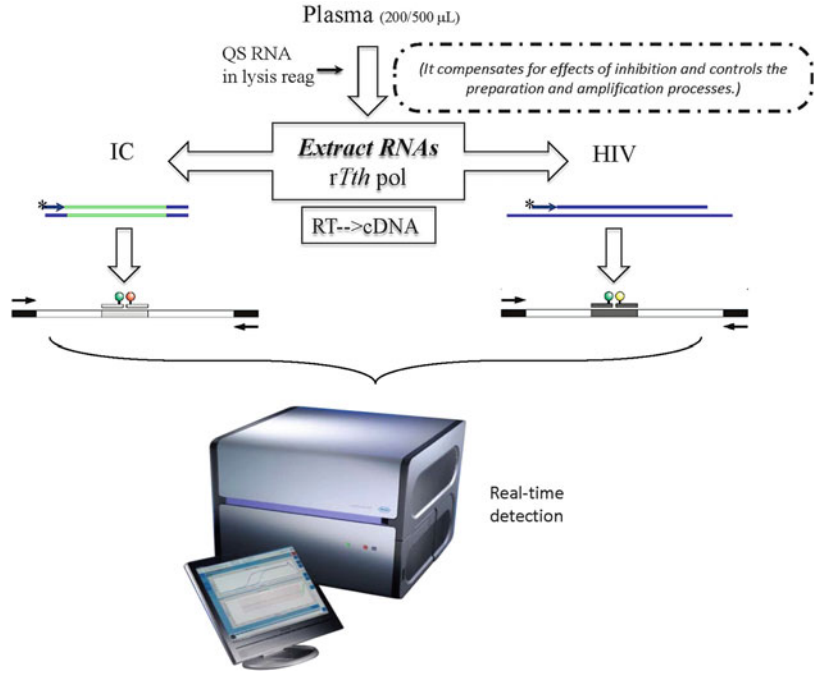


Fig. 25.8 HIV detection: quantitative RT-PCR (Roche TaqMan)

relatively insensitive, to third generation bDNA tests that are accurate, reproducible, highly sensitive, and automated laboratory tests for more optimal patient management. The results of this evolution in bDNA technology are FDA-approved methods for the detection and quantification of HIV1 (VERSANT HIV1 RNA 3.0 Assay)
  - HIV1 present in patient blood is disrupted to release viral RNA
- Nucleic acid sequence-based assay (NASBA), NucliSens HIV1 QT assay, NucliSens EasyQ HIV1 assay
  - NASBA is a sensitive, isothermal, transcription-based amplification system designed specifically for the detection of RNA targets. NucliSens HIV1 QT is both FDA approved and CE marked. The NucliSens HIV1 QT is a NASBA with endpoint electrochemiluminescence (ECL) detection. Further refinement has led to the development of the CE marked NucliSens EasyQ HIV1 – a combination of NASBA amplification and real-time detection using molecular beacons utilizing the NucliSens EasyQ analyzer
  - The NucliSens EasyQ HIV1 assay (bioMérieux, Boxtel, The Netherlands) is a quantitative, next generation amplification assay designed to overcome the labor and throughput obstacles in the earlier assays
  - Real-time detection of amplicons occurs during the NASBA reaction following primer exhaustion and continued production of antisense RNA from the T7 promoter. During the exponential phase of production, amplicons are detected using molecular beacon oligonucleotide probes that fluoresce upon binding to the specific target sequence
  - The published performance characteristics of the NucliSens EasyQ

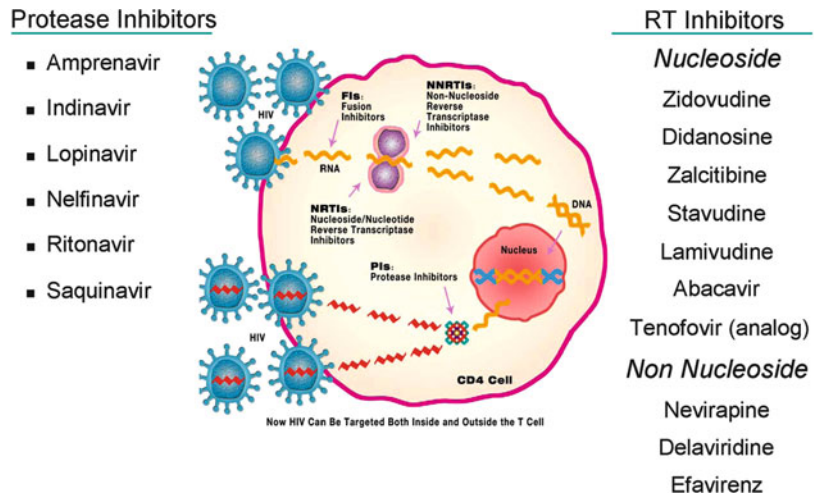


Fig. 25.9 Sixteen drugs approved for HIV treatment

HIV1 v1.1 assay show that it is comparable to the NucliSens HIV1 QT, the Versant HIV1 RNA, and the COBAS Amplicor HIV1 Monitor v1.5 assays in linearity, specificity, and reproducibility
  - Although only plasma is FDA approved, can use CSF, lymph tissue, genital secretions, and cells
  - Purified nucleic acid may be used for other molecular testing, such as sequencing
- Genotyping
  - Genotypic resistance testing uses RT-PCR and DNA sequencing techniques to identify the presence or absence of resistance related mutations in the viral genome
    - 16 ARV approved for the treatment of HIV (Fig. 25.9). ARV therapy slows the replication of HIV in the body. If patients are infected with HIV and are currently on active ARV therapy but the patients' viral load continues to increase, then the patients may undergo either genotypic or phenotypic resistance testing
    - Phenotypic testing measures the ability of the HIV1 virus to grow in different concentrations of drug under artificial conditions in the laboratory. Although phenotyping is a direct measure of resistance, it is more complex than genotyping and therefore slower and more costly to perform
    - Genotyping testing detects viral sequence change and sequence entire HIV protease (codon 1–99) and HIV-RT (codon 40–335) to cover all known drug resistance mutations
    - No consensus on genotyping versus phenotyping; however, it is anticipated that genotyping will be used more often because of its greater accessibility, lower cost, and faster turnaround time
- Mechanism of resistance
  - HIV is a highly polymorphic G virus (quasispecies) which during replication converts RNA to DNA by the action of the viral reverse transcriptase enzyme (RT)
  - The RT enzyme has very little proofreading (correction) capacity, and therefore errors are incorporated into the proviral DNA during replication. Over time, these errors, at concise drug binding sites, can provide a selection advantage for the virus in the presence of ARV drugs

  - The resistant virus predominates with a subsequent increase in viral load. However, the extent of such resistance and the implications for choice of therapy can be determined by reading the sequence of the genes encoding the protease and the RT enzymes
  - Taking medication exactly as prescribed is a very important part of avoiding resistance. Missing doses or not taking them on time lowers the amount of ARV chemicals in the body, which means the virus is not properly suppressed. The virus is then able to replicate faster, increasing the chance of it becoming resistant
  - Resistance to some ARVs can limit future treatment options. If HIV is resistant to one drug, it will sometimes be resistant to similar drugs in the same group. This is called cross-resistance and it means that some ARV drugs will not work even if they have not been used before
- Indications for drug-resistant testing
  - Drug-naïve patients with acute or recent infection
  - Therapy failure, including suboptimal treatment response, when treatment change is considered
  - Pregnant HIV1-infected women and pediatric patients with detectable viral load when treatment initiation or change is considered
  - Transmitted drug-resistant virus is common in some areas and is more likely to be detected earlier in the course of HIV infection; consider resistance testing earlier in the course of infection
- One of the FDA-approved assay for DNA sequencing: TRUGENE™ HIV1 Genotyping and OpenGene DNA Sequencing System, Bayer HealthCare, Berkeley, CA
  - It is a two-step procedure which first amplifies the protease and reverse transcriptase regions of the HIV1 genome using RT-PCR
  - The amplified DNA is then sequenced to yield to the nucleotide profile of the virus using a sequencing gel
  - Once the sequence has been generated, it is compared to the wild-type HIV1 sequence and any differences that confer drug resistance are highlighted
- Another FDA-approved assay for DNA sequencing is ViroSeq™ HIV1 genotyping system, Celera Diagnostics, Alameda, CA (distributed by Abbott Laboratories, Abbott Park, Il)
  - It is a two-step procedure which first amplifies the protease and reverse transcriptase regions of the HIV1 genome using RT-PCR and cycling sequencing
  - The amplified DNA is then sequenced to yield to the nucleotide profile of the virus using a capillary electrophoresis
  - The minimum input of viral RNA to the assay should be 1,000 copies/ml when using 1 ml of plasma to be successful in genotyping
- Pitfalls of genotyping
  - Genotypic variants comprising less than 20–30% of the sample may not be detected as genotyping results reflect the predominate subtype
  - Interpretation of genotyping results is based on the HIV1 clade B, the most prevalent clade in the developed world. However, other subtypes and recombinants of HIV1 may be undetected
  - Assessing HIV1 resistance is complicated by the replication kinetics of resistant mutants. Resistant mutants are often less fit than wild-type virus and may become undetectable with selective drugs. Nevertheless, these mutants persist in the patient and when the selective drug pressure is reapplied, the mutants replicate and a resistant population quickly predominates

25.2.4 Clinical Utility

- Plasma HIV RNA is a surrogate marker of HIV disease progression that is used to guide and monitor therapy and management
- ARV therapy should be implemented in patients with any of the following clinical findings: symptomatic HIV infection or AIDS-defining condition, CD4 count $\leq$350 cells/mm3 or viral load $\geq$100,000 copies/ml (pregnant mothers: $\geq$1,000 copies/ml)
- The initial highly active antiretroviral therapy (HAART) goal in the ARV therapy-naïve patient should be to attain a viral load of <50 copies/mL and should include the rational sequencing of ARV agents to achieve the maximum possible viral replication suppression
- In ARV treatment-naïve patients or patients who are on a successful treatment regimen, monitoring of viral loads should be measured at baseline, every 2–4 weeks after initiation and every 3–4 months once maximal suppression is attained, although patients with CD4 counts >500 cells/mm^3 may require less frequent viral load monitoring
- Typically, in patients beginning therapy or in those changing therapy as a result of virologic failure, viral load measured 2–4 weeks after therapy initiation. A decrease by at least 1 log (10-fold) indicates effective therapy. Most patients reach the goal of <50 copies/mL within 6 months. An absent or incomplete response of the viral load to ARV therapy should raise concerns about poor patient adherence to therapy and/or viral resistance
- If significant increase (3-fold increase or more) in viral load without clear explanation, viral load should be repeated to confirm virologic failure
- Genotypic resistance testing should be performed prior to initiating treatment in ARV therapy-naïve patients and in patients with >1,000 copies/ml, or nonresponsive to ARV
- Genotypic resistance testing is not recommended in patients with 500–1,000 copies/mL or less and has discontinued ARV therapy for more than 1 year

25.3 Hepatitis C Virus

25.3.1 General Characteristics

- Hepatitis C virus (HCV) was first recognized in 1974 as a non-A, non-B hepatitis virus (NANBH) and first identified in 1989 using molecular methods. HCV is the major cause of non-A, non-B hepatitis (91%) affecting about 3% of the world's population
- The most common route of transmission is via blood and blood products, i.e., immune globulin, surgery, and intravenous drug abuse which has significantly reduced with the advent of routine blood screenings. Sexual transmission as well as vertically from mother to infant occurs; the rate of vertical transmission of HCV is 6%
- HCV is a positive sense, single-stranded RNA virus that represents the third genus of the family *Flaviviridae*. The genome encodes for a single open reading frame coding structural (one core and two envelopes) proteins as well as a series of nonstructural proteins (Fig. 25.10)
  - 5′ untranslated region: most constant, used for HCV RNA assays and genotyping
  - Core region: constant, used in some genotype assays, core protein assay, PCR–RFLP, and RIBA tests
  - Envelope region: hypervariable region, associated with high rate of mutation in quasispecies
  - NS2 region: codes for protease
  - NS3 region: codes for protease/helicase, RIBA tests found in this region
  - NS4 region: c100p antigen used in anti-HCV, RIBA tests targeted this region
  - NS5a region: codes for interferon response element
  - NS5b region: codes for RNA polymerase, NS5 antigen used in anti-HCV, RIBA tests target this area
- HCV consists of a heterogeneous group of genotypes based on the sequence homology of 5′ untranslated region. Currently, there are 6 types and over 90 subtypes. Types 1, 2, and 3

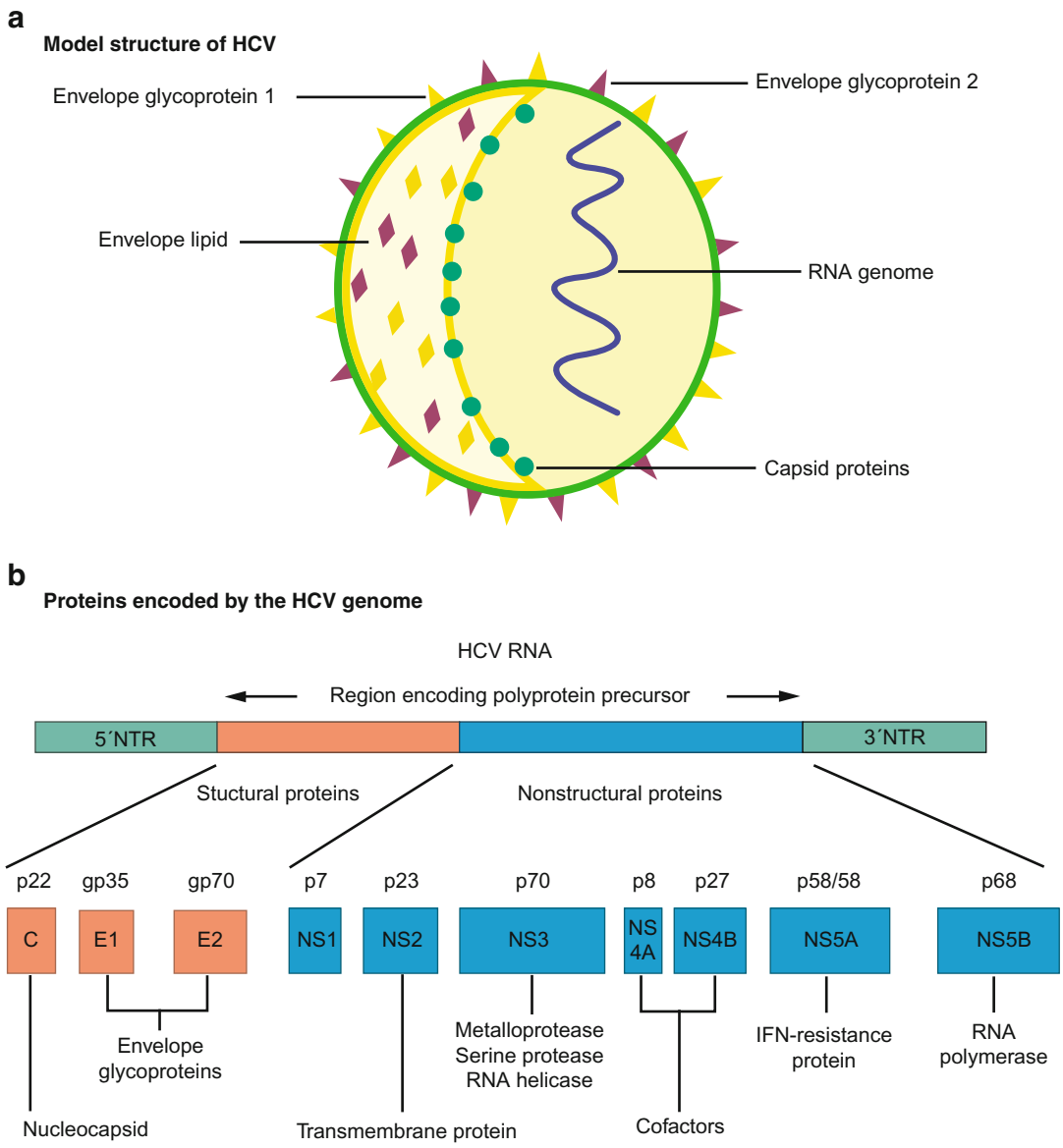


Fig. 25.10 Hepatitis C virus (*HCV*) genome structure (Adapted from Anzola et al. Expert Rev Mol Med 2003)

distributed worldwide, with types 1a and 1b responsible for approximately 60% of infections. Type 4 occurs primarily in the Middle East, type 5 in South Africa, and type 6 in Hong Kong. In the United States (US), approximately 72% of people infected with HCV have genotype 1, and most others are types 2 or 3 (genotypes 4, 5, and 6 are not common in the US)

- There is little difference in the mode of transmission or natural history of infection among the different genotypes
- Cure rates with antiviral therapy are notably higher with genotypes 2 and 3, and the duration of HCV therapy is shorter for these genotypes
- Infection with HCV is curable by therapy, with the current standard treatment based on

the combination of pegylated interferon alpha (IFNa) and ribavirin. Virus eradication, characterized by the sustained virological response, i.e., an undetectable HCV RNA 24 weeks after treatment completion, is achieved in 40–50% of patients infected with HCV genotype 1 and approximately 80% of patients infected with genotypes 2 and 3. The outcome of therapy is influenced by several parameters, including the treatment schedule, disease characteristics, viral factors, and recently identified genetic factors that include single nucleotide polymorphisms located upstream of the gene encoding IFNλ3 (IL28B)

- MicroRNA-targeted therapy has been suggested as a potential means of combatting the virus but is not currently in use

25.3.2 Clinical Presentation

- Prior to the isolation of the virus in 1989, hepatic infection with HCV was previously known as non-A, non-B hepatitis
- In the US, approximately four million have been exposed to the virus; three million are chronic carriers
  - Acute infection is usually asymptomatic. 25% of patients develop acute hepatitis with jaundice and abnormal liver function (Fig. 25.11)
    - HCV RNA can be detected in blood within 1–3 weeks and is present at the onset of symptoms
    - Antibodies to HCV are detected by enzyme immunoassay (EIA) in only 50–70% of patients at the onset of symptoms, increasing to more than 90% after 3 months RNA
- Chronic infection: 60–85% patients eventually develop chronic infection and/or chronic hepatitis (Fig. 25.11)
  - Persistence of HCV infection is diagnosed by the detection of HCV RNA in the blood for at least 6 months
  - The most important sequelae of chronic HCV infection are progressive liver fibrosis leading to cirrhosis, end-stage liver disease, and hepatocellular carcinoma (HCC)
  - Patients are often asymptomatic or have nonspecific symptoms such as fatigue, malaise, and abdominal discomfort
    - Mild to moderate elevations of ALT (SGPT) or AST (SGOT)
  - Some asymptomatic patients have normal liver enzymes
    - As many as 44% have normal levels at initial evaluation
    - May have positive HCV antibody test despite normal liver enzymes
      - 20% patients eventually develop cirrhosis which takes decades to occur. The severity of liver cirrhosis does not correlate with liver enzymes and can only be evaluated by liver biopsy
      - All etiologies of cirrhosis increase the risk, and patients with cirrhosis have an annual risk of 1–6% of developing HCC
      - Approximately 250,000 patients die each year from HCV-related liver disease

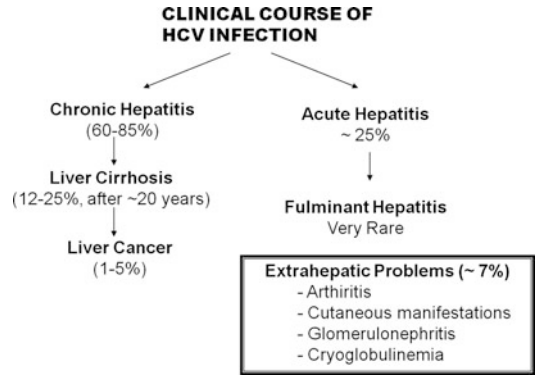


Fig. 25.11 Natural history of hepatitis C virus infection (Adapted from National Institutes of Health 2002)

25.3.3 Diagnostic Methods

- Specimens: blood plasma or serum
  - Collection of samples in EDTA plasma

  - Rapid separation of serum or plasma from cells is recommended by centrifugation within 1 h of collection
  - Unseparated EDTA plasma is stable at room temperature up to 24 h after collection
  - Separated serum or plasma is stable at room temperature for up to 3 days, at refrigerator temperatures for up to 1 week, and frozen at –70 °C for years
- Conventional tests and problems
  - Serological studies
    - Enzyme immunoassay (EIA)
      - The detection of HCV antibodies is recommended as the initial test for the identification of HCV and is useful for screening at risk populations
      - EIA is comparatively inexpensive, reproducible, and carries a high sensitivity (99%) and specificity (99%)
      - EIA can detect antibodies average 2–10 weeks after infection. However, during this "window period," a patient will have detectable viral RNA, but have undetectable levels of antibodies
      - A negative enzyme immunoassay is usually sufficient to exclude the diagnosis of HCV infection in immunocompetent patients
      - However, the test can be falsely negative in those with immunodeficiencies or end-stage renal disease
      - Conversely, false-positive EIAs may occur in patients with autoimmune disorders. In these patients, an assay for HCV RNA is necessary for diagnosis of chronic infection
      - Once patients seroconvert, they usually remain positive for HCV antibody. Thus, the presence of HCV antibody may reflect remote or recent infection
      - The immunoblot assay is still useful as a supplemental assay for persons screened in nonclinical settings and in persons with a positive EIA who test negative for HCV RNA
      - Other HCV antibody assays include anti-HCV IgM assays and avidity tests. The significance of the presence of anti-HCV IgM during HCV infection is unclear. Anti-HCV IgMs have been reported in 50–93% of patients with acute HCV and 50–70% of patients with chronic HCV
      - Therefore, anti-HCV IgM cannot be used as a reliable marker of acute HCV infection and, so far, IgM assays have not been used in clinical practice. However, increasing serial measurements of anti-HCV IgM titers early after the onset of the symptoms may help to identify patients with acute HCV. Alternatively, an increase in the anti-HCV IgG avidity index within a week after the onset of clinical symptoms has also been reported to indicate acute HCV infection. Both parameters could be used together before anti-HCV seroconversion occurs, or when no baseline sample is available to confirm the diagnosis of acute infection
    - A total HCV core antigen enzyme-linked immunosorbent assay (ELISA HCV 3.0) and ORTHO® trak-C™ assay (Ortho Diagnostics) for detection and quantification of total core antigen in blood
      - The HCV core protein is highly antigenic, induces specific cellular and humoral responses, and probably plays a pivotal role in the pathogenesis of HCV infection. The availability of an anticore monoclonal antibody allowed the development of an ELISA to detect HCV core Ag in peripheral blood of patients with HCV
      - It tests positive for anti-HCV antibodies and for prospective low-risk population screening
      - Total HCV core antigen ELISA (quantitative, Ortho Clinical Diagnostics) has sensitivity close to PCR assays in diagnosing acute HCV infection in windows period (before HCV antibodies developed)

  - It is also used in monitoring response to antiviral treatment
  - The ELISA HCV 3.0 uses three recombinant antigens (c22–3, c200, and NS5) originating from four regions of the viral genome (core, NS3, NS4, and NS5)
  - ORTHO® trak-C™ assay is a quantitative immunoassay for total (both free and antibody bound) HCV nucleocapsid core antigen. The assay uses a capture ELISA format. The difference between the two assays (ELISA HCV 3.0 and ORTHO® trak-C™ assay) is a step in this assay that disrupts immune complexes present in the sample. The assay's intended use is testing for HCV core antigen either during preseroconversion (acute) or postseroconversion (chronic) phases of HCV infection
  - Less expensive and less prone to carryover than PCR testing and can be used as a rough screen as it correlates with HCV RNA load. However, it is less sensitive than PCR-based assays and individual variation is higher
  - Similar technology is used by the Architect HCV assay by Abbot Laboratories. The analytical sensitivity of the recently developed Architect HCV assay (Abbott Laboratories, Abbott Park, IL) varies according to the HCV genotype from 500 to 3,000 international units of HCV RNA per milliliter (IU/mL)
  - A drawback of the total HCV core Ag ELISA is its lower limit of detection, and HCV-positive patients on dialysis often have low HCV RNA levels
- Recombinant immunoblot assay (RIBA)
  - RIBA has been developed for the simultaneous detection of anti-HCV antibodies and HCV core antigen
  - RIBA was used to confirm EIA results since the early generation. It had a high rate of false positives
  - Third generation of RIBA (RIBA 3.0) was developed to test HCV (which includes NS5 protein) after earlier generations. It has high specificity
  - A new combination assay has been developed and licensed in Europe (Monolisa HCV Ag/Ab ULTRA; Bio-Rad, Marnes la Coquette, France). Monolisa HCV Ag/Ab ULTRA is based on the combination of an indirect test for the antibodies and a sandwich test for Ag detection
- Molecular methods
  - Qualitative (Table 25.3)
    - The qualitative HCV RNA test is used to confirm HCV diagnosis following a positive or indeterminate antibody test result. It differentiates between resolved and active infection and may be useful for detecting acute infection prior to seroconversion. It is especially useful for confirming diagnosis in people with indeterminate HCV immunoblot (RIBA) results, as well as in immunosuppressed or immunoincompetent individuals
    - Recommended sensitivity for testing is 50 IU/ml
    - APTIMA® HCV RNA Qualitative Assay (Gen-Probe)
      - Target amplification based on sequences of the 5′ NC region of the HCV genome
      - Amplification of HCV RNA via transcription-mediated amplification method (TMA)
      - The detection limit of TMA is 10 IU/ml
    - Amplicor HCV test and COBAS Amplicor HCV test, v2.0 (Roche)
      - Use the primers KY78 and KY80 to amplify a 244-bp sequence of within the highly conserved 5′ UTR of the HCV genome
      - Limit of detection (200 ul): 25–50 IU/ml depending genotypes (i.e., 1b = 25, 1a = 50)

Table 25.3 Characteristics of current HCV RNA assays

Assay	Manufacturer	Technique	Lower limit of detection (qualitative assay)	Dynamic range of quantification (quantitative assay)
Amplicor® HCV v2.0	Roche Molecular Systems	Manual RT-PCR	50 IU/ml	NA
COBAS® Amplicor® HCV v2.0	Roche Molecular Systems	Semiautomated RT-PCR	50 IU/ml	NA
Versant® HCV RNA qualitative assay	Bayer HealthCare	Manual TMA	10 IU/ml	NA
Amplicor HCV Monitor® v2.0	Roche Molecular Systems	Manual RT-PCR	600 IU/ml	600–500,000 IU/ml
COBAS® Amplicor HCV Monitor v2.0	Roche Molecular Systems	Semiautomated RT-PCR	600 IU/ml	600–500,000 IU/ml
LCx HCV RNA Quantitative assay	Abbott Diagnostic	Semiautomated RT-PCR	25 IU/ml	25–2,630,000 IU/ml
Versant® HCV RNA 3.0 assay	Bayer HealthCare	Semiautomated bDNA	615 IU/ml	615–7,700,000 IU/ml
COBAS® TaqMan HCV Test	Roche Molecular Systems	Semiautomated real-time PCR	15 IU/ml	43–69,000,000 IU/ml
Abbott RealTime	Abbott Diagnostic	Semiautomated real-time PCR	30 IU/ml or 12 IU/ml[a]	12–100,000,000 IU/ml

Note: *RT* reverse transcriptase, *PCR* polymerase chain reaction, *TMA* transcription-mediated amplification, *bDNA* "branched DNA." *NA* not applicable
[a]For 0.2 ml or 0.5 ml of plasma analyzed, respectively

  - Use of centrifugation or Ultracolumn (QIAGEN) to process a large volume (1 ml), the LOD can be further improved
- Quantitative (Table 25.3)
  - On average 1–2 log10 units/ml less sensitive than qualitative tests
  - Used to establish baseline viral load (prior to therapy) and to monitor changes in viral load during therapy
  - PCR – Two real-time PCR platforms are currently available for the detection and quantification of HCV RNA: the COBAS TaqMan platform, which can be used together with automated sample preparation with the COBAS AmpliPrep system (CAP–CTM; Roche Molecular System, Pleasanton, CA), and the Abbott platform (Abbott Diagnostic, Chicago, IL), which uses the m2000RT amplification platform together with the m2000SP device for sample preparation (ART). The lower limits of detection were 12 IU/ml for ART and 15 IU/ml for CAP/CTM
  - VERSANT HCV RNA 3.0, quantiplex assay (bDNA) (Bayer)
    - Signal amplification directed to the 5′ NC region and core regions of the HCV genome
    - Microwell plate format
    - Equivalent detection of genotypes 1–6
    - LOD: 3200 HCV RNA copies/mL (5.2 HCV RNA copies/IU)
    - Broad dynamic range (615–7,690,000 IU/mL)
    - Comparative evaluations between Bayer bDNA and Roche PCR viral load assays demonstrated that PCR

reported significant lower viral load (as much as 1 log10) at the higher range
- HCV RNA detection and quantification are useful in clinical practice to diagnose chronic HCV infection, identify patients who need antiviral therapy, monitor the virological responses to antiviral therapy, and document treatment failure

– Genotyping
  - TRUGENE HCV 5′ NC Genotyping Kit (Bayer HealthCare LLC, Berkeley, California)
    – This technique utilizes PCR fragments previously generated by the diagnostic Roche Amplicor HCV test
    – Simultaneous PCR amplification and direct sequencing (CLIP sequencing) of the 5′ noncoding region (5′ NCR)
  - VERSANT HCV genotype assay (LiPA; Bayer HealthCare LLC)
    – The INNO-LiPA HCV II method uses 19 type-specific oligonucleotide probes attached to nitrocellulose strips to detect sequence variations found in the 5′ NC region of HCV
    – The biotin-labeled PCR product is hybridized to the probes on the strip under stringent conditions. After hybridization and washing, streptavidin-labeled alkaline phosphatase is added; followed by incubation with a chromogen, which results in the development of a purple-brown precipitate when there is a match between the probe and the biotinylated PCR product
    – Hybridization of the amplicon with one or more lines on the strip allows the classification of six major genotypes and their most common subtypes
  - Third Wave Technologies' Invader assay (Third Wave Technologies, Madison, Wisconsin) is a new DNA-scanning method application, which has been termed cleavase fragment length polymorphism (CFLP)
    – Relies on formation of unique secondary structure that results when DNA is allowed to cool following brief heat denaturation and serves as substrates for structure-specific cleavase I enzyme generating a set of cleavage products
    – Formation of secondary structures is sensitive to nucleotide sequences
    – The presence of sequence polymorphisms results in the generation of unique collections of cleavage products or structural fingerprints
    – It targets the well-conserved 5′ NCR of HCV
  - Determination of HCV genotype is needed before the initiation of therapy with pegylated IFNa and ribavirin because it determines both the dose of ribavirin and the treatment duration required

– HCV resistance testing
  - Several amino acid substitutions that confer resistance to directly acting antiviral molecules, such as protease inhibitors, have been identified. In case of a failure of the triple combination of pegylated IFNa, ribavirin, and either telaprevir or boceprevir, HCV variants that are resistant to these compounds are selected
  - Direct sequence analysis or reverse hybridization methods can be used to identify amino acid substitutions that confer resistance to antiviral drugs
  - Ultradeep sequencing methods, such as pyrosequencing, can detect minor resistant populations down to $<1\%$
  - HCV GenoSure NS3/4A represents the first in a series of HCV drug resistance assays that have been developed at Monogram Biosciences to support the clinical evaluation of HCV direct-acting antiviral (DAA) agents and their use in the management of HCV infection. HCV GenoSure NS3/4A analyzes the genetic sequence for the nonstructural proteins NS3 and NS4A of HCV

genotypes 1a and 1b that encode for an enzyme essential to viral replication. The assay detects mutations in NS3 and NS4A and specifically identifies those associated with boceprevir and telaprevir resistance

- Pitfalls
  - It is important to note that a "genotype bias" is possible for all HCV molecular assays because of the extensive genetic heterogeneity of the virus
  - False-positive results due to contamination (detected by negative control)
  - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of bacteria during specimen preparation
  - "Home brew" PCR assays are not standardized and variations in sample handling and laboratory methods can affect the sensitivity of the assay

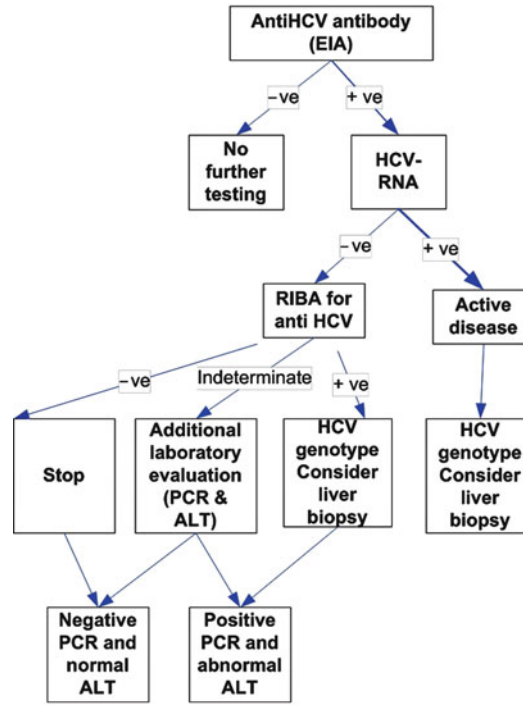


Fig. 25.12 Algorithm of HCV testing (Adapted from www.cdc.gov/hepatitis)

25.3.4 Clinical Utility

- HCV tests should be used in high-risk patients, such as intravenous drug users, children born to HCV-positive mothers, and HIV-positive patients. Figure 25.12 shows the algorithm of HCV testing
- Patients suspected of having chronic HCV infection should be tested for HCV antibodies. Patients suspected of having an acute infection should be tested for both HCV antibodies and also HCV RNA with a real-time PCR analysis. HCV RNA should be repeated as the patient is undergoing treatment, to better adjust the use of therapeutic agents
- HCV RNA testing should be performed in
  - Patients with a positive anti-HCV test
  - Patients for whom antiviral treatment is being considered, using a quantitative assay
  - Patients with unexplained liver disease whose anti-HCV test is negative and who are immune compromised or suspected of having acute HCV infection
- HCV genotype should be determined in all HCV-infected individuals prior to treatment in order to determine the duration of therapy and likelihood of response. Genotypic analysis has shown some amino acid substitutions to correlate with resistance to therapy. No commercial test is yet available, although home brew kits have been made
- The treatment of choice is peginterferon plus ribavirin

25.4 Hepatitis B Virus

25.4.1 General Characteristics

- Hepatitis B virus (HBV) is an enveloped dsDNA hepadnavirus. It is 47-nm spherical virus with three important polypeptides: HBV surface antigen (HBsAg) is an envelope protein, HBV core antigen (HBcAg) is a core protein, and HBV e antigen (HBeAg) is an early protein and a nonstructural protein coded by core gene. The envelope protein is involved in viral binding and released into

susceptible cells. The inner capsid relocates the DNA genome to the cell's nucleus where viral mRNAs are transcribed

- HBV is a circular, partially double-stranded DNA virus of approximately 3,200 nucleotides. This highly compact genome contains four open reading frames encoding the envelope (PreS1, PreS2, S), core (core, precore), polymerase, and X protein (Fig. 25.13)
- Although hepatocytes are most susceptible to infection, other cell types may be affected to a lesser extent. The life cycle of HBV begins when it attaches to the cell surface. In the cytoplasm, the DNA is still in the core but then capsid is removed and DNA passes into nucleus, where it forms a covalently closed circular DNA (cccDNA)
- HBV uses the host transcription machinery to replicate its genes and uses RNA polymerase II of the host. The (–) strand of the cccDNA will act as the template for this transcription. After transcription, the mRNAs are translated by the host's protein synthesis machinery to form viral proteins in the endoplasmic reticulum. The proteins are then assembled into virions that are secreted
- HBV is recognized as endemic in China and other parts of Asia. Over one-third of the world's population has been or is actively infected by HBV
- HBV strains are classified into at least 10 HBV genotypes (A to J) and several subtypes and based on the nucleotide homology of the surface gene. Except for the newly identified genotypes I and J, the geographic and ethnic distributions of HBV genotypes and subtypes are well characterized. Genotype A is mainly found in Northwestern Europe (subtype A2), North America, and Africa (subtype A1 or A3), whereas genotypes B and C have been described in Southeastern Asian populations. At present, genotype B is divided into B1–B6 subtypes. Among them, B1 is isolated in Japan, B2–5 are found in East Asia, and B6 is found in indigenous populations living in the Arctic, such as Alaska, Northern Canada, and Greenland. Genotype C, including subtypes C1–C5, mainly exists in East and Southeast Asia. Genotypes E and F are seen in East Africa and the New World, respectively. Genotype D is most often found in southern Europe, parts of Central Asia, India, Africa, and the Middle East. Genotype G is a recently determined genotype in France, America, and Germany while genotype H has been reported in patients from Central America. Recently, genotype I, a novel intergenotypic recombination among genotypes A, C, and G, was isolated in Vietnam and Laos. The newest HBV genotype, J, was identified in the Ryukyu islands in Japan, and this genotype has a close relationship with gibbon/orangutan genotypes and human genotype C
- Acute infection with genotypes A and D results in higher rates of chronicity than genotypes B and C. Compared to genotype A and B cases, patients with genotypes C and D have lower rates of spontaneous HBV e antigen (HBeAg) seroconversion; when this occurs, it tends to be delayed. Hepatitis genotypes C and D have lower response rates than genotypes A and B. Genotype C is also more associated with severe liver disease, including to hepatocellular carcinoma

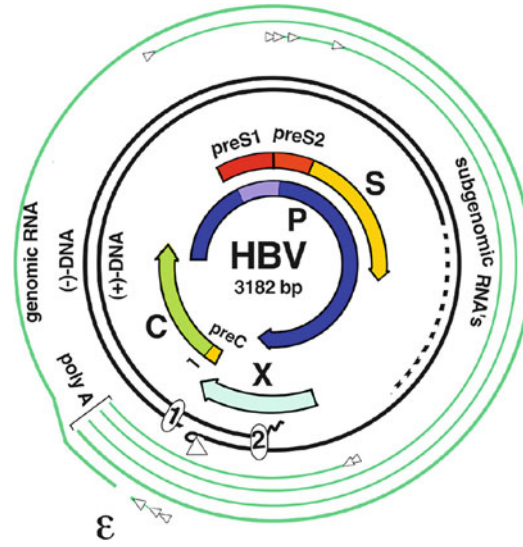


Fig. 25.13 HBV genome (Courtesy of Stephan Urban and Stefan Seitz, University of Heidelberg Dept. of Molecular Virology)

- The rate of new HBV infections has declined by approximately 82% since 1991, when a national strategy to eliminate HBV infection was implemented in the US. The decline has been greatest among children born since 1991, when routine vaccination of children was first recommended

25.4.2 Clinical Presentation

- Transmitted parenterally and sexually by contaminating open cuts or mucous membranes and has a long incubation period (45–120 days) (Fig. 25.14). Asian patients are more likely to be vertically infected (mother to child) than African or Western patients
- Majority of affected patients recover from the illness, characterized by
  - Anorexia, nausea, vomiting, headache, fever, abdominal pain, dark urine, and sometimes jaundice
  - Elevated transaminases, hyperbilirubinemia, and elevated alkaline phosphatase may also occur
  - Extrahepatic manifestations include arthralgias, arthritis, nephritis, and dermatitis
- 10% of patients continue to carry the virus or markers of the active viral infection greater than 6 months after initial infection
  - Small percentage may develop chronic persistent hepatitis with sequence fibrosis and cirrhosis
  - Incidence of HCC is increased with the viral genome found integrated in the cellular DNA in 75% of cases
  - May be associated with polyarteritis and cryoglobulinemia
- Recently, several clinical scoring systems, or nomograms, consisting of previously confirmed independent risk predictors such as sex, age, family history of HCC, alcohol consumption, serum alanine aminotransferase (ALT) level, HBeAg status, serum HBV DNA level, and/or HBV genotype have been introduced. These easy-to-use nomograms are based on noninvasive clinical characteristics and have been found to accurately predict HCC risk in either community- or hospital-based HBV-infected persons

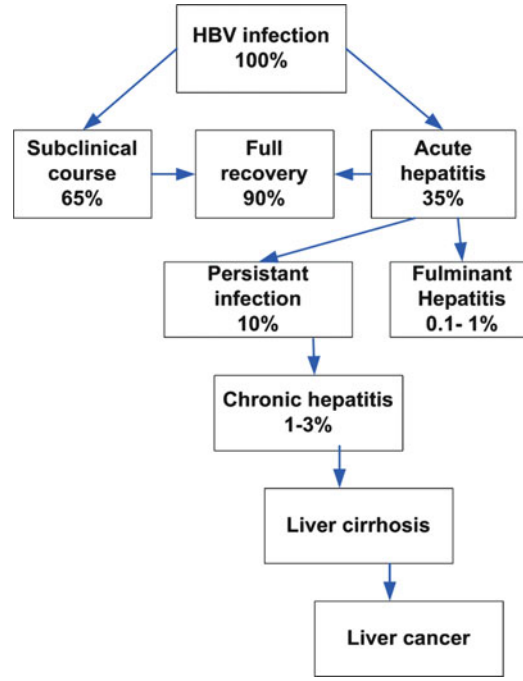


Fig. 25.14 Natural history of HBV

25.4.3 Diagnostic Methods

- Specimens: whole blood, serum, or plasma
- Conventional tests and problems
  - Serological studies
    - Viral antigens and particles (Fig. 25.15)
      - Dane particle
        - dsDNA bilayered sphere
        - 42 mm diameter; 22 nm core
        - Rarely identified in infectious serum
        - Thought to be infectious virus particle
      - HBsAg
        - Indicative of prior HBV exposure
        - Located on surface of Dane particle
        - Previously known as Australia antigen

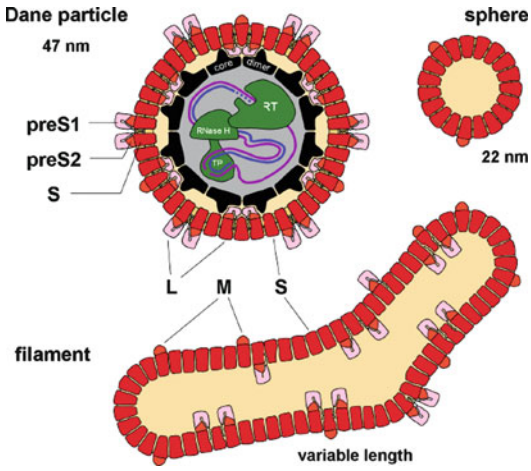


Fig. 25.15 HBV viral particle and antigens (Courtesy of Stephan Urban and Stefan Seitz University of Heidelberg Dept. of Molecular Virology)

  - HBcAg
    - Represents acute or chronic infection
    - 28 nm core of the Dane particle
  - HBeAg
    - Marker of HBV infection
    - Present in HBsAg-positive patients
    - Strong correlation with large serum concentrations of Dane particle and HbsAg
    - HBeAg is associated with high infectivity
- Antibodies (Fig. 25.16)
  - Anti-HBs
    - Antibody to surface antigen
    - Detected after disappearance of HbsAg
    - Protective properties
  - Anti-HBc
    - Antibody to core antigen
    - Detected after appearance of HbsAg
    - Used to confirm HBV infection when HBsAg and Anti-HBs are absent (window phase)
  - Anti-HBe
    - Antibody to HBeAg antigen protective properties
    - Associated with low risk of infectivity in presence of HBsAg
  - Interpretation of HBV serologic test results (Table 25.4)
- Molecular methods
  - Qualitative
    - COBAS AmpliScreen HIV1/HCV/HBV Tests (Roche Molecular Diagnostics)
      - It detects HBV DNA in human plasma
      - It is intended to be used to screen donors for HBV DNA
      - Detection limit is 100 copies/ml
      - It targets the S gene
  - Quantitative
    - Used to establish baseline viral load (prior to therapy) and to monitor changes in viral load during therapy
    - Digene HBV DNA hybrid capture II
      - Detection and quantitation of HBV DNA in serum
      - Limit of detection: 4,700 HBV DNA copies/ml
      - Quantitative range: 1.4×10^5 and 1.7×10^9 HBV copies/ml
    - PCR – Amplicor HBV Monitor and its semiautomated COBAS HBV Amplicor Monitor test (Roche)
      - Detection and quantitation of HBV DNA in serum or plasma
      - Use the primers HBV-104UB and HBV-104D to amplify a 104-bp sequence within the highly conserved precore/core region of the HBV genome
      - Amplify genotypes A to E equally and reduced amplification of genotypes F and G
      - Limit of detection: 200 copies/ml
      - Quantitative range: 10^3–4×10^7 copies/mL
    - Real-time PCR LightCycler/FRET hybridization probes
      - It targets 259-bp fragment of S gene,
      - Quantitative range: 250–5×10^8 copies/ml

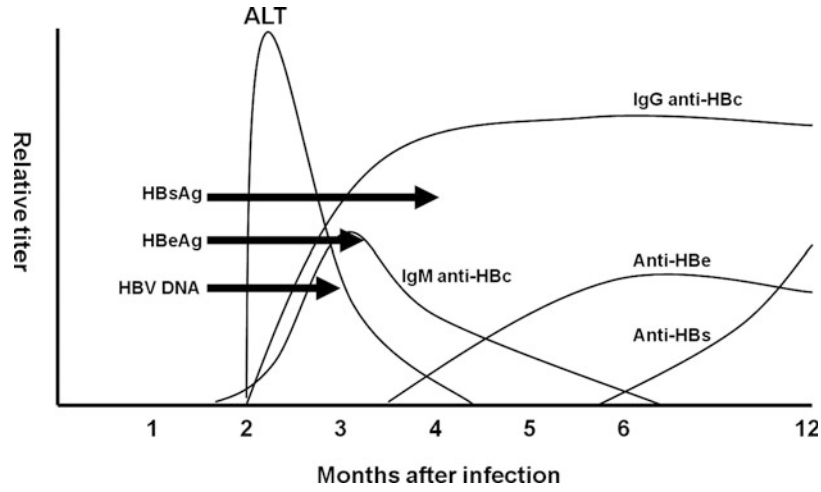


Fig. 25.16 Time course for appearance of viral antigens and antibodies in acute hepatitis B infection

Table 25.4 Hepatitis B antibody/antigen interpretation

Tests	Results	Interpretation
HBsAg	Negative	Susceptible
anti-HBc	Negative	
anti-HBs	Negative	
HBsAg	Negative	Immune due to natural infection
anti-HBc	Positive	
anti-HBs	Positive	
HBsAg	Negative	
anti-HBc	Negative	Immune due to vaccination
anti-HBs	Positive	
HBsAg	Positive	
anti-HBc	Positive	Acutely infected
IgM- Anti HBc	Positive	
anti-HBs	Negative	
HBsAg	Positive	
anti-HBc	Positive	Chronically Infected
IgM- Anti HBc	Negative	
anti-HBs	Negative	
HBsAg	Negative	*Results inconclusive:*
anti-HBc	Positive	Resolved infection (most common)
anti-HBs	Negative	False-positive anti-HBc "Low-level" chronic infection Resolving acute infection

- Real-time PCR – Roche TaqMan Assay
  - Utilizes FRET technology and probes based on the detection of amplicon during temperature cycling
  - It targets S gene
  - Limit of detection: 50 copies/ml
  - Quantitative range: 5 to 2 × 10^8 HBV IU/mL (30–10^7 copies/ml; 1 IU = 5.82 copies)
- bDNA assay – (VERSANT Hepatitis B Virus DNA 3.0 Assay) (Bayer Corporation)
  - Signal amplification directed to the 5′ NC region and core regions of the HCV genome
  - Microwell plate format
  - Limit of detection: 2,000 copies/ml
  - Quantitative range: 2.0 × 10^3 to 1.0 × 10^8 HBV DNA copies/ml
  - Equivalent detection of genotypes A through F

– Genotyping and mutation analysis – currently used mainly for epidemiological purposes, rarely needed for clinical purposes
- Line probe assay-LiPA; INNO-LiPA HBV Genotyping assay, Innogenetics N.V., Ghent, Belgium
  - This method is based on the reverse hybridization principle, such that biotinylated amplicons hybridize to specific oligonucleotide probes that are immobilized as parallel lines on membrane-based strips. The amplified region analyzed overlaps the sequence encoding the major hydrophilic region of HbsAg

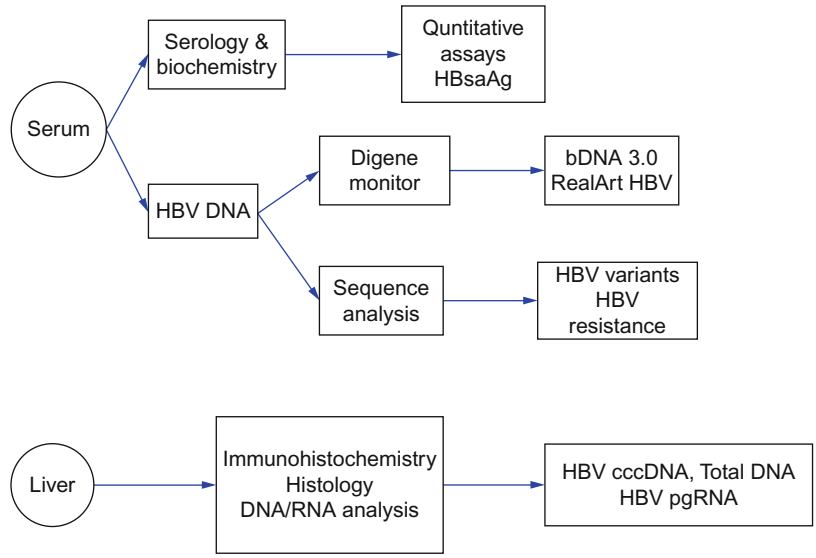


Fig. 25.17 An outline of the HBV assays available for testing of serum and liver biopsy samples

  - TRUGENE® HBV Genotyping Kit (Bayer Corporation)
    - Sequencing and phylogenetic analysis of the pre-S1/pre-S2 region of the HBV genome
    - Identifies HBV genotype, drug resistance mutations, and anti-HBs escape mutations based on comparison of DNA sequence
- Pitfalls
  - The analytical sensitivity and specificity of current real-time PCR assays allow for accurate quantification over a range of approximately 7–8 logs. They are not sufficient to quantify the very high HBV DNA levels that can be found in certain HBV-infected patients, which necessitates retesting these samples after dilution, a factor of quantification errors
  - Equal quantification of all HBV genotypes and robustness of quantification in case of nucleotide polymorphisms has not been validated for the current commercial real-time PCR assays
  - There is currently no uniform tendency to report HBV DNA levels in standardized units (such as copies/ml or genome equivalents/ml or IU/ml)
  - Not all assays are currently registered for use with plasma and serum
  - No precise thresholds of HBV DNA have been established that could guide medical decisions

25.4.4 Clinical Utility

- Viral load testing is used for the assessment and monitoring of responses to therapy in HBV infection (Fig. 25.17)
- In HBV carriers with active liver disease, HBV DNA loads are measured not only to assess patients regarding the need for either interferon alpha or lamivudine (a DNA polymerase inhibitor) antiviral therapy but also to monitor their effectiveness
- An increase in HBV viral load is also used as a marker of the emergence of lamivudine-resistant viral mutants
- Active chronic infections with HBV treated with lamivudine require surveillance for the emergence of lamivudine-resistant viral mutants. During lamivudine monotherapy, point mutations at the active site of the polymerase gene (YMDD variants) occur with a frequency of 14–32% after 1 year in phase III studies, and in 42% and 52% of Asian patients after 2 and 3 years of therapy, respectively. The emergence of lamivudine resistance is detected by a rise in HBV viral load and confirmed by sequencing of the active site of the DNA polymerase gene

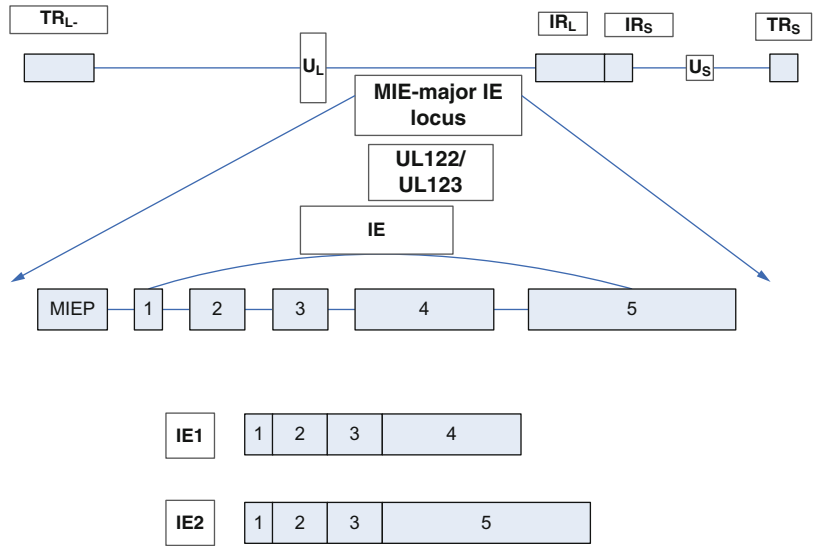


Fig. 25.18 CMV genome. *MIEP* major IE promoter, U_L unique long region, U_S unique short region, *TR* terminal repeat sequence, *IR* inverted repeat sequence

- The presence of HBV precore mutants may cause active liver disease despite the absence of HBeAg, the common marker for active hepatitis in HBV infection. This may be due to either a premature stop codon point mutation in the precore gene (G1896A) or a mutation in the basal core promoter region downregulating HBeAg production, both of which can only be reliably detected genotypically

25.5 Cytomegalovirus

25.5.1 General Characteristics

- Cytomegalovirus (CMV) is a viral genus of the viral group known as *Herpesviridae* or herpesviruses, sharing a number of characteristics with chicken pox, the various herpes viruses, and the Epstein–Barr virus (EBV). The species that infects humans is commonly known as human CMV (HCMV) or human herpesvirus (HHV) 5. It is the largest of the HHVs characterized by 240 kb double-stranded linear DNA virus (Fig. 25.18) with 162 hexagonal protein capsomeres surrounded by three distinct layers: a matrix or tegument, a capsid, and an outer envelope
- CMV can reside latent in the salivary gland cells, endothelium, macrophages, and lymphocytes. CMV infection is asymptomatic
- The virus acts by blocking cell apoptosis via the mitochondrial pathway and causing massive cell enlargement, which is the source of the virus name
- Clinically symptomatic patients are infants and immunocompromised adults. For infants, the mode of transmission is from the mother via the placenta, during delivery or during breast-feeding
- For adults, CMV transmission occurs from close contact with individuals excreting virus in saliva, urine, and other bodily fluids. Transmission of CMV has been reported from blood transfusion and organ transplant
- By the age of 30, approximately 40% of individuals are infected by CMV; by the age of 60, 80–100% of the population has been exposed to the virus
- Human CMV is difficult to culture as it slowly grows only in human fibroblasts since CMV is very specific in term of species and cell type to be infected

25.5.2 Clinical Presentation

- CMV elicits both humoral and cellular immune responses. CMV presents as primary, latent, reactivated, and reinfection

- CMV is transmitted by direct person-to-person contact with a person excreting the virus in their saliva, urine, or other body fluids; contact with infectious body fluids (e.g., urine on a diaper); or contact with fomites (e.g., a child's toy that has infectious virus on it). CMV can be sexually transmitted and can also be transmitted via transplanted organs and blood transfusions. Infectious CMV may be shed in the bodily fluids of any previously infected person, and thus may be found in urine, saliva, blood, tears, semen, and breast milk. The shedding of virus may take place intermittently, without any detectable signs
- The incidence of primary CMV infection in pregnant women in the US varies from 1% to 3%. Healthy pregnant women are not at special risk for disease from CMV infection. When infected with CMV, most people have no symptoms and very few have a disease resembling mononucleosis. But can be severe and life-threatening in immunocompromised patients including organ recipients and AIDS patients. It is their developing unborn babies that may be at risk for congenital CMV disease and occurs in about 1% of all newborns. 10–20% of these infants will develop complications before school age. CMV remains the most important cause of congenital viral infection in the US
- In infants and young children, typical features of the infection include hepatosplenomegaly, extramedullary cutaneous erythropoiesis and thrombocytopenia, and petechial hemorrhages. Encephalitis often leads to severe mental and motor retardation
- For immunocompromised patients, CMV disease is an aggressive condition. CMV hepatitis can cause fulminant liver failure. CMV infection can also cause CMV retinitis and CMV colitis

25.5.3 Diagnostic Methods

- Specimens
  - Whole blood, urine, CSF, amniotic fluid, bone marrow, biopsies
- Conventional tests (Table 25.5)
  - CMV cytology and histology tests
    - General
      - The CMV cytologic and histologic examination detects within cells evidence of CMV in urine or other body fluids or tissues. It also called CMV inclusion body detection. It is used to evaluate and manage CMV infection
      - The CMV cytologic and histological diagnosis is often considered the "gold standard" for diagnosing end-organ disease. Conventional H&E stains reveal enlarged (cytomegalic) cells that are often two- to fourfold larger than surrounding cells, usually with large eosinophilic intranuclear inclusions, sometimes surrounded by a clear halo, and smaller cytoplasmic inclusions. Its sensitivity shows ranging widely from 10% to 87%. It is important to note that 37.5% of patients with gastrointestinal CMV disease fail to demonstrate any inclusions
      - Immunohistochemistry with monoclonal antibodies directed against CMV immediate early antigen increases diagnostic yield of CMV compared to routine H&E staining. Sensitivity of immunohistochemistry for detecting CMV infection can approximate 93%
    - Advantages
      - Specific and definite diagnosis
      - Confirm end-organ disease along with virus infection diagnosis
    - Pitfalls
      - Invasive procedure required
      - Insensitive
  - Viral culture
    - Conventional culture
      - Viral culture is the traditional method of virus detection for demonstrating viremia and sometimes for establishing end-organ disease. Conventional CMV culture involves isolation of the virus in fibroblast tissue culture where it produces a distinctive

Table 25.5 CMV diagnostic tests

Assay	Specimen	Diagnostic utility and significance
CMV IgG antibody	Serum	Indication: to determine past exposure status. Diagnose primary infection Positive: indicates CMV infection in the past Negative: indicates absence of previous CMV infection Caveat: A negative result does not rule out recent CMV infection
CMV IgM antibody	Serum	Indication: to determine past exposure status. Diagnose primary infection Positive: indicates current primary CMV infection or possible reactivation of latent infection Negative: indicates absence of acute infection
CMV antigenemia	Whole blood	Indication: to diagnose and monitor CMV disease in immunocompromised patients Positive: indicates the presence of CMV viremia. May indicate impending CMV episode in absence of current disease Negative: CMV viremia not detected
Centrifuge-enhanced CMV Culture (shell vial)	BAL Bronch Gastric biopsy	Indication: to detect the presence of CMV infection Positive: indicates the presence of CMV infection Negative: does NOT rule out CMV disease, due to the low sensitivity of this test
Centrifuge-enhanced CMV Blood culture (shell vial)	Whole blood in heparin or EDTA	Indication: to diagnose and monitor CMV disease in immunocompromised patients Positive: indicates the presence of CMV viremia, but does not necessarily indicate clinical disease. May indicate impending CMV episode in the absence of current disease Negative: does NOT rule out CMV disease, due to the low sensitivity of this test
CMV DNA by hybrid capture	Whole blood in EDTA Refrigerate up to 48 h.	Indication: to diagnose and monitor CMV disease in immunocompromised patients Positive: indicates the presence of CMV DNA Negative: CMV DNA not detected
CMV DNA by PCR amplification	Serum/plasma Peripheral blood	Indication: to diagnose and monitor CMV disease in immunocompromised patients Positive: indicates the presence of CMV DNA Negative: CMV DNA not detected
CMV mRNA by Nucleic acid sequence-based Amplification (NASBA)	Whole blood in heparin or EDTA	Indication: to diagnose and monitor CMV disease in immunocompromised patients Positive: indicates the presence of CMV mRNA Negative: CMV mRNA not detected

cytopathogenic effect that is easily confirmed by fluorescent antibody stains

– Presence of the virus (positive cultures) can often be determined in as little as 1–2 days, but cultures that are negative for the virus must be held for 3 weeks to confirm the absence of CMV because the virus may be present in very low numbers in the original sample and/or the CMV strain may be slow-growing
– CMV culture can be performed on blood, tissue, urine, saliva, or respiratory swabs. Blood culture has a sensitivity of 45–78% but a very high specificity, approaching 89–100% for detecting disease. CMV detection in

blood is more strongly associated with disease than detection in urine or saliva, with a positive predictive value of 60%
  - Advantages
    - Gold standard test for CMV detection
    - Able to recover other viruses from the same specimen
  - Pitfalls
    - Low sensitivity compared to newer techniques, such as antigenemia and PCR and nucleic acid probe
    - Lack of virus quantitation
    - A long incubation period (1–3 weeks), lack of virus quantitation
    - False-negative results if cell culture inoculation is delayed
- Shell vial assay
  - Another CMV culture (shell vial culture) method is rapid viral culture. The shell vial culture with immunofluorescence staining is used for the early diagnosis of CMV infection
  - In immunocompromised patients, a sensitivity of 78% and a specificity of 100% have been claimed
  - Specimens are centrifuged into fibroblast monolayers, incubated briefly, and then stained with a fluorescent monoclonal antibody to early antigens, before the appearance of cytopathogenic effect. Results are available much earlier than with conventional CMV culture, usually within 24–48 h
  - The cells are read under a fluorescent microscope
  - Advantages
    - Higher sensitivity than conventional methods
    - Quantitative shell vial cultures are available, but not widely used
  - Pitfalls
    - May need large amount of biomass for virus recovery
- Serological studies
  - CMV antibody testing
    - Antibody testing can be used to determine if someone has had recent or past exposure. There are two types of CMV antibodies that are produced in response to a CMV infection, IgM and IgG, and one or both may be detected in the blood
    - IgM antibodies are the first to be produced by the body in response to a CMV infection. CMV IgM antibodies are detected within a week or two after primary infection and lasts 3–4 months. After several months, the level of CMV IgM antibody usually falls below detectable levels
    - It is not detectable in recurrent infection except in immunocompromised patients where it is detectable in about a third of the cases
    - CMV IgG antibody is produced early in primary infection and provides protection from primary infections. Levels of IgG rise during the active infection then stabilize as the CMV infection resolves and the virus becomes inactive. After a person has been exposed to CMV, he or she will have some measurable amount of CMV IgG antibody in their blood for the rest of their life
    - CMV IgG avidity test to distinguish primary CMV infection from past or recurrent infection (reactivation or reinfection). CMV IgG avidity is low (<30%) in primary infection
    - Prenatal diagnosis of congenital CMV infection is performed only in the case of primary maternal infection as transplacental transmission of CMV is higher in 40% of primary maternal CMV infection. Whereas it is low in the case of recurrent infection 1–4%
    - CMV IgG antibody testing can be used, along with IgM testing, to help confirm the presence of a recent or previous CMV infection
  - CMV antigenemia test
    - This test is based upon the detection of pp65, a structural protein expressed

on the surface of infected polymorph nuclear lymphocytes
  - The specimen must be processed within 6–8 h of collection, but results can be available within 8–24 h
  - The number of infected leukocytes present had been reported to correlate with the severity of infection
  - This test is usually applied to blood and cerebrospinal fluid, with a sensitivity of 60–100% and a specificity of 83–100%
  - Commercial kit
    - The CMV Brite™ Turbo Antigenemia Kit uses the well-defined C10/C11 antibody cocktail to detect the CMV lower matrix phosphoprotein (pp65), an early antigen in virus replication, which is abundantly present in antigen-positive polymorphonuclear cells
    - The CMV Brite Turbo Kit is a rapid new version of the first FDA-registered immunofluorescence antigenemia kit for in vitro CMV diagnosis
  - Advantages
    - Inexpensive kits are commercially available
    - May be able to detect CMV before development of symptoms
  - Pitfalls
    - Labor-intensive
    - Require skilled personnel
    - Subjective interpretation
    - Although an improvement over CMV viral culture, this technique is only semiquantitative, and the reading of results is somewhat subjective
    - Poor sensitivity in urine samples. The assay is adversely affected by low leukocyte counts
- Molecular methods (Table 25.5)
  - General
    - CMV DNA tests can be qualitative or quantitative. The two major techniques used are hybrid capture and PCR. PCR appears to be more sensitive than hybrid capture
    - Different tissues and body fluid compartments can be used for PCR including whole blood, plasma, leukocytes, buffy coat specimens, bronchoalveolar lavage (BAL) fluid, target organ tissue, or stool. Whole blood PCR testing is more sensitive than plasma PCR testing
    - Higher CMV viral loads seem to correlate with symptomatic disease
    - Stool CMV DNA testing is a technique that is noninvasive and would detect infection even when located in the right colon, out of the reach of a sigmoidoscopic examination. Stool CMV DNA also is possibly more sensitive than blood CMV DNA for detecting colonic disease, as it is more organ specific
    - CMV DNA testing in blood is a very promising technique for diagnosing CMV disease. Its advantages include quick results (6–48 h) and high sensitivity
  - Qualitative
    - CMV DNA by hybrid capture assay (Digene): FDA cleared
      - The Digene Hybrid Capture CMV test is a molecular assay that detects the presence of CMV DNA in white blood cells from blood collected in EDTA. It is more sensitive than CMV culture and has a sensitivity and specificity comparable to the CMV antigenemia assay for the detection of CMV viremia
      - Unlabeled CMV probes hybridized with viral DNA, then immobilized on a solid phase before being measured by conjugated antihybrid antibody
    - CMV mRNA by NASBA
      - Assay of CMV mRNA by nucleic acid sequence-based amplification (NASBA) has been investigated for the detection of active viral gene expression and replication
      - CMV-specific mRNA may be a more specific marker for CMV than culture or DNA PCR since it would

theoretically be indicative of CMV replication
  - Since late transcripts such as pp67 mRNA reflect a complete replication cycle of CMV, they could be indicative of disease
  - NucliSens CMV pp67 Assay (Organon Teknika Inc., Durham, NC) is an FDA-cleared test. NucliSens CMV pp67 measures replication of CMV in blood using NASBA RNA amplification technology
  - This assay detects messenger RNAs coding for the matrix tegument protein pp67 of CMV, a true late protein, which is only expressed during viral replication
  - The NASBA technology selectively amplifies RNA in a DNA background and allows direct testing in whole blood
  - It is a direct route for diagnosing an active CMV infection and monitoring treatment efficacy
- Quantitative
  - PCR
    - Amplicor CMV MONITOR test (Roche Molecular Systems) is a quantitative microtiter-based PCR assay
      - CMV viral DNA in the specimen was quantitated by coamplifying a region of the CMV DNA polymerase gene in the presence of a known quantity of quantitative standard
      - The primers used were specific for the CMV polymerase gene and amplified a 362-bp fragment of the gene
      - An internal quantitation standard (QS) that is added at a known concentration during specimen processing so that extraction and recovery of DNA, in addition to amplification and detection, can be monitored
      - The lower limit of sensitivity of the assay is 400 copies/ml of plasma. The linear range of the assay is 400–400,000 copies of CMV DNA per mL
      - The inherent sensitivity of molecular detection of CMV poses a problem since latent CMV genomes, present in most seropositive individuals, may be detected. Therefore, it is critical to adjust the sensitivity of the PCR so that latent genomes are not detected
    - Advantages
      - Sensitive enough to detect virus before symptom development
      - Specimen could be stored and transported
      - Rapid and less expensive
    - Pitfalls
      - False positive
      - Contamination must be prevented
  - Real-time PCR
    - The COBAS® AmpliPrep/COBAS® TaqMan® CMV Test is a real-time PCR-based quantitative assay for CMV DNA detection in human plasma using the COBAS® AmpliPrep Instrument for automated specimen processing and the COBAS® TaqMan® Analyzer or COBAS® TaqMan® 48 Analyzer for automated amplification and detection. The test can quantitate CMV DNA over the range of 150–10,000,000 copies/mL
    - Various home brew methods were developed either using real-time TaqMan PCR (ABI prism 7700) or real-time LightCycler PCR
    - Advantages
      - Quick result: turnaround time for the real-time PCR assay is 6–48 h
      - High sensitivity
      - CMV DNA testing also has greater than 80% concordance with CMV antigen test results, but also has the added benefits of increased specimen stability, smaller required

specimen volume, and with the ability to be performed in patients with depressed white blood cell counts results
  - Pitfalls
    - False-positive results due to contamination (detected by negative control)
    - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of bacteria during specimen preparation
    - There is a lack of standardization of this process, with different techniques and assays, different quantitation methods, and different tissues and blood compartments being tested, thus making interpretation of results across studies challenging. Quantitative PCR is more sensitive than qualitative PCR, and CMV DNA values obtained with "in-house" quantitative assays are 3- to 10-fold higher than the commercial assay
    - Different quantitation methods include reporting results as genomic copies/mL, copies/mL, copies/microgram of total DNA, copies/106 leukocytes, and copies/2 $\times$ 10^5 leukocytes; such varied quantitation methods can make comparison of results almost impossible
    - No clearly defined cutoff values for determining CMV disease, this method has low specificity and low positive predictive value. Most of CMV cutoff on plasma ranging from 400 to 10,000 copies/mL of CMV DNA
- Antiviral susceptibility testing of CMV isolates
  - General
    - Resistance of CMV to antivirals was a major clinical problem in patients with AIDS. Currently marketed anti-CMV drugs, namely, ganciclovir (GCV), its oral prodrug valganciclovir (vGCV), foscarnet (FOS), and cidofovir (CDV), all target the viral DNA polymerase
    - CMV infection remains a major problem in transplantation, and resistance to antivirals is encountered in all forms of transplantation
    - In general, it takes weeks to months for CMV to develop resistance to antivirals. In patients with AIDS, some studies showed a 10% prevalence of resistance to ganciclovir by 3 months of therapy; similar time courses were found for foscarnet and cidofovir
  - Phenotypic methods
    - Plaque reduction assay
      - The gold standard for antiviral susceptibility testing of CMV is plaque reduction assay
      - In this assay, a standardized inoculum of a stock virus is inoculated into cultures and incubated in the presence of the antiviral agent
      - The cultures are then observed for the presence of viral plaques
      - The IC50 of the agent for the isolate is defined as the concentration of agent causing a 50% reduction in the number of plaques produced
      - Plaque reduction assays are labor-intensive
      - Plaque reduction assays are limited by the excessive time required completing the assay (4–6 weeks) and the lack of a standardized method validated across different laboratories
      - In addition, repeated passage of isolates to prepare viral stocks may influence the results of assays by selecting CMV strains that are not representative of the original population of the viruses
    - DNA hybridization assay
      - Whole genomic DNA is extracted and transferred by capillary action onto

negatively charged nylon membranes after incubation with a specific agent
  - The membranes are hybridized to a 125I-labeled human CMV probe (Diagnostic Hybrids, Athens, Ohio), rinsed, washed, and counted in a gamma counter
  - Mean hybridization values (in counts per minute [cpm]) for each concentration of antiviral agent are calculated and expressed as a percentage of the cpm in control cultures
  - The IC50 is defined as the concentration of antiviral agent resulting in a 50% reduction in viral nucleic acid hybridization values (i.e., DNA synthesis) compared with the hybridization values of controls
  - Disadvantage of DNA hybridization assays is that they require the use of radiolabeled probes
  - DNA hybridization assays have the advantage over plaque reduction assays of eliminating the variation due to subjective errors resulting from plaque counting by different individuals
- Viral load assays (e.g., antigenemia or quantitative DNA)
  - CMV viral load assay may rise as an indicator of antiviral resistance, but other factors (including compliance and declining immune function) may be responsible
  - The assay can measure viral DNA concentration exposed to a range of drug concentrations and is grown for 4 days in the presence or absence of drug then the IC50 is determined
  - Quantitative antigenemia assays are less exact than quantitative polymerase chain reaction
  - In many patients, certain CMV diseases (e.g., gastrointestinal disease or retinitis) are not always associated with measurable viral loads
- Other phenotypic methods: Viral production is measured by using immunofluorescence-, immunoperoxidase-, ELISA-, or flow cytometry-based methods for detection and quantitation of cells expressing CMV antigens (immediate–early, early, or late)

- Genotypic methods
  - General
    - Viral UL97 kinase and UL54 DNA polymerase gene mutations are well-documented mechanisms of resistance to current antivirals
    - The mutation of the viral phosphotransferase gene (UL97) coding sequence, which may confer resistance only to ganciclovir (GCV). One of seven canonical UL97 mutations (M460V/I, H520Q, C592G, A594V, L595S, and C603W) is found in over 80% of GCV-resistant clinical CMV strains
    - Although UL97 mutations do not affect susceptibility to FOS or CDV, UL54 mutations can confer resistance to all current drugs and may emerge after prolonged GCV therapy to increase the level of resistance conferred by a preexisting UL97 mutation. Mutations in UL54 are often accompanied by mutations in UL97, showing higher levels of resistance to ganciclovir with possible cross-resistance to foscarnet and/or cidofovir
    - The viral polymerase gene UL54 occurs in regions between codons 300 and 1,000 (Fig. 25.19)
  - Genotypes for CMV antiviral resistance
    - Detection of mutations is based on PCR amplification of the specific region of the genome followed by restriction enzyme analysis or direct sequencing of the amplification product
    - Standard dideoxy sequencing can detect an emerging resistance mutation when it exceeds approximately 20% of the sequence population

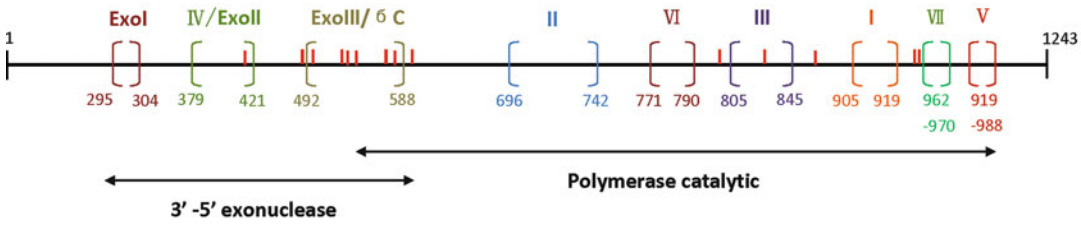


Fig. 25.19 Map of the CMV UL54 DNA polymerase gene showing conserved functional domains, ranges of codons containing drug resistance mutations, and the locations of drug resistance mutations and nonresistant sequence variants newly described (Adapted from Hakki et al. Curr Opin Infect Dis 2011)

- Pyrosequencing has been reported to detect mutant subpopulations at an approximately 6% level at codons 460, 520, and 592–607 of UL97, but the short sequence reads are unsuitable for analyzing the full range of codons needed for genotypic diagnosis (e.g., UL97 codons 335–670 and UL54 codons 300–1,000)
- An additional method "real-time PCR" with melting curve assay can also detect drug resistance mutations. The advantages are that low copy numbers without cell culture can be detected, mixed virus populations can be analyzed semi-quantitatively, and multiplex reaction to different mutations can be detected simultaneously. The disadvantages are that polymorphisms near known mutations may affect the melting curve and give false-positive result and different probes are needed to identify each mutated codon. Besides, the assay may not be able to differentiate different point mutations that occur in the same codon, such as M460I (ATT/ATA) and M460V (GTG)

25.5.4 Clinical Utility

- Quantitative PCR determination of CMV viral load in solid organ transplant recipients can predict CMV disease and relapse (Table 25.5) as well as for initiating antiviral therapy
- Viral load testing in patients with HIV infection is currently used to predict CMV disease and to monitor the efficacy of treatment
- Pitfalls
  - Using restriction enzyme analysis, not all of the presently confirmed resistance mutations are accompanied by alteration of known restriction enzyme recognition sites which lead to false-negative results, and base changes not associated with drug resistance can produce new restriction sites which lead to false-positive results
  - PCR assays are not standardized, and variations in sample handling and laboratory methods can affect the sensitivity of the assay
  - Well-defined CMV DNA standards are needed to avoid variation of viral load values obtained with commercial and home brew assays
- Clinical utility: The standardization of automated sequencing methods and the characterization of mutations associated with drug resistance will offer routinely genotypic resistance testing in a time frame that impacts clinical care

25.6 Epstein–Barr Virus

25.6.1 General Characteristics

- EBV is a dsDNA virus of the herpes family (HHV type 4) characterized by icosahedral capsid and a glycoprotein-containing envelope. It is one of most common viruses in human

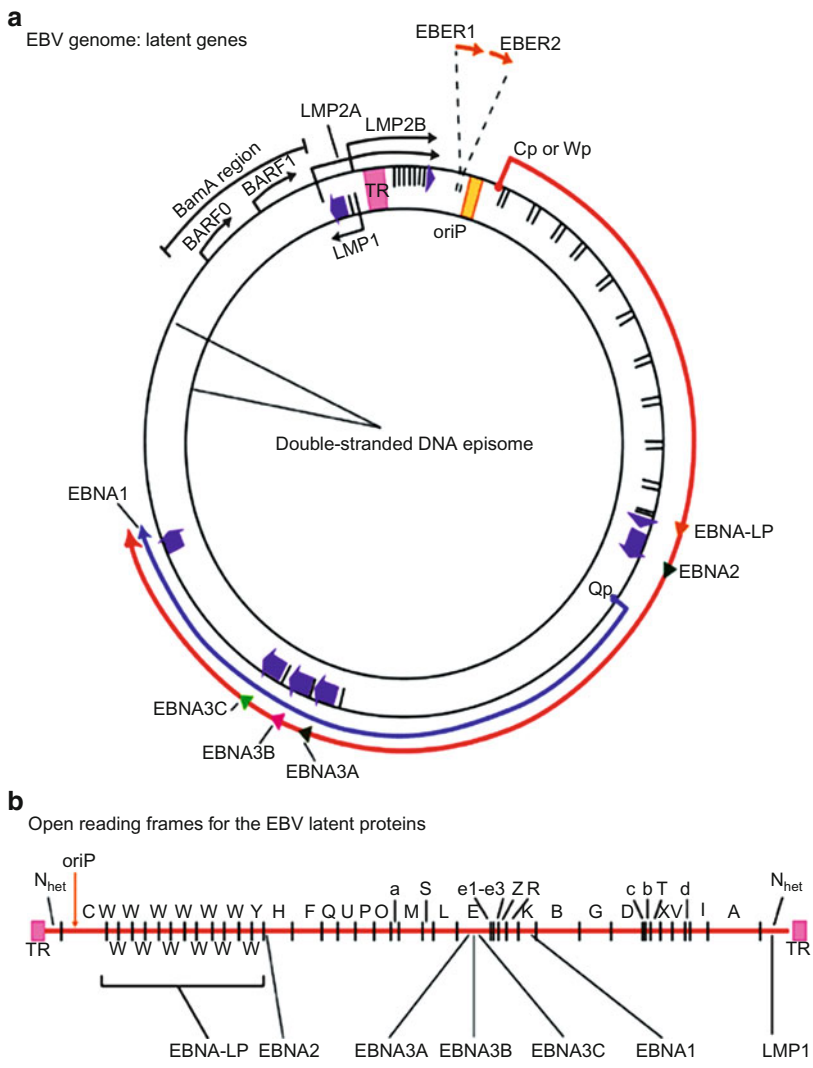


Fig. 25.20 Epstein–Barr virus genome structure (Adapted from Murray et al. 2001)

- The genome is a linear dsDNA molecule with 172 kbp. The viral genome does not normally integrate into the cellular DNA but forms circular episomes which reside in the nucleus
- Most common mode of transmission of EBV is through exposure to infected saliva from asymptomatic individuals. Virus is relatively fragile and does not survive long outside the human host fluids. Primary infection of the virus occurs by oral transmission from asymptomatic individuals. EBV preferentially infects B lymphocytes and remains latent, affecting more than 95% of population
- The genome is large enough to code for 100–200 proteins but only a few have been identified. The proteins characterized thus far fall into the following group: (1) Latent proteins, including EB viral nuclear antigen complex (EBNA), latent membrane protein (LMP), terminal protein, and lymphocyte-detected membrane antigen (LYDMA); and (2) lytic cycle proteins, including membrane antigen (MA), early antigen complex (EA), and viral capsid antigen complex (VCA). Critical viral target genes: EBNA1, LMP1, and LMP2 (Fig. 25.20)

25.6.2 Clinical Presentation

- EBV causes infectious mononucleosis, an acute but self-limiting disease affecting children and young adults. After primary infection, the virus persists indefinitely in B lymphocytes, only to reactivate when cellular immunity is impaired
- In infected young children remain asymptomatic or developed nonspecific viral illness
- EBV causes infectious mononucleosis, associated with PTLD, linked to several malignancies such as nasopharyngeal carcinoma (NPC) among Chinese males and Burkitt lymphoma (BL) in children of Central Africa/New Guinea. In HIV-infected individuals, EBV is associated with diseases such as oral hairy leukoplakia and AIDS-related non-Hodgkin lymphoma; EBV might imply a higher risk of some autoimmune diseases such as systemic lupus erythematosus (SLE) and multiple sclerosis
- EBV is a ubiquitous virus which causes persistent, latent infection that can be reactivated. More than 90% of the adult population is estimated to demonstrate serologic evidence of prior exposure with EBV
- Primary infection in young children is often asymptomatic or causes nonspecific minor illness
- For adolescents and young adults, primary infection is typically manifested as infectious mononucleosis (IM), usually a self-limiting condition characterized by fever, sore throat, myalgias, lymphadenopathy, and hepatosplenomegaly
- A strong association between EBV and Burkitt lymphoma in children of Central Africa/New Guinea and nasopharyngeal carcinoma among Chinese males
- In HIV-infected individuals, EBV is associated with diseases such as oral hairy leukoplakia and AIDS-related non-Hodgkin lymphoma, i.e., oral hairy leukoplakia and CNS lymphoma
- Patients undergoing transplantation are prone to develop posttransplant lymphoproliferative disease (PTLD)
- Although the correlation between EBV burden and disease status is incompletely understood, several studies have shown an association between symptomatic infection and elevated DNA loads in clinical samples. Increasing virus burden is also believed to be a rapid indicator of immunopathological changes preceding and/or underlying the B lymphocyte-driven changes caused by EBV. Therefore, determining EBV DNA loads in EBV-related disorders in immunocompromised populations is an important step toward disease diagnosis, management, and treatment

25.6.3 Diagnostic Methods

- Specimens
  - Whole blood, plasma, CSF, biopsy
- Conventional tests and problems (Table 25.6)
  - EBV antibodies: EBV antibodies are used to help diagnose Mono if people are symptomatic but have a negative Mono test. EBV antibodies include (1) viral capsid antigen (VCA)-IgM, VCA-IgG, and D early antigen (EA-D) – to detect a current or recent infection; and (2) VCA-IgG and Epstein–Barr nuclear antigen (EBNA) – to detect a previous infection
    - Heterophile IgM antibody
      - Present in 90% of adults during the course of illness
      - Nonspecific serologic response to EBV infection
    - Classic Paul–Bunnell test
      - Measures agglutination of sheep RBCs by patient serum; limited by false-positive agglutinins in sera of normal individuals (Forssman agglutinins), patients with serum sickness, etc
    - Monospot test: detects agglutinins to formalized horse RBCs not removed by prior absorption with guinea pig kidney
    - Viral capsid antigen antibody (Fig. 25.21)
      - IgM – indicates recent infection, lasts only 4–8 weeks

Table 25.6 EBV diagnostic tests

EBV test	Diagnostic utility and significance	Comments
Infectious mononucleosis slide test	Initial testing to confirm infectious mononucleosis or recent EBV infection (Monospot test) Negative Monospot test is common in children and immunocompromised adults	If test results are negative but a strong clinical suspicion exists, repeat testing in 7–14 days
Epstein–Barr virus (EBV) antibody to viral capsid antigen, IgM and IgG	Clarify or confirm equivocal or negative Monospot test Discriminate EBV from other IM-like diseases (e.g., CMV, toxoplasmosis)	Repeat testing in 10–14 days may be helpful if results are equivocal
EBV antibody to nuclear antigen, IgG	Confirm previous infection with EBV	Repeat testing in 10–14 days may be helpful if results are equivocal
EBV by PCR	Detect EBV in cerebrospinal fluid and serum specimens Diagnose EBV-related diseases in immunocompromised patients or patients with lymphoproliferative tumors Do not use to confirm acute mononucleosis	Negative result does not rule out the presence of PCR inhibitors in patient specimen or EBV DNA in concentrations below assay detection
EBV, quantitative PCR	Monitor disease (whole blood, plasma, serum, or CSF specimens) Do not use to confirm acute mononucleosis	
EBV antibody to early D antigen (EA-D), IgG	Confirm chronic active mononucleosis, posttransplant lymphoproliferative disease, and nasopharyngeal carcinoma This antibody test is more useful and appropriate than early antigen R for mononucleosis assessment	Repeat testing in 10–14 days may be helpful if results are equivocal
EBV by in situ hybridization	Virus identification of EBV	

  - IgG – peaks during week 3–4 of infection, can persist for more than 1 year or entire lifetime
- Early antigen antibody, Anti-D
  - Diffusely nuclear and cytoplasmic staining of infected cells
  - Present in 40% of infectious mononucleosis patients
  - Persists for 3–6 months
  - Detected in patients with nasopharyngeal carcinoma
- Early antigen antibody, Anti-R
  - Stains cytoplasmic aggregates
  - Found in atypical protracted cases of infectious mononucleosis
  - Found in patients with African Burkitt lymphoma
- Epstein–Barr nuclear antigen antibody
  - Appears 3–4 weeks after infection
  - Persistent for life
  - Found in patient with Burkitt lymphoma
- Molecular methods (Table 25.6)
  - In situ hybridization (Biogenex, San Ramon, CA)
    - Used for tissue biopsy
    - The EBV EBER Probe is specific for EBER RNA transcripts and is intended for the detection of latent EBV infection
    - The EBV Not I/Pst I DNA Probe is specific for the Not I/Pst I repeat sequence of EBV and is intended for the detection of active EBV infection
  - Quantitative competitive PCR
    - Specific primers are specifically designed to the EBV viral latent membrane protein 2a (LMP2a) and

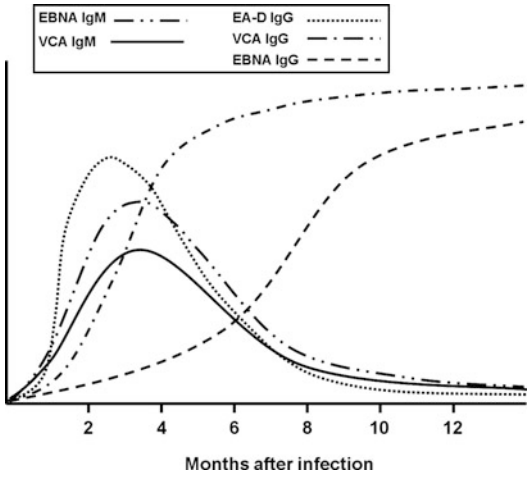


Fig. 25.21 Time course for appearance of antibodies in EBV infection

internal competitor DNA (ssDNA) that is confirmed against a known number of Namalwa cells (B cell lymphoma cell line containing two integrated copies of the EBV genome per cell)
- Four separate PCR reaction tubes each containing internal competitor DNA (8 copies/μl, 40 copies/μl, 200 copies/μl, or 1,000 copies/μl) are placed in competition with EBV-specific primers for amplification of patient DNA
- PCR amplicons are examined by electrophoresis through a 2% agarose gel and visualized using a gel imaging documentation system. The band densities are quantitatively measured using Bio-Rad Quantity One software and used to calculate EBV copies
- Although highly accurate and reproducible, such assays are rather laborious and require intensive post-PCR handling. Each sample has to be spiked with different amounts of internal standard to achieve precise quantification

– Real-time PCR
  - LightCycler® EBV quantitative kit (Roche)
    – Detection of LMP gene in EBV viral genome
    – EBV is amplified with specific primers in a PCR reaction. The amplicon is detected by fluorescence using a specific pair of hybridization probes
    – A melting curve analysis is performed after the PCR run to differentiate positive samples from non-EBV species; i.e., other herpes virus family
    – The internal control is a synthetic double-stranded DNA molecule with primer binding sites identical to the EBV target sequence, comprising a unique hybridization probe binding region that differentiates the internal control from the target-specific amplicon. It is added already to the lysed sample before the purification step and copurified/amplified with the EBV DNA from the specimen in the same PCR reaction (dual color detection)
    – The kit allows quantification in a range of 10^2–106 copies per reaction. The lower detection limit of the kit is $\leq$10 copies per reaction (95% confidence interval; probit analysis)
- Other commercial assays and reagents
  – Other commercial assays and reagents included Nanogen EBV Q-PCR Alert, Argene EBV R-gene™, QIAGEN artus® EBV (LC and RG) PCR Kits, Cepheid affigene® EBV trender and SmartEBV™, ELITech/Epoch EBV ASR, and Quantification of HHV4 PrimerDesign™ Ltd
  – The range of EBV PCR targets included EBNA1, EBNA2, BNRF1 p143, BXLF1, EBER1, BALF5, and BamHI-W
  – Amplification platforms included Roche LightCycler® 1.5, 2.0, and 480 systems, Applied Biosystems™ 7300, 7500, 7500 Fast, and 7900 HT Fast Real-Time PCR Systems, Agilent Mx3000P® quantitative PCR System, QIAGEN Rotor-Gene™ Q and Rotor-Gene™ 3000, and Cepheid SmartCycler™ II

- Real-time PCR using SYBR Green I dye
  - To maximize detection rates and reduce false-negative results, two primer sets targeting the highly conserved EBV regions, (1) Epstein–Barr nuclear antigen 1 (EBNA1) and (2) BamHI fragment H rightward open reading frame 1 (BHRF1), used to detect and measure absolute EBV DNA load in clinical settings with different EBV-associated diseases. Two separate real-time quantitative PCR assays using SYBR Green I dye and a single quantification standard containing the two EBV genes
  - PCR products analyzed by an amplification curve, melt analyses, and amplification efficiency
  - The lower limit of detection for both EBV regions was 2.0×10^3 copies/ml
  - Sensitive and cost-effective
- Pitfalls
  - Quantitative PCR requires analysis of absolute lymphocyte count, which inversely affects viremia; real-time PCR does not
  - Important to note that real-time PCR assay requires sequential analysis of run data prior to result reporting to prevent false positives (i.e., pseudoamplification and amplification of non-EBV species) and false negatives (i.e., shifted melting curve for EBV variants)
  - PCR assays are not standardized and variations in sample handling and laboratory methods can affect the sensitivity of the assay

25.6.4 Clinical Utility

- Serial viral load testing can be used to monitor disease burden and assess efficacy of immunosuppressive therapy in posttransplant patients
- Detection of EBV in tissue biopsy assists the diagnosis of EBV-related malignancies, including lymphoma and nasopharyngeal carcinoma

25.7 Herpes Simplex Virus

25.7.1 General Characteristics

- Family of enveloped icosahedral nucleocapsid viruses with total nine members
- Herpes simplex virus (HSP), type 1 and type 2, demonstrates an 83% DNA homology in protein-coding regions
- The genetic map of the two viruses is colinear, and the genomes are of approximately the same size, HSV1 of 152 kbp and HSV2 of 155 kbp
- Humans are the only known reservoir
- Direct contact with lesion or secretions necessary for transmission. After direct exposure to infectious material (i.e., saliva, genital secretions), initial viral replication occurs at either the skin or mucous membrane entry site, typically of epithelial cells
- HSV1 and HSV2 most common. HSV1 acquired early in life, usually associated with oral lesions. HSV2 acquired after onset of sexual activity, associated with genital lesions. Both viral types can cause oral–facial and genital infections and maybe clinically indistinguishable
- Risk of transmission of HSV from HSV-infected mother during vaginal delivery to infant is 50%, estimated to be between 1 in 2,000 and 1 in 5,000 births
- Beyond the neonatal period, most childhood HSV infections are caused by HSV1. The seroprevalence of HSV1 antibodies increases with age and is 20% by age 5 years. No increase occurs until age 20–40 years, when 40–60% of individuals are HSV1 seropositive
- Latent infections reside in neurons (trigeminal, sacral, and vagal ganglia) and can be activated by a stimulus (e.g., physical or emotional stress, fever, ultraviolet light) causes reactivation of the virus in the form of skin vesicles or mucosal ulcers, with symptoms less severe than primary infection

25.7.2 Clinical Presentation

- Primary infection usually with a 2–20 days incubation period
- Cutaneous vesicles characterized by ulcers that eventually pustulate, dry, and crust; mucosal vesicles appear as shallow punctuate ulcers that often coalesce
- Primary herpetic gingivostomatitis/pharyngotonsillitis (HSV1): Most cases are asymptomatic. Most cases are between 6 months and 5 years. Characterized by generalized malaise, fever, linear gingivitis, and lymphadenopathy
- Primary herpes genitalis (HSV2): Genital HSV2 infection is twice as likely to reactivate and recurs 8–10 times more frequently than genital HSV1 infection. A classic vesicular rash may be noted or progressive lesions (pustules or painful ulcerative lesions). Lesions may persist for as many as 3 weeks. Painful inguinal lymphadenopathy, dysuria, and vaginal discharge are frequent complaints. Most primary genital HSV infections are asymptomatic, and 70–80% of seropositive individuals have no history of symptomatic genital herpes. HSV can be transmitted in the presence or absence of symptoms
- Primary cutaneous herpetic infections can occur in wrestlers and rugby players with contaminated abrasions (herpetic gladiatorum or scrumpox)
- HSV keratitis presents with an acute onset of pain, blurring of vision, chemosis, conjunctivitis, and characteristic dendritic lesions of the cornea
- HSV meningitis
  - 1–7% of all cases of aseptic meningitis
  - HSV2 > HSV1
  - 20–45% with meningitis have recurrent episodes
  - HSV accounts for 10–20% of all cases of sporadic viral encephalitis in the US. The clinical hallmark of HSV encephalitis has been the acute onset of fever and focal neurologic (especially temporal lobe) symptoms. Clinical differentiation of HSV encephalitis from other viral encephalitides, focal infections, or noninfectious processes is difficult
  - Neonates (<6 weeks) have the highest frequency of visceral and/or CNS infection of any HSV-infected patient population
- HSV infection of visceral organs usually results from viremia, and multiple organ involvement is common
- Recurrent infection at sites of primary infection
  - Activation of latent virus form neurons of cervical ganglia (herpes labialis, HSV1) or sacral ganglia (HSV2)
  - Self-inoculation of fingers and thumbs (herpetic whitlow) can occur in children with orofacial herpes, although less common
  - Antiviral prophylaxis recommended for persistent recurrent cases
  - Some cases of erythema multiforme (EM) are believed to represent an allergic response to recurrent HSV infection

25.7.3 Diagnostic Methods

- Specimens
  - Vesicular fluid, ulcerated lesions, pharyngeal and throat swabs, urine, CSF, autopsy and biopsy material, ocular exudates, and vaginal swabs
  - Specimen is best collected within the first 3 days after appearance of lesion but no more than 7 days
- Conventional tests and problems
  - Viral culture
    - Conventional
      - Cell culture requires the collection of live virus samples that require special care in transport to the laboratory to retain viability. When viable samples are used, culture can be highly specific (if typing is performed) and positive results are generally reliable
      - The sensitivity of culture declines rapidly as lesions begin to heal,

and for this reason, frequently nonpositive result can be falsely negative. Type-specific serology tests should be used in these cases to confirm a clinical diagnosis of genital herpes
    - Many commercial cell lines are used (A549, RK, ML, HNK, MRC5, etc.)
    - Diagnosed by observation of cytopathic effect (CPE) induced by virus which usually occurs in 1 week after initial inoculation
  - Shell vial assay
    - A centrifugation-enhanced culture technique used to obtain rapid culture results. Generally less sensitive than conventional culture
    - The test can detect HSV in shell vial cultures (MRC5 cells) before the development of cytopathic effect (pre-CPE)
    - Immunofluorescent assay (IFA) staining of shell vial for viral detection and typing
- Cytology
  - Intranuclear inclusion bodies
  - Multinucleated, molded giant cells
  - Margination of nuclear chromatin
  - Indicates the presence of herpesvirus (HSV1/2 or varicella zoster virus [VZV]) and excludes coxsackievirus and nonviral entities
- Serological studies
  - Limited by cross-reacting antigens between HSV1 and HSV2
  - ELISA
    - Performed on fluids or other samples using HSV-specific antibody that is bound to a solid surface
    - Antibody captures antigen to which anti-HSV antibodies labeled with enzymes are added. These attach to the bound antigen and cause a color change
  - IFA and immunoperoxidase (IPA) assay
    - Detect HSV antigen in smears or tissues. HSV-specific antibodies are labeled with fluorescent dyes or enzymes (peroxidase)
    - Labeled antibodies are incubated with the specimen and bind to HSV antigens in the specimen, if present
    - Attached fluorescent dye or enzyme can be visualized in appropriate regions of infected cells under a microscope
    - Used in conjunction with shell vial culture
  - ELVIS® (enzyme-linked virus inducible system)
    - Technique combines cell culture amplification with HSV-activated reporter genes
    - The test produces results that are equal to conventional culture
- Molecular methods
  - Real-time PCR (LightCycler® – HSV1/2 detection kit, Roche)
    - Detection and differentiation of HSV1/2
    - HSV1/2 is amplified with specific primers in a PCR reaction. The amplicon is detected by fluorescence using a specific pair of hybridization probes
    - A melting curve analysis is performed after the PCR run to differentiate positive samples in HSV1 or HSV2. Melting points for HSV1 and HSV2 are significantly different (HSV1 at 53.9 °C, whereas HSV2 at 67.1 °C) and allow clear determination of the HSV type
    - The internal control is added already to the lysed sample before the purification step and copurified/amplified with the HSV DNA from the specimen in the same PCR reaction (dual color detection)
  - Real-time PCR (HSV1/2 PCR kit, Abbott Molecular)
    - The HSV1/2 PCR kit is for the detection and differentiation of HSV 1 and 2 DNA using PCR in the ABI PRISM® 7000 and 7900HT sequence detection system (Applied Biosystems)

    - The HSV TM Master contains reagents and enzymes for the specific amplification of a 148-bp region of the HSV genome
    - The amplicon is detected by measuring the FAM fluorescence (HSV1) and NED fluorescence (HSV2) in the ABI PRISM® SDS
    - In addition, the HSV1/2 PCR kit contains a second heterologous amplification system to identify possible PCR inhibition. This is detected as an internal control (IC) by measuring the VIC fluorescence
    - External positive controls (HSV1 LC/RG/TM QS 1–4 and HSV2 LC/RG/TM QS 1–4) are used to allow the determination of the pathogen load
  - Advantages
    - HSV real-time PCR-based assay is highly sensitive and specific, and it can detect the virus even during the low viral shedding
    - PCR-based technique was shown to increase the overall rate of HSV detection by 61–71%. Even in patients with visible genital ulcerations, PCR detected 88% more infections than virus culture
- Sensitivity and specificity
  - The lower limit of detection (analytical sensitivity) for HSV qualitative PCR is 25 copies/reaction (~1,250 copies/ml)
  - The sensitivity of PCR
    - HSV in skin lesions: sensitivity 83–100% and specificity 100%
    - CSF: sensitivity 70–100%
  - HSV was detected more frequently by PCR than by viral culture regardless of whether samples were obtained from HSV lesions, or from genital or oral secretions during a period of subclinical shedding. Yield of virus positivity is four times greater by PCR than by culture, and the results are more reliable, especially in settings in which transport or climate may interfere with the yield from viral culture
  - Due to the sensitivity of PCR, many labs now only offer PCR tests and culture is used only when sensitivity test is needed
  - CSF culture
    - ~80% positive with first attack
    - 0% with recurrent episodes
    - Multiplex PCR could be performed for HSV1, HSV2 and VZV, CMV, and HSV6 from the same sample
- Pitfalls
  - Important to note that PCR assays are not standardized and variations in sample handling and laboratory methods can affect the sensitivity of the assay
  - PCR cannot always diagnose HSV encephalitis in the first few days of illness, and serial evaluations of CSF by PCR during the first week of illness are necessary

25.7.4 Clinical Utility

- Diagnosis of herpes encephalitis in neonates and immunocompromised patients by detection of HSV in CSF
- CSF PCR for HSV DNA should be performed in patients with febrile encephalopathy even in the absence of focal features, initial CSF pleocytosis, or abnormal CT. Mild or atypical HSV encephalitis may be associated with infection from HSV1 or HSV2
- In addition to CSF, other specimens can be used for PCR, including mucosal secretion, skin lesion, etc
- Current treatment guidelines for herpes include three antiviral therapies, acyclovir, famciclovir, and valacyclovir, and should begin as soon as possible after symptoms begin. Antiviral therapy may be effective when taken during onset of prodromal symptoms; i.e., tingling
- Antiviral therapy will reduce the duration of outbreak by approximately 2 days
- Suppressive therapy is highly effective and can dramatically reduce the frequency of recurrences. Suppression can be continued for years with very low risk of toxicity or development of drug-resistant HSV. Suppressive therapy will also reduce the frequency of asymptomatic HSV shedding

Table 25.7 Varicella zoster virus (VZV) diagnostic tests

Test name	Diagnostic utility and significance	Comments
VZV – viral culture	Culture considered gold standard	Not recommended for CSF samples
VZV – DFA (direct fluorescent antibody)	Rapid confirmation of VZV High sensitivity and specificity	
VZV – PCR	VZV test need be performed by PCR when vesicle fluid specimens negative for VZV by culture and/or DFA	PCR is the most sensitive and rapid test
VZV – antibodies, IgG and IgM	Diagnose clinical infections with varicella or herpes zoster Identify hospitalized children with varicella Assess immune status of individuals exposed to varicella, especially pregnant women	

25.8 Varicella Zoster Virus

25.8.1 General Characteristics

- VZV is a human alpha herpes virus, a member of herpes family (HHV3) with a linear, double-stranded DNA genome
- The VZV genome is 125kb
- Isolated in patients with chicken pox (primary), subsequent latency followed by reactivation of virus, known as shingles (recurrent). The virus is spread via contact with vesicular fluid or inhalation of respiratory droplets
- Multiple recurrences are common and can be triggered by immunosuppression, exposure to cytotoxic drugs, radiation, and malignancy

25.8.2 Clinical Presentation

- Varicella (chicken pox)
  - Chicken pox, which is caused by the VZV, is one of the most contagious childhood diseases. Nearly every unvaccinated child becomes infected with it
  - Mild self-limited illness common in school-aged children with fever followed by vesicular eruption on skin and mucous membranes
  - Spreads by respiratory secretions with a 10–14 days incubation period
  - More severe in adults, pneumonia common
- Herpes zoster (shingles)
  - Recurrent infection, usually in adults that may be activated by trauma, neoplasm, or immunosuppression
  - Virus remains latent in sensory ganglia of spinal or cranial nerves causing dermatomal pain and vesicular eruptions, fever, and malaise. Commonly occurs in trunk, but may affect any dermatome
  - Associated with encephalitis and delayed cerebral vasculitis
- Zoster sine herpete occurs in the event of recurrence in the absence of vesicle formation
- Post herpetic neuralgia: pain lasting longer than 1 month after an episode, occurs in as many as 14% of affected individuals, particularly those over 60 years of age. Most neuralgias resolve within one year with 50% experiencing resolution within 2 months
- Ramsay Hunt syndrome: combination of cutaneous involvement of herpes zoster infection of external auditory canal and ipsilateral facial and auditory nerve. Syndrome can cause facial paralysis, hearing deficits, and vertigo

25.8.3 Diagnostic Methods

- Specimens
  - Skin vesicle fluid, cerebrospinal fluid, nasopharyngeal secretion, bronchial washings, blood, amniocentesis fluid, and urine
- Conventional tests and problems (Table 25.7)
  - Viral culture
    - Conventional

  - Virus is difficult to grow in cell culture. Unlike other HSVs (HSV1 and HSV2), VZV manifests a very narrow range of hosts
  - Viral isolation should be attempted in cases of severe disease, especially in immunocompromised persons
  - The best results are obtained from vesicular fluid with lower yield from other sites (nasopharyngeal secretion, blood, urine, bronchial washings, and cerebrospinal fluid)
  - Diagnosed by observation of cytopathic effect (CPE) induced by virus which usually occurs in 1 week after initial inoculation
- Shell vial assay
  - A centrifugation-enhanced culture technique used to obtain rapid and more sensitive culture results
  - It provides results within 2–3 days
- Cytology
  - Intranuclear inclusion bodies
  - Multinucleated, molded giant cells
  - Margination of nuclear chromatin
- Serological studies
  - Enzyme-linked immunosorbent assays (ELISA) range in sensitivity from 86% to 97% and range in specificity from 82% to 99%
  - Latex agglutination (LA) is a rapid, simple to perform assay to detect antibodies to VZV glycoprotein antigen
    - 96% is positive in convalescent-phase serum specimens
    - 61% is positive in persons after vaccination
  - Fluorescent antibody to membrane antigen (FAMA) test
    - It is highly sensitive and is the gold standard for screening for immune status for VZV
    - 100% positive in convalescent-phase serum specimens
    - 77% positive in persons after vaccination
  - Direct fluorescent antibody (DFA)
    - Using fluorescein-labeled monoclonal antibodies specific for either HSV or VZV antigens
    - Results are obtained within several hours
    - Specimen is best collected from the base of a skin lesion, preferably a fresh fluid-filled vesicle
    - The use of DFA may be positive when viral cultures are negative because infected cell viral proteins persist after cessation of viral replication
- Molecular methods (Table 25.7)
  - Conventional PCR targets VZV orf 29 gene, and detection limit is 500 copies/ml
  - Real-time PCR (artus® VZV LC PCR kit, QIAGEN)
    - artus® PCR Kits are used for identification of pathogens including the herpes viruses CMV, EBV, HSV1/2, and VZV
    - Harmonized amplification profiles allow the parallel testing of these four different herpes viruses in a single run on LightCycler® Instruments
    - Analytical sensitivity: five copies/reaction
    - Specificity: 100%
  - LightCycler® VZV Qualitative Kit
    - It targets orf gene 28, DNA polymerase, orf gene 29, orf gene 38, or DNA binding protein
    - It is 91% more sensitive than the shell vial cell culture assay from dermal specimens
  - Real-time quantitative PCR (TaqMan®) technique
    - It targets orf gene 28, orf gene 38, or glycoprotein B
    - Assay results range from 10 copies/ml to 1×10^{10} copies/ml
    - It is 53.8% more sensitive than cell culture from dermal specimens
- Pitfalls
  - Important to note that PCR assays are not standardized and variations in sample handling and laboratory methods can affect the sensitivity of the assay

25.8.4 Clinical Utility

- Although the varicella vaccine is routinely given as a childhood vaccine, certain high-risk groups and scenarios require careful monitoring (i.e., immunocompromised patients)
- Intrauterine infection of the fetus with VZV can be detected by PCR testing of amniocentesis fluid
- It can be applied on different specimens including mucosa secretion, skin lesion, etc
- Diagnosis of encephalitis in immunocompromised patients by detection of VZV in CSF
- Early initiation of VZV-specific antiviral therapy may prevent serious morbidity among HIV-infected patients

25.9 Human Papillomavirus

25.9.1 General Characteristics

- Human papillomavirus (HPV) is a member of the *Papillomaviridae* family that can completely integrate with the DNA of the host cell. Humans are the only known reservoir for HPV
- Papillomaviruses are nonenveloped viruses of icosahedral symmetry with 72 capsomeres that surround a genome containing double-stranded circular DNA with approximately 8,000 base pairs
- The expression of viral genes is closely associated with an epithelial localization and linked to the state of cellular differentiation. Most viral genes are not activated until the infected keratinocyte leaves the basal layer. Production of virus particles can occur only in highly differentiated keratinocytes. Therefore, virus production only occurs at the epithelial surface where the cells are ultimately sloughed into the environment
- Over 100 genotypes of HPV have been identified based on DNA sequence heterology. A specific group, termed high-risk genital HPV types (especially 16, 18, 31, 45, and 58, but also 33, 35, 39, 51, 52, 56, 59, 68, 73, 82), is recognized as a necessary factor for the development of cervical cancer
- The genome HPV virus is circular (Fig. 25.22). The genome has eight open reading frames that encode ten proteins. The genes for these are divided into an early region that are expressed in the skin's infected basal cells that have yet to differentiate, and a late region with two genes whose protein products exist only in cells after cell differentiation
- The E5 (changes the cellular responses to programmed cell death or apoptosis), E6 (binds to tumor suppressor protein, p53), and E7 (binds and inactivates retinoblastoma protein, Rb) proteins are early viral proteins expressed upon infection and cause destabilization of the infected cell and induces replication
- As the cell differentiates, it migrates upward and induces expression of the E1, E2, and E4 genes; E1 and E2 cause viral replication and E4 destabilizes the cytoskeleton and prevents cellular differentiation
- In the upper epithelial cell layers, the late viral proteins L1 (major capsid protein) and L2 (minor capsid protein) are expressed. They bind the viral DNA and autoassemble, giving rise to the complete virions, ready for a new infection that is released as the keratinocytes desquamate
- The most common mode of transmission is via contact; i.e., sexual or autoinoculation

25.9.2 Clinical Presentation

- HPV is by far the most common sexually transmitted disease. An estimated 80% of sexually active adults have been infected with one or more genital HPV strains. The vast majority of infected adults experience transient infectivity and are unaware of the condition; however, they may be able to infect others
- However, most women infected with high-risk HPV, especially women under 30 years of age, do not develop cervical cancer. Their immune system effectively clears the infection over the course of several months
- Specific factors that determine which HPV infections persist and develop into squamous intraepithelial lesions currently are unknown.

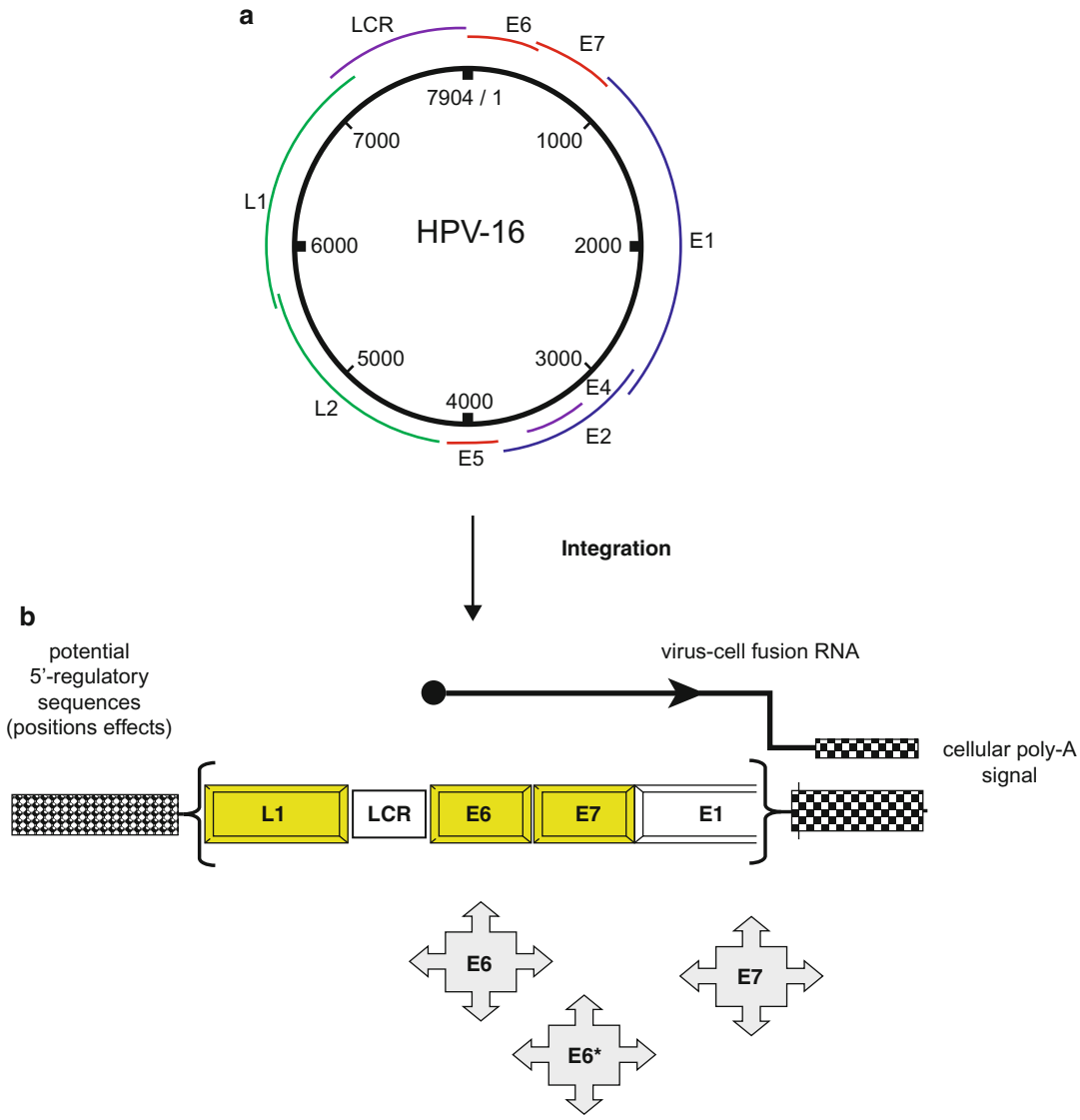


Fig. 25.22 The genome HPV16 virus (Adapted from Finzer et al. Cancer Lett 2002)

Cigarette smoking, ultraviolet radiation, pregnancy, folate deficiency, and immune suppression have been implicated as possible cofactors

- Low-risk HPV types (6, 11, 42, 43, and 44) produce benign epithelial tumors of the skin and mucous membranes. Infection with certain types of HPV (high risk) can also increase the risk of developing cervical and other cancer types. Conditions associated with HPV
  - Verruca vulgaris (common wart) – associated with HPV2, HPV4, and HPV40. Highly contagious and can spread to other sites of skin or mucous membranes via autoinoculation
  - Condyloma acuminatum (venereal wart) – associated with HPV6, HPV11, HPV16, and HPV18 and is considered a sexually transmitted disease with lesions occurring in sites of sexual contact or trauma

(i.e., mucous membranes of genitalia, perianal region, oral cavity, and larynx)
- Flat warts are most commonly found on the face or forehead and are most common in children and teens
- Plantar warts are found on the soles of the feet
- Subungual and periungual warts form under the fingernail (subungual) and around the fingernail or on the cuticle (periungual) and are a subtype of the common skin wart. They may be more difficult to cure than warts in other locations
- Butcher warts are caused by HPV7 and occur in people handling meat, poultry, and fish
- Focal epithelial hyperplasia (Heck disease) is caused by HPV13 (and possibly HPV32) and commonly occurs in Native American and Inuit populations. A childhood condition characterized by multiple soft, nontender flat papules and plaques of the oral mucous membrane
- Laryngeal papillomatosis – frequently recurs and may require repetitive surgery when interferes with breathing. Rare cases can progress to laryngeal cancer (HPV30 and HPV40)
- HIV-associated papillomatosis – HPV7 and immunocompromised states
- Cervical cancer – History of HPV (high-risk types) infection is strongly associated with development of cervical cancer. However, most HPV infections do not progress to cervical cancer. Because the progression of transforming normal cervical into cancerous cells is a slow process, cancer occurs in people who have been infected with HPV for a long time, usually over a decade. High-risk HPV types 16 and 18 are together responsible for over 70% of cervical cancer cases; type 16 alone causes 41–54% of cervical cancers
- Anal/rectal cancer – Although rare, anal/rectal cancer is becoming more prevalent in the US. Similar to cervical cancer, the main cause of anal cancer is HPV and is most commonly acquired through anal intercourse. However, anal cancer can also be acquired from other genital areas that are infected with HPV, particularly from the vulva or penis. High-risk HPV types (16, 18, 31, 33 and 35) are associated with anal squamous intraepithelial lesions and account for approximately 80% of cervical and anal cancers
- Head and neck squamous cell cancer – HPV16 is the most common type detected in head and neck squamous cell cancers (HNSCCs). HPV16 accounts for 78.6–100% of HPV-positive oropharyngeal cases. HPV type 18 accounts for 1% of oropharyngeal, 8% of oral cavity, and 4% of laryngeal HPV-positive SCCs. HPV16 and HPV18 genotypes are associated with a better prognosis in squamous cell carcinomas of the head and neck. Furthermore, the HPV-positive tumors are also more sensitive to radiation and chemotherapy. Coinfections are possible and most frequently include HPV16. HPV33 is often reported and has been identified in up to 10% of HPV-positive HNSCCs. Numerous other HPV types have been rarely detected in HNSCCs and include types 6, 11, 35, 45, 51, 52, 56, 58, 59, and 68
- Other cancers – HPV (including 16, 18, and 31) may also cause vulvar, nonmelanoma skin cancers, and (rarely) penile cancer. High-risk types of HPV can cause intraepithelial neoplasias, or abnormal and precancerous cell growth, in the vulva and cervix, which can progress to cancer

25.9.3 Diagnostic Methods

- Specimens – cervical washings/brushings collected in liquid media (i.e., PreservCyt), Digene specimen collection tube (Hybrid Capture only), or biopsies
- Conventional tests and problems
  - Viral culture
    - HPV cannot be reliably cultured and is not identified using this technique

- Cytology
  - Koilocytosis describes the combination of perinuclear clearing (halo) with a pyknotic or shrunken nucleus
- Serological studies are not useful for diagnosis

- Molecular methods and genotyping
  - Nucleic acid hybridization
    - In situ hybridization (INFORM HPV DNA test, Ventana Medical Systems Inc., Tucson, AZ)
      - Use tissue sections, liquid-based cytology specimens, and cervical smears
      - On slide detection of high- and low-risk HPV genotypes
      - 16 probe cocktail for high-risk HPV genotypes 16, 18, 31, 33, 35, 39, 51, 52, 56, 58, and 66
      - 6 probe cocktail for low-risk HPV genotypes 6 and 11
  - Nucleic acid hybridization
    - Digene Hybrid Capture II (Digene Corporation, Gaithersburg, MD)
      - Method utilizes an RNA probe mix for the detection of the L1 gene of HPV. Assay can identify HR HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68. In addition, a kit detecting low-risk virus (6, 11, 42, 43, 44) is also available
      - Manual, semiautomated (rapid capture), and automated (utilizing QIAGEN's QiaSymphony extraction system) platforms available
      - Signal amplification is based on immunocapture of DNA/RNA hybrids that are immobilized on a 96-well microplate, reacted with alkaline phosphatase-conjugated antibodies specific for the RNA: DNA hybrids and detected with a chemiluminescent substrate
      - Can detect 5,000 viral copies per sample, or one picogram of HPV DNA per sample
  - Signal amplification
    - Invader Assay Cervista HPV assays (Hologic, Madison, WI)
      - The Invader assay is an isothermal linear signal amplification using structure-specific oligonucleotide cleavage and has been applied to DNA-based genotyping. This method uses two types of isothermal reactions: a primary reaction that occurs on the targeted DNA sequence and a secondary reaction that produces a fluorescent signal
      - Invader utilizes an internal control for human HIST2HBE to assure DNA quality and quantity in each reaction
      - Two diagnostic qualitative assay formats that utilize the Invader chemistry are available: Cervista HPV HR (detects pool of 13 HPV genotypes) and Cervista HPV16/18 (detects HPV16 and/or HPV18)
      - Cervista HPV HR uses isothermal signal amplification to detect 13 h HPV types utilizing three probe pools based on phylogenic relatedness. 3 probe pools include A5/A6 [51, 56], A7 [18, 39, 45, 59, 68], and A9 pool [16, 31, 33, 35, 52, 58]. Assay cannot determine the specific HPV genotype present
      - The Cervista HPV16/18 test is a diagnostic test to genotype HPV16 and HPV18 in cervical specimens
      - Can detect 1,250–5,000 viral copies per specimen
  - Target amplification
    - TMA – APTIMA HPV assay (Gen-Probe; San Diego, CA)
      - The APTIMA HPV assay is an in vitro nucleic acid amplification test for the qualitative detection of E6/E7 viral messenger RNA (mRNA) from 14 high-risk types of HPV in cervical specimens
      - Detects HPV subtypes: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68, but does not distinguish between the 14 high-risk types
      - The assay is used with the TIGRIS DTS system, one of the first

diagnostic instruments to truly automate nucleic acid testing (NAT) from start to finish by from sample preparation, amplification, and detection to reporting results
  - The TIGRIS system can process approximately 450 samples in an 8-h shift, and up to 1,000 samples in approximately 13.5 h
  - The APTIMA HPV assay involves three main steps, mainly: target capture, target amplification by transcription-mediated amplification (TMA), and detection of the amplification products (amplicon) by the hybridization protection assay (HPA) that measures the emitted relative light units (RLU) in a luminometer
  - Can detect 100–300 viral copies per sample
- Real-time PCR – COBAS® HPV test (Roche Diagnostics, Indianapolis, IN)
  - The COBAS HPV test is a clinical diagnostic qualitative assay to detect HPV in patient samples using the COBAS 4800 system that automates specimen extraction, amplification, and detection. The assay utilizes amplification of HPV DNA by real-time PCR and nucleic acid hybridization to detect 14 high-risk HPV genotypes in a single reaction tube that target the polymorphic L1 region of the HPV genome
  - The assay specifically genotypes HPV16 and HPV18 while concurrently detecting the other high-risk genotypes in a pooled fashion (31, 33, 35, 39, 45, 51, 51, 56, 58, 59, 66, 68)
  - The COBAS HPV test is based on two major processes: (1) automated specimen preparation to simultaneously extract HPV and cellular DNA and (2) PCR amplification of target DNA sequences
  - Using both HPV- and β-globin-specific complementary primer pairs and real-time detection of cleaved fluorescent-labeled HPV- and β-globin-specific oligonucleotide detection probes
  - The master mix reagent for the COBAS HPV test contains primer pairs and probes specific for the 14 high-risk HPV types and β-globin DNA
  - The detection of amplified DNA (amplicon) is performed during thermal cycling using oligonucleotide probes labeled with four different fluorescent dyes. The amplified signal from 12 high-risk HPV types (31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66 and 68) is detected using the same fluorescent dye, while HPV16, HPV18, and β-globin signals are each detected with their own dedicated fluorescent dye
  - An additional primer pair and probe target the human β-globin gene (330-bp amplicon) to provide a process control
  - Can detect 300–1,200 viral copies/ml
- Genotyping
  - Roche linear array (Fig. 25.23)
    - Qualitative test that utilizes amplification of HPV target DNA by PCR and nucleic acid hybridization bases on four major steps: (1) sample preparation, (2) PCR amplification of target DNA using HPV-specific complementary primers, (3) hybridization of the amplified products to oligonucleotide probes, and (4) colorimetric detection of the probe-bound amplified products
    - Uses a pool of biotinylated primers to define a sequence of nucleotides for the L1 region of the HPV genome designed to amplify HPV DNA from 37 HPV genotypes, including 13 high-risk genotypes (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68)
    - β-globulin gene is concurrently isolated and ensures adequacy of cellularity, extraction, and amplification for each processed sample

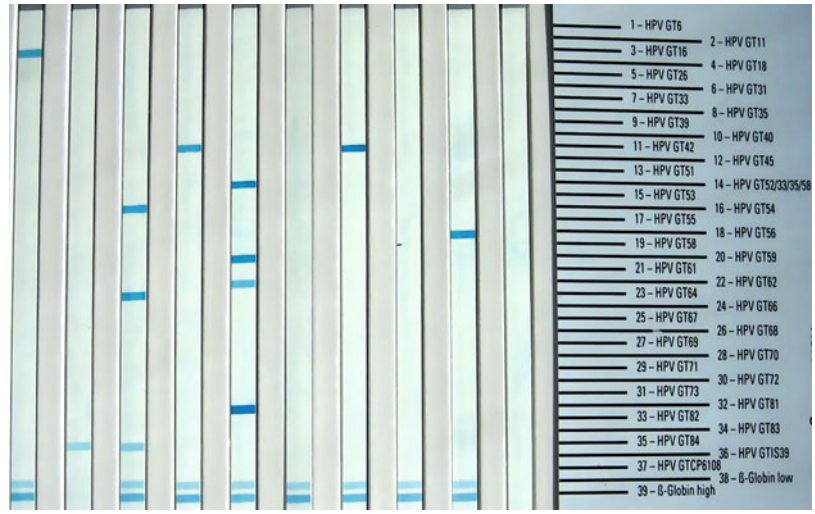


Fig. 25.23 Roche linear array HPV genotyping assay with reference guide utilized for interpretation. *P* positive control, *N* negative control, *ßgL* ß globulin low, *ßgH* ß globulin high

Table 25.8 hc2 high-risk HPV DNA test performance versus consensus histology results (CIN 2–3). Age-specific characteristics. Kaiser study data

	Age <30	Age 30–39	Age >39
Number of cases	287	233	365
Prevalence of disease (%)	12.2	11.2	2.7
Sensitivity (%)	100.00 (35/35)	88.46 (23/26)	80.00 (8/10)
95% confidence interval	90.0–100	69.9–97.6	44.4–97.5
Specificity (%)	31.4 (79/252)	66.2% (137/207)	79.15 (281/355)
95% confidence interval	25.7–37.5	59.3–72.6	74.6–83.3
Negative predictive value (%)	100 (79/79)	97.86 (137/140)	99.29 (281/283)
Positive predictive value (%)	16.83 (35/208)	24.73 (23/93)	9.76 (8/82)

Adapted from Digene package insert

- Sensitivity and specificity
  - Overall, the sensitivity for cytology for detecting HGSIL ranges from 50% to 70% and specificity of 86–98%
  - Overall, the sensitivity of HPV DNA test for detecting HGSIL is about 80–98% and specificity of 64–95%
  - Overall, the sensitivity of COBAS HPV assay for detecting HGSIL is about 69–71% and specificity of 90–94%
  - Overall, the sensitivity of Cervista HPV assay for detecting HGSIL is about 93–100% and specificity of 43–44%
  - Overall, the sensitivity of APTIMA HPV test for detecting HGSIL is about 87–93% and specificity of 60–63%
  - However, the sensitivity and specificity are influenced by the age and prevalence (Table 25.8)
- Pitfalls
  - Presently available assays provide only qualitative results and do not correlate the magnitude of the positive assay signal to meaningful quantitative results
  - The effects of other potential variables such as vaginal discharge, use of tampons, douching, personal lubricants, topical medicaments, and specimen collection variables may affect the performance of the assay
  - A negative HPV result does not exclude the possibility of present or future cytologic abnormalities
  - COBAS HPV assay

Table 25.9 Summary of recommendations for cervical cancer screening

Variable	ACS-ASCCP-ASCP Draft 2011	ACOG 2009	USPSTF Draft 2011
Age to start HPV testing	21 years	21 years	21 years
Testing frequency			
Age 21–29 year (PAP alone)	Every 3 years	Every 2 years	Every 3 years
Age 30 year and older			
PAP alone	Every 3 years	Every 3 years	Every 3 years
PAP and HPV cotesting	Recommended – no more than every 3 years	Allowed – no more than every 3 years	No recommendation – insufficient data available
Age to stop HPV testing	65 years after 3 negative Pap tests or two negative HPV tests in past 3 years	65–70 after 3 negative tests in a period	65 years after adequate screening
After hysterectomy	Discontinue, if not dysplasia or neoplasia	Discontinue, if not dysplasia or neoplasia	Discontinue, if not dysplasia or neoplasia
Screening after HPV vaccination	Same as if unvaccinated	Same as if unvaccinated	Unaddressed

Note: *ACOG* American College of Obstetricians and Gynecologists, *ACS* American Cancer Society, *ASCCP* American Society for Colposcopy and Cervical Pathology, *ASCP* American Society for Clinical Pathology, *HPV* human papillomavirus, *PAP* Papanicolaou, and *USPSTF* US Preventive Services Task Force

    - Strong laboratory information system needed to support automation of bar coding
    - Invalid results frequent
- APTIMA HPV assay
    - Assay does not detect E6/E67 mRNA of HPV low-risk types
    - Performance characteristics based only on ThinPrep 2000 processor
- Cervista HPV assay
    - Indeterminate test results caused by insufficient mixing, pipetting error, or inadequate genomic DNA
    - Exhibits cross-reactivity to HPV67 and HPV70 genotypes
- Digene hybrid capture assay
    - Limited sensitivity (1 pg/ml)
    - Mixed high-risk and low-risk probes, cannot distinguish specific HPV types
    - Manual platform is labor-intensive

25.9.4 Clinical Utility

- To screen patients with ASCUS (atypical squamous cells of undetermined significance), Pap smear results to determine the need for referral to colposcopy. The results of this test are not intended to prevent women from proceeding to colposcopy (Table 25.9)
- In women 30 years and older, the hc2 high-risk HPV DNA test can be used with Pap smear to adjunctively screen to assess the presence or absence of high-risk HPV types. This information, together with the physician's assessment of cytology history, other risk factors, and professional guidelines, may be used to guide patient management (Table 25.9)
- Recently, a new test scheme was proposed (Fig. 25.24)

25.10 Influenza A, B, and C Viruses

25.10.1 General Characteristics

- Influenza is part of the Orthomyxoviridae family and can be classified into three basic types, influenza A, B, or C (Table 25.10). Each influenza virus type is an enveloped single-stranded RNA virus that shares

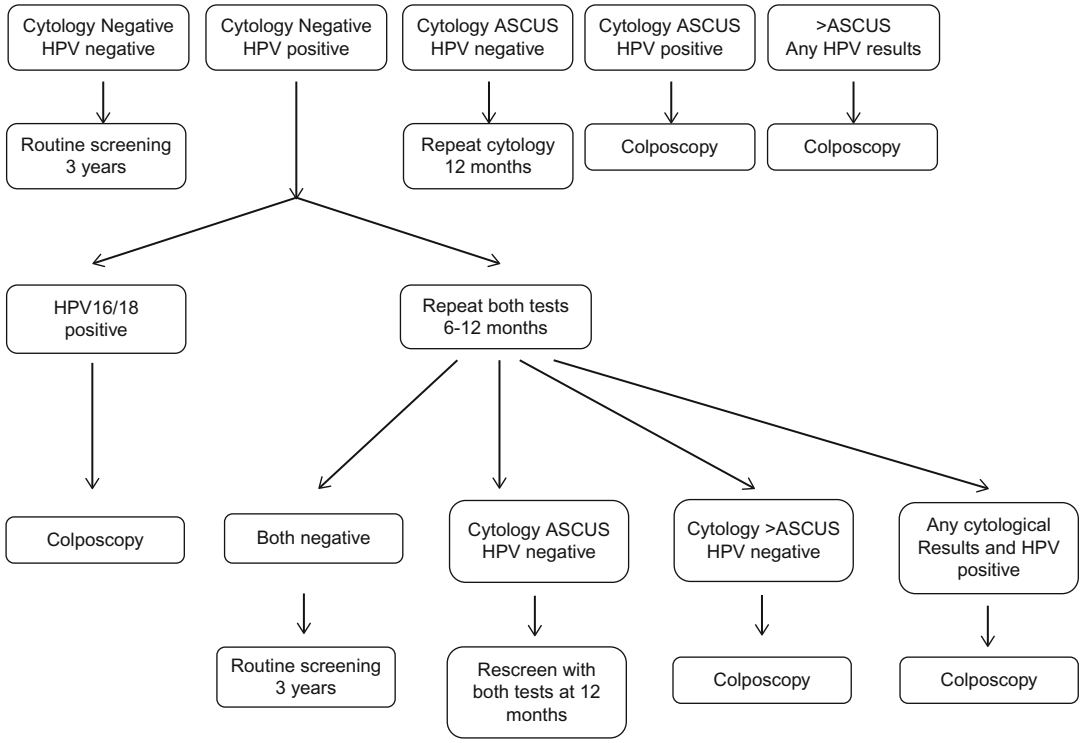


Fig. 25.24 Proposed management scheme of ASCUS based on cytology and/or high-risk HPV DNA test (Adapted from Wright et al. Obstet Gynecol 2004)

Table 25.10 Comparison of influenza A, B, and C

	Type A	Type B	Type C
Severity of illness	++++	++	+
Animal reservoir	Yes	No	No
Human pandemics	Yes	No	No
Human epidemics	Yes	Yes	No (sporadic)
Antigenic changes	Shift, drift	Drift	Drift
Segmented genome	Yes	Yes	Yes
Amantadine, rimantadine	Sensitive	No effect	No effect
Zanamivir (Relenza)	Sensitive	Sensitive	
Surface glycoproteins	2	2	(1)

structural and biological similarities but differs antigenically. Type A influenza virus which causes pandemic is found in a variety of warm-blooded animals. Types A and B are predominantly human pathogens. Type C is found in humans and pigs

- Influenza viruses have a segmented RNA genome (Fig. 25.25). Influenza A and B contain eight distinct segments and are covered with surface glycoproteins, hemagglutinin (HA), neuraminidase (NA), and matrix 2. Influenza C has seven segments and one surface glycoprotein. The viruses are typed based on these proteins. For example, influenza A (H3N2) expresses hemagglutinin 3 and neuraminidase 2

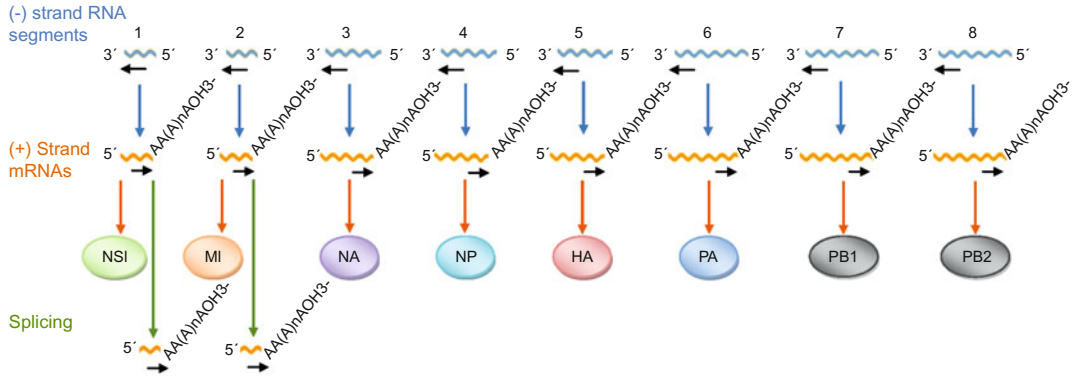


Fig. 25.25 Influenza virus genome: The virus contains 7–8 single-stranded RNAs (influenza A and B contains 8 RNAs, and influenza C contains 7 RNAs). The RNAs code for 9–11 viral proteins: hemagglutinin *HA*, neuraminidase *NA*, polymerase complex *PA, PB1*, and *PB2*, nucleoprotein *NP*, matrix protein *M*, and nonstructural protein *NS*. PCR primers usually target HA and NA consensus region

- Influenza is a dynamic virus that may evolve in two different ways via antigenic drift and antigenic shift resulting in genetic diversity. Antigenic shift occurs when two different strains of influenza viruses combine with antigenically different HA and NA by reassortment of viral RNA segments; this process occurs every 10–40 years. Antigenic drift occurs by random point mutation in viral RNA leading to amino acid substitutions in HA glycoproteins. Influenza type A viruses undergo both antigenic shift and drift; influenza type B viruses undergo antigenic drift
- Each influenza RNA segment is further encapsulated by nucleoproteins to form ribonucleotide–nucleoprotein complexes surrounded by matrix proteins
- Influenza virus infections rank as one of the most common infectious diseases in humankind. However, influenza may potentially cause severe epidemics and kills an average of 20,000 individuals in the US
- Influenza virus infection occurs after transmission of respiratory secretions from an infected individual to a person who is immunologically susceptible
- A number of different subtypes are classified within the influenza A type based on two viral surface proteins: hemagglutinin and neuraminidase. The most common prevailing human influenza A subtypes are H1N1 and H3N2. Each year, the distributed vaccine contains A strains from H1N1 and H3N2, along with an influenza B strain. A novel influenza A H1N1 strain emerged in March 2009 and is the causative agent of the current public health emergency
- In March 2009, an outbreak of influenza-like illness occurred in Mexico and the US; the CDC reported seven cases of novel A/H1N1 influenza. Preliminary genetic characterization found that the hemagglutinin gene was similar to that of swine flu viruses present in US pigs since 1999, but the neuraminidase and matrix protein genes resembled versions present in European swine flu isolates. The six genes from American swine flu are themselves mixtures of swine flu, bird flu, and human flu viruses
- Recently, several studies reported the novel A/H1N1 strains carried a prominent amino acid change at position 222 (D222N) within the primary hemagglutinin receptor binding site. Previously, enhanced virulence associated with the change, D222G, has been clinically linked to severe morbidity and mortality

25.10.2 Clinical Presentation

- Although the presentation of influenza virus infection is variable, typical symptoms may

Table 25.11 Comparison of available influenza diagnostic tests

Influenza diagnostic tests[a]	Types detected	Method	Typicalprocessing time[b]	Sensitivity[d] for 2009 H1N1 influenza	Distinguishes 2009 H1N1 influenza from other influenza A viruses?
Viral cell culture (conventional)	A and B[c]	Virus isolation	2–10 days	N/A	Yes[f]
Rapid cell culture (shell vials; cell mixtures)	A and B[c]	Virus isolation	1–3 days	N/A	Yes[f]
Immunofluorescence, direct (DFA) or indirect (IFA) antibody staining	A and B[c]	Antigen detection	2–4 h	47–93 %	No
RT-PCR[d] (single plex and multiplex; real-time and other RNA-based)	A and B[c]	RNA detection	Varied (generally 6–8 h)	86–100%	Yes
Rapid influenza diagnostic tests[e]	A and B	Antigen detection	<30 min.	10–70%	No

[a]Serologic (antibody detection) testing is not recommended for routine patient diagnosis
[b]The amount of time needed from specimen collection until results are available
[c]May be adapted to identification of specific subtypes
[d]Compared with rRT-PCR tests; rRT-PCR tests are compared to other testing modalities including other rRT-PCR assays
[e]Rapid influenza diagnostic tests include tests that are CLIA waived (can be performed in an outpatient setting) and tests that are moderately complex (can be performed only in a laboratory). Clinical specimens approved for RIDTs vary by test and may not include all respiratory specimens
[f]Requires additional testing on the viral isolate

include the following: fever, sore throat, myalgia, headache, rhinitis, fatigue, and coughing. Onset of illness may be abrupt
- Patients with a preexisting immunity or received vaccination may have mild and less severe symptoms
- Acute encephalopathy has recently been described to be associated with influenza A virus. Clinical features included altered mental status, coma, seizures, and ataxia
- In young infants, influenza may produce a sepsis-like picture with shock; occasionally, influenza viruses can cause croup or pneumonia
- In April 2009, a novel influenza A (H1N1) virus was identified in Mexico and has since spread rapidly worldwide. The unique genetic and antigenic features of this virus have resulted in a high incidence of infection, with an epidemiologic profile that is different from that of previous seasonal influenza infections. As a consequence, a surge of pediatric patients has been presenting to emergency departments and physician's offices across the country during this 2009–2010 flu season
- Vomiting and diarrhea have been reported more often with 2009 H1N1 influenza virus infection than with seasonal influenza
- The incubation period typically ranges from 18 to 72 h

25.10.3 Diagnostic Methods

- Specimens (Table 25.11)
  - Nasopharyngeal aspirate/swab/washing, tracheal aspirate, or bronchoalveolar lavage
  - Transport
    - Culture/DFA – 3 mL (minimum 1 ml) of respiratory sample in viral transport media (Microtest M4) or in sterile leakproof container at 2–8 °C
    - Serologic – 1 mL (minimum 0.5 ml) serum in an SST tube at 2–8 °C
  - Unacceptable specimens: dry swabs or wood and calcium alginate swabs that may inactivate the virus for culture. Plasma or hemolyzed, lipemic, icteric, turbid, bacterially contaminated, or heat-inactivated serum is inadequate for serological test

  - All respiratory specimens should be kept at 4°C for no longer than 72 h before testing and ideally should be tested within 24 h of collection. If storage longer than 72 h is necessary, clinical specimens should be stored at –70 °C. Freezing at higher temperatures (e.g., –20 °C) can reduce the likelihood of virus detection
- Conventional tests and problems (Table 25.11)
  - Viral culture
    - The criterion standard for diagnosing influenza A and B is via viral propagation in embryonated hen eggs or Madin–Darby canine kidney (MDCK) cells
    - Laboratory diagnosis of influenza is established once specific cytopathic effect is observed or hemadsorption testing findings are positive
    - After culture isolation, final identification via immunoassays or immunofluorescence
    - Staining the infected cultured cell lines with fluorescent antibody confirms the diagnosis
    - The viral culture process requires 2–10 days to complete
    - Primary method for vaccine production
  - DFA testing
    - DFAs are widely available, have variable sensitivity (range 47–93%) for 2009 H1N1 influenza virus, and a high specificity (≥96%). DFAs detect and distinguish between influenza A and B viruses but do not distinguish among different influenza A subtypes
    - The technique is more rapid (24 h) to result; it is less sensitive than culture methods
  - Serologic studies
    - Two samples should be collected per person. One sample within the first week (acute) of symptoms and a second sample (convalescent) 2–4 weeks later. If antibody levels increase from the first to the second sample, influenza infection likely occurred
    - Because of the length of time needed for a diagnosis of influenza by serologic testing, other diagnostic testing should be used if a more rapid diagnosis is needed
    - Inability to differentiate between current and previous infection. Cannot be used for rapid diagnosis
  - Rapid influenza diagnostic test (RIDT) (Table 25.12)
    - RIDTs are widely available but have variable sensitivity (range 10–70%) for detecting 2009 H1N1 influenza when compared with real-time reverse transcriptase polymerase chain reaction (rRT-PCR), and a negative RIDT result does not rule out influenza virus infection
    - RIDTs have a high specificity (>95%). Depending on which commercially available RIDT is used, the test can either (1) detect and distinguish between influenza A and B viruses or (2) detect both influenza A and B but not distinguish between influenza A and B viruses
    - Fastest method of currently available diagnostic tools. Result may be obtained in <30 min
    - However, the technique has a sensitivity of 70–80%
- Molecular methods (Table 25.11)
  - General information
    - Nucleic acid amplification tests, including rRT-PCR, are the most sensitive and specific influenza diagnostic tests
    - But they may not be readily available, obtaining test results may take one to several days, and test performance depends on the individual rRT-PCR assay. As with any assay, false negatives can occur
    - Not all nucleic acid amplification assays can specifically differentiate 2009 H1N1 influenza virus from other influenza A viruses
    - Several rRT-PCR assays have been evaluated and authorized by the FDA under an emergency use authorization (EUA) to diagnose 2009 H1N1 influenza virus infection, such as Roche

Table 25.12 Commercially available rapid point of care influenza detection kits (Adapted from http://www.cdc.gov/flu/professionals/diagnosis/testing_algorithm.htm)

Procedure (manufacturer/distributor)	Influenza virus types detected	Approved specimens[a]	Test time
3M™ Rapid Detection Flu A + B Test[d, f] (3M)	A and B	NP[b] swab/aspirate Nasal wash/aspirate	15 min
BinaxNOW® Influenza A and B[e, f] (Alere)	A and B	NP[b] swab Nasal wash/aspirate/swab	15 min
BioSign® Flu A + B[d, f] (Princeton BioMedtech)	A and B	NP[b] swab/aspirate/wash, nasal swab	15 min
Clearview® Exact Influenza A and B[d, f] (Alere)	A and B	Nasal swab	15 min
Directigen™ EZ Flu A and B[d, f] (Becton–Dickinson)	A and B	NP[b] wash/aspirate/swab Throat swab	15 min
OSOM® Influenza A and B[d, f] (Genzyme)	A and B	Nasal swab	10 min
QuickVue® Influenza Test[c, e] (Quidel)	A or B	Nasal wash/aspirate/swab	10 min
QuickVue® Influenza A and B Test[e, f] (Quidel)	A and B	NP[b] swab Nasal wash/aspirate/swab	10 min
SAS™ FluAlert A and B[d, f] (SA Scientific)	A and B	Nasal wash/aspirate	15 min
SAS™ FluAlert A[c, e] (SA Scientific)	A only	Nasal wash/aspirate	15 min
SAS™ FluAlert B[c, e] (SA Scientific)	B only	Nasal wash/aspirate	15 min
TRU FLU®[d, f] (Meridian Bioscience)	A and B	NP[b] aspirate/swab Nasal wash	15 min
XPECT™ Flu A and B[d, f] (Remel/Thermofisher)	A and B	Nasal wash/swab Throat swab	15 min

[a]List may not include all test kits approved by the US Food and Drug Administration. Discontinued tests not included. Approved respiratory specimens according to manufacturer's package insert. Note that test performance may vary if other respiratory specimens are used
[b]*NP* nasopharyngeal
[c]Does not distinguish between influenza A and B virus infections when used alone
[d]Moderately complex test – requires specific laboratory certification
[e]CLIA-waived test. Can be used in any office setting. Requires a certificate of waiver or higher laboratory certification
[f]Distinguishes between influenza A and B virus infections

real-time-ready influenza A/H1N1 test, Cepheid Xpert® flu A panel for the diagnosis of 2009 H1N1 Influenza virus infection, ELITech Molecular Diagnostics 2009 H1N1 influenza A virus real-time RT-PCR test, Focus Diagnostics influenza A H1N1 (2009) real-time RT-PCR, IMDx 2009 influenza A H1N1 real-time RT-PCR assay from IntelligentMDx, and QIAGEN *artus*® influenza A H1N1 2009 LC RT-PCR kit

– RT-PCR: artus® influenza A H1N1 2009 LC RT-PCR kit (QIAGEN Diagnostics)
  - The assay is for the detection of influenza A viral RNA and the detection and differentiation of 2009 H1N1 influenza virus RNA in nasopharyngeal swabs of symptomatic patients
  - The assay utilizes the EZ1 Advanced instrument with the EZ1 DSP Virus Card v. 2.0 (QIAGEN) and the EZ1 DSP virus kit (QIAGEN) for viral nucleic acid extraction. The LightCycler 2.0 instrument with software v. 4.1 (Roche) is used for amplification and detection
  - The limit of detection for influenza A (seasonal H1N1) is 57.0 TCID 50 (tissue culture infectious dose 50) /ml, and the limit of detection for 2009 H1N1 influenza is 4.217 PFU (plaque-forming unit)/ml
– RT-PCR: Cepheid Xpert® flu panel (ASR)
  - Accurate detection and differentiation of Influenza A from Influenza B infection and simultaneous identification of 2009 H1N1 flu strain

- Less than 2 minutes hands-on time improves lab workflow efficiencies
- Accepts nasal aspirate/washes (NA/W) or nasopharyngeal (NP) swab to accommodate wide range of specimen types

– RT-PCR: Roche real-time-ready influenza A/H1N1 detection set
  - In mid-November 2009, the test received emergency use authorization (EUA) from the US FDA
  - The test for the 2009 H1N1 influenza virus detects RNA from the 2009 H1N1 influenza A virus and provides a rapid means of identification of patients infected with this virus duration of the declaration of emergency
– Influenza antiviral resistance test
  - In the US, four antiviral drugs are FDA approved for use against influenza: amantadine, rimantadine, zanamivir (Relenza®), and oseltamivir (Tamiflu®). The adamantane drugs (amantadine and rimantadine) are approved for influenza A, while the neuraminidase inhibitor drugs (zanamivir and oseltamivir) are approved for both influenza A and influenza B
  - To date, 2009 H1N1 influenza, influenza A (H3N2), and influenza B viruses have been detected by surveillance. All three virus strains typically are sensitive to oseltamivir and zanamivir, but the influenza A strains have been resistant to the adamantanes. While, sporadic oseltamivir-resistant 2009 influenza A (H1N1) virus infections were identified
  - The oseltamivir-resistant 2009 influenza A (H1N1) is commonly described point mutation in the virus neuraminidase gene (histidine to tyrosine at position 275 of the N1 neuraminidase, commonly referred to as H274Y in N2 numbering), which is known to confer high-level resistance to oseltamivir
  - Influenza viruses A (H1N1) carrying the H274Y mutation have reduced ability to replicate and transmit efficiently when compared with parental, susceptible virus, but the clinical implications of infection with these viruses have been largely unknown
  - RT-PCR along with sequencing (Sanger sequence or next gen sequence such as pyrosequence) is usually used for influenza antiviral resistance test

- Sensitivity and specificity
  – Sensitivity of DFA methodology is dependent upon adequacy of the specimen, i.e., >20 cells. Otherwise, specimen may be inadequate for accurate interpretation resulting in false negatives
  – Rapid diagnostic testing is approximately >70% sensitive for detecting influenza and approximately >90% specific. Thus, as many as 30% of samples that would be positive for influenza by viral culture may give a negative rapid test result; some rapid test results may indicate influenza when a person is not infected with influenza
- Pitfalls
  – Due to the length of time required to perform viral culture, the assay has poor efficacy as results are obtained much after the patient has left the office or well past the time when drug therapy could be effective
  – Development of PCR-based assays must always consider antigenic drift and random mutations due to viral evolution that may result in false negatives

25.10.4 Clinical Utility

- Because of cost, availability, and sensitivity issues, diagnosis of influenza is often based on clinical criteria and presentation
- RT-PCR assays provide a rapid and specific diagnosis of influenza to allow for early therapeutic intervention and prophylactic treatment in high-risk patients, i.e., geriatric care facility
- Molecular diagnosis will play a large role in epidemiologic surveillance, vaccine strain selection, and surveillance of emergent novel influenza viruses, i.e., the Hong Kong H5N1 outbreak with sequence analysis

Fig. 25.26 Influenza A epidemic information

25.11 Avian Influenza (Bird) and Influenza (Flu) A Viruses

25.11.1 General Characteristics

- Influenza viruses that infect birds are called avian influenza viruses, and commonly known as "bird flu." Only influenza A viruses and subtypes infect birds
- There are substantial genetic differences between the subtypes that typically infect both people and birds. Within subtypes of avian influenza A viruses, there also are different strains
- The incubation period of avian influenza A virus is typically 2–5 days but can be as long as 8–17 days
- These influenza viruses occur naturally among birds
  - Wild birds worldwide carry the viruses in their intestines, but usually do not get sick from them
  - However, avian influenza is very contagious among birds and can make some domesticated birds, including chickens, ducks, and turkeys, very sick and kill them
- There are many different subtypes of type A influenza viruses (Fig. 25.26)
  - These subtypes differ because of changes in certain proteins on the surface of the influenza A virus (hemagglutinin and neuraminidase proteins)
  - There are 16 known hemagglutinin subtypes and 9 known neuraminidase subtypes of influenza A viruses
  - Many different combinations of hemagglutinin and neuraminidase proteins are possible. Each combination represents a different subtype
  - Avian influenza A H5 and H7 viruses can be distinguished as "low pathogenic" and "high pathogenic" forms on the basis of genetic features of the virus and the severity of the illness they cause in poultry; influenza H9 virus has been identified only in a "low pathogenicity" form
  - Each of these three avian influenza A viruses (H5, H7, and H9) theoretically can be partnered with any one of nine neuraminidase surface proteins; thus, there are potentially nine different forms of each subtype (e.g., H5N1, H5N2, H5N3, H5N9)

25.11.2 Clinical Presentation

- The reported signs and symptoms of avian influenza in humans have ranged from eye infections (conjunctivitis) to influenza-like illness symptoms (e.g., fever, cough, sore throat, muscle aches) to severe respiratory illness (e.g., pneumonia, acute respiratory distress, viral pneumonia) sometimes accompanied by nausea, diarrhea, vomiting, and neurologic changes
- CDC and WHO recommend oseltamivir, a prescription antiviral medication, for treatment and prevention of human infection with avian influenza A viruses. Analyses of available H5N1 viruses circulating worldwide suggest that most viruses are susceptible to oseltamivir. However, some evidence of resistance to oseltamivir has been reported in H5N1 viruses isolated from some human H5N1 cases. Monitoring for antiviral

resistance among avian influenza A viruses is important and ongoing
- The first avian influenza virus to infect humans occurred in Hong Kong in 1997. The epidemic was linked to chickens and classified as avian influenza A (H5N1). Human cases of avian influenza A (H5N1) have since been reported in Asia, Africa, Europe, Indonesia, Vietnam, the Pacific, and the Near East. Hundreds of people have become sick with this virus. Slightly more than 60% of those who became ill have died

25.11.3 Diagnostic Methods

- Specimen: see Influenza A, B, and C section
- Viral culture: see Influenza A, B, and C section
- Serological test: see Influenza A, B, and C section
- Molecular test
  - Real-time RT-PCR
    - On February 3, 2006, the FDA announced clearance of the influenza A/H5 (Asian lineage) virus real-time RT-PCR
    - Primer and probe set and inactivated virus as a source of positive RNA control for the in vitro qualitative detection of highly pathogenic influenza A/H5 virus (Asian lineage)
    - Two genetic lineages of influenza A/H5 viruses exist: Eurasian (Asian) and North American. The primer and probe set, developed at CDC, is designed to detect highly pathogenic influenza A/H5 viruses from the Asian lineage associated with recent laboratory-confirmed infections of avian influenza in humans in east Asia and, most recently, in Turkey and Iraq
    - The test is limited to laboratories designated by the Laboratory Response Network (LRN), which consists of approximately 140 US laboratories in 50 states
- Limitation
  - Due to the limitation of the assay, negative results do not preclude influenza virus infection and should not be used as the sole basis for treatment or other patient management decisions

25.12 Adenovirus

25.12.1 General Characteristics

- Adenovirus is ubiquitous in humans and is endemic
- Adenoviruses are medium-sized (90–100 nm), nonenveloped icosahedral viruses containing double-stranded DNA (Fig. 25.27)
- There are at least 52 known immunologically distinct types (seven subgenera: A–G) that can cause human infections. Types 4 and 7 are common in military recruit outbreaks; type 14 is commonly associated with severe and sometimes fatal respiratory illnesses
- Adenovirus transcription can be defined as a two-phase event, early and late, occurring before and after DNA replication
- Early transcription is accompanied by a complex series of splicing events, with four early "cassettes" of gene termed E1, E2, E3, and E4. Early genes facilitate DNA replication and result in the transcription and translation of the late genes (Fig. 25.28)
- Adenovirus produces cytolysis in different tissues and induces host inflammatory responses and cytokine production
- The route of entry is usually by droplet nuclei or by oral infection. Transmission of adenovirus is via direct contact, the fecal–oral route, and occasionally waterborne

25.12.2 Clinical Presentation

- Most adults have measurable titers of antiadenovirus antibodies, implying prior infection. Most adenovirus infections occur early in life, and by age 10, most children

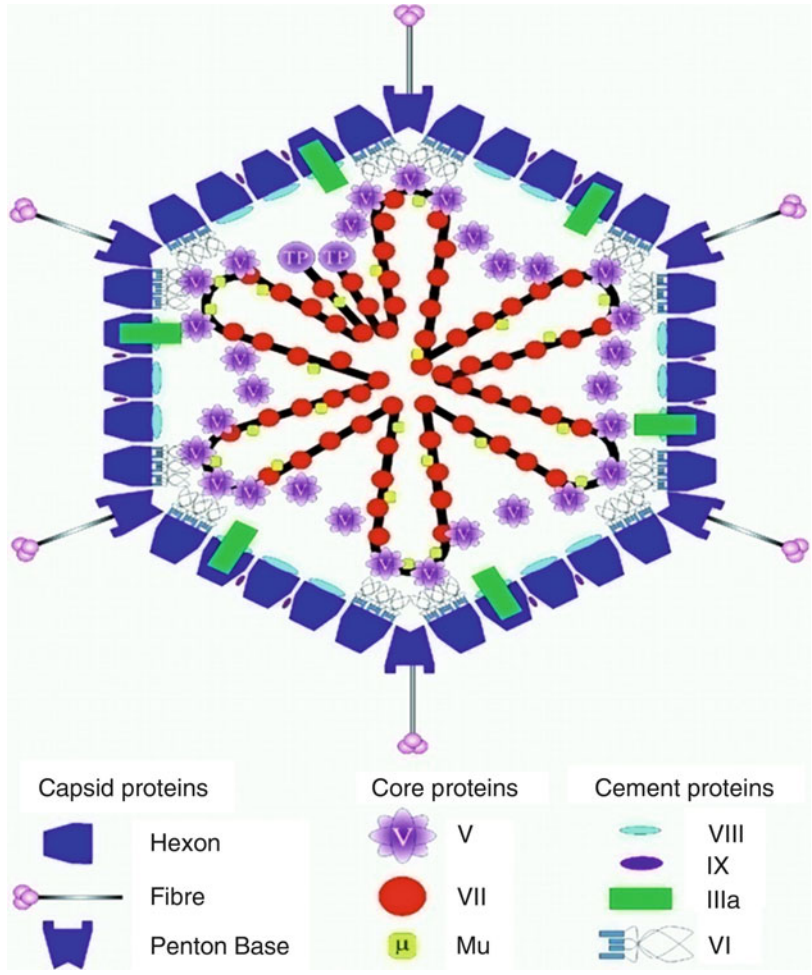


Fig. 25.27 Structure of adenovirus (Adapted from Russell WC. J Gen Virol 2008)

have been infected by at least one serotype. However, most infections are asymptomatic

- Some adenovirus types can establish persistent subclinical infections in tonsils, adenoids, and intestines of infected hosts with viral shedding occurring for as long as several months to years
- Adenovirus may infect multiple organ systems and is recognized as the etiologic agent of a variety of diverse syndromes: acute respiratory disease (ARD), pharyngoconjunctival fever, epidemic keratoconjunctivitis, acute hemorrhagic cystitis, gastroenteritis, and adenoviral infections in immunocompromised hosts
- Infection with serotypes 1, 2, and 5 is most frequent in the first years of life, and all serotypes can occur during any season, but infections are most frequent in late winter and early spring. Historically, 1–5% of all respiratory infections have been caused by adenovirus. The prevalence of adenovirus respiratory infection in children has ranged from 2% to 14%. Adenovirus also causes 5–15% of acute diarrheal infections in children

25.12.3 Diagnostic Methods

- Specimens
  - Respiratory, stool, and blood
- Conventional tests and problems
  - Viral culture
    - The "gold" standard for adenovirus testing is viral culture

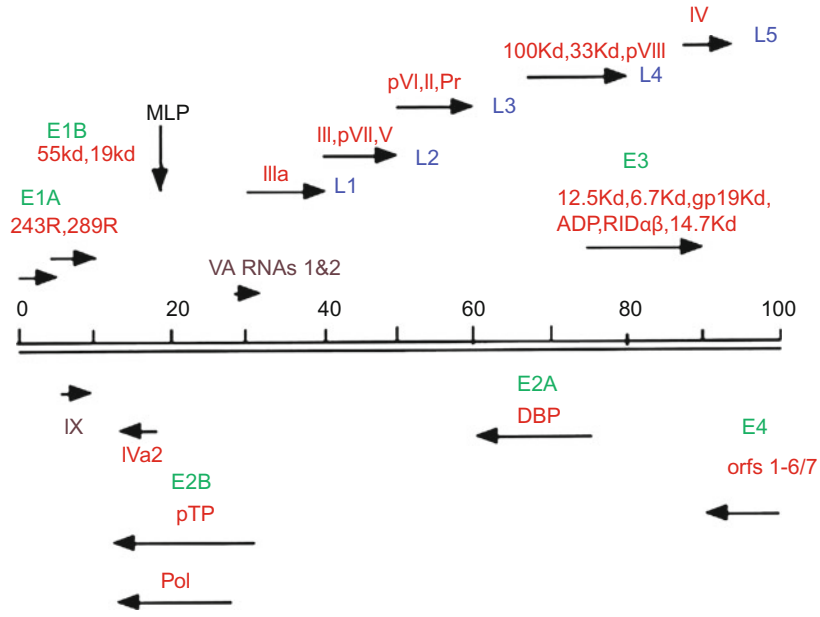


Fig. 25.28 Transcription of the adenovirus genome. The early transcripts are outlined in *green*, the late in *blue*. *Arrows* indicate the direction of transcription. The gene locations of the VA RNAs (nontranslated RNAs) are denoted in *brown*. MLP, major late promoter (Adapted from Russell WC. J Gen Virol 2008)

- Collected samples are layered on a monolayer of cells in test tubes. If present, adenovirus will present as cytopathic effect of rounded cells
- The best cell line for testing adenovirus is A549. This is a human carcinoma continuous cell line
- When samples are collected within 1–3 days of clinical onset, cell culture generally is positive within 4–7 days. Samples collected after 3 days may take 1–3 weeks to produce cytopathic effect
- Slow growth – not ideal for acute identification: The viral culture can confirm an adenovirus diagnosis, but it may not provide timely results for immediate patient care

– Shell vial culture
  - Shell vial is another cell culture test but the results are ready in 3 days
  - Vials of A549 cells are inoculated with collected samples and centrifuged. The vials are then incubated and stained at day three with immunofluorescent antibodies specific to adenovirus
  - The cells infected with adenovirus will light up with examination under a fluorescent microscope
  - It is recommended that the shell vial testing is for diagnosing ocular adenovirus infection

– Serologic studies
  - Seroreactivity to adenovirus is common; measuring adenovirus-specific IgM or rising level of IgG antibody can be helpful. Positive adenovirus titers occur in 50% of individuals >4 years old
  - Serology is less useful in the acute clinical setting
  - For a serologic diagnosis, serum should be obtained as early as possible in the clinical course, followed by a second titer 2–4 weeks later. A 4-fold rise in acute titers to convalescent titers is diagnostic
  - Serological assays have been used over the years to measure adenovirus-specific IgM or rising levels of IgG antibody; however, serology is no longer for diagnosis because it is slow and less specific than antigen or DNA detection

– Immunofluorescence
  - DFA or indirect IFA may be used for direct examination of tissue. It uses a mouse antibody against an adenovirus group-specific hexon antigen

  - This method has shown to have poor sensitivity compared to other tests, and shell viral cultures are run concurrently to improve sensitivity
  - Several commercial kits are (1) Light Diagnostics™ IFA and DFA respiratory viral screening kits for respiratory syncytial virus (RSV), Flu A, B, parainfluenza 1, 2, 3, and adenoviruses (Millipore Corp.); and (2) MonoFluo® RSV, Flu A, B, parainfluenza 1,2, 3, and adenoviruses (Bio-Rad Laboratory)
- Molecular methods
  - PCR: Marked improvement in sensitivity when compared to viral culture
  - Most of these assays amplify conserved regions of the hexon gene
  - Most of real-time quantitative PCRs are designed to detect adenovirus DNA from all major subgroups of the virus
  - Four common multiplex assays are the MultiCode-Plx assay (EraGen), the ResPlex III assay (Millipore), ADENOVIRUS R-gene™ CE Marked PCR Assay, which detects and quantifies all the 52 adenovirus serotypes, and the xTAG RVP assay (Luminex Molecular Diagnostics). The xTAG RVP is the only FDA-approved test and has a reported sensitivity of 78.3% and a specificity of 100%
- Pitfalls
  - Development of PCR-based assays must always consider antigenic drift and random mutations due to viral evolution that may result in false negatives
  - Important to note that PCR assays are not standardized and variations in sample handling and laboratory methods can affect the sensitivity of the assay

25.12.4 Clinical Utility

- PCR assays provide a rapid and specific diagnosis of adenovirus to allow for early therapeutic intervention and prophylactic treatment in high-risk patients, i.e., geriatric care facility and immunocompromised patients
- Detection of high viral load in blood and monitoring of viral load during treatment can correlate with disseminated adenovirus disease in immunosuppressed patients. Adenovirus is increasingly important in solid organ and bone marrow transplant patients; infections in pediatric transplant cases can exceed 22% with mortality rates as high as 60%

25.13 Respiratory Syncytial Virus

25.13.1 General Characteristics

- RSV is a negative sense, enveloped RNA virus. The virion is variable in shape and size (120–300 nm), is unstable in the environment (surviving only a few hours on environmental surfaces), and is readily inactivated with soap and water and disinfectants
- RSV is a member of the paramyxovirus family that produces a characteristic fusion of human cells (syncytial effect) in tissue culture
- RSV is a medium-sized, enveloped virus with an antisense single-stranded RNA genome
- RSV has two heterotypic strains of viruses that are antigenically distinct, and are classified as subgroups A and B
- The major difference between these subgroups is the antigenic properties of the G surface glycoprotein
- Transmission is from aerosolized respiratory droplets via close contact with infected persons or contact with contaminated surfaces
- Most prevalent in infants aged 2–6 months, but children of any age with underlying cardiac or pulmonary disease or who are immunocompromised are at risk for serious complications from RSV infection

25.13.2 Clinical Presentation

- RSV infections typically occur in temperate climates during late fall through early spring and account for 5–15% of community-acquired pneumonias. 50%

of children ≤1year are infected and 100% are infected by 3 years. Immunity wanes with age. Disease may reoccur in patients >65 years
- Two subtypes have been identified. Subtype A involves a severe clinical presentation and predominates in most outbreaks. Subtype B predominates in most asymptomatic strains of the virus that the majority of the population experiences
- RSV bronchiolitis presents with a 2–3-day "prodromal" phase, which resembles a common viral upper respiratory tract infection. Additional symptoms include rhinorrhea, wheezing, coughing, low-grade fever, and pneumonia. Circumoral and nail bed cyanosis may occur in severely affected infants
- In the majority of patients with RSV bronchiolitis, symptoms resolve within 5–7 days

25.13.3 Diagnostic Methods

- Specimens – respiratory swabs and bronchoalveolar lavage
- Conventional tests and problems
  - Viral culture
    - RSV has a high liability and any specimens should be transported to the laboratory promptly and inoculated into cell cultures
    - Nasopharyngeal aspirates, nasal washes, or tracheal secretions are generally the best specimens for isolation
    - Specimens should not be subjected to major temperature changes such as freezing and thawing
    - Human heteroploid cells, such as HEP2 and HeLa, generally provide the best tissue culture for the isolation of RSV
    - RSV produces a characteristic CPE consisting of syncytia formation and appears in 4– 5 days
    - Decreased sensitivity in adults from reduced viral shedding during acute infections as compared to adolescents
  - Rapid antigen detection
    - Direct and indirect immunofluorescence (IF) methods
      - Ability to perform direct screening with low cost
      - Both direct and indirect IF utilizing either polyclonal or monoclonal antibodies are available which possess a high degree of sensitivity and specificity
      - The general sensitivity of IF is 80–90% and for monoclonal antibody 95–100%. The specificity is at least 94%
      - IF techniques are fast and easy to perform, but the interpretation of results is subjective and the specimen must contain adequate nasopharyngeal cells
      - Incorrect and indeterminate results may occur for specimens with few epithelial cells or when nonspecific antibody reagents are used
    - ELISA (BD Directigen™ RSV)
      - ELISA assays do not require expensive laboratory equipment, take only 15–20 min, and are inexpensive compared to cell culture
      - ELISA techniques offer the advantages of objective interpretation, speed, and the possibility of screening a large number of specimens. Disadvantages include a generally poorer sensitivity and a "gray zone" of equivocal results, which requires confirmation by a time-consuming blocking ELISA procedure
    - Serology
      - Acute and convalescent phase sera are required for the serologic diagnosis of RSV
      - A fourfold increase in antibody titer or the appearance of specific IgM antibody is required for serologic confirmation of infection
      - It includes complement fixation (CF) antibody titers, ELISA, neutralization to specific A and B subtypes, and indirect IF

  - It is unlikely to be of help in the management of the patient because of the length of time required. Furthermore, the serological response in young infants may be poor and not detectable by some antibody assays
  - Seroconversion does not occur for at least 2 weeks and may require 4–6 weeks. CFTs are less sensitive than neutralization and ELISA assays
- Molecular methods
  - NASBA-beacon
    - NucliSens EasyQ RSV A + B assay (bioMérieux)
    - Real-time PCR-based assay utilizing NASBA technology containing internal control and specific molecular beacon mix targeting fusion protein of RSV
    - Limit of detection is 22 input copies of RSV
    - Improved time to result, <4 h
  - Cepheid SmartCycler® System – RSV (ASR)
    - The RSV Type A or B Primer ASR contains primers designed to detect RSV type A or B
    - The RSV Type A or B Probe ASR contains FAM-labeled probes designed to detect RSV type A or B
  - xTAG® Respiratory Viral Panel (RVP) (Luminex)
    - The RVP is a qualitative nucleic acid multiplex test intended for the simultaneous detection and identification of multiple respiratory virus nucleic acids in nasopharyngeal swabs from individuals suspected of respiratory tract infections
    - The xTAG RVP has been designed to simultaneously probe for 12 viral targets in a single patient specimen: influenza A, influenza A subtype H1, influenza A subtype H3, influenza B, RSV subtype A, RSV subtype B, parainfluenza 1, parainfluenza 2, and parainfluenza 3, human metapneumovirus, rhinovirus, and adenovirus. All viruses are probed for in a single multiplex reaction
    - xTAG RVP incorporates multiplex Reverse RT-PCR and multiplex target-specific primer extension (TSPE) with Luminex Molecular Diagnostic proprietary Universal Tag sorting system on the Luminex xMAP platform. xTAG RVP is compatible with both the Luminex 100/200 IS systems
- Sensitivity and specificity
  - Improved sensitivity and specificity when compared to conventional tests, particularly in the adult population
  - No cross-reactivity was shown for PIV13; influenza A and B; measles; adenovirus types 1 and 5; hMPV A1, A2, B1, and B2 indicating that the assay is specific for RSV
- Pitfalls
  - Although the RSV virus can be cultured from an infected individual, the delay in definitive diagnosis of 3–5 days decreases the clinical utility

25.13.4 Clinical Utility

- Ease of assay, rapid turnaround time, and improved sensitivity have enhanced clinical utility in early detection of respiratory illness

25.14 Severe Acute Respiratory Syndrome

25.14.1 General Characteristics

- Severe acute respiratory syndrome (SARS) is a recently identified respiratory illness that first infected individuals in parts of Asia, North America, and Europe in late 2002 and early 2003
- The SARS-associated coronavirus belongs to the *Coronaviridae* family, a family of large, enveloped positive-stranded RNA viruses. It is the first example of a coronavirus causing serious disease in humans
- The SARS-coronavirus (SARS-CoV) genome is 29,272 nucleotides in length with 41% being G/C residues

- SARS is spread mainly through contact with infected saliva or droplets from coughing. Vertical transmission from mother to infant does not appear to occur

25.14.2 Clinical Presentation

- The SARS virus produces an atypical pneumonia that often leads to respiratory failure, with pulmonary edema and hyaline membrane formation similar to that seen with adult respiratory distress syndrome (ARDS)
- During the early phase of the disease, fever greater than 38 °C (100.4 °F) is the hallmark symptom. This finding is often associated by myalgia, rigors, and other flu-like symptoms
- During the second week, patients develop a dry, nonproductive cough, shortness of breath, and lung infiltrates with rapid progression to respiratory distress
- The cause of death is respiratory failure, with the best predictor of mortality being old age. The fatality of SARS is less than 1% for people aged 24 or younger, 6% for those 25–44, 15% for those 45–64, and more than 50% for those over 65
- Except for ventilation, no effective treatment is currently available

25.14.3 Diagnostic Methods

- Specimens
  - Respiratory sample: nasal wash, nasopharyngeal swab, BAL, bronchial wash, or sputum
  - Transport: 1 ml (minimum volume 0.5ml for adults and pediatrics) respiratory sample in viral transport media (Microtest M4) or in sterile leakproof container at 2–8 °C
  - Unacceptable conditions: Dry swabs are not acceptable. Respiratory aspirates collected in containers with tubing as samples tend to leak, compromising the specimen
- Conventional tests and problems
  - Viral culture
    - Requires BSL3 facility
    - Difficulty in culturing the virus from infected individual late in the outbreak during late stages of illness due to possible genetic drift of virus
  - Serologic studies
    - Utility of serologic testing is poor due to late seroconversion of infected patients, i.e., 2–4 weeks
    - Antigen detection with monoclonal antibodies or monospecific polyclonal antibody against the N protein was found to be a sensitive and specific test for the diagnosis of SARS
    - For antibody testing, the indirect immunofluorescent antibody test is more commonly performed than the neutralizing antibody test since the former involves minimal manipulation of infectious virus and therefore carries less risk of a biohazard. The test is generally not useful during the first week of illness
- Molecular methods
  - RT-PCR
    - Multiple RT-PCR assays have been developed to detect SARS RNA in clinical specimens utilizing nested, nonnested, one-step or two-step conventional, or real-time RT-PCR assays
    - Most nucleic acid amplification tests are designed with the Orf1b or nucleoprotein gene. The latter gene has the theoretical advantage of being more abundant in infected cells and therefore of higher sensitivity, but this has not been clearly proven in clinical studies
    - LightCycler SARS-CoV quantitation kit [Roche Diagnostics Corporation] for use with the LightCycler instrument
      - Ready to use which amplifies a 180-bp target sequence of the replicase 1AB/polymerase gene of SARS CoV
      - The analytical sensitivity for the SARS-CoV test is about 20 copies/PCR reaction
    - RealArt HPA-coronavirus RT-PCR kits [artus] for use with the LightCycler

instrument, the ABI PRISM 7000, 7700, and 7900H instruments
  - Amplifies an 80-bp region of the SARS-CoV genome
- EraGen Biosciences MultiCode-RTx (research only)
  - EraGen's platform increases size of the genetic "alphabet" from the two DNA base pairs to six pairs with the development of eight new synthetic bases
  - It is a new multiplexed real-time PCR platform
  - Only standard PCR primers need to be designed. Since reporters are placed directly onto the primers and not on probes
  - It targets nucleocapsid *(nuc)* or polymerase *(pol)* gene
- Sensitivity and specificity
  - Sensitivity of commercial assays ranged from 36% to 80% and specificity ranged from 80% to 100%
  - The absolute sensitivity of the RT-PCR assays ranged from 10 to 100 genome equivalents per reaction
- Pitfalls
  - When present, SARS antibodies can be detected in serum at any point during the course of the disease. However, most patients do not seroconvert until after the second week, highlighting the importance of an RT-PCR assay for early diagnosis of the virus
  - Positive results must be confirmed by repeat testing using an aliquot of the original specimen and/or another laboratory before reporting. Alternatively, testing of a second gene region may be helpful. Furthermore, testing of one sample from a single source does not rule out the presence of SARS-associated coronavirus
  - A negative result does not rule out SARS as the presence of PCR inhibitors in the patient specimen, poor RNA quality, or nucleic acid concentrations below the level of detection of the assay may occur

25.14.4 Clinical Utility

- During the first week, serum and plasma are preferred for RT-PCR. Between 1 and 3 weeks, these sample types are less effective; stool and respiratory samples are the preferred types. After 3 weeks, stool is the preferred sample type for RT-PCR. Viral load in the upper respiratory tract and feces is low during the first 4 days of infections and peaks at approximately the 10th day of illness
- During the 10th–15th day of illness, high viral loads are independent predictors of poor clinical outcomes

25.15 Enterovirus

25.15.1 General Characteristics

- Enteroviruses represent one of the most common human viruses, affecting an estimate 50 million individuals in the US and potentially one billion worldwide. Enterovirus infections most commonly occur in temperate zones during the summer and early fall
- Enteroviruses are a diverse group of small, nonenveloped ssRNA viruses of 6–7 kb that are transmitted by the fecal–oral route. Enteroviruses comprise a group of human viruses that includes polioviruses, echoviruses, coxsackie A viruses, coxsackie B viruses, and various enterovirus subtypes. The original classification of human enteroviruses has been substituted by a taxonomic scheme based on molecular and biological properties of the viruses. This revised classification recognizes at least 90 subtypes and separates them into four species
- Several modes of transmission exist for these viruses, including fecal–oral, respiratory, transplacental, perinatal, and self-inoculation modes, but the majority are fecal–oral
- Although enteroviruses undergo rapid replication in the GI tract, they rarely cause significant GI disease. Instead, they travel via the bloodstream to target organs where they further replicate and induce pathologic alteration

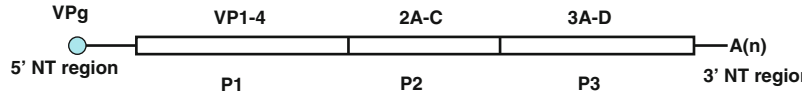


Fig. 25.29 Enterovirus genome: 7,450 nucleotide long single-stranded RNA virus with a 5′ NT region of 743 nt, a 6625 coding region, and 3′ polyA region (*VP* viral protein, *P* polypeptide, *NT* nontranslational region). PCR primers usually design to target to 5′ NT region

25.15.2 Clinical Presentation

- Most infections are subclinical, although may cause a variety of acute and chronic diseases
  - Acute: mild upper respiratory illness (common cold), febrile rash (hand, foot, and mouth disease and herpangina), aseptic meningitis, pleurodynia, encephalitis, acute flaccid paralysis, and neonatal sepsis-like disease
  - Chronic: myocarditis, cardiomyopathy, type 1 diabetes mellitus, and neuromuscular disease
- The highest incidence of enterovirus infection is in infants and young children

25.15.3 Diagnostic Methods

- Specimens
  - Non sterile sites: nasal/throat swabs, and feces where the presence of the virus might merely indicate coincidental carriage
  - Sterile sites: vesicular fluid, CSF, serum, urine, or gathered at autopsy, are more reliable
  - Samples transported in viral transport media, were either transported directly to the laboratory or were stored at 4 °C for a maximum of 24
- Conventional tests and problems
  - Viral culture and shell vial culture
    - Gold standard to detect enterovirus
    - Time-consuming methods and insensitive methods, relying on the presence of viable virus
    - Inability to fully characterize some enterovirus strains associated with late inadequate collection, handling and processing of samples, or because of intrinsic insensitivity to cell lines used
    - It can take up to 8 days for CPE to appear when virus is present in low titers (e.g., in CSF specimens), and some type A coxsackievirus do not grow in cell culture
    - Although shell vial culture using monoclonal antibodies has decreased the culture time compared with that for tube culture, it is less sensitive than conventional culture
  - Serology
    - Serotype is usually irrelevant to individual management
    - The absence of a widely shared antigen has hampered the development of immunoassays for the enterovirus
    - Reports of monoclonal antibodies that cross-react with multiple enterovirus serotypes are promising, but further testing is required to determine the clinical relevance of those observations
- Molecular methods
  - Real time RT-PCR – ABI Prism (Applied Biosystems, Foster City, Calif.)
    - Improved speed and accuracy using TaqMan assay platform
    - Targets conserve sequences of the 5′ NTR and VP 1 and 2 (capsid protein). The 5′ NTR is the most highly conserved region and is involved in viral protein translation (Fig. 25.29)
    - An enterovirus real-time TaqMan PCR analysis of serum or plasma may be a good alternative for the enterovirus culture of feces, particularly in neonates with sepsis

- AnDiaTec® Enterovirus real time RT-PCR Kit
  - The kit is a screening assay for the detection of enteroviruses (coxsackie A, coxsackie B, and echovirus) in the capillary system of the LightCycler (Roche)
  - Targets conserve sequences of the 5′ NTR
- Cepheid SmartCycler® System
  - It detects a 115-bp region of the 5′ NTR
- NASBA
  - NASBA-electrochemiluminescently (ECL) and NASBA-beacon are not significantly different in sensitivity and specificity
    - Targets conserve sequences of the 5′NTR
  - NASBA-ECL
    - NucliSens basic kit has proved of equal or greater sensitivity for detection of enteroviruses
    - In the NucliSens basic kit, amplified RNA products are detected by hybridization using ECL-labeled probes, a highly sensitive methodology
  - NASBA-beacon
    - NucliSens EasyQ Enterovirus Test (bioMerieux, Durham, NC), which utilizes real-time molecular beacons as probes (NASBA-beacon)
    - Real-time RT-PCR using TaqMan to shorten both technical hands-on time and time to result
- Enterovirus consensus, Argene Biosoft (for research use only in the US)
  - One-step RT-PCR of all enterovirus serotypes in one single reaction tube
  - Amplified region is in the 5′ NCR of the genome
  - Detection is performed with a biotinylated enterovirus generic probe

- Sensitivity and specificity
  - NASBA-ECL and NASBA-beacon were similar in sensitivity, 100% and 94.5%, respectively
  - RT-PCR sensitivity is 97%, while culture sensitivity is 54.5%
  - Real-time RT-PCR sensitivity is 100% and the specificity is 96.2%
- Pitfalls
  - Parechoviruses may cause similar clinical illnesses but are not detected by enterovirus testing
  - Poor handling of CSF or CSF collected during late infection can yield false-negative results
  - Enteroviruses can be shed in high titers in stool for prolonged periods. Therefore, a positive result in stool alone may not correlate with current disease

25.15.4 Clinical Utility

- Enteroviral meningitis is common in the US and leads to a large number of hospitalizations per year due to an inability to distinguish from bacterial meningitis. Therefore, enterovirus differentiation from bacterial illness can significantly reduce hospitalizations, antimicrobial use, and diagnostic testing
- Rarely, dual infections (enteroviral and bacterial) can occur. Therefore, a positive enterovirus result with clinical features incompatible with benign viral meningitis should not lead to discontinuation of antibiotics

25.16 JC/BK Virus

25.16.1 General Characteristics

- The BK virus (BKV) and JC virus (JCV) are small, nonenveloped, closed circular, double-stranded DNA virus and belong to human polyomaviruses, members of the Papovaviridae family
- They were first isolated in 1971 and named JC and BK after the initials of the patients in which they were first discovered. JCV was isolated from the brain tissue of a patient with progressive multifocal leukoencephalopathy (PML); BKV was isolated from the urine of a renal transplant

patient who developed ureteral stenosis postoperatively
- BKV and JCV share 75% homology at the level of nucleotide sequence. Each is 70% homologous to simian virus 40 (SV40), which belongs to polyomaviruses and was introduced in the human population, between 1955 and 1963, by contaminated polio vaccines produced in SV40-infected monkey cells
- The two are not cross-reactive serologically, and serologic tests for antibodies are able to distinguish between BKV and JCV
- More than 70% of the adult population has antibodies to BKV and JCV, with primary infections typically occurring in childhood
- After an initial infection, polyomaviruses establish latency in various tissues. The primary sites of latency are uroepithelial cells for BK virus and B lymphocytes and renal tissue for JCV. Additional sites of latency for both viruses include the ureters, brain, and spleen

25.16.2 Clinical Presentation

- Both BKV and JCV have been associated with human tumors. The recent evidence that SV40 may be a cofactor in the etiology of specific human tumor types has raised again the interest on the two human polyomaviruses as possible agents involved in human oncogenesis
- In immunocompetent individuals, primary BKV infections usually cause a mild respiratory illness and, rarely, cystitis, whereas primary JCV infections are typically asymptomatic. BKV seroprevalence peaks at ~90% at age 5–9 years old
- Reactivation of latent as well as primary BKV and JCV infections may occur in immunocompromised individuals, i.e., organ transplantation, AIDS, and leukemia. BKV infections can lead to interstitial nephritis, tubulitis, hemorrhagic cystitis, and kidney allograft rejection
- JCV is responsible for progressive multifocal leukoencephalopathy (PML), a fatal demyelinating disease of the central nervous system seen in up to 70% of AIDS patients

25.16.3 Diagnostic Methods

- Specimens: urine, plasma, CSF, and tissue biopsy
- Conventional tests and problems
  - Cytology and immunohistocytochemistry for urine sample
    - It can be used to confirm BKV reactivation diagnosis
    - The urine sample can be stained with JC/BK and SV40 monoclonal antibodies
  - Viral culture
    - JC virus is difficult to culture
    - The most sensitive cell type for JCV is primary human fetal glial cells, which is not an easy reagent to acquire
    - BK virus will grow in common cell lines, such as human diploid fibroblasts, but several days and weeks are required before CPE is evident
  - Serologic studies
    - Hemagglutination inhibition or enzyme-linked immunosorbent assay methods can measure titers of antibodies to JCV and BKV
    - Serological tests of blood and CSF for anti-JCV and BKV antibodies are not useful in the diagnosis of PML and immunosuppressed individuals because antibodies to JCV and BKV are common and many patients with PML or immunosuppressed patients fail to develop a significant rise in antiviral antibody titers in serum or CSF
- Molecular methods (Fig. 25.30)
  - PCR, quantitative
    - Seminested PCR to measure serum BKV DNA has been shown to have a higher specificity of 88% (50% positive predictive value) to detect BKV nephropathy. It was also shown to have a sensitivity of 100% for detecting BKV nephropathy
    - Serial quantification of BKV DNA levels through PCR can aid in managing BKV nephritis by measuring increasing or decreasing activity levels noninvasively

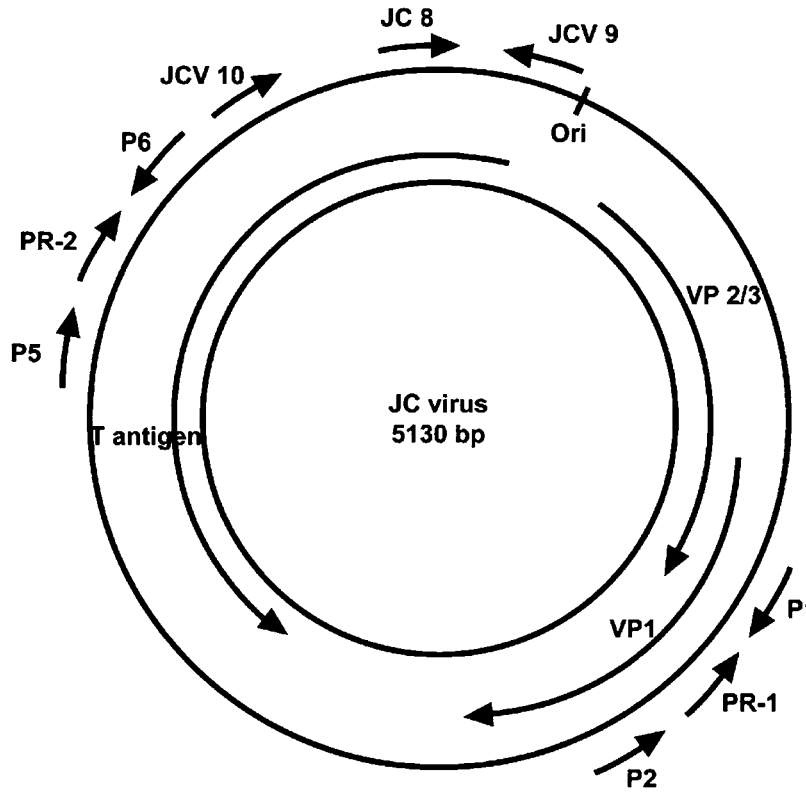


Fig. 25.30 JC virus genome structure

- PCR analysis of JCV DNA in spinal fluid is a noninvasive method of detecting active JCV infection which is 95% specific and 80% sensitive
- JC/BK Consensus Complete kit, Argene Biosoft (US: For research use only). The kit is for the detection and typing of JC and BK viruses by PCR and hybridization on microwell plates, and offers high sensitivity, up to 1 copy per PCR for BKV and 10 copies per PCR for JCV primers/probe product is designed to amplify JCV/BKV using 5′ nuclease real-time assay. The targeted sequence corresponds to a fragment of 197/198 bp located in the gene of large T antigen
- Real-time TaqMan PCR and LightCycler Probes (Homebrew) targets highly conserved sequence of JCV/BKV genomes (VP2 gene)

- Sensitivity and specificity
  - Analytical specificity: no cross-reactivity with HSV family viruses, simian virus, adenovirus, and HIV. Absolute sensitivity: 10 JC/BK virus detection
  - PCR has been able to detect JCV in CSF in 80–90% of PML patients
  - The specificity of diagnosis is influenced by the choice of primers and extraction methods but can approach 100%
- Pitfalls
  - Sequence variation of polyomavirus genome and within various JC/BK subtypes that may cause difficulty in primer and probe design
  - Competition between JC and BK viruses due to sensitivity may lead to false-negative PCR result

25.16.4 Clinical Utility

- Detection of the virus by PCR may be indicative of an active infection. Therefore, the identification of viral DNA may warrant the institution of antiviral therapies

and/or a decrease of immunosuppressive therapies
- Determination of viral DNA presence or concentration in transplant patients is useful in establishing the cause of allograft rejection. Viral load may also be useful in immunocompromised patients
- BKV nephropathy is associated with BK viremia of >5,000 copies/mL (plasma) and BK viremia >107 copies/mL and is seen in approximately 8% of kidney transplant recipients
- Though latency is typically associated with the absence of viremia, low levels (<2,000 copies/mL plasma) are seen in some asymptomatic individuals

25.17 HHV6

25.17.1 General Characteristics

- HHV6: two variants (A and B)
  - Type B is most common as primary infections in US, Japan, and Europe
  - Type A is most common among asymptomatic children in Africa
- Nearly 100% of population infected by 3 years of age
- Virus becomes latent or persistent until reactivated (immunocompromised hosts). Can be present in mononuclear cells, saliva, and the central nervous system
- 1% of population shows vertical transmission
- Congenital infection also reported

25.17.2 Clinical Presentation

- Most infants develop an undifferentiated febrile illness
- 20% present as roseola infantum
  - High fever, followed by maculopapular rash on face, neck, and trunk
  - Febrile seizures are most common neurologic complication
  - Encephalitis and encephalopathy is also reported
- Reactivated HHV6 presents as fever, rash, and encephalitis
- Differentiating reactivated from latent infection is very difficult, especially following stem cell or organ transplant

25.17.3 Diagnostic Methods

- Specimens: HHV6 can be isolated from peripheral blood mononuclear cells
- Conventional tests and problems
  - Serologic testing
    - Primary infection causes an IgM spike in the first week of illness followed by IgG rise approximately 2 weeks later. IgG antibody persists indefinitely
    - Tests include indirect immunofluorescence, immunoblot, and enzyme immunoassays. Viral culture can also be used but is time-consuming
    - However, gap between onset of infection and rise of titer does not allow a timely diagnosis
- Molecular methods
  - PCR
    - Qualitative nested polymerase chain reaction
    - Can be performed on plasma, mononuclear cells, whole blood, saliva, urine, cervical swabs, and placenta
    - Studies have shown results in plasma showed a sensitivity to diagnose reactivated or primary HHV6 infection with 75–95% sensitivity
    - Additionally, DNA can be detected in absence of viral replication
    - Specificity was reported to be 84% compared with viral culture
  - RT-PCR
    - Increased sensitivity and specificity (90% and 98%, respectively)
  - Quantitative PCR
    - Viral load may be helpful
    - Primary infection usually ranges from 4.4 to 4.9 log (10) genome equivalent copies (gec)/106 PBMCs

    - Asymptomatic individuals are lower (1.5–3.7)
    - Viral load with individuals with latent infection does overlap with immunocompromised host. However, plasma can be used to distinguish between the two (as latent individuals do not have plasma positive viral DNA)
    - Viral loads can be separated into very high (integrated virus), intermediate (primary, reactivated, or integrated virus) and low (latent infection)
- Pitfalls
  - Some overlap does occur with latent and immunocompromised hosts (as listed above)
  - Contamination is an issue, given that the techniques use signal amplification

25.18 Clinical Utility

- HHV6 can be especially dangerous in immunocompromised hosts
- HHV6 is present in a very large percentage of the population (100% of individuals greater than 3 years of age)
- Testing is required to determine if the host has a primary or latent infection
- Standard testing (viral capture or serology) is too time-consuming
- Quantitative PCR and RT-PCR can especially be useful

Further Reading

Anzola M, Burgos JJ. Hepatocellular carcinoma: molecular interactions between hepatitis C virus and p53 in hepatocarcinogenesis. Exp Rev Mol Med. 2003;5:1–16.

Bautista E, Chotpitayasunondh T, Gao Z, Harper SA, Shaw M, Uyeki TM, et al. Clinical aspects of pandemic 2009 influenza A (H1N1) virus infection. N Engl J Med. 2010;362:1708–19.

Boeckh M, Boivin G. Quantitation of cytomegalovirus: methodologic aspects and clinical applications. Clin Microbiol Rev. 1998;11:533–54.

Cockerill 3rd FR, Smith TF. Response of the clinical microbiology laboratory to emerging (new) and reemerging infectious diseases. J Clin Microbiol. 2004;42:2359–65.

Comanor L, Hendricks D. Hepatitis C virus RNA tests: performance attributes and their impact on clinical utility. Expert Rev Mol Diagn. 2003;3:689–702.

Ellermann–Eriksen S. Macrophages and cytokines in the early defence against herpes simplex virus. Virol J. 2005;2:59.

Erice A. Resistance of human cytomegalovirus to antiviral drugs. Clin Microbiol Rev. 1999;12:286–97.

Espy MJ, Uhl JR, Sloan LM, Buckwalter SP, Jones MF, Vetter EA, et al. Real-time PCR in clinical microbiology: applications for routine laboratory testing. Clin Microbiol Rev. 2006;19:165–256.

Finzer P, Aquilar-Lemarroy A, Rosl F. The role of human papillomavirus oncoproteins E6 and E7 in apoptosis. Cancer Lett. 2002;188:15–24.

Hakki M, Chou S. The biology of cytomegalovirus drug resistance. Curr Opin Infect Dis. 2011;24:605–11.

Kesli R, Polat H, Terzi Y, Kurtoglu MG, Uyar Y. Comparison of a newly developed automated and quantitative hepatitis C virus (HCV) core antigen test with the HCV RNA assay for clinical usefulness in confirming anti-HCV results. J Clin Microbiol. 2011;49:4089–93.

Landry ML, Garner R, Ferguson D. Comparison of the nuclisens basic kit (nucleic acid sequence-based amplification) and the argene biosoft enterovirus consensus reverse transcription-PCR assays for rapid detection of enterovirus RNA in clinical specimens. J Clin Microbiol. 2003;41:5006–10.

Landry ML, Garner R, Ferguson D. Real-time nucleic acid sequence-based amplification using molecular beacons for detection of enterovirus RNA in clinical specimens. J Clin Microbiol. 2005;43:3136–9.

Lurain NS, Chou S. Antiviral drug resistance of human cytomegalovirus. Clin Microbiol Rev. 2010;23:689–712.

Mahony JB, Richardson S. Molecular diagnosis of severe acute respiratory syndrome: the state of the art. J Mol Diagn. 2005;7:551–9.

Moore C, Valappil M, Corden S, Westmoreland D. Enhanced clinical utility of the NucliSens EasyQ RSV A + B Assay for rapid detection of respiratory syncytial virus in clinical samples. Eur J Clin Microbiol Infect Dis. 2006;25:167–74.

Moret H, Brodard V, Barranger C, Jovenin N, Joannes M, Andreoletti L. New commercially available PCR and microplate hybridization assay for detection and differentiation of human polyomaviruses JC and BK in cerebrospinal fluid, serum, and urine samples. J Clin Microbiol. 2006;44:1305–9.

Murray PG, Young LS. Epstein–Barr virus infection: basis of malignancy and potential for therapy. Expert Rev Mol Med. 2001;3:1–20.

National Institutes of Health. National institutes of health consensus development conference statement: Management of hepatitis C. Hepatology. 2002;36:S3–20.

Neville BW, Waldron CA, Herschraft EE. Oral and maxillofacial pathology. 1st ed. Philadelphia: Saunders; 1995.
Oberste MS, Pallansch L. Enterovirus molecular detection and typing. Rev Med Microbiol. 2005;16:163–71.
Ooi MH, Wong SC, Lewthwaite P, Cardosa MJ, Solomon T. Clinical features, diagnosis, and management of enterovirus 71. Lancet Neurol. 2010;9:1097–105.
Posthuma CC, van der Beek MT, van der Blij-de Brouwer CS, van der Heiden PL, Marijt EW, Spaan WJ, et al. Mass spectrometry-based comparative sequencing to detect ganciclovir resistance in the UL97 gene of human cytomegalovirus. J Clin Virol. 2011;51:25–30.
Russell WC. Update on adenovirus and its vectors. J Gen Virol. 2000;81:2573–604.
Shafer RW. Genotypic testing for human immunodeficiency virus type 1 drug resistance. Clin Microbiol Rev. 2002;15:247–77.
Solomon T, Lewthwaite P, Perera D, Cardosa MJ, McMinn P, Ooi MH. Virology, epidemiology, pathogenesis, and control of enterovirus 71. Lancet Infect Dis. 2010;10:778–90.
Taubenberger JK, Layne SP. Diagnosis of influenza virus: coming to grips with the molecular era. Mol Diagn. 2001;6:291–305.
Whiley DM, Mackay IM, Sloots TP. Detection and differentiation of human polyomaviruses JC and BK by LightCycler PCR. J Clin Microbiol. 2001;39:4357–61.
Wright TC Jr, Schiffman M, Solomon D, Cox JT, Garcia F, Goldie S, et al. Interim guidance for the use of human papillomavirus DNA testing as an adjunct to cervical cytology for screening. Obstet Gynecol. 2004;103:304–9.
Zoulim F. New nucleic acid diagnostic tests in viral hepatitis. Semin Liver Dis. 2006;26:309–17.

Molecular Bacteriology and Mycobacteriology

26

Jeong Hwan Shin, Mona Sharaan, Josephine Wu, Miao Cui, David Y. Zhang, and Yi-Wei Tang

Contents

26.1 ***Chlamydia trachomatis*** 734
26.1.1 General Characteristics of *Chlamydia* spp. 734
26.1.2 Clinical Presentation of *C. trachomatis* 734
26.1.3 Diagnostic Method 734

26.2 ***Neisseria gonorrhoeae*** 739
26.2.1 General Characteristics 739
26.2.2 Clinical Presentation 739
26.2.3 Diagnostic Method 740

26.3 ***Clostridium difficile*** 743
26.3.1 General Characteristics 743
26.3.2 Clinical Presentation 744
26.3.3 Diagnostic Method 745

26.4 ***Borrelia burgdorferi*** 746
26.4.1 General Characteristics 746
26.4.2 Clinical Presentation (Lyme Disease) 746
26.4.3 Diagnostic Methods 746

26.5 ***Streptococcus pyogenes*** 748
26.5.1 General Characteristics 748
26.5.2 Clinical Presentation 748
26.5.3 Diagnostic Method 748

26.6 ***Streptococcus agalactiae*** 749
26.6.1 General Characteristics 749
26.6.2 Clinical Presentations 749
26.6.3 Diagnostic Method 750

26.7 ***Enterococci*, Vancomycin-Resistant** 752
26.7.1 General Characteristics 752
26.7.2 Clinical Presentation 753
26.7.3 Diagnostic Method 754

26.8 ***Staphylococcus aureus*, Methicillin-Resistant** 756
26.8.1 General Characteristics 756
26.8.2 Clinical Presentation 756
26.8.3 Diagnostic Method 756

26.9 ***Legionellaceae*** 759
26.9.1 General Characteristics 759
26.9.2 Clinical Presentation 760
26.9.3 Diagnostic Method 760

26.10 ***Mycobacterium tuberculosis*** 761
26.10.1 General Characteristics 761
26.10.2 Clinical Presentation 762
26.10.3 Diagnostic Method 762

26.11 ***Nontuberculous mycobacteria*** 769
26.11.1 General Characteristics 769
26.11.2 Clinical Presentation 769
26.11.3 Diagnostic Method 770

26.12 **Molecular Diagnosis of Sepsis** 774
26.12.1 General Characteristics 774
26.12.2 Clinical Presentation 775
26.12.3 Diagnostic Method 775

Further Reading 777

J.H. Shin, MD, PhD
Department of Laboratory Medicine, Inje University College of Medicine, Busanjin-gu, Busan, South Korea

M. Sharaan, MD
Molecular Pathology Laboratory, Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

J. Wu, DDS, CLSP(MB), CLDIR
Mount Sinai School of Medicine, New York, NY, USA

M. Cui, MD • D.Y. Zhang, MD, PhD, MPH
Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

Y.-W. Tang, MD, PhD, F(AAM), FIDSA (✉)
Memorial Sloan-Kettering Cancer Center, Professor of Laboratory Medicine and Pathology, Cornell University Weil College of Medicine, New York, NY, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_26,

26.1 *Chlamydia trachomatis*

26.1.1 General Characteristics of *Chlamydia* spp.

- Obligate intracellular bacterial parasites
- Develop glycogen-containing microcolonies or inclusions called Halberstadter–Prowazek bodies
- Life cycle (Fig. 26.1)
- Group-specific antigens associated with the cell wall
  - Can be detected by complement fixation testing
- Clinically relevant species
  - *Chlamydia trachomatis, C. pneumoniae*, and *C. psittaci*
- Genome of *C. trachomatis*
  - 1,042,519-bp long with 894 predicted protein-coding sequences
- Genome of *C. pneumoniae*
  - 1,230,230-bp long with 1,073 open reading frames
- 186 genes of the *C. pneumoniae* genome are not homologous to sequences of the *C. trachomatis* genome, and 70 genes of the *C. trachomatis* genome are unrepresented in the *C. pneumoniae* genome
- *C. trachomatis* is transmitted sexually. Vertical transmission may also occur, related to maternal genital infection
- *C. pneumoniae* and *C. psittaci* are spread by the respiratory route

26.1.2 Clinical Presentation of *C. trachomatis*

- The most prevalent bacterial sexually transmitted infection in the United States (Table 26.1)
- In 2006, the total rate of chlamydia for the United States was 345.0 per 100,000 population
- Site of infection: cervix, urethra, oropharynx, conjunctiva, and rectum
- Most of infections are asymptomatic
- Ocular trachoma (leading cause of blindness in the underdeveloped world), keratitis, pannus formation, conjunctival scarring, trichiasis, and entropion
- Nongonococcal urethritis (Table 26.2)
- Cervicitis or urethritis
  - In female, pelvic inflammatory disease, salpingitis, tubo-ovarian abscess, and perihepatitis
  - In male, epididymitis
- Inclusion conjunctivitis appears within first week after birth (Fig. 26.2)
- Neonatal pneumonia
  - Neonate becomes ill with pneumonitis 4–16 weeks after birth, conjunctivitis may precede, and eosinophilia is common
- Lymphogranuloma venereum
  - Caused by genotypes L1–L3
  - 1–4 weeks incubation period
  - Culminating in headache, myalgia, fever, and the formation of a painless herpetiform ulceration at the site of inoculation
  - As the organism spreads, the inguinal lymph nodes swell, become tender, and may rupture and drain through the skin
- Women may experience elephantiasis of the vulva

26.1.3 Diagnostic Method

- Specimens
  - Cervical swab, urethral swab, rectal swab, nasopharyngeal swab, ocular specimen collection from cervix, urethra, oropharynx, conjunctiva, and rectum by scraping the mucosa (Table 26.3)
  - Specimens should be taken using a Dacron- or rayon-tipped swab without a wooden shaft
  - Swabs with wooden shafts may contain substances that are toxic to the growth of *C. trachomatis* and/or inhibitory to nucleic acid amplification
  - For urine specimens, first-voided morning urine specimens are preferred for molecular testing
- Conventional method

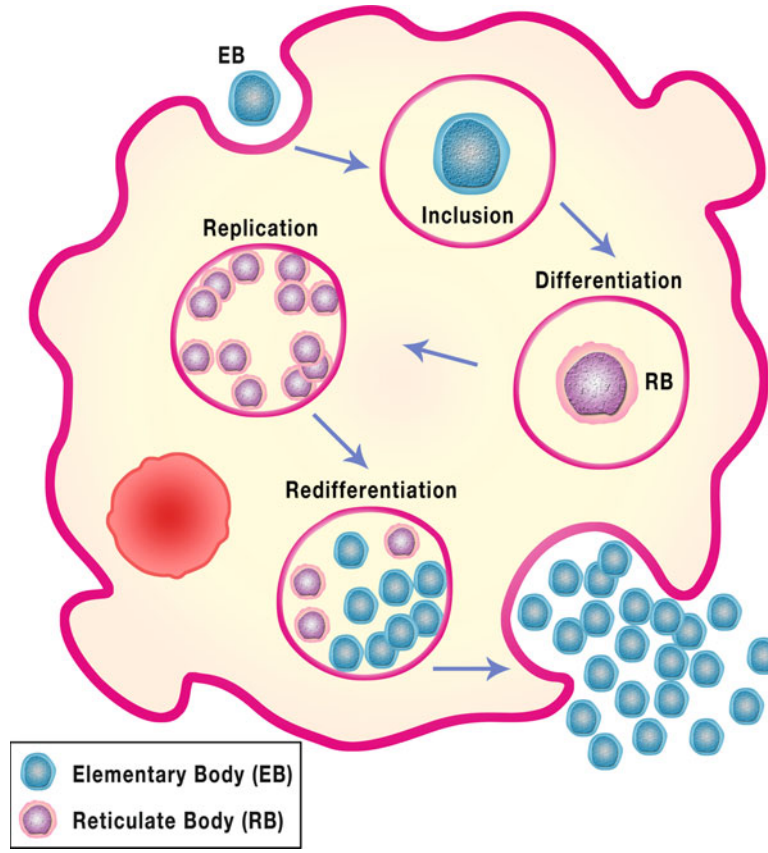


Fig. 26.1 Life cycle of *Chlamydia*. Metabolically inert elementary bodies are taken up by host cells, within which they develop into large replicative particles called reticulate bodies. Following replication, the organism redifferentiates into elementary bodies, which are infectious particles released from the cell

- Cell culture
  - The most specific test available to diagnose chlamydial infection
  - Expensive and technically difficult
  - Rigorous transport requirements: requires refrigeration and rapid lab processing (within 24 h)
  - Culture media to transport specimens include 2-sucrose phosphate, sucrose-glutamate phosphate, or other synthetic media
  - Adequate numbers of columnar epithelial cells must be obtained
  - Cycloheximide-treated McCoy cells are most common cell lines
  - HeLa 229 cells treated with DEAE-dextran and cycloheximide can also be used
  - After 48–72 h, intracytoplasmic inclusions can be detected by fluorescent staining antibodies that bind specifically to the major outer membrane protein (MOMP) of *C. trachomatis*
  - Advantages: most specific, thus the test of choice in cases of suspected sexual abuse. It allows for isolation of an organism that can be tested for antibiotic susceptibility. It is the preferred method for diagnosing chlamydial infections of the rectum, conjunctiva, or oropharynx
  - Limitations: low sensitivity. It cannot be used to test urine specimens or vaginal specimens in postpubertal females
- Stain
  - Giemsa or Gimenez methods
  - Chlamydial elementary bodies are gram negative
- Immunofluorescence – direct immunofluorescence test
  - Urethral (males), cervical (females), rectal (symptomatic), conjunctival

Table 26.1 Chlamydia trachomatis screening recommendations

Routinely screen all sexually active women aged ≤25 years for *C. trachomatis* infection, whether or not they are pregnant. Women and adolescents aged ≤20 years are at highest risk for chlamydial infection, but the majority of reported data indicate that infection is prevalent among women aged 20–25 years. More targeted screening might be indicated by local prevalence data
Screening of women aged >25 years for *C. trachomatis* should also be considered if they are at increased risk
Prevalence of *C. trachomatis* infection varies widely among communities and patient populations. Knowledge of the patient population is the best guide to developing a screening strategy. Certain risk factors should be considered, including
Having new or multiple sex partners
Having a prior history of a sexually transmitted disease
Not using condoms consistently and correctly
Personal risk depends on the number of risk markers and local disease prevalence. Specific risk-based protocols need to be tested locally
The optimal timing of screening in pregnancy is uncertain
Screening early in pregnancy provides increased opportunities to improve pregnancy outcomes, including low birth weight and premature delivery
Screening and treatment in the third trimester might be more effective at preventing transmission of chlamydial infection to the infant during birth by reducing the risk for reinfection
The incremental benefit of repeated screening is unknown
The optimal interval for screening women with a previous negative screening test is uncertain. The interval for rescreening should be based on changes in sexual partners, young age, and other *C. trachomatis* risk factors. If evidence exists that a woman is at low risk for infection (e.g., in a mutually monogamous relationship with a previous history of negative screening tests for chlamydial infection), screening frequently might be unnecessary

Adapted from CDC (2002)

(symptomatic), and nasopharyngeal specimens (symptomatic)

- Uses antibodies directed against lipopolysaccharide (LPS) or major outer membrane protein
- For best results, store/transport at 20–30 °C or refrigerate at 2–8 °C and stain within 7 days of collection
- Turnaround time: 1–2 days

Table 26.2 Factors that attempt to distinguish male nongonococcal urethritis from those cases caused by *Neisseria gonorrhoeae*

Factor	Gonococcal urethritis	Nongonococcal urethritis
Typical incubation period	2–5 days	7–14 days
Range of symptom onset	1–10 days	2–35 days
Frank purulent discharge (% of cases)	75	11–33
Mucopurulent discharge (% of cases)	25	50
Clear or moderately viscid discharge (% of cases)	4	10–50
Dysuria (% of cases)	73–88	53–75
Combination of dysuria and discharge (% of cases)	71	38

Adapted from Harkins and Munson (2011)

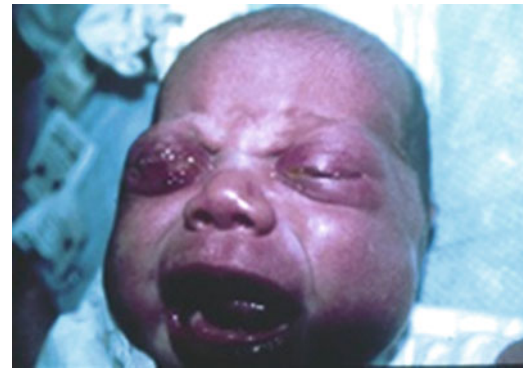


Fig. 26.2 Inclusion conjunctivitis of *C. trachomatis*

- DFA utilizing monoclonal antibodies targeting MOMP has a sensitivity of 80–90% and a specificity of 98–99% for detecting *C. trachomatis*
- DFA targeted to the LPS can cross-react with other *Chlamydia* species

– Enzyme immunoassay

- Sensitivity varies from 70% to 100%
- It uses antibodies directed against LPS
- Chlamydial LPS antibodies may also cross-react with the LPS of other gram-negative bacteria to give false-positive results
- Advantages: Specimens do not require refrigeration

Table 26.3 Recommended diagnostic modalities and specimen sources for evaluation of child sexual abuse victims

	Recommended specimen sources for STD etiology			
	Chlamydia trachomatis		*Neisseria gonorrhoeae*	
Diagnostic modality	Female	Male	Female	Male
Culture			Oropharynx	Oropharynx
	Anus	Anus	Anus	Anus
				Urethra
		Urethral discharge		Urethral discharge
NAAT	Urine		Urine	
	Vagina		Vagina	

Adapted from Harkins and Munson (2011)
Abbreviation: *NAAT* nucleic acid amplification testing

- Disadvantages: It lacks sensitivity as screening assay, especially for asymptomatic men. It cross-reacts with the LPS of other *Chlamydia* species. Enzyme-linked immunosorbent assay (ELISA) can be used with only endocervical, urethral, or conjunctival specimens

– Serologic diagnosis
  - Group-specific antigen, associated with the cell wall, can be detected by complement fixation test
  - It is less widely available and multiple antigens must be tested

- Molecular method
  – Nucleic acid hybridization method
    - General characteristics
      – This is a nonculture, nonamplification nucleic acid probe hybridization-based test
      – Uses DNA probes that hybridize to *C. trachomatis* ribosomal RNA
      – The DNA probe can be stored for up to 7 days without refrigeration before processing
      – Specimens: can be used for endocervical, urethral, and conjunctival specimens, but not recommended for use on urine, rectal, oropharyngeal, or vaginal specimens
    - PACE® 2 System Assay (Gen-Probe)
      – Uses a single-stranded DNA (ssDNA) probe with a chemiluminescent label that is complementary to the ribosomal RNA of the target organism (16S rRNA)
      – A stable DNA: RNA hybrid is formed
      – The labeled DNA: RNA hybrid is isolated and luminescence is measured in a Gen-Probe Leader® luminometer (Gen-Probe, Inc., San Diego, CA)
      – The only assay approved for use with conjunctival specimens
      – Fair sensitivity about 70% and excellent specificity 99.8%
    - Signal amplification
      – Hybrid Capture II (Digene)
        - US food and drug administration (FDA) approved for endocervical and male urethral specimens, but not rectal, respiratory, or vaginal specimens
        - Target: both genomic DNA and cryptic plasmid DNA
        - Specimens can be shipped at room temperature and are stable when stored at room temperature for 14 days or at 20 °C for up to 3 months
        - RNA–DNA hybrids are captured by hybrid-specific antibodies
        - Sensitivity 92.3–97.7% and specificity 98.2–98.6%
  - Nucleic acid amplification method
    – Polymerase chain reaction (PCR) (Cobas Amplicor, Roche Diagnostic Systems)
      - FDA-cleared (approved) collection sites
        – Urine from males and females
        – Endocervical swab specimens
        – Male urethral swab specimens (symptomatic or asymptomatic)

- The target for the Roche Amplicor test for *C. trachomatis* is a 207-bp segment of the cryptic plasmid DNA sequence
- Uses PCR to amplify target DNA and nucleic acid hybridization
- Urine specimens are stable for 24 h, swab specimens are stable for 1 h at room temperature, urine can be stored at −20 °C for 1 month, and swabs can be stored at 2–8 °C for 7 days
- The sensitivity and specificity have varied across studies and by specimen type
  - Sensitivity in endocervical samples has ranged from 51.9% to 96.8%: 44.4–82.5% for female urine, 98% for male urethral swab, and 92% for male urine
  - Specificity for samples from all sites has been reported to be 98.4–100%
- The use of an internal amplification control enables the laboratory to determine the presence of inhibitors
- Can detect new variant of *C. trachomatis* (nvCT) (common in Swedish countries)

– Real-time PCR (Cobas 4800 CT/NG assay and Cobas TaqMan CT assay, Roche Molecular Systems)
  - Cobas 4800 CT/NG assay
    - Fully automated system
    - Multiplex real-time PCR assay
    - Dual target: cryptic plasmid and major outer membrane protein gene
    - Can detect new variant
  - Sensitivity 92.0–94.5% and specificity 99.5–100%

– Real-time PCR (Abbott RealTime CT/NG assay, Abbott *m*2000 System, Abbott Molecular)
  - They are required to be used together as a system for the detection of CT/NG from multiple specimen types
  - FDA-cleared (approved) collection sites
    - Urine from males and females (symptomatic and asymptomatic)
    - Male urethral swab specimens
    - Vaginal and endocervical swab specimens
  - The target for *C. trachomatis* is the 102-base cryptic plasmid DNA sequence
  - The use of an internal amplification control enables the laboratory to determine the presence of inhibitors
    - 136-base amplicon of the hydroxypyruvate reductase gene from the pumpkin plant
  - 93.7–100% in sensitivity and 98.2–100% in specificity
  - Uses real time to amplify target DNA

– Transcription-mediated amplification (TMA) (APTIMA Assay, Gen-Probe)
  - FDA-cleared (approved) collection sites: endocervical and male urethral swab specimens, female and male urine specimens
  - Target gene: 16S rRNA
  - Transport swab specimens to the laboratory and store at 2–25 °C until tested
  - Urine specimens can be stored at 2–8 °C for up to 7 days from collection
  - Urine and swab specimens should be assayed within 7 days of collection
  - Sensitivity/specificity: 92.8–94.1%/97.6–99.4% in endocervical swabs, 95.5%/97.5% in male urethral swabs, and 97.9%/98.5% in urine

– Strand displacement amplification (SDA) (BD ProbeTec ET System, Becton–Dickinson)
  - FDA-cleared (approved) collection sites: endocervical swabs, male urethral swabs, and male and female urine specimens
  - This technology is similar to PCR but uses restriction enzymes instead of synthetic oligopeptides to target specific regions of *C. trachomatis* for amplification
  - Targets multicopy cryptic plasmid
  - Swab must be stored and transported to the laboratory and/or test site at 2–27 °C within 4–6 days of collection

  - Urine is stable for 4–6 days at 2–8 °C and 2 days at 15–27 °C
  - Sensitivity: 92.8% for endocervical swabs and 80.5% for female urine. For male, 92.5% from urethral swabs and 93.1% from urine
  - A high throughput system (the BD Viper System) is available
- Clinical utility of molecular diagnostics
  – Convenient and acceptable samples such as initial stream urine and self-collected vaginal specimens are used in molecular tests, which increase compliance with testing
  – Improved performance including high sensitivity and early detection of the organism
- Pitfalls of molecular diagnostics
  – False-positive results due to carryover contamination (detected by negative control)
  – False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
  – Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure

26.2 *Neisseria gonorrhoeae*

26.2.1 General Characteristics

- Sexually transmitted or vertically transmitted during parturition
- Fastidious, gram-negative diplococci
- Usually in pairs (diplococci) with flattened adjacent sides, immotile, cell envelope is present
- Aerobic or facultatively anaerobic; very sensitive to drying, chilling, and pH change; iron required for growth
- Best grown in environment with 2–8% CO_2
- Thayer–Martin medium (chocolate agar plus vancomycin, colistin, and nystatin) permits isolation in otherwise contaminated specimen
- Gonococci can be presumptively detected as grayish-white oxidase-positive colonies

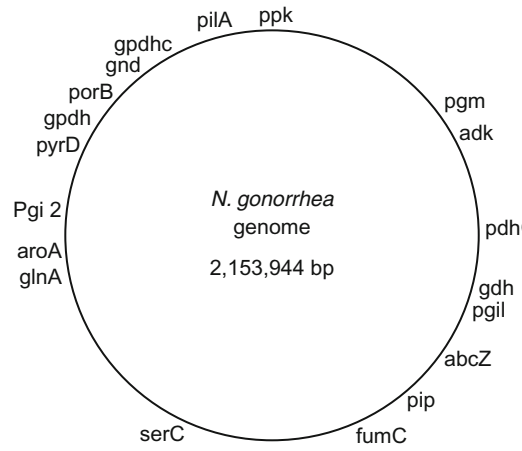


Fig. 26.3 Location of 18 housekeeping genes and *PORB* on chromosomal map of *N. gonorrhoeae* strain FA 1090

- Speciation is based on carbohydrate metabolism; *N. gonorrhoeae* produces acid from glucose only
- The *N. gonorrhoeae* genome is circular, consists of 2,153,922 basepairs, and contains 2002 protein-encoding genes (Fig. 26.3)

26.2.2 Clinical Presentation

- Site of infection: cervix, urethra, oropharynx, conjunctiva, and rectum
- Most of infections are asymptomatic
- Primarily restricted to the mucus membranes of the endocervix, urethra, rectum, and pharynx
- In males: acute anterior urethritis, urethral stricture, epididymitis, and prostatitis
- In females: 20–80% asymptomatic, but complications include pelvic inflammatory disease and generalized peritonitis, tubal infertility, and ectopic pregnancy
- Proctitis
- Bacteremia, disseminated disease, and subacute bacterial endocarditis
- Arthritis–dermatitis syndrome, neonatal gonococcal arthritis
- Meningitis
- Ophthalmia neonatorum
- Vulvovaginitis in prepubescent females

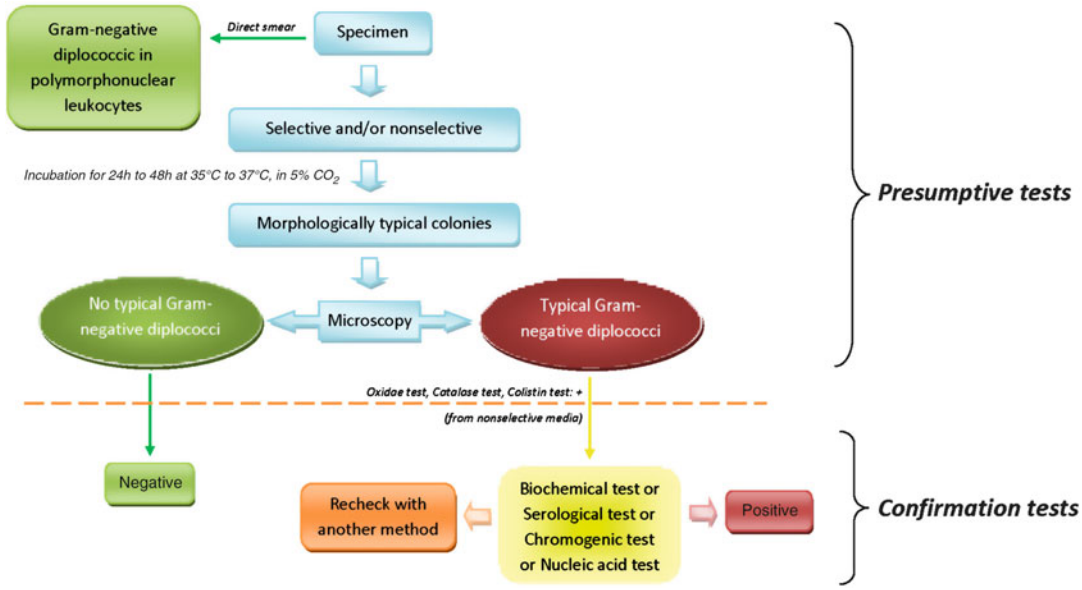


Fig. 26.4 Algorithm for culture and identification of *Neisseria gonorrhoeae*

26.2.3 Diagnostic Method

- Specimens
  - Urethral swab, cervical swab, conjunctival discharge, skin lesion scrapings, synovial fluid, pharyngeal swab, rectal swab, and blood
  - For best results, store/transport at 20–30 °C or refrigerate at 2–8 °C and process within 7 days of collection
  - Swabs with wooden shafts may contain substances that are toxic to the growth of *N. gonorrhoeae* and/or inhibitory to nucleic acid amplification
  - For urine specimens, first-voided morning urine specimens are preferred for molecular testing
- Conventional methods
  - Gram stain
    - Rapid tool
    - Comparable sensitivity to bacterial culture for symptomatic urethral gonorrhea in men
    - Relatively insensitive for specimens collected from women and for specimens from extragenital sites
  - Culture (Fig. 26.4)
    - Thayer–Martin chocolate agar from infectious material
    - Culture is both sensitive and specific if specimen was promptly and correctly collected and processed
    - Gold standard for definitive diagnosis
    - Advantages: near 100% specificity, the optimal test for use in populations with low gonorrhea prevalence. It allows for isolation of an organism for antimicrobial sensitivity testing
    - Limitations: the need to store the culture medium in refrigeration, the need for transportation under appropriate conditions to maintain organism viability, and the 24–72-h time delay
    - Biochemical identification tests
    - Acid production from glucose only differentiates from other nonurogenic *Neisseria* species
  - Immunologic methods for culture confirmation
    - Fluorescent antibody tests
    - Specimen types: urethral (males), cervical (females), rectal (symptomatic),

conjunctival (symptomatic), and nasopharyngeal specimens (symptomatic)
- Turnaround time: 1–2 days
- Sensitivity of the test for detection of genital disease can range from 60% to 100%

– Molecular techniques
  - Nucleic acid hybridization method
    – General characteristics
      - The most sensitive and most specific nonculture, nonamplification tests
      - Uses DNA probes that hybridize to *N. gonorrhoeae* ribosomal RNA
      - The DNA probe can be stored for up to 7 days without refrigeration before processing
      - Specimens: can be used for endocervical, urethral, and conjunctival specimens, but not recommended for use on urine, rectal, oropharyngeal, or vaginal specimens
    – PACE 2 Assay System (Gen-Probe)
      - Uses a ssDNA probe with a chemiluminescent label that is complementary to the ribosomal RNA of the target organism (16S rRNA)
      - Stable DNA: RNA hybrid is formed
      - The labeled DNA: RNA hybrid is isolated and measured in a Gen-Probe Leader luminometer
      - The only assay approved for use with conjunctival specimens. Rapid turnaround time (3–5 h) for results
      - Sensitivity 97.8% and specificity 98.9%
– Signal amplification method
  - Hybrid Capture II CT/GC DNA test (Digene)
    – The test is FDA cleared (approved) for endocervical, male urethral, and vaginal swab specimens and in male and female urine specimens
    – RNA–DNA hybrids are captured by hybrid-specific antibodies
    – Target: both genomic DNA and cryptic plasmid DNA
    – Sensitivity 92.6–95.2% and specificity 98.5–98.9%
    – Specimens (endocervical) can be shipped at room temperature and are stable when stored at room temperature for 14 days or at −20 °C for up to 3 months
– Nucleic acid amplification method
  - Advantages
    – Improved sensitivity compared with bacterial culture
    – Require the organism to be viable for detection and require less stringent transport conditions
    – NAATs can be used effectively on noninvasive specimens
  - Disadvantages
    – High cost
    – Carryover contamination
    – Cross-reaction
    – Inhibition of the reaction
    – Inhibitory substances: beta-HCG, crystals, hemoglobin, and nitrites
    – Common in urine
    – High quality control requirements
  - PCR (Cobas Amplicor, Roche Molecular)
    – Common in urine
    – FDA-cleared (approved) collection sites: urine from males, endocervical swab specimens, and male urethral swab specimens
    – The target for the Roche Amplicor test for *N. gonorrhoeae* is cytosine methyltransferase gene homologue
    – This test uses PCR to amplify target DNA and nucleic acid hybridization
    – Urine specimens are stable for 24 h, swab specimens are stable for 1 h at room temperature, urine can be stored at −20 °C for 30 days, and swabs can be stored at 2–8 °C for 7 days
    – The sensitivity and specificity have varied across studies and by specimen type

  - Sensitivity in endocervical samples has ranged from 84% to 100%: 100% for male urethral swab, 66.7% for female urine, and 95.2% for male urine
  - Specificity was 98.7%
  - Highly sensitive in swab specimens from symptomatic and asymptomatic women and men (>96.4% sensitivity/>97.9% specificity)
  - Low sensitivity was observed in urine (42.3%) from asymptomatic men
  - Higher sensitivities and specificities in other genital specimen types (>92.4% sensitivity/98% specificity)
- Real-time PCR (Cobas 4800 CT/NG assay, Roche Molecular Systems)
  - Cobas 4800 CT/NG assay
    - Fully automated system
    - Multiplex real-time PCR assay
    - Target: a direct repeat region called DR-9
  - Sensitivity 92.9–100% and specificity 99.4–100%
- Real-time PCR (Abbott RealTime CT/NG assay, Abbott *m*2000 System, Abbott Molecular)
  - They are required to be used together as a system for the detection of CT/NG from multiple specimen types
  - FDA-cleared (approved) collection sites
    - Urine from males and females (symptomatic and asymptomatic)
    - Male urethral swab specimens
    - Vaginal and endocervical swab specimens
  - The target for *N. gonorrhoeae* is the 122-base DNA amplicon of opacity (Opa) gene
  - The use of an internal amplification control enables the laboratory to determine the presence of inhibitors
    - 136-base amplicon of the hydroxypyruvate reductase gene from the pumpkin plant
  - 91.4–100% in sensitivity and 99.3–100% in specificity
  - Uses real time to amplify target DNA
- Transcription-mediated amplification (TMA) (APTIMA Assay, Gen-Probe)
  - FDA-cleared (approved) collection sites: endocervical and male urethral swab specimens and female and male urine specimens
  - Target gene – 16S rRNA
  - Transport swab specimens to the laboratory and store at 2–25 °C until tested
  - Store urine specimens at 2–8 °C for up to 7 days from collection
  - Urine and swab specimens should be assayed within 7 days of collection
  - Sensitivity/specificity: 98.6–99.2%/98.7–99.8% in endocervical swabs, 99.1%/97.8% in male urethral swabs, and 98.5%/99.6% in urine
- Strand displacement amplification (SDA) (BD ProbeTec ET System)
  - FDA-cleared (approved) collection sites: endocervical swabs, male urethral swabs, and male and female urine specimens
  - This technology is similar to PCR but uses restriction enzymes instead of synthetic oligopeptides to target specific regions of *N. gonorrhoeae* and for amplification
  - Swab must be stored and transported to the laboratory and/or test site at 2–27 °C within 4–6 days of collection
  - Storage up to 4 days has been validated with clinical specimens
  - The SDA is currently not FDA cleared (approved) for testing asymptomatic males for *N. gonorrhoeae*
  - It targets multicopy chromosomal pilin gene
  - Urine is stable for 4–6 days at 2–8 °C and 2 days at 15–27 °C
  - Sensitivity of 96.5% for endocervical specimens, 84.9% for female urine,

98.5% from male urethral swabs, and 97.9% from male urine
- Clinical utility of molecular diagnostics
  - Convenient and acceptable samples such as initial stream urine and self-collected vaginal specimens are used in molecular tests, which has increased compliance with testing
  - Improved performance in early detection and medical intervention of *N. gonorrhoeae*
- Pitfalls of molecular diagnostics
  - False-positive results due to carry-over contamination (detected by negative control)
  - False-positive results due to cross-reactivity with nongonococcal *Neisseria* sp. that rarely cause genital disease
  - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
  - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure
  - Species cross-reactiveness (false-positives) may happen when nonurogenic specimens are tested

– Antimicrobial susceptibility method
- Conventional methods
  - Nitrocefin method
    - Detects β-lactamase production
    - Disk diffusion *Neisseria gonorrhoeae* (GC) agar—incubated for 20–24 h
    - Minimum inhibitory concentrations (MIC) detected by
      - Agar dilution (reference method): complex to perform
      - E-test (MIC on a strip)
- Molecular method
  - Amplification methods (PCR)
    - Detect mutated *GYR*A and *PAR*C genes in bacteria with high resistance to quinolones (chromosomal resistance)
    - Detect mutated *MTR* and *PENB*, which reduces susceptibility to penicillin (chromosomal resistance)
    - Detect *TETM* in bacteria with high resistance to tetracycline (plasmid-borne resistance)
  - Pitfalls of molecular diagnostics
    - Resistance can be due to multiple genetic changes, for which there is no single probe
    - New mutations affecting the level of resistance are being continually discovered

26.3 *Clostridium difficile*

26.3.1 General Characteristics

- An obligate anaerobic, gram-positive, spore-forming rod
- Saccharolytic and proteolytic
- Metabolically dormant, survive for long periods, and resistant to harsh physical or chemical treatments
- Genome: circular chromosome (4,290,252 bp) and a plasmid (7,881 bp)
- The pathogenicity locus (PaLoc) (Fig. 26.5)
  - 19.6 kb in size
  - Contains five genes, TCDA, *TCDB*, *TCDC*, *TCDR*, and *TCDE*
  - Responsible for the synthesis and regulation of toxin A (TCDA) and B (TCDB)
  - Nontoxigenic strains lack the PaLoc; however, isolates with a defective PaLoc can still cause disease
  - The genes encoding the binary toxin are not found on the PaLoc
- Toxins A and B
  - Both toxins are glucosyltransferases
  - They transfer a UDP-glucose to small GTPases, such as RHO, RAC, and CDC42
  - TCDA: 308KDa, an enterotoxin
  - TCDB: 270 kDa, a cytotoxin, 100–1,000-fold more toxic to culture cells than TCDA

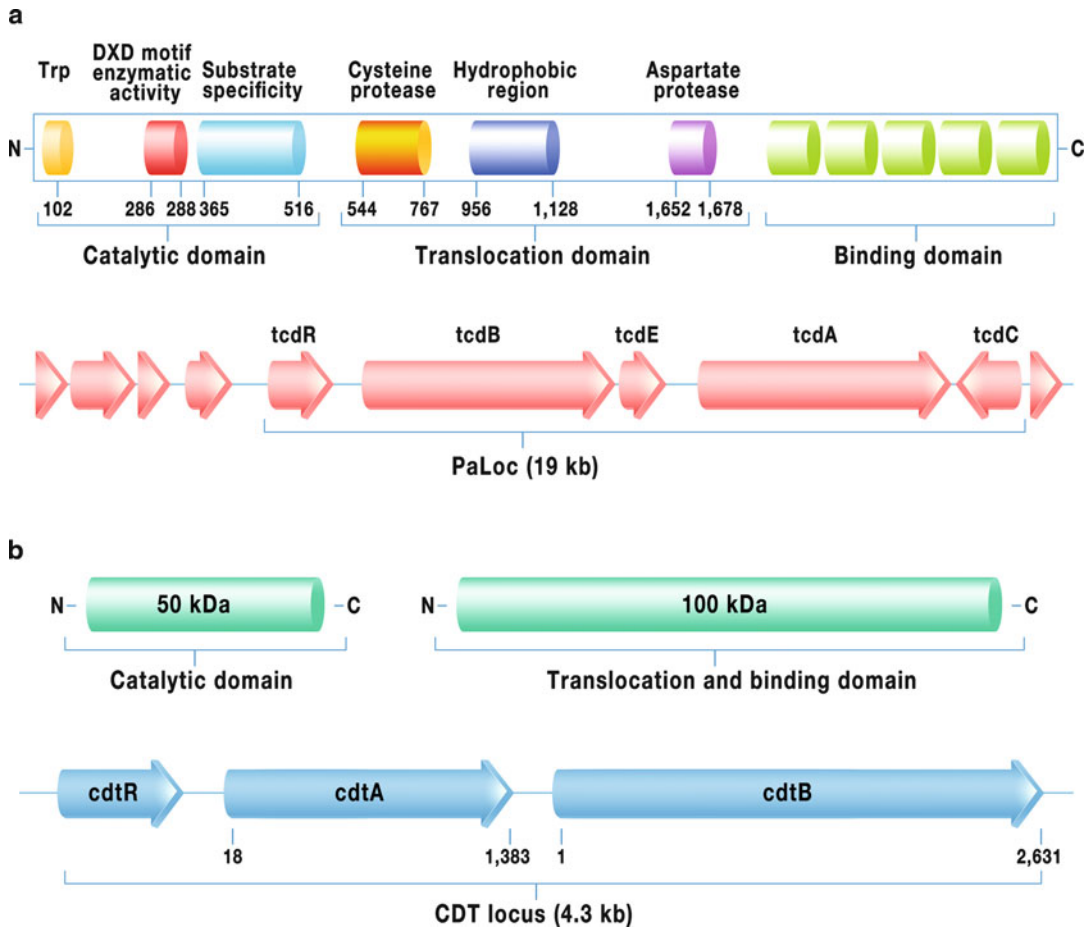


Fig. 26.5 The pathogenicity locus of *C. difficile*

- *C. difficile* transferase (CDTa)
  - Binary toxin
  - Encoded by the CDT locus (LdtLoc)
  - ADP-ribosylating toxin
  - Two subunits: CDTa and CDTb. Each component alone is not cytotoxic, but together they cause cytotoxicity in vitro
  - Approximately 6–12.5% of strains overall
- Only toxigenic strains cause disease, due to the production of toxins A and/or B
- All patients with diarrhea who have been hospitalized more than 72 h should be tested for *C. difficile* infections (CDI)
- A unique, highly virulent genotype (BI/NAP1/027) is responsible for recent high epidemic in North America

26.3.2 Clinical Presentation

- Risk factors
  - Antibiotic exposure: clindamycin, cephalosporins, beta-lactamase, and fluoroquinolones
  - Use of combination antibiotic therapy and long-term receipt of antibiotic therapy
  - Hospitalization, prolonged hospital stay, and residence in an intensive care unit
  - Advanced age
  - Suppressed immune system
- Clinical features
  - The primary cause of nosocomial diarrhea
  - Diarrhea associated with a history of antibiotic use
  - Watery stool with a characteristic foul odor

- Abdominal cramps, fever, leukocytosis, and hypoalbuminemia
- Antibiotic-associated diarrhea
- Pseudomembranous colitis

26.3.3 Diagnostic Method

- Specimens
  - CDI test should only be performed on uniformed stools
- Conventional method
  - Culture
    - Selective media: cycloserine-cefoxitin-fructose agar (CCFA)
    - Alternative: pretreatment with alcohol shock
      - To decrease the normal fecal flora
      - To prevent overgrowth
    - Toxigenic culture
      - Culture on selective media followed by in vitro toxin detection (by enzyme immunoassay, cell culture cytotoxicity assay, or PCR) to determine the toxigenicity of the isolated strain
    - Incubation in an anaerobic atmosphere for at least 48 h
  - Cell culture cytotoxicity assay
    - The reference standard for the detection of *C. difficile* toxins
    - Stool filtrates are inoculated onto a monolayer of a cell culture
    - Observation for a toxin-B-induced cytopathic effect after 24 h and 48 h
    - Cell lines used for cell culture cytotoxicity assay
      - Vero cells
      - HeLA cells
      - Human foreskin fibroblasts
      - Hep2 cells
    - Neutralization is executed to determine the specificity of the cytopathic effect
  - Enzyme immunoassay
    - Rapid and easy to perform assays
    - Detects glutamate dehydrogenase (GDH) and/or toxins A and B
    - Detection formats
      - Microwell-type enzyme immunoassays: Results are displayed as a color change which can be detected visually or photospectrometically. Suitable for testing solitary samples, e.g., Premier tox A (ToxinA, Meridian)
      - Membrane-type enzyme immunoassays: Results can be visually read for a membrane. Suitable for testing samples in batches, e.g., Techlab C. DIFF QUIK CHEK COMPLETE (GDH and ToxinA, Meridian)
    - Diagnostic performance: sensitivity
      - 72–93% in comparison with cell culture cytotoxicity assay, 52–86% in comparison with toxigenic culture, and 88–89% in comparison with culture
    - Diagnostic performance: specificity
      - 89–98% in comparison with cell culture cytotoxicity assay, 76–98% in comparison with toxigenic culture, and 91–97% in comparison with culture
- Molecular method
  - Superior to all methods except toxigenic culture
  - FDA-cleared NAATs
    - BD GeneOhm *C. diff.* assay (BD)
      - Target: conserved regions of *TCDB*
      - Sensitivity 84–96%, specificity 94–100%
    - Prodesse ProGastro CD assay (Gen-Probe)
      - Target: conserved regions of *TCDB*
      - Sensitivity 77–92%, specificity 95–99%
    - Cepheid GeneXpert *C. difficile* (Cepheid)
      - Target: conserved regions of *TCDB* and *TCDC*
      - Sensitivity 94–100%, specificity 93–99%
    - Differentiation between BI/NAP1/027 and non-BI/NAP1/027 strains

- Illumigene *C. difficile* assay (Meridian Diagnostics)
  - Uses loop-mediated isothermal amplification
  - Detects the PaLoc at a conserved region of *TCDA* that is present even in TCDA–, TCDB + strains
  - Sensitivity 99%, specificity 98%
- Clinical utility of molecular diagnostics
  - NAAT testing is rapidly replacing other methods in clinical microbiology laboratories
  - They are superior to all methods except toxigenic culture
  - Should not be used as a test of cure
  - Currently no testing strategy is optimally sensitive and specific; therefore, clinical suspicion and consideration of the patient risk factors are important in making clinical decisions about whom to treat

26.4 *Borrelia burgdorferi*

26.4.1 General Characteristics

- Spirochete; etiologic agent of Lyme disease, transmitted by ticks
  - *Ixodes dammini*
  - *Amblyomma americanum*
- Ten different *Borrelia* species have been described within the *B. burgdorferi sensu lato complex: B. burgdorferi sensu stricto, B. garinii, B. afzelii, B. japonica, B. andersonii, B. valaisiana, B. lusitaniae, B. tanukii, B. turdi, and B. bissettii*
- Only *B. burgdorferi sensu stricto, B. garinii*, and *B. afzelii* are implicated in human disease
- *B. burgdorferi* has a genome of 910,725 bp, with at least 17 linear and circular plasmids with a combined size of more than 533,000 bp
- Morphology: coarse, irregular coils
- Visualized by Wright or Giemsa stains
- Difficult to culture

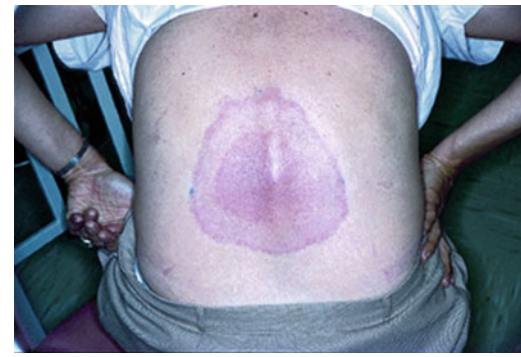


Fig. 26.6 Erythema migrans as seen in Lyme disease

26.4.2 Clinical Presentation (Lyme Disease)

- Stage 1 (erythema chronicum migrans)
  - Red macule at site of tick bite, progressing to annular erythema with central clearing (Fig. 26.6)
  - Can appear at sites other than the initial bite location
  - Fades in 3–6 weeks
  - Constitutional syndrome – fever, headache, malaise, adenopathy, and mild meningeal irritation
- Stage 2
  - Neurologic disease – follows rash, associated with severe headaches and cranial nerve palsies, and resolves after several months
  - Cardiac disease – fluctuating cardiac arrhythmias, resolves after several weeks, and can recur
- Stage 3
  - Arthritis
  - Chronic central nervous system (CNS) disease

26.4.3 Diagnostic Methods

- Specimens
  - Cerebrospinal fluid (CSF), urine, whole blood, serum, joint fluid, ticks, or skin biopsy

  - Storage instructions: refrigerate CSF, urine, synovial fluids, and blood
  - Sample collection and storage for molecular studies
    - Bacterial culture: Barbour–Stoner–Kelly (BSK) H medium in screw cap tubes at 32 °C
    - Ticks: Identify *Ixodes* ticks – live ticks are suitable for all methods, store dead ticks in 70% alcohol
    - CSF: 2 mL or more in sterile tubes. Store aliquots at −20 °C until use
    - Joint fluid: 2 mL or greater in a red-top tube (no anticoagulant) or a lavender-top tube with ethylenediaminetetraacetic acid (EDTA additive). Store aliquots at −20 °C
    - Whole blood: 5–10 mL in a purple-top tube. Store at 4 °C and process as soon as possible after receipt
    - Serum: 1–3 mL in sterile tube. Store aliquots at 20 °C until use
    - Urine: use 15 mL preferably collected prior to antibiotic therapy. Store at −70 °C or add equal volume of 95% ethanol prior to storage
    - Tissue: use a single-standard 3–5-mm skin punch biopsy or equivalent-sized tissue specimen. Stored at −70 °C
- Conventional method
  - Warthin–Starry silver stain – tissue sections
  - Acridine orange or Giemsa – blood and CSF
  - Culture: culture through inoculation into a tube of modified Kelly medium (BSK II) – yield is low (Table 26.4)
    - On skin biopsies from confirmed erythema migrans, the success rate is 40–70%
    - On blood samples from confirmed Lyme disease cases, the sensitivity is <4%
    - On CSF samples from confirmed neuroborreliosis, the positivity rate is <10%
  - Serologic diagnosis
    - ELISA – primary screening – quick, reproducible, and relatively inexpensive; false-positive results due to cross-reactivity
    - Immunofluorescent antibody (IFA) – peak IgM in third to sixth week, followed by IgG
    - Western blot – confirms positive ELISA or IFA tests
- Molecular method
  - Standard PCR
    - Chromosomal and plasmid targets: chromosomal-flagellin, the *POLD3* (also known as *p66*) gene segment encoding a 66-kDa protein (Oms66), 16S rRNA, *RECA*, and plasmid *OSPA*/*B*
  - LightCycler real-time PCR (*Artus*® Borrelia PCR Kits)
    - Offers advantages of rapid detection, PCR product quantitation, and closed system with minimal to no contamination
    - Targets the *RECA* gene
  - Ribosequencing of Lyme disease-associated *Borrelia*
    - Based on the analysis of 16S rRNA signature sequences representing different genospecies
  - Sensitivity and specificity
    - PCR is more sensitive than culture
    - Depending on the stage of the disease and type of tissue and sensitivity and specificity (Table 26.5)
  - Clinical utility of molecular diagnostics

Table 26.4 Selected features of various modifications of Kelly medium

Name	Description
Stoenner modification	Addition of Yeastolate and CMRL 1066, without glutamine and without sodium bicarbonate
Barbour modification (BSK II)	Use of CMRL 1066 without glutamine, Yeastolate, neopeptone as the peptone preparation, and HEPES as a buffer
Preac-Mursic modification (MKP*a*)	Removal of Yeastolate, different proportions of certain ingredients
BSK-H	Deletion of gelatin, different proportions of certain ingredients

Adapted from Aguero-Rosenfeld et al. (2005)

Table 26.5 Sensitivities and specificities of PCR assays for detection of *B. burgdorferi* DNA in different clinical specimens from patients with LB

Clinical specimen and region	No. of studies included	Median % sensitivity (range)	Reported % specificity (range)
Skin biopsy			
EM	16	69 (36–88)	98–100
United States	4	64 (59–67)	98–100
Europe	12	73 (36–88)	100
ACA, Europe	8	76 (54–100)	100
Blood, plasma, serum	6	14 (0–100)	
United States	3	18 (0–59)	100
Europe	3	10 (4–100)	NA
CSF	16	38 (12–100)	93–100
United States	6	73 (25–93)	93–100
Europe	10	23 (12–100)	98–100
Synovial fluid	8	78 (42–100)	100
United States	4	83 (76–100)	100
Europe	4	66 (42–85)	100

Adapted from Aguero-Rosenfeld et al. (2005)

- Rapid primary diagnosis for efficient and early medical treatment
- Improved performance of nucleic acid assays permits early detection and medical intervention for meningitis

– Pitfalls of molecular diagnostics
  - False-positive results due to carryover contamination (detected by negative control)
  - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
  - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure

26.5 *Streptococcus pyogenes*

26.5.1 General Characteristics

- Group A, gram-positive, β-hemolytic Streptococcus
- Colonize the throat or skin
- Group A carbohydrate: composed of N-acetylglucosamine linked to a rhamnose polymer backbone

26.5.2 Clinical Presentation

- The most common cause of bacterial pharyngitis
- Scarlet fever
- Impetigo
- Pyoderma and cellulitis
- Streptococcal toxic shock syndrome
- Necrotizing fasciitis
- Puerperal sepsis or childbed fever
- Poststreptococcal infection sequelae
- Acute rheumatic fever and rheumatic heart disease
- Acute glomerulonephritis
- Reactive arthritis
- Tourette syndrome, tics, and movement and attention deficit disorders

26.5.3 Diagnostic Method

- Specimens
  – Throat swab and other infectious sites
  – General guidelines to handle specimens
- Conventional methods
  – Rapid antigen test
  – Target: group A carbohydrate antigen
  – Agglutination technique
  – Enzyme immunoassay
  – Immunochromatographic test
  – Sensitivity and specificity
    - Moderate sensitivity
    - Excellent specificity: commonly higher than 95%
  – Culture
    - Throat culture
      – The most reliable method for detecting the presence of group A *Streptococcus* in throat
      – Optimal recovery: may be achieved by use of blood agar plates containing sulfamethoxazole–trimethoprim to

inhibit some of the normal flora and growth under anaerobic conditions to enhance streptolysin O activity
    - Beta-hemolytic on 5% sheep blood agar
  - Confirmatory biochemical identification tests
    - Catalase negative
    - Susceptibility to bacitracin
    - Positive pyrrolidonyl arylamidase (PYR) test
  - Grouping
    - Lancefield serological grouping
    - Based on the immunological differences in their cell wall polysaccharides (groups A, B, C, F, and G) or lipoteichoic acids (group D)
    - Confirmation by Lancefield capillary precipitin technique and the slide agglutination procedure
  - Serological diagnosis
    - Extracellular product: streptolysin O, DNase B, hyaluronidase, NADase, and streptokinase
    - Antistreptolysin O (ASO)
- Molecular method
  - Nucleic acid amplification method
    - Real-time PCR: LightCycler *Strep* A assay (Roche Diagnostics)
    - Target gene is *PTSI* gene
    - Sensitivity 93% and specificity 98%
    - Offers advantages of rapid detection, PCR product quantitation, and closed system with minimal to no contamination
    - Direct detection from throat swab
  - Clinical utility of molecular diagnostics
    - Significantly faster turnaround time compared to traditional methodologies
    - Improved newborn care and reduced patient hospital stays
    - Reduces unnecessary antibiotic use in uncolonized women and thereby reduces emergence of antibiotic-resistant group B *Streptococcus* (GBS) strains
  - Pitfalls of molecular diagnostics
    - False-positive results due to carryover contamination (detected by negative control)
    - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
    - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure

26.6 *Streptococcus agalactiae*

26.6.1 General Characteristics

- Facultative gram-positive β-hemolytic GBS
- GI tract is the natural human reservoir for GBS and is the likely source of vaginal colonization
- Approximately 10–30% of pregnant women are colonized with GBS in the vagina or rectum
- Possible sexual transmission (controversial)
- Vertical transmission can occur in utero or during parturition
- Pregnant women who are colonized require antibiotic prophylaxis (Fig. 26.7)

26.6.2 Clinical Presentations

- Early onset (first 5 days of life) neonatal infection
- 0.34–0.37 cases per 1,000 live births in recent years
- Bacteremia
- Late-onset (7 days–3 months of age) neonatal infection
- Bacteremia, fulminant meningitis, osteomyelitis, and septic arthritis
- Postpartum women
- Endometritis, cesarean section wound infection, and bacteremia
- Immune-compromised hosts
- Pyelonephritis, pneumonia, tracheobronchitis, cellulitis, septic arthritis, meningitis, endocarditis, and bacteremia

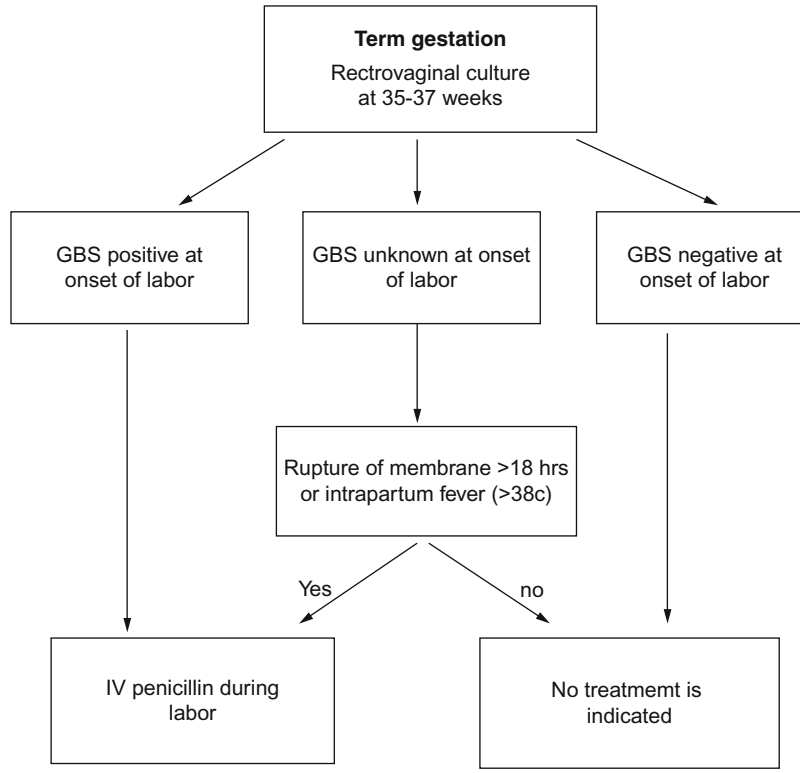


Fig. 26.7 Algorithm for group B *Streptococcal* (GBS) prophylaxis

26.6.3 Diagnostic Method

- Specimen (Fig. 26.8)
  - CSF, amniotic fluid, whole blood, vaginal, and rectal swab
  - Specimens (for molecular studies): must be tested within 24 h if at room temperature
  - Specimens stored between 2 °C and 8 °C are stable for up to 6 days
- Conventional method
  - Culture
    - Selective enrichment broth
      - Todd–Hewitt broth supplemented with either colistin (10 ug/ml) and nalidixic acid (15 ug/ml) or gentamicin (10 ug/ml) and nalidixic acid (15 ug/ml)
    - Granada medium
      - Selective and differential agar medium
      - Growth under anaerobic condition
      - Orange carotenoid pigmented colonies for 93–98.5% of GBS human isolates
    - Optimal culturing technique includes culturing both vaginal and anal samples (combined) at 35–37 weeks of gestation
    - Sensitivity and specificity, 77% and 97%, respectively
  - Biochemical identification tests
    - Hippurate hydrolysis – positive
    - CAMP tests – positive
    - Catalase negative, which differs from Listeria monocytogenes (catalase positive)
    - Specific identification of GBS
      - Slide agglutination tests
        - Not sufficiently sensitive for direct detection of GBS from clinical samples
      - GBS antigen detection assays
  - Immunological assays
    - Latex agglutination test
    - Immunoassay – rapid immunoassay

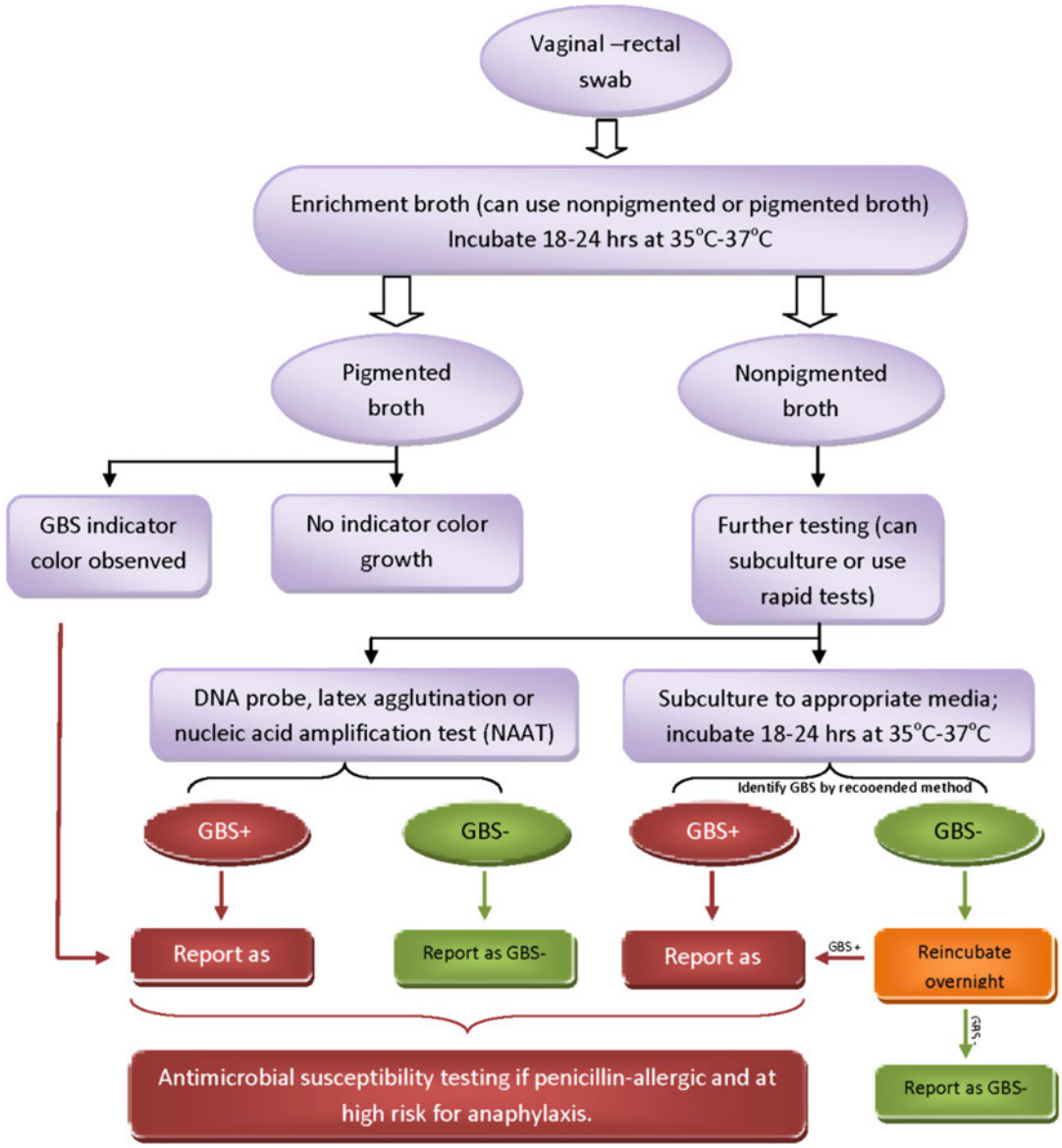


Fig. 26.8 Algorithm for recommended laboratory testing for prenatal screening for group B *Streptococcal* (GBS) colonization

- Low sensitivity
  - GBS colonization in pregnant women directly from vaginorectal swabs ranged from only 4% to 37%
- Molecular method
  - Nucleic acid hybridization method
    - AccuProbe group B *Streptococcus* culture identification test (Gen-Probe)
      - Targets specifically the GBS ribosomal RNA
      - Suitable to identify GBS from 18-h to 24-h cultures in selective enrichment broth
      - The test offers a rapid, nonsubjective method for the definitive identification of GBS

  - Sensitivity 94.7–100%, specificity 96.9–99.5%
- Nucleic acid amplification methods
  - Standard PCR
    - In-house PCR
    - Sample: CSF, amniotic fluid, whole blood, vaginal, and rectal swabs
    - Sensitivity, specificity, and positive and negative predictive values of 96.0%, 99.4%, 88.9%, and 99.8%, respectively
    - Targets 16S rRNA, 16S–23S spacer region, or *CFB* gene, which encodes the CAMP factor
  - Real-time PCR-based commercial assays
    - GeneXpert *Strep* B test (Cepheid)
      - The only FDA-cleared replacement for standard culture testing
      - Sensitivity 94% and specificity 96%
      - Automated DNA amplification and real-time detection in a single step
      - Test results in less than 1 h and may be used at the time of labor
      - Target gene is *CFB*
    - LightCycler *Strep* B assay (Roche Diagnostics)
      - Sensitivity 95% and specificity 98%
      - Target gene is *CFB*
      - Offers advantages of rapid detection, PCR product quantitation, and closed system with minimal to no contamination
- Clinical utility of molecular diagnostics
  - Significantly faster turnaround time compared with traditional methodologies (Fig. 26.9)
  - Improved newborn care and reduced patient hospital stays
  - Reduces unnecessary antibiotic use in uncolonized women and thereby reduces emergence of antibiotic-resistant GBS strains
- Pitfalls of molecular diagnostics
  - Available probe hybridization methods are suitable for GBS identification from overnight cultures in selective broth but are poorly sensitive for direct detection and identification of GBS from vaginorectal swabs obtained from pregnant women
  - False-positive results due to carryover contamination (detected by negative control)
  - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
  - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure

26.7 *Enterococci*, Vancomycin-Resistant

26.7.1 General Characteristics

- γ-Hemolytic, but some may show α- or β-hemolysis
- Normal component of fecal flora
- Mechanisms of glycopeptide resistance in *Enterococcus* spp. (Table 26.6)
  - Resistance is conferred by synthesis of different peptidoglycan precursors with decreased capability to bind vancomycin and/or teicoplanin
  - Six different genes (*vanA, vanB, vanC, vanD, vanE*, and *vanG*) confer resistance
  - They can be distinguished on the basis of the level, inducibility, and transferability of resistance to vancomycin and teicoplanin
  - *vanA* type glycopeptide resistance
    - Characterized by acquired inducible resistance to both vancomycin and teicoplanin
    - It is mediated by transposon Tn1546 or closely related genetic elements

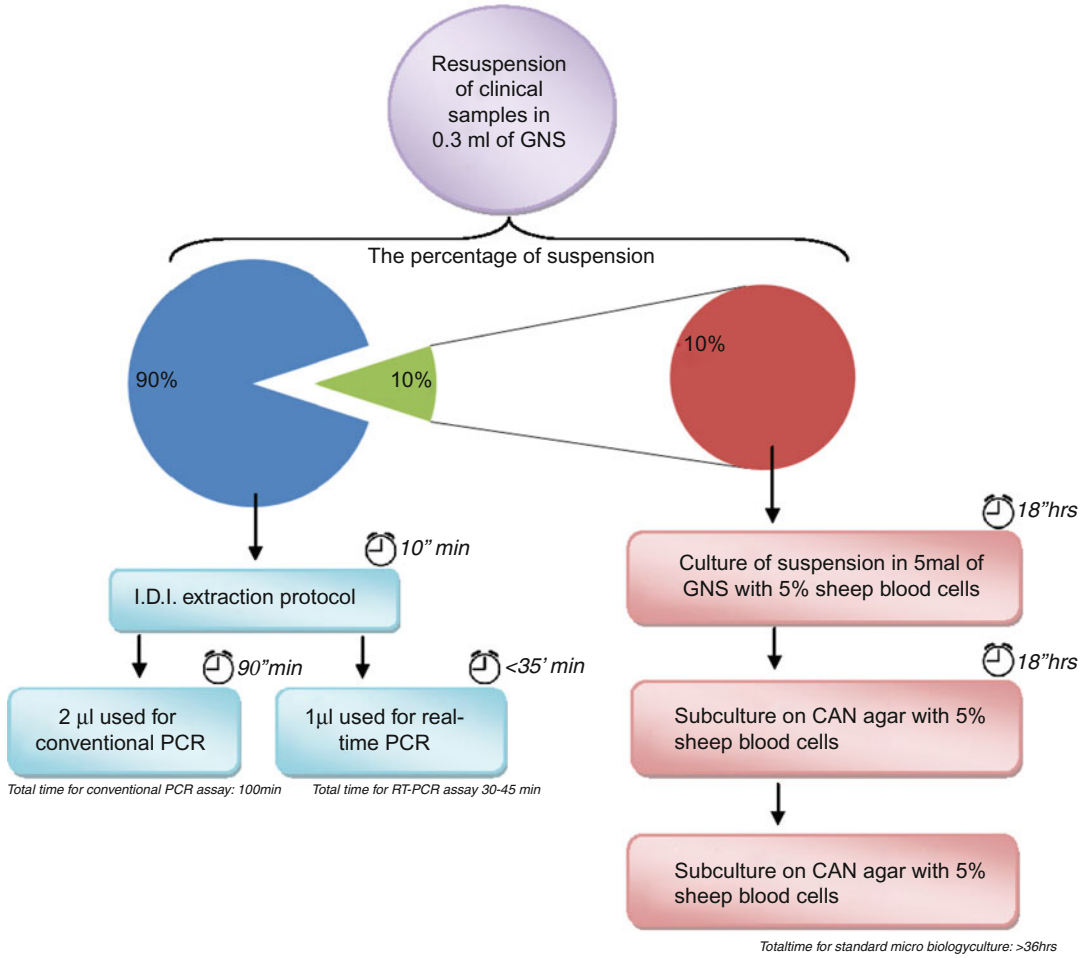


Fig. 26.9 Time from clinical sampling to results using three different methods to detect group B *Streptococcal* colonization in pregnant women at delivery

  - *vanB* type glycopeptide resistance
    - Characterized by acquired inducible resistance to various concentrations of vancomycin but typically not to teicoplanin
- High risk for vancomycin-resistant *enterococci* (VRE) colonization
  - Severely ill patients
  - Those receiving multiple and prolonged courses of antimicrobial agents
  - Long-term care facilities and urban referral hospitals
  - Solid organ transplant recipients and hematology patients
  - Healthcare worker
- Risk factor for VRE bacteremia
  - Hemodialysis
  - Organ transplantation
  - Receipt of corticosteroids, chemotherapy
  - Surgery
  - Severe illness
  - Long-term antibiotic administration
  - Neutropenia

26.7.2 Clinical Presentation

- Urinary tract infection
- Subacute endocarditis

Table 26.6 Characteristics of the types of resistance to glycopeptide antibiotics found in *Enterococcus* species

Phenotype	VANA	VANB	VANC	VAND	VANE	VANG
Genotype	*vanA*	*vanB*	*vanC1* and *vanC2/vanC3*	*vanD*	*vanE*	*vanG*
Genetic characteristics	Acquired	Acquired	Intrinsic	Acquired	Acquired	Acquired
Transfer resistance	Yes	Yes	No	No	No	No
Expression	Inducible	Inducible	Constitutive	Constitutive	Inducible	Inducible
MIC (μg/mL)						
Vancomycin	64 to >1,000	4 to >1,000	2–32	16–256	16	16
Teicoplanin	16–512	0.25–2	0.12–2	2–64	0.5	0.5
Species	*E. faecalis*	*E. faecalis*	*E. gallinarum*	*E. faecium*	*E. faecalis*	*E. faecalis*
	E. faecium	*E. faecium*	*E. casseliflavus*	*E. faecalis*		
	E. avium	*E. gallinarum*	*E. flavescens*			
	E. gallinarum	*E. durans*				
	E. durans					
	E. mundtii					
	E. casseliflavus					
	E. raffinosus					

Adapted from Zirakzadeh and Patel (2006)
Abbreviation: *MIC* minimum inhibitory concentrations

- Intraabdominal abscess
- Soft tissue infection
- Neonatal infections

26.7.3 Diagnostic Method

- Specimens
  - Perineal/rectal swab, stool, urine, drain sites, and open wounds. Environmental sampling for surveillance purpose
  - Specimens (molecular studies): can be tested within 24 h at room temperature
  - If not, it is recommended that specimens are refrigerated
  - Specimens stored between 2 °C and 8 °C are stable for up to 6 days
- Conventional Method
  - Gram stain
  - Culture
    - Nonselective broths and agar such as trypticase soy, heart infusion, or Todd–Hewitt, which may be supplemented with blood
    - Chromogenic agar is commercially available
  - Identification by biochemical methods (Table 26.7)
    - Hippurate and CAMP tests: negative
    - Growth in bile esculin
    - Growth in 6.5% NaCl
    - PYR – positive
    - Pigment and acid production
    - Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectrometry
      - MALDI-TOF MS technique has been demonstrated as a reliable tool for isolation identification in minutes
- Phenotypic methods for susceptibility testing
  - Disk diffusion (30 μg)
    - Breakpoints for determining VRE are susceptible (17 mm), intermediate (15–16 mm), and resistant (14 mm)
    - Plates require full 24 h
    - Unreliable for detecting resistance in strains with intermediate- or low-level resistance to vancomycin
  - MIC detected by
    - Broth dilution or agar dilution
    - Requires incubation for a full 24 h

Table 26.7 Differentiating features of various *Enterococcus* spp.

Species	Motility	Pigment	Xylose	Arabinose	Pyruvate
E. faecalis	Negative	Negative	Negative	Negative	Positive
E. faecium	Negative	Negative	Negative	Positive	Negative
E. casseliflavus	Positive[a]	Positive[b]	Positive	Positive	Variable
E. mundtii	Negative	Positive	Positive	Positive	Negative
E. raffinosus	Negative	Negative	Negative	Positive	Positive
E. gallinarum	Positive[a]	Negative	Positive	Positive	Negative

[a]Nonmotile strains have been found
[b]Nonpigmented strains have been found

    - Breakpoints for determining VRE are susceptible (<4 μg/mL), intermediate (8–16 μg/mL), and resistant (>32 μg/mL)
  - E-test (MIC on a strip)
  - Automated instruments
    - Vitek (bioMerieux)
    - MicroScan autoSCAN (Dade Behring)
    - Detect bacterial growth and metabolic reactions in the microwells of plastic test cards by measuring fluorescence
  - Agar screening plates (6 μg/mL of vancomycin)
    - An acceptable method of determining vancomycin susceptibility in the absence of a reliable MIC method
    - Cannot differentiate between intermediate- and high-level resistance
- Molecular method
  - DNA hybridization assay
    - AccuProbe *enterococcus* culture identification test (Gen-Probe)
      - For the identification of *Enterococcus avium, E. casseliflavus, E. durans, E. faecalis, E. faecium, E. gallinarum, E. hirae, E. mundtii, E. pseudoavium, E. malodorous*, and *E. raffinosus* isolated from culture
      - Sensitivity and specificity 100%
    - GPCPC-*E. faecalis*/OE PNA FISH (peptide nucleic acid fluorescent in situ hybridization) (AdvanDx)
      - PNA molecules: DNA mimics in which the negatively charged sugar-phosphate backbone of DNA is replaced with a noncharged polyamide or "peptide" backbone
      - A multicolor, qualitative nucleic acid hybridization assay
      - Rapid identification of *E. faecalis* and other enterococci directly from positive blood culture
      - Targets the species-specific rRNA in *E. faecalis* and other enterococci
      - Advantages of rapid hybridization kinetics
      - VRE EVIGENE Detection assay (AdvanDx) for detection of *vanA* and *vanB*
  - Amplification methods
    - Multiplex PCR-RFLP: targets *vanA* and *vanB*
    - Real-time PCR – LightCycler (Roche Molecular Diagnostics)
      - Dual fluorescence resonance energy transfer (FRET) hybridization probes: targeting *vanA, vanB*, and *vanC*
      - The analytical sensitivity was determined to be <10 targets/μL (50 copies/reaction tube)
    - Real-time PCR – SmartCycler (Cepheid) – FDA approved
      - Detects *vanA* and *vanB*
      - Sensitivity of 100% and a specificity of 96.8%
  - Clinical utility of molecular diagnostics
    - Rapid detection of colonization and associated infection with high mortality rate in high-risk individuals and

interventional isolation to prevent spread of the disease
- Pitfalls of molecular diagnostics
  - False-positive results with organisms other than enterococci (enteric anaerobes), which carry the *van* genes
  - False-positive results due to carryover contamination (detected by negative control)
  - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
  - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure

26.8 *Staphylococcus aureus*, Methicillin-Resistant

26.8.1 General Characteristics

- Aerobic, gram-positive cocci in clusters (grapelike)
- Catalase positive
- Coagulase positive
- May produce golden yellow colonies on agar
- Normal flora of humans found in nasal passages, skin, and mucous membranes
- Have numerous virulence factors including toxins like *tsst-1* and PVL
- *S. aureus* has a genome size of 2.8 Mb and possesses approximately 253 open reading frames encoding putative transport pumps
- *S. aureus* N315 was isolated as a methicillin-resistant *S. aureus* (MRSA), which contains a plasmid of 25 kb in size. Its genome size is 2.81 Mb with a 32.8% G + C content, 20 copies of insertion sequences, and five transposons
- Methicillin resistance is conferred by acquisition of a staphylococcal cassette chromosome (SCC*mec*), containing the *mecA* gene (Fig. 26.10)
- The *mecA* protein product is PBP2a, a penicillin-binding protein with reduced affinity for β-lactam rings (Fig. 26.11)
- Resistance is also conferred by the *BLAZ* gene, which encodes a β-lactamase

26.8.2 Clinical Presentation

- Cause a variety of community- and health-care-associated infections
- Furuncles and carbuncles
- Necrotizing pneumonia
- Wound infections
- Food poisoning
- Septicemia
- Acute endocarditis
- Osteomyelitis
- Septic arthritis
- Toxic shock syndrome
- Scalded skin syndrome

26.8.3 Diagnostic Method

- Specimens
  - Specimens include nasal or skin swabs, urine, and material from drain sites, skin, and open wounds
  - Specimens (for molecular studies) that can be tested within 24 h can be kept at room temperature; if not, refrigeration is recommended. Specimens stored between 2 °C and 8 °C are stable for up to 6 days
- Conventional methods
  - Gram stain
  - Culture
    - Nonselective media
    - Selective chromogenic agar
      - CHROMagar MRSA II (CMRSAII) (BD BBL)
      - ChromID MRSA (bioMérieux)
      - MRSA Select (Bio-Rad) Oxoid Chromogenic MRSA agar (Oxoid)
    - Broth media

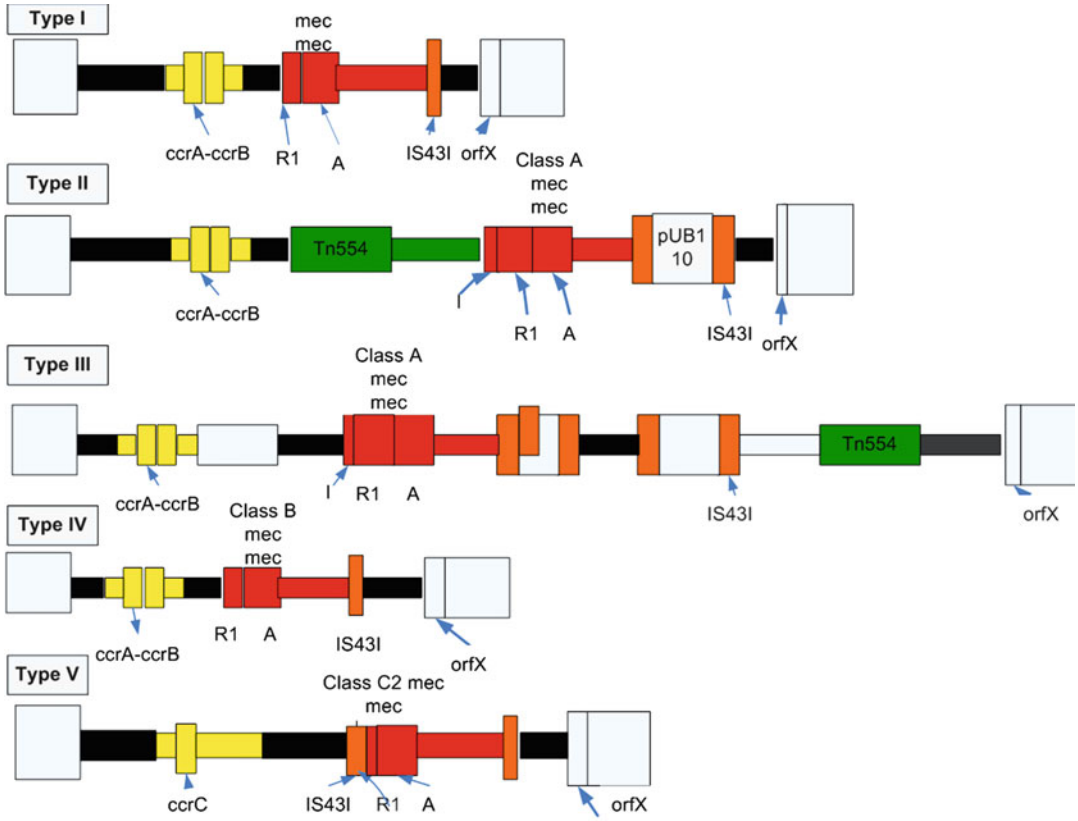


Fig. 26.10 Types of SCC*mec* regions in MRSA. The *mec* gene complex containing *mec*A and regulatory genes is drawn in *red* and the cassette chromosome recombinase (*CCR*) gene complex in *yellow*. Insertion sequences are drawn in *orange* and transposon Tn554 and areas surrounding it and homologous between the types in *green*, *ORFX*, a gene of unknown function containing the insertion site for SCC*mec*

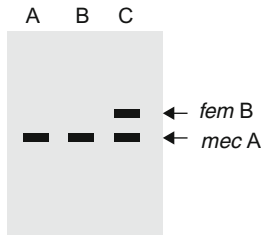


Fig. 26.11 Detection of *mec*A gene by PCR visualized by gel electrophoresis. *Lanes A* and *B*: bacterial isolates with *mec*A gene. *Lane C*: bacterial isolate with *mec*A and *fem*B genes

- Identification by biochemical method
  - Coagulase test
  - DNase and heat-stable nuclease tests (sensitivity 93% and specificity 96%)
- Immunological method
  - Direct latex agglutination reaction with a monoclonal antibody (sensitivity 96% and specificity 98%)
  - Passive hemagglutination
- MALDI-TOF mass spectrometry
  - Characterizes mostly proteins but, increasingly, nucleic acid
  - The organic molecule is ionized and subsequently identified based on the mass-to-charge ratio
  - Advantages: inherent accuracy and the high speed of signal acquisition
  - A powerful method for rapid identification of clonal strains of *S. aureus*
  - Useful for tracking nosocomial outbreaks of MRSA and for epidemiologic studies

- Phage amplification KeyPath MRSA/MSSA blood culture test (MicroPhage)
  - Does not require technical expertise or sophisticated equipment
  - KeyPath MRSA/MSSA blood culture test (MicroPhage)
    - FDA-approved test
    - Identify and determine methicillin resistance or susceptibility of *S. aureus* directly from gram-positive blood cultures
  - MRSA/MSSA screening test
    - A nasal screening test for use in the rapid and direct detection of *S. aureus* and determination of methicillin resistance/susceptibility using nasal swabs from patients

- Phenotypic antimicrobial susceptibility methods
  - Disk diffusion
    - On Mueller Hinton agar plate
    - 1-μg oxacillin disk or 30-μg cefoxitin disk
    - Breakpoints for determining MRSA
    - Oxacillin: susceptible (13 mm), intermediate (11–12 mm), and resistant (10 mm)
    - Cefoxitin: susceptible ($\geq$22 mm) and resistant ($\leq$21)
    - Plates require a full 24-h incubation
  - MIC detected by
    - Broth dilution or agar dilution
    - Requires incubation for a full 24 h
    - Breakpoints for determining MRSA
      - Oxacillin: susceptible (<2 μg/mL) and resistant (>4 μg/mL)
      - Cefoxitin: susceptible (<4 μg/mL) and resistant (>8 μg/mL)
  - E-test
  - Automated instruments
    - Vitek, MicroScan, Phoenix
    - Detect bacterial growth and metabolic reactions in the microwells of plastic test cards by measuring fluorescence
  - Agar screening plates (6 μg/mL of oxacillin)
  - Latex agglutination test
    - Based on detection of PBP2a by agglutination with latex particles coated with monoclonal antibodies to PBP2a
    - A penicillin-binding protein 2a latex agglutination (PBP-LA) assay (Denka Seiken, Japan/Oxoid)
      - Rapid, simple, and inexpensive
      - FDA-approved test for the identification of methicillin resistance in *S. aureus* bacteria from culture plates
      - Lower sensitivity than that of PCR assay
- Molecular methods
  - Nucleic acid hybridization method
    - AccuProbe test (Gen-Probe)
      - Culture identification test for *S. aureus*
      - Specimen (broth culture or solid media method)
    - GPCC-*S. aureus*/CNS PNA FISH (peptide nucleic acid fluorescent in situ hybridization), (AdvanDx)
      - PNA molecules: DNA mimics in which the negatively charged sugar-phosphate backbone of DNA is replaced with a noncharged polyamide or "peptide" backbone
      - A multicolor, qualitative nucleic acid hybridization assay
      - Rapid identification of *S. aureus* and CoNS from gram-positive cocci in clusters of (GPCC) positive blood cultures
      - Targets the species-specific rRNA in *S. aureus* and CoNS
      - Advantages of rapid hybridization kinetics
  - Nucleic acid amplification method
    - PCR – conventional (in-house method)
      - Genome targets are nuclease (*NUC*), coagulase (*COA*), protein A (*SPA*), *FEMA* and *FEMB*, or 16S rRNA
    - LightCycler SeptiFast test (Roche Diagnostics)
      - Targets ITS region of *S. aureus*
      - Quick and highly sensitive
      - Direct detection from blood samples

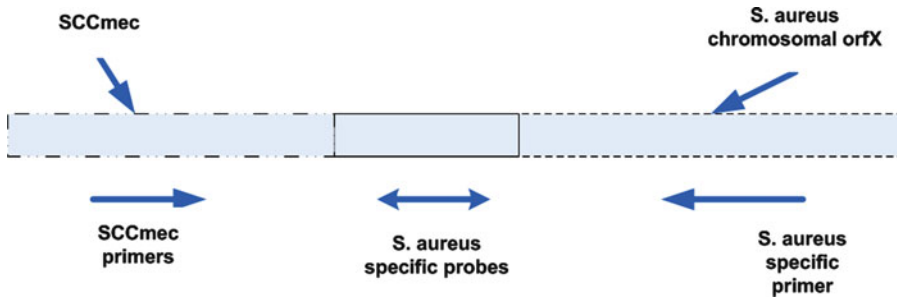


Fig. 26.12 Primer- and probe-binding sites for real-time PCR detection of *mecA*

- Cepheid analyte-specific reagent – GeneXpert (Cepheid) (Fig. 26.12)
  - Automated DNA amplification and real-time detection in a single step
  - Test results in <1 h
  - Target gene is 98-bp region of the *Staphylococcus* protein A (*SPA*) gene
- BD GeneOhm StaphSR assay (BD)
  - Multiplex real-time PCR method
  - Amplifies a specific target sequence of *S. aureus* and a specific target near the staphylococcal cassette chromosome (SCC) *mec* insertion site and the *orfX* junction in MRSA
  - Provides distinctive results for each target and differentiates between MSSA and MRSA

– Clinical utility of molecular testing for MRSA
  - Rapid surveillance method for infection control
  - Rapid detection of MRSA permits early institution of appropriate antibiotic therapy and control measures

– Pitfalls of molecular testing for MRSA
  - False-positive results due to carryover contamination (detected by negative control)
  - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
  - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure

26.9 *Legionellaceae*

26.9.1 General Characteristics

- Previously undescribed human pathogenic bacteria until major outbreak of fulminant pneumonia in 1976 in Philadelphia, PA
- Fifty species of *Legionella* and 15 serotypes of *Legionella pneumophila* are recognized
- Widely distributed in environment, particularly water towers, lakes, and water supplies, in association with algae
- Faint gram-negative motile rods, do not stain by hematoxylin and eosin
- Best identified in tissue with the Dieterle silver impregnation method, seen both intracellularly and extracellularly
- Buffered charcoal yeast extract agar (BCYE) is best for primary isolation
- Slow growth, up to 5 days
- Grows best in 2–5% carbon dioxide environment
- Weakly oxidase positive, produces catalase, and hydrolyzes starch, gelatin, and hippurate
- Up to 80–90% of fatty acids are branched chains

- Antigenic structures have six distinct serogroups defined by immunofluorescent staining of whole bacterial cells
- Serogroup 1 detected in clinical disease more often

26.9.2 Clinical Presentation

- Pneumonia
  - May have extrapulmonary manifestations
  - Usually presents with malaise, myalgia, headache, fever, chills, and diarrhea
  - May progress to respiratory failure
- Pontiac fever
  - Self-limited illness
  - Fever, malaise, and headache
  - No respiratory disease
- Other infections
  - Bacteremia
  - Hemodialysis-associated
  - Pericarditis
  - Meningoencephalitis

26.9.3 Diagnostic Method

- Specimens
  - Blood, lower respiratory tract tissue, or respiratory secretions
  - For molecular testing – BAL fluid, bronchial wash, sputum (minimum 2 mL), and throat swab in 1-mL transport media
  - Transported to the lab within 24 h post collection
- Conventional tests (Table 26.8)
  - Culture (preferred method for the diagnosis of *Legionella*)
    - Can culture organism from blood, lower respiratory tract tissue, or secretions
    - Buffered charcoal yeast extract agar supplemented with ketoglutarate
    - Indicated in patients with severe pneumonia, outbreak scenarios, elderly persons, smokers, immunosuppressed individuals, those with chronic lung disease, and patients who reside in hospitals with *Legionella*-colonized water supplies
  - Direct fluorescent antibody staining
    - It can be done on lower respiratory tract secretions or tissue
    - It uses fluorescein isothiocyanate-conjugated anti-*L. pneumophila* monoclonal antibodies to detect *L. pneumophila* surface antigen
    - It can provide results within 2–4 h, but it is technically demanding and has low sensitivity
    - It is now generally not accepted as sufficient for the diagnosis of *Legionella* infection in the absence of other supporting evidence
  - Urinary antigen detection
    - Commercial kits that use both radioimmunoassay and enzyme immunoassay methodologies are available to detect *L. pneumophila* urine antigen serogroup 1
    - An immunochromatographic assay (Binax) is commercially available
    - It is easy to perform and can provide a result within 15 min
    - *Legionella* antigenuria can be detected as early as 1 day after onset of symptoms and persists for days to weeks
  - Serologic testing
    - Indirect immunofluorescence is the standard reference test
    - A fourfold or greater increase in antibody (IgG or IgM) titer to >128 is considered diagnostic. Sensitivity 40–60%, specificity 96–99%
    - Cross-reactive antibody formation among members of the family of *Legionellaceae* and non-*Legionella* bacteria can make it difficult to determine the infecting species
  - MALDI-TOF mass spectrometry
    - MALDI-TOF MS technique has been demonstrated a reliable tool for isolation identification and subtyping in minutes
- Molecular methods
  - Conventional PCR with gel identification (home brew or commercial kit provided

Table 26.8 Diagnostic tests for *Legionella* infection

Test	Turnaround time	Sample type	Sensitivity (%)	Specificity (%)	Comments
Culture	3–7 days	LRT[a]	<10–80	100	Detects all species
		Blood	<10	100	Too insensitive for clinical use
Direct fluorescent antibody	<4 h	LRT	25–70	>95	Technically demanding
Antigen detection	<1 h	Urine	70–90	>99	Only reliable for detection of *L. pneumophila* serogroup 1
Serologic testing	3–10 weeks	Serum	60–80	>95	Must test both acute- and convalescent-phase serum samples; single titer results can be misleading
PCR	<4 h	LRT	80–100	>90	No commercially available assay for testing clinical samples; detects all species and serogroups
		Serum	30–50	>90	
		Urine	46–86	>90	

Adapted from Murdoch (2003)
[a]*LRT* lower respiratory tract

by Minerva Biolabs [Berlin, Germany]) – targets 16S rRNA, 5S rRNA, or *MIP* gene
  - Conventional PCR and reverse dot blot hybridization to probes immobilized on nylon membranes with biotinylated primers (home brew)
  - LightCycler (Roche) real-time PCR – targets *MIP* gene or 16S rRNA
  - Real-time PCR (ABI prism 7700) – targets 16S rRNA or 23S–5S spacer region
  - SDA with an energy transfer (ET) detection method (BD ProbeTec ET) – cleared by FDA and targets *MIP* gene
- Clinical utility
  - The urine antigen test combined with *Legionella* PCR is the best initial testing strategy that will detect *Legionella* species and provide results within a short time frame
- Pitfalls
  - *Legionella* species exist widely in environment; false-positive results happen due to contaminated DNA in reagents and specimens
  - False-positive results due to carryover contamination (detected by negative control)
  - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
  - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure

26.10 *Mycobacterium tuberculosis*

26.10.1 General Characteristics

- Acid-fast, nonspore-forming, slowly growing, gram-positive rod
- True branching occurs in vitro under special culture conditions
- Cell wall backbone contains two polymers, peptidoglycan and arabinogalactan, covalently linked by phosphodiester bonds
- Cell wall lipids account for 60% of dry weight of cell wall
  - Mycolic acids
  - Cord factor
  - Wax D
- Obligate aerobe
- Slow growing, requiring 10–20 days at 37 °C before colonies can be visualized

Table 26.9 Comparison of characteristic acid-fast stain, culture, and PCR-based methods for *M. tuberculosis*

Method	Advantage	Disadvantages
Acid-fast stain	Rapid detection	Inadequate sensitivity and specificity
Culture	Allows susceptibility test High specificity with nucleic acid-based identification	Inadequate sensitivity Prolonged time to result Requires further identification after positive culture
PCR	High sensitivity and specificity Rapid detection time No transport requirement Allows detection from noninvasive specimens Best for diagnosis and screening	Potential contamination Unable to assess viability Limited ability for genotype and susceptibility testing

Adapted from Yang and Rothman (2004)

- Primary isolation requires complex media containing either egg–potato base or serum–agar base
- Glycerol is preferred carbon and energy source
- Catalase and peroxidase are present
- Iron assimilation via one hydrophilic and one lipophilic transport system
- Mycobacteriophages
  - Double-strand DNA phages, unassociated with virulence
- Highly resistant to drying
- *M. tuberculosis* genome has been fully sequenced
  - First major pathogen to be sequenced
  - 4,411,522 bp
  - 3,924 open reading frames
  - GC content of 65.6%
  - +/−70% of the genes can be identified at this stage, the remainder are unique and encode proteins with unknown functions
  - 59% of genes are transcribed in the same direction as chromosomal replication

26.10.2 Clinical Presentation

- 85% of cases are pulmonary
- Symptoms of active TB disease include chronic fatigue, a bad cough that lasts longer than 2 weeks, chest pain, hemoptysis, increased sputum production, loss of weight, loss of appetite, chills, fever, and night sweats
- Nonspecific constitutional symptoms
- In disseminated disease (miliary TB), lesions may develop in any organ, with corresponding organ-specific symptoms

26.10.3 Diagnostic Method

- Specimens (Table 26.9)
  - Sputum, respiratory specimens, CSF, blood, urine, and other nonrespiratory specimens
  - Gastric lavage fluid is an acceptable surrogate for sputum in pediatric and neurologically compromised patients
  - Sputum specimens should be collected early in the morning on three occasions
  - Midstream urine specimens should be collected in a sterile plastic container on three early mornings
  - CSF requires a high volume of aspirates, at least 5 mL
  - Specimens should be sent to the laboratory within 24 h after collection
  - Specimens can be stored at 2–8 °C for up to 7 days before processing
- Conventional method
  - Staining and microscopy
    - Traditional Ziehl–Neelsen
    - Kinyoun cold procedure
    - Auramine–rhodamine fluorochrome stain
    - Fluorescent light-emitting diode (LED) microscopy
    - High specificity but sensitivity as low as 22–78% compared with culture

Table 26.10 Classification of mycobacteria and their biochemical reactions

Mycobacterial group	Species	Key biochemical tests
M. tuberculosis complex	*M. tuberculosis, M. bovis*, *M. africanum, M. microti*, *M. canetti*	Niacin, nitrate reduction, susceptibility to TCH if *M. bovis* is suspected
NTM		
Photochromogens	*M. kansasii*, *M. marinum*	Tween-80 hydrolysis, nitrate reduction, pyrazinamidase, 14-day arylsulfatase, urease, niacin
Scotochromogens	*M. scrofulaceum*, *M. gordonae*	Permissive growth temperature (*M. xenopi*: optimal growth 45 °C), Tween-80 hydrolysis, nitrate reduction, semiquantitative catalase, urease, 14-day arylsulfatase
Nonphotochromogens	*M. avium*, *M. intracellulare*, *M. xenopi, M. terrae*	Heat-resistant and semiquantitative catalase activity, nitrate reduction, Tween-80 hydrolysis, urease, 14-day arylsulfatase, tellurite reduction, acid phosphatase activity
Rapidly growing	*M. fortuitum*, *M. perigrinum*, *M. abscessus*, *M. chelonae*	Growth on MacConkey agar, nitrate reduction, Tween-80 hydrolysis, 3-day arylsulfatase, iron uptake

- Mycobacterial culture using solid media
  - NaOH–NALC decontamination and concentration is needed for nonsterile specimens (e.g., urine, sputum) processing prior to culture
  - Lowenstein–Jensen media
    - Egg-based media
    - The traditional solid media for culture of mycobacteria
    - Contains malachite green dye to inhibit growth of contaminating organisms
  - Middlebrook media (7H10 and 7H11)
    - Agar-based media
    - Easy detection of colony growth
    - Preferred for nontuberculous mycobacterium (NTM)
  - Advantages of solid media
    - Growth can be quantified
    - Colony morphology and pigmentation can be examined
    - Biochemical tests can be performed if warranted (Table 26.10)
    - Longer-term storage
- Mycobacterial culture systems
  - Gold standard
  - Automated or semiautomated
    - MB/BacT (Organon Teknika)
      - It is fully automated, nonradiometric system
      - Carbon dioxide is released into the medium by actively metabolizing microorganisms and is detected by a gas-permeable sensor
      - Color changes are monitored by a reflectometric detection unit
    - BACTEC MGIT 960 system (BD)
      - Fully automated, high-capacity, nonradiometric noninvasive instrument
      - It contains a modified Middlebrook 7H9 broth in conjunction with fluorescence quenching-based oxygen sensor to detect the growth of mycobacteria
      - In the presence of mycobacterial growth, fluorescence is detected using 365-nm ultraviolet transilluminator
- Identification method (nonmolecular)
  - Biochemical tests for *M. tuberculosis*
    - Positive for niacin accumulation
    - Positive for nitrate reduction
    - Growth inhibited by thiophene-2-carboxylic hydrazide
  - High-performance liquid chromatography
    - Detects the spectrum of mycolic acids present in the cell wall by comparing with in-house databases
    - Rapid direct testing
    - Inexpensive cost of consumables
    - Not as sensitive as nucleic acid amplification

Table 26.11 Updated guidelines for the use of nucleic acid amplification tests in the diagnosis of tuberculosis

NAA testing should be performed on at least one respiratory specimen from each patient with signs and symptoms of pulmonary TB for whom a diagnosis of TB is being considered but has not yet been established and for whom the test result would alter case management or TB control activities. The following testing and interpretation algorithm is proposed
Revised testing and interpretation algorithm
• Routinely collect respiratory specimens (e.g., sputum), process (liquefy, decontaminate, and concentrate), and test by AFB smear microscopy and culture as previously recommended. Specimen collection and microbiologic testing should not be delayed to await NAA test results • At least one specimen, preferably the first diagnostic specimen, from each patient to be tested by NAA should be processed, suspended in a sufficient volume of buffer to ensure adequate sample volume for all planned tests (e.g., microscopy, culture, and NAA), and tested using an NAA test for TB. NAA testing should be performed in accordance with the manufacturer's instructions or a validated standard operating procedure • Interpret NAA test results in correlation with the AFB smear results – If the NAA result is positive and the AFB smear result is positive, presume the patient has TB and begin anti-TB treatment while awaiting culture results. The positive predictive value of FDA-approved NAA tests for TB is >95% in AFB smear-positive cases – If the NAA result is positive and the AFB smear result is negative, use clinical judgment whether to begin anti-TB treatment while awaiting culture results and determine if additional diagnostic testing is needed. Consider testing an additional specimen using NAA to confirm the NAA result. A patient can be presumed to have TB, pending culture results, if two or more specimens are NAA positive – If the NAA result is negative and the AFB smear result is positive, a test for inhibitors should be performed and an additional specimen should be tested with NAA. Sputum specimens (3–7%) might contain inhibitors that prevent or reduce amplification and cause false-negative NAA results • If inhibitors are detected, the NAA test is of no diagnostic help for this specimen. Use clinical judgment to determine whether to begin anti-TB treatment while awaiting results of culture and additional diagnostic testing • If inhibitors are not detected, use clinical judgment to determine whether to begin anti-TB treatment while awaiting culture results and determine if additional diagnostic testing is needed. A patient can be presumed to have an infection with nontuberculous mycobacteria if a second specimen is smear positive and NAA negative and has no inhibitors detected – If the NAA result is negative and the AFB smear result is negative, use clinical judgment to determine whether to begin anti-TB treatment while awaiting results of culture and additional diagnostic tests. Currently available NAA tests are not sufficiently sensitive (detecting 50–80% of AFB smear-negative, culture-positive pulmonary TB cases) to exclude the diagnosis of TB in AFB smear-negative patients suspected to have TB

Source: CDC (2009)

  - Requires experienced technician
  - Large amount of biomass required
  - Expensive instrumentation
- Molecular methods (Table 26.11)
  - Direct detection from clinical specimen
    - PCR: Cobas Amplicor *M. tuberculosis* PCR (Amplicor, Roche Molecular Diagnostics)
      - It is a colorimetric method and is FDA approved for AFB smear-positive respiratory specimens
      - The target is a 584-bp segment of the 16S rRNA gene
      - Includes an internal PCR control
      - Sensitivity: overall sensitivity for respiratory specimens 79.4–91.9%, smear-negative specimens 40.0–73.1%
      - Specificity 99.6–99.8%
    - Transcription-mediated amplification (TMA): Amplified *M. tuberculosis* Direct test (AMTD, Gen-Probe)
      - A chemiluminescent method, FDA approved for AFB smear-positive and smear-negative respiratory samples
      - Autocatalytic, isothermal synthesis of RNA
      - RNA target: 16S rRNA
      - >1 billion copies of RNA amplicon are produced
      - The amplicon is detected by the hybridization protection assay (HPA)
      - Measured in a luminometer
      - Bloody specimens are not suitable
      - Lack of internal amplification control
      - Respiratory specimen: sensitivity 90.0–95.2%, specificity 97.6–100%

  - Nonrespiratory specimens: 86.8% sensitivity and 100% specificity
  - Sensitivity for AFB smear-positive specimens is 91–100% and for smear-negative specimens is 40.0–92.9%
- Strand displacement amplification (SDA): BDProbe Tec ET (energy transfer) *M. tuberculosis* direct detection assay (DTB) (BDProbe Tec: Becton–Dickinson)
  - Isothermal amplification
  - Fluorescent energy transfer detection
  - FDA approved
  - Target: a 95-bp region in IS6110 and 16S rRNA gene
  - An internal amplification control to verify no inhibition of the SDA reaction
  - The sensitivity for AFB smear-positive specimens (98.4%) is higher than for smear-negative specimens (40.3%)
- Nucleic acid sequence-based amplification (NASBA): GenoType Mycobacteria Direct assay for detection of *M. tuberculosis* complex and four atypical mycobacteria (Hain Lifescience)
  - Based on the nucleic acid sequence-based amplification applied to DNA strip technology
  - Three steps
    - Isolation of 23S rRNA
    - Amplification of RNA by NASBA method
    - The reverse hybridization on membrane strips
  - Simultaneous detection of *M. avium, M. intracellulare, M. kansasii, M. malmoense*, and MTBC
  - Sensitivity 92% and specificity 100%
- Real-time PCR: Xpert MTB/RIF (GeneXpert, Cepheid)
  - Automated molecular test for MTB and resistance to RIF
  - A single-use sample processing cartridge system that holds all required sample preparation and real-time PCR reagents and hosts the whole PCR reaction
  - Amplify an MTB-specific sequence of the *RPOB* gene
  - Internal extraction control assures extraction performance
  - Sensitivity: 95.6–99.1% (all culture positive), 99.0–100% (smear positive), and 83.3–92.8% (smear negative). Specificity 97.1–100%
  - Identification of mycobacteria from culture
    - See next section on nontuberculous mycobacterium (Sect. 26.11)
  - Pitfalls of molecular tests
    - False-positive results due to carryover contamination (detected by negative control)
    - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
    - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure
  - Clinical utility of molecular tests (Fig. 26.13)
    - Early diagnosis of TB and determination of drug resistance is important for the initiation of treatment and interruption of the chain of transmission
    - Can easily distinguish *M. tuberculosis* from NTM
  - Antimicrobial susceptibility method for MTB
    - Drug susceptibility tests must be performed in the following circumstances
      - All initial isolates of *M. tuberculosis*
      - Isolates from patients who remain culture positive after 3 months of treatment
      - Isolates from patients who are clinically failing treatment
      - An initial isolate from a patient relapsing after previously successful TB treatment
- Mechanisms of resistance (Table 26.12)

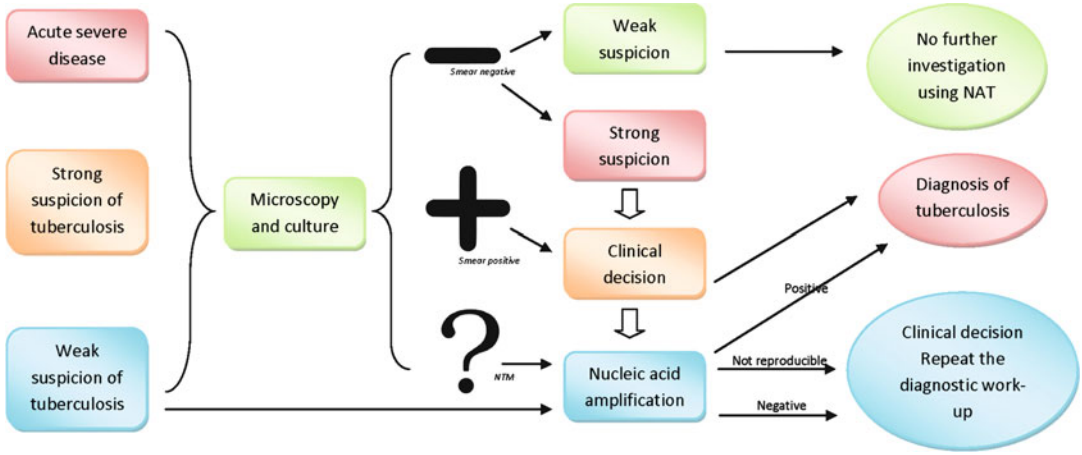


Fig. 26.13 Clinical settings and applications of nucleic acid test (NAT) for tuberculosis diagnosis

Table 26.12 Mechanism of drug resistance in *M. tuberculosis*

Drug	Gene	Gene function	Role	Action mechanism	Mutation frequency (%)
Isoniazid	*KATG*	Catalase–peroxidase	Prodrug conversion	Inhibition of mycolic acid biosynthesis and other multiple effects	50–95
	INHA	Enoyl ACP reductase	Drug target		8–43
Rifampicin	*RPOB*	B subunit of RNA polymerase	Drug target	Inhibition of RNA synthesis	95
Pyrazinamide	*PNCA*	Nicotinamidase/ pyrazinamidase	Prodrug conversion	Depletion of membrane energy	72–97
Ethambutol	*EMBCAB*	Arabinosyl transferase	Drug target	Inhibition of arabinogalactan synthesis	47–65
Streptomycin	*RPSL*	S12 ribosomal protein	Drug target	Inhibition of protein synthesis	52–59
	RRS	16S rRNA	Drug target		8–21
	GIDB	rRNA methyltransferase (G527 in 530 loop)	Drug target		?
Amikacin/ kanamycin	*RRS*	16S rRNA	Drug target	Inhibition of protein synthesis	76
Capreomycin	*TLYA*	2′-O-methyltransferase			
Quinolones	*GYRA*	DNA gyrase subunit A	Drug target	Inhibition of DNA gyrase	75–94
	GYRB	DNA gyrase subunit B			
Ethionamide	*ETAA/ ETHA*	Flavin monooxygenase	Prodrug conversion	Inhibition of mycolic acid synthesis	
	INHA		Drug target		
PAS	*THYA*	Thymidylate synthase	Drug activation	Inhibition of folic acid and iron metabolism	

Adapted from Zhang and Yew (2009)

– Resistance of an antituberculosis drug
  - Spontaneous chromosomal mutations at a frequency of 10^{-6} to 10^{-8} mycobacterial replications
  - No mobile genetic elements such as plasmids and transposons
  - Not all resistance mechanisms have been discovered

- Isoniazid
  - The most common
  - INH monoresistant TB can be treated easily with standard chemotherapy using INH, RMP, PZA, and EMB
  - Activated by catalase–peroxidase enzyme (KATG) encoded by the *KATG* gene
  - Primary target of INH inhibition
    - The INHA enzyme (enoyl-acyl carrier protein reductase)
    - Involved in elongation of fatty acids in mycolic acid synthesis
  - Mutation in *KATG* is the main mechanism of INH resistance
    - KATG S315T: the most common
  - Mutations in the promoter region of *MABA/INHA* operon
  - Overexpression of INHA
    - Low-level resistance
  - Mutations at the INHA active site
    - Lowering the INHA affinity to the INH–NAD adduct
    - Low-level resistance
  - Mutations in the promoter region of *AHPC*
  - Mutations in *MSHA*
- Rifampicin
  - Much more ominous prognosis
  - Recommendation for treatment: INH, PZA, and EMB for 18–24 months
  - Interferes with RNA synthesis by binding to the beta subunit of the RNA polymerase
  - Mutations in *RPOB*
    - Generally high-level resistance
- Pyrazinamide
  - Prognostic value in MDR-TB
  - *PNCA* mutations
    - Highly diverse and scattered along the gene
    - Unique to PZA resistance
- Ethambutol
  - Prognostic value in MDR-TB
  - Mutations in the *EMB*CAB operon: *EMBB*, *EMBC*
- Aminoglycoside
  - Resistance to streptomycin
    - Mutations in *RPSL* and *RRS* gene
  - Resistance to kanamycin
    - Mutations in *RRS* gene
- Fluoroquinolone
  - A pivotal position and a poor outcome in the treatment of MDR-TB
  - Mutations in quinolone resistance-determining region
  - Expression of MFPA

- Conventional methods
  - Phenotypic method
    - Absolute concentration method
      - Uses a standardized inoculum grown on drug-free media containing graded concentrations of the drug to be tested
      - Resistance is expressed in terms of the lowest concentration of the drug that inhibits growth, MIC
    - Resistance ratio method
      - Compares the growth of unknown strain of tubercle bacilli with that of a standard laboratory strain
      - Resistance is expressed as the ratio of the MIC of the test strain to the MIC of the standard strain in the same set
    - Proportion method
      - Enables a precise estimation of the proportion of mutants resistant to a given drug
      - For each drug tested, several dilutions of standardized inoculums are inoculated onto control and drug-containing agar media
      - The extent of growth in the absence or presence of drug is compared and expressed as a percentage
      - If growth at the critical concentration of a drug is $>1\%$, the isolate is considered clinically resistant
      - Many rapid testing methods are used to determine drug susceptibility of *M. tuberculosis*: BACTEC 460, MGIT 960, MB/BacT system, and ES II system
    - Radiometric method
      - Uses liquid medium containing 14 °C-labeled growth substrate

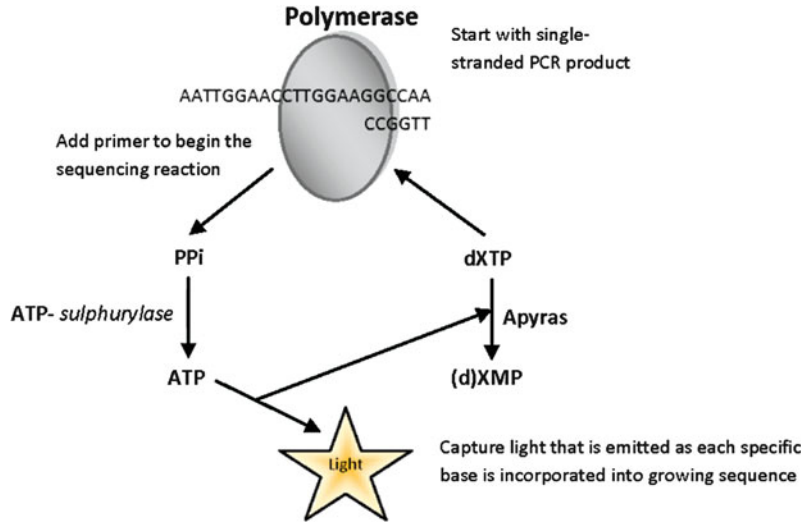


Fig. 26.14 Schematic diagram of a pyrosequencing reaction: The light is emitted as each specific base is incorporated into the growing DNA sequence and the light signals are translated into sequence data. Abbreviations: *ATP* adenosine triphosphate, *(d) XMP* deoxynucleotic monophosphate, *dXTP* deoxyxanthine triphosphate, *PPi* inorganic pyrophosphate

  - Growth is indicated by the amount of 14 °C-labeled carbon dioxide (CO_2) released, as measured by the BACTEC 460 instrument
- Microscopic observation drug susceptibility (MODS) assay
  - A broth microtiter method designed to detect *M. tuberculosis* complex and to detect resistance to isoniazid and rifampin
- Colorimetric assay
  - Based on the observation that growing tubercle bacilli convert a yellow dye to a purple color
- MDR-XDR TB color test
  - Thin-layer agar (TLA) technology
  - Both culture and direct AST methods on a single agar plate
  - Radiometric method

– Genotypic method
- PCR and DNA sequencing
  - Considered as the reference method for detection of drug resistance mutations
  - The most widely used genotypic method
  - To determine the presence or absence of a specific mutation
- Line probe assay
  - A series of steps including DNA extraction, amplification of DNA, hybridization, and colorimetric development
  - INNO-LiPA Rif TB (Innogenetics NV)
    - Ten oligonucleotide probes: one specific for *M. tuberculosis* complex, five wild-type probes (S1–S5), and four probes (R) for the detection of the common mutations that cause resistance to RMP
    - Sensitivity 95%, specificity 100%
    - Insufficiently sensitive to be used directly on unselected specimens
  - GenoType MTBDR (Hain Lifescience)
    - Detection of resistance mutation in *RPOB* (RMP resistance)
    - Did not detect INH resistance
  - GenoType MTBDR plus (Hain Lifescience)
    - Detection of the most common resistance mutations in *RPOB* (RMP resistance), *KATG*, and *INHA* gene (INH resistance)
    - Strains or directly in smear-positive specimens
- Pyrosequencing (Fig. 26.14)
  - Rapid detection of rifampicin resistance
  - Target: 180-bp region of the *RPOB* gene

  - Full agreement with other molecular methods (line probe assay and cycle sequencing)
- DNA microarray
  - Used for rapid detection of mutations associated with TB drug resistance
  - Can simultaneously detect different drug-resistant mutations
  - CombiChip Mycobacteria (GeneIn): *KATG*, *INHA*, and *RPOB* genes
- Real-time PCR-based Cepheid assay
  - Detection of resistance to INH and RMP
  - Sensitivity 85% and 98% for INH and RMP, specificity 100%
- PCR/ESI-MS (PCR followed by amplicon characterization using electrospray ionization mass spectroscopy)
  - A rapid, high throughput method
  - T5000 biosensor (Ibis) and PLEX-ID (Abbott)
  - 16 primer pairs in eight multiplexed reactions for multilocus PCR
  - Detect and identify gene mutations associated with INH, RIF, EMB, and FQ resistance
  - Limitations
    - Restricted usable amplicon size (preferably less than 160 nucleotides long)
    - The limited information content compared to traditional sequencing
    - Since this uses aggregate base composition signatures and not sequencing, potentially ambiguous results might occur if two mutations that result in no mass change are simultaneously present within the same amplicon

26.11 *Nontuberculous mycobacteria*

26.11.1 General Characteristics

- Free-living organisms and ubiquitous in the environment
- Important reservoirs include water, soil, animals, and others
- Recognized as true pathogens and important causes of human infection (Table 26.13)
- Most mycobacteria grow optimally between 35 °C and 37 °C in 5–10% CO_2 except *M. marinum, M. ulcerans, M. chelonae*, and *M. haemophilum* (25–33 °C)
- No known primary animal host; usually present in the soil
- No evidence for human–human transmission
- Classification (Table 26.10)
  - Four categories according to Runyon system
  - Divided by growth rates and pigment production
    - Group 1: slow grower, photochromogens (pigment producers in the presence of light)
    - Group 2: slow grower, scotochromogens (pigment producers in the absence of light)
    - Group 3: slow grower, nonchromogens
    - Group 4: rapid grower
- Must follow biosafety level 2 or 2-plus protocols
  - All specimen processing should be performed in a biosafety cabinet

26.11.2 Clinical Presentation

- Clinical manifestation of NTM
  - Pulmonary infection
  - Lymph node infection
  - Soft tissue infection
  - Bone and joint infection
  - Disseminated infection
- Clinical presentation and general features associated with different species of NTM (Table 26.14)
  - *M. avium* complex
    - Includes two established species, *M. avium* and *M. intracellulare*
    - Isolated from water, soil, plants, house dust, and dairy products
    - Causes pulmonary disease

Table 26.13 Clinical and microbiological criteria for diagnosing nontuberculous mycobacterial lung disease

Clinical (both required)	Microbiological
1. Pulmonary symptoms, nodular or cavitary opacities on chest radiograph, or a high-resolution computed tomography scan that shows multifocal bronchiectasis with multiple small nodules	1. Positive culture results from at least two separate expectorated sputum samples: If the results from one are nondiagnostic, repeating sputum smears and culture should be considered or
2. Appropriate exclusion of other diagnosis	2. Positive culture results from at least one bronchial wash or lavage or
	3. Transbronchial or other lung biopsy with mycobacterial histopathological features (granulomatous inflammation or acid-fast bacilli) and positive culture for NTM or biopsy showing mycobacterial histopathological features (granulomatous inflammation or acid-fast bacilli) and one or more sputum or bronchial washings that are culture positive for NTM
	4. Expert consultation should be obtained when NTM are recovered that are either infrequently encountered or that usually represent environmental contamination
	5. Patients who are suspected of having NTM lung disease but do not meet the diagnostic criteria should be followed until the diagnosis is firmly established or excluded
	6. Making the diagnosis of NTM lung disease does not, per se, necessitate the institution of therapy, which is a decision based on the potential risks and benefits of therapy for individual patients

Adapted from Griffith et al. (2007)

  - Clustering of cases in AIDS patients, also associated with emphysema
  - Highly resistant to antituberculous drugs
- *M. ulcerans*
  - Produces chronic ulcerating skin disease in the tropics, near rivers and swamps
  - Usually on legs and arms
  - Treatment requires wide excision
- *M. kansasii*
  - Tap water is a major reservoir of *M. kansasii* associated with human disease
  - Causes pulmonary or disseminated disease
- *M. xenopi*
  - Almost exclusively from hot water and hot water taps within hospitals
  - Causes pulmonary disease
- *M. marinum*
  - Salt water, water tanks, and swimming pools are major reservoirs
  - Causes cutaneous ulcers
- *M. fortuitum–M. chelonae* complex
  - Isolated from natural water sources and tap water as well as from soil and dust
  - Requires temperatures of about 28–30 °C for primary isolation
  - Rapid growers
  - Causes cutaneous ulcers and lymphadenopathy

26.11.3 Diagnostic Method

- Specimens
  - Sputum, respiratory specimens, CSF, blood, and other nonrespiratory specimens
  - Gastric lavage fluid is an acceptable surrogate for sputum in pediatric and neurologically compromised patients
  - Properly labeled in sterile containers and immediately transported to the laboratory
  - Stored at 2–8 °C (for up to 7 days) until processed
- Conventional methods
  - Staining and microscopy

Table 26.14 Nontuberculous mycobacteria recovered from humans

Clinical disease	Common etiologic species	Morphologic features
Pulmonary disease	*M. avium* complex	Slow growth, not pigmented
	M. kansasii	Pigmented
	M. abscessus	Rapid growth, no pigment
	M. xenopi	Slow growth, pigmented
	M. malmoense	Slow growth, not pigmented
Lymphadenitis	*M. avium* complex	Slow growth, not pigmented
	M. malmoense	Slow growth, not pigmented
	M. scrofulaceum	Pigmented
Cutaneous disease	*M. marinum*	Photochromogen requires low temperatures (28–30 °C) for isolation
	M. fortuitum	Rapid growth, no pigment
	M. chelonae	Rapid growth, no pigment
	M. abscessus	Rapid growth, no pigment
	M. ulcerans	Slow growth, not pigmented
Disseminated disease	*M. avium* complex	Slow growth, not pigmented
	M. kansasii	Photochromogen
	M. chelonae	Not pigmented
	M. haemophilum	Not pigmented, requires hemin, low temperature, and CO_2 to grow

Adapted from Guidelines for Tuberculosis Control in New Zealand (2010)

- Traditional Ziehl–Neelsen
- Kinyoun cold procedure
- Auramine–rhodamine fluorochrome stain (preferred)
- High specificity but sensitivity as low as 22–78% compared with culture
- Organisms may be difficult to distinguish from *M. tuberculosis* on microscopic examination

– Mycobacterial culture using solid media
  - Selective and nonselective media
  - Lowenstein–Jensen media
    – Egg-based media
    – The traditional solid media for culture of mycobacteria
    – Contains malachite green dye to inhibit growth of contaminating organisms
  - Middlebrook media (7H10 and 7H11)
    – Agar-based media
    – Easy detection of colony growth
    – Preferred for NTM
  - Advantages of solid media
    – Growth can be quantified
    – Colony morphology and pigmentation can be examined
    – Biochemical tests can be performed if warranted
    – Longer-term storage
– Mycobacterial culture using liquid media
  - Gold standard
  - Automated or semiautomated
    – MB/BacT (Organon Teknika Corp., Durham, NC)
      - It is fully automated, nonradiometric system
      - Carbon dioxide is released into the medium by actively metabolizing microorganisms and is detected by a gas-permeable sensor
      - Color changes are monitored by a reflectometric detection unit
    – Mycobacteria growth indicator tube (MGIT)
      - It contains a modified Middlebrook 7H9 broth in conjunction with fluorescence quenching-based oxygen sensor to detect the growth of mycobacteria
      - In the presence of mycobacterial growth, fluorescence is detected using 365-nm ultraviolet transilluminator
    – BACTEC system (Becton–Dickinson) radiometric and nonradiometric

- It uses the same fluorescence quenching-based oxygen sensor as the MGIT (Becton–Dickinson) system to detect growth
- Nonautomated (manual)
  - Septi-check blood cultures
  - Isolator tube

– Identification method (nonmolecular)
- Pigmentation and growth characteristics
- Biochemical tests: require weeks of subcultures
- High-performance liquid chromatography
  - Detects the spectrum of mycolic acids present in the cell wall by comparing with in-house databases
  - Rapid direct testing
  - Inexpensive cost of consumables
  - Not as sensitive as nucleic acid amplification
  - Requires experienced technician
  - Large amount of biomass required
  - Expensive instrumentation
- MALDI-TOF mass spectrometry

- Molecular method
  – Direct detection from clinical specimen
    - Nucleic acid sequence-based amplification (NASBA)
      – GenoType Mycobacteria Direct assay for detection of *M. tuberculosis* complex and four atypical mycobacteria (Hain Lifescience)
      – Based on the nucleic acid sequence-based amplification applied to DNA strip technology
      – Three steps
        - Isolation of 23S rRNA
        - Amplification of RNA by NASBA method
        - The reverse hybridization on membrane strips
      – Simultaneous detection of *M. avium, M. intracellulare, M. kansasii, M. malmoense*, and MTBC
      – Sensitivity 92% and specificity 100%
    - DNA microarrays
      – Quick examination of multiple DNA targets in a single hybridization step
      – Oligonucleotide probes based on the 16S rRNA, the DNA gyrase subunit B (*GYRB*), or the *RPOB* gene
    - PCR–electrospray ionization (ESI) mass spectrometry
      – Has been reported to directly detect and differentiate mycobacterial species from clinical specimens
  – Identification of NTM from culture
    - DNA probe technology: AccuProbe
      – Nucleic acid probes for culture confirmation
      – Sonication lyses microbes, releases 16S rRNA
      – DNA probes are ssDNA probes labeled with acridinium ester for hybridization
      – The hybridization is detected by light emission in a luminometer
      – Can be used for both solid and liquid cultures
      – An ability to identify a series of clinically important mycobacteria
        - MTBC, *M. avium* complex, *M. avium, M. kansasii*, and *M. gordonae*
      – Sensitivity 99.2%, specificity 99.0%
      – Pitfalls
        - Not sensitive enough for direct detection on clinical specimens
        - Misidentification of *M. celatum* as *M. tuberculosis*
        - Cross-reactivity has also occurred with *M. terrae* complex
        - No commercial probe for majority of NTM
    - Line probe technology
      – Characteristics
        - PCR with biotinylated primers
        - Reverse hybridization with specific DNA probes
        - Immobilization on a strip
        - Colorimetric detection in an automated instrument
      – INNO-LiPA Mycobacteria V2 (Innogenetics)

    - Based on the amplification of the mycobacterial spacer region 16S–23S rRNA and simultaneous identification of 17 mycobacterial species in one strip test
    - Mixed population are easily identified
    - Sensitivity 100%, specificity 94.4–100%
  - GenoType Mycobacterium (Hain Lifescience)
    - Rapid, easy to perform, and easy to interpret
    - GenoType MTBC for *M. tuberculosis* complex
      - Based on the *GYRB* gene polymorphism
    - GenoType *Mycobacterium* CM (common mycobacteria) and GenoType *Mycobacterium* AS (additional species) for NTM
      - Use 23S rDNA as their target
      - Can distinguish almost 30 different NTM
    - Sensitivity 97.9% (CM) and 99.3% (AS), specificity 92.4% (CM) and 99.4% (AS)
- PCR-based sequencing
  - The gold standard for identification of mycobacteria
  - Comparison of the nucleotide sequence with a library of known sequences
  - Targets most commonly 16s rRNA region; alternate assays target *RPOB, GYRB, HSP65*, *RECA*, and 32-kDa protein genes or the 16S–23S rRNA gene spacer
  - Database for sequences
    - GenBank
    - EzTaxon
    - Noncommercial DB for 16S rRNA gene of type strain
    - RIDOM (Ribosomal Differentiation of Medical Microsystems database)
    - RDP (Ribosomal Database Project)
    - MicroSeq system (Applied Biosystems)
      - Commercial 16S ribosomal DNA sequencing system
      - Pitfalls Problematic identification if no matching sequence found in known database
      - Unable to differentiate *M. tuberculosis* from the other species in the complex, *M. chelonae* from *M. abscessus, M. simiae* from *M. genavense*, or *M. kansasii* from *M. gordonae*
      - High expense
- PRA method (PCR and restriction enzyme analysis)
  - Based on the amplification of the gene encoding the 65-kDa heat shock protein followed by restriction fragment length polymorphism
  - Two restriction enzymes: *BSTEII* and *HAEIII*
  - Isolates from both solid and liquid cultures can be used
  - Pitfalls: misidentification due to intraspecies genetic variability
- Pyrosequencing
  - A novel method of nucleic acid sequencing by synthesis
  - Based on the detection of released pyrophosphate (PPi) during DNA synthesis
  - Optimal for determining short sequences
  - For mycobacteria: targets 30-bp sequence of the hypervariable A region of the 16S rRNA gene
- DNA microarray
  - Based on hybridization of fluorescently labeled PCR amplicons of an unknown strain to a DNA array
  - Allows the identification of a large number of strains in one reaction
  - Quick examination of multiple DNA targets in a single hybridization step

  - Oligonucleotide probes based on the 16S rRNA, the DNA gyrase subunit B (*GYRB*), or the *RPOB* gene
- Clinical utility
  - Early diagnosis of infection and determination of drug resistance is important for the initiation of treatment and interruption of the chain of transmission
  - Can identify many different NTM species
- Pitfalls
  - False-positive results due to carryover contamination (detected by negative control)
  - False-negative results due to amplification inhibition (detected by internal control) or due to a loss of target during specimen preparation
  - Nucleic acid amplification assays cannot distinguish between living and dead organisms, and should not be used in patients treated in the previous 3 weeks as a test of cure
- Antimicrobial susceptibility method for NTM
  - Susceptibility testing of NTM is a controversial issue
  - There are no data to show that drug susceptibility test results predict clinical outcome for many NTM infections
  - CLSI has recently released recommendations to standardize the performance of NTM drug susceptibility tests
  - Indications for clarithromycin susceptibility testing of *Mycobacterium avium* complex (MAC)
    - Clinically significant isolate from a patient who has received previous macrolide therapy (i.e., clarithromycin or azithromycin)
    - Patients who have developed MAC bacteremia on macrolide preventative therapy
    - Patients failing or relapsing on macrolide therapy
    - Baseline isolates from significant MAC infections may also be tested (or stored and tested retrospectively if the patient does not respond to treatment)
  - *M. kansasii*
    - All initial isolates of *M. kansasii* should be tested against rifampicin
    - For patients failing or relapsing on treatment
    - For rifampicin-resistant isolates, the following antibiotics should be tested: isoniazid, ethambutol, rifabutin, clarithromycin, ciprofloxacin, streptomycin, and cotrimoxazole
  - Rapidly growing NTM
    - All clinically significant rapid growers should be subjected to tests against amikacin, cefoxitin, ciprofloxacin, clarithromycin, doxycycline, imipenem, and sulphonamide
    - Tobramycin should also be tested for *M. chelonae* isolates only
    - Conventional methods
      - See section on *M. tuberculosis* (Sect. 26.10)
    - Phenotypic methods
      - See section on *M. tuberculosis* (Sect. 26.10)

26.12 Molecular Diagnosis of Sepsis

26.12.1 General Characteristics

- Current gold standard of bloodstream microbial detection and identification
- Automated, continuous-monitoring culture systems are commercially available
- Followed by gram stain, subculture, and use of biochemical phenotypic methods
- Major limitations
  - Time-consuming process
  - Culture methods miss fastidious organisms that are difficult or impossible to culture
  - Cultures can be confounded if antibiotics are administered before the blood is sampled
- Bloodstream infections are associated with high rates of morbidity and mortality

- Sepsis and septic shock are the tenth leading cause of death in the US, accounting for 6% of all deaths

26.12.2 Clinical Presentation

- Bacteremia
  - The microbial load may be as low as one colony-forming unit (CFU)/ml
  - Making the microbiological diagnosis difficult
  - Transient
    - Lasting for a few minutes or a few hours
    - Associated with procedures involving anatomic sites colonized by normal microbial flora or with a manipulation of localized infected sites
  - Intermittent
    - Associated with closed-space infections
    - Recurrent episodes of bacteremia due to the same microorganism intermittently detected in blood
  - Continuous (persistent low grade)
    - Commonly associated with an intravascular focus of infection such as infective endocarditis or vascular graft infection
    - Recurrent episodes of bacteremia due to the same microorganism intermittently detected in blood
- Sepsis
  - A clinical syndrome related to an infectious process with important alterations in the inflammatory response and coagulation
  - Sepsis
    - SIRS due to suspected or confirmed infection
  - Severe sepsis
    - Sepsis complicated by organ dysfunction, hypoperfusion, or hypotension
  - Septic shock
    - Sepsis with hypotension despite adequate fluid resuscitation along with presence of perfusion abnormalities
  - Multiple organ dysfunction syndrome (MODS)
    - Sign and symptoms of severe multiple organ dysfunction

26.12.3 Diagnostic Method

- Conventional method
  - Blood culture
    - The current gold standard of bloodstream infections
    - The advantage of allowing the evaluation of antimicrobial susceptibility
    - Manual method
    - Continuous-monitoring blood culture systems (CMBCS)
      - BACTEC (Becton–Dickinson): analysis of CO_2 release using fluorescent sensors
      - BacT/Alert (bioMerieux): analysis of CO_2 release using colorimetric sensors
      - VersaTREK (TREK Diagnostic Systems): measuring pressure changes
- Molecular method
  - Species identification and antibiotic resistance determination from culture samples
    - Peptide nucleic acid (PNA) FISH (AdvanDx)
      - Fluorescence in situ hybridization with PNA probes
      - Use fluorescent-labeled peptide nucleic acid (PNA) probes targeting the rRNA genes of bacteria (*S. aureus*, *E. faecalis*, *E. coli*, or *P. aeruginosa*) and *Candida* species (*C. albicans*/*C. parapsilosis, C. tropicalis*, or *C. glabrata*/*C. krusei*)
      - Approved by the FDA
      - A new format (QUICHFISH) is available for detection and differentiation of S. aureus and CoNS, which can be completed within 30 minutes
    - StaphPlex system (Qiagen): multiplex PCR method
      - Simultaneous species-level identification and detection of Panton–Valentine leukocidin (PVL) and several antimicrobial resistance determinants of staphylococci
      - Directly from blood culture medium in which gram-positive cocci in clusters were seen by gram staining

  - A unique target-enriched multiplex PCR
  - Amplify and detect 18 *Staphylococcus*-specific genes simultaneously in one reaction
    - The *tuf* gene target provides identification and differentiation of coagulase-negative staphylococci
    - The *nuc* gene target is specific for *S. aureus*
    - The *mecA* gene confers resistance to methicillin
    - The *ermA* and *ermC* genes contribute to resistance to macrolides, lincosamides, and streptogramins
    - The *tetM* and *tetK* genes are responsible for resistance to tetracycline
  - The amplified products are further characterized by using a Luminex suspension array
- Hyplex BloodScreen (BAG): multiplex PCR
  - Multiplex PCR with the subsequent hybridization on an ELISA plate fluorescence in situ hybridization with PNA probes
  - The assay is available to detect resistance markers, such as *van* genes and beta-lactamase genes
- Prove-it Sepsis (Mobidiag): multiplex PCR
  - Multiplex PCR with the subsequent hybridization on a microarray
  - The microarray format allows the detection of a wider panel of bacterial species and of the *mecA* gene
- MALDI-TOF mass spectrometry (MS) (Bruker Daltonics; bioMerieux)
  - Colonies from an agar plate or liquid culture are mixed with a MALDI matrix and rapidly analyzed by MS
  - MALDI-TOF system works by comparison of the mass spectral signals obtained from postculture specimens with a database of spectra from reference standard spectra
  - A significant advantage: rapid with a minimal amount of labor compared with conventional methods
  - Disadvantage: Identification of mixed populations of bacteria will probably be difficult owing to dynamic range issues in the mass spectrometer
- Direct amplification and detection methods from blood samples
  - SeptiFast (Roche): multiplex real-time PCR
  - Identify the 25 organisms that account for more than 90% of the culturable pathogens associated with sepsis
  - Identify 10 bacteria at the species level, several more at the genus level, and five *Candida* species and *Aspergillus fumigatus*
  - SeptiFast and blood culture results were usually in agreement, suggesting that SeptiFast can add value as an adjunct to blood culture
  - SeptiTest (Molzym): broad-range PCR
    - Broad-range PCR with subsequent sequencing
    - Targeting the 16S rRNA genes of bacteria and the 18S rDNA of fungi
    - Unique sample preparation methodology that reduces the burden of human DNA in an extracted blood sample
  - Vyoo (SIRS-Lab): multiplex PCR
    - Multiplex PCR with subsequent gel electrophoresis
    - Addresses approximately 35 bacterial species and 6 fungal species
    - Detects a few genetic markers of antibiotic resistance
    - Allows the selective removal of human DNA by a patented affinity chromatography
  - PCR/electrospray ionization (ESI) mass spectrometry (MS)
    - Commercial product: TIGER, Ibis T5000/Abbott PLEX-ID

- BAC Spectrum: assay for direct analysis of bloodstream infections that runs on the PLEX-ID
- Multiple pairs of primers are used to amplify carefully selected regions of bacterial or fungal genomes
- The primer target sites are broadly conserved, but the amplified region carries information on the microbe's identity in its nucleotide base composition
- PCR/ESI-MS method detects pathogens with no bias due to culturability: Aerobic, anaerobic, culturable, fastidious, and unculturable organisms are identified in the same way
- An important aspect of PCR/ESI-MS is that the method is semi-quantitative

Further Reading

Aguero-Rosenfeld ME, Wang G, Schwartz I, et al. Diagnosis of Lyme borreliosis. Clin Microbiol Rev. 2005;18:484–509.

Attar ZJ, Chance ML, el-Safi S, et al. Latex agglutination test for the detection of urinary antigens in visceral leishmaniasis. Acta Trop. 2001;78:11–6.

Bartlett JG, Gerding DN. Clinical recognition and diagnosis of *Clostridium difficile* infection. Clin Infect Dis. 2008;46 Suppl 1:S12–8.

Bergeronand MG, Ke D. New DNA-based PCR approaches for rapid real-time detection and prevention of group B streptococcal infections in newborns and pregnant women. Expert Rev Mol Med. 2001;3:1–14.

Bernardo K, Pakulat N, Macht M, et al. Identification and discrimination of *Staphylococcus aureus* strains using matrix-assisted laser desorption/ionization-time of flight mass spectrometry. Proteomics. 2002;2:747–53.

Black CM. Current methods of laboratory diagnosis of *Chlamydia trachomatis* infections. Clin Microbiol Rev. 1997;10:160–84.

Boehme CC, Nabeta P, Hillemann D, et al. Rapid molecular detection of tuberculosis and rifampin resistance. N Engl J Med. 2010;363:1005–15.

Brown DFJ, Edwards DI, Hawkey PM, et al. Guidelines for the laboratory diagnosis and susceptibility testing of methicillin resistant *Staphylococcus aureus* (MRSA). J Antimicrob Chemother. 2005;56:1000–18.

Brown-Elliott BA, Griffith DE, Wallace Jr RJ. Diagnosis of nontuberculous mycobacterial infections. Clin Lab Med. 2002;22:911–25.

Carroll KC, Bartlett JG. Biology of *Clostridium difficile*: implications for epidemiology and diagnosis. Annu Rev Microbiol. 2011;65:501–21.

CDC. Sexually transmitted diseases treatment guidelines. MMWR. 2002;51 (No. RR-15).

CDC. Updated guidelines for the use of nucleic acid amplification tests in the diagnosis of tuberculosis. 2009;58:7–10.

Cetinkaya Y, Falk P, Mayhall CG. Vancomycin-resistant enterococci. Clin Microbiol Rev. 2000;13:686–707.

Chapin K, Musgnug M. Evaluation of three rapid methods for the direct identification of *Staphylococcus aureus* from positive blood cultures. J Clin Microbiol. 2003;41:4324–7.

Chemlal K, Portaels F. Molecular diagnosis of nontuberculous mycobacteria. Curr Opin Infect Dis. 2003;16:77–83.

Clerc O, Greub G. Routine use of point-of-care tests: usefulness and application in clinical microbiology. Clin Microbiol Infect. 2010;16:1054–61.

Cockerill 3rd FR, Smith TF. Response of the clinical microbiology laboratory to emerging (new) and reemerging infectious diseases. J Clin Microbiol. 2004;42:2359–65.

Cunningham MW. Pathogenesis of group A streptococcal infections. Clin Microbiol Rev. 2000;13:470–511.

Ecker DJ, Sampath R, Li H, et al. New technology for rapid molecular diagnosis of bloodstream infections. Expert Rev Mol Diagn. 2010;10:399–415.

Eshoo MW, Crowder CD, Li H, et al. Detection and identification of *Ehrlichia* species in blood by use of PCR and electrospray ionization mass spectrometry. J Clin Microbiol. 2010;48:472–8.

Espy MJ, Uhl JR, Sloan LM, et al. Real-time PCR in clinical microbiology: applications for routine laboratory testing. Clin Microbiol Rev. 2006; 19:165–256.

Fagan EJ, Jenkins C, Walton C, et al. Screening methods for methicillin-resistant *Staphylococcus aureus*. J Med Microbiol. 2010;59(Pt 10):1153–5.

Fosheim GE, Carey RB, Limbago BM. Evaluation of the Advandx VRE Evigene assay for detection of *vanA* in vancomycin-resistant *Staphylococcus aureus*. J Clin Microbiol. 2007;45:1611–3.

French GL. Methods for screening for methicillin-resistant *Staphylococcus aureus* carriage. Clin Microbiol Infect. 2009;15 Suppl 7:10–6.

Gegia M, Mdivani N, Mendes RE, et al. Prevalence of and molecular basis for tuberculosis drug resistance in the Republic of Georgia: validation of a Qiaplex system for detection of drug resistance-related mutations. Antimicrob Agents Chemother. 2008; 52:725–9.

Griffith DE, Aksamit T, Brown-Elliott BA, et al. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. Am J Respir Crit Care Med. 2007;175: 367–416.

Guidelines for Tuberculosis Control in New Zealand; 2010, Ministry of Health, Welington, New Zealand.

Guntupalli R, Sorokulova I, Krumnow A, et al. Real-time optical detection of methicillin-resistant *Staphylococcus aureus* using lytic phage probes. Biosens Bioelectron. 2008;24:151–4.

Hall TA, Sampath R, Blyn LB, et al. Rapid molecular genotyping and clonal complex assignment of *Staphylococcus aureus* isolates by PCR coupled to electrospray ionization-mass spectrometry. J Clin Microbiol. 2009;47:1733–41.

Harkins AL, Munson E. Molecular diagnosis of sexually transmitted *Chlamydia trachomatis* in the United States. ISRN Obstet Gynecol. 2011;2011: 279149.

He Y, Chang TC, Li H, et al. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry and database for identification of *Legionella* species. Can J Microbiol. 2011;57:533–8.

Jarzembowski JA, Young MB. Nontuberculous mycobacterial infections. Arch Pathol Lab Med. 2008; 132:1333–41.

Kaatz GW, Moudgal VV, Seo SM, et al. Phenothiazines and thioxanthenes inhibit multidrug efflux pump activity in *Staphylococcus aureus*. Antimicrob Agents Chemother. 2003;47:719–26.

Kaul KL. Molecular detection of *Mycobacterium tuberculosis*: impact on patient care. Clin Chem. 2001;47:1553–8.

Ke D. Development of conventional and real-time PCR assays for the rapid detection of group B *Streptococci*. Clin Chem. 2000;46:324–31.

Li H, Turhan V, Chokhani L, et al. Identification and differentiation of clinically relevant *Mycobacterium* species directly from acid-fast bacillus-positive culture broth. J Clin Microbiol. 2009; 47:3814–20.

Mancini N, Carletti S, Ghidoli N, et al. The era of molecular and other non-culture-based methods in diagnosis of sepsis. Clin Microbiol Rev. 2010;23:235–51.

Marshall R, Chernesky M, Jang D, et al. Characteristics of the M2000 automated sample preparation and multiplex real-time PCR system for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae*. J Clin Microbiol. 2007;45:747–51.

Massire C, Ivy CA, Lovari R, et al. Simultaneous identification of mycobacterial isolates to the species level and determination of tuberculosis drug resistance by PCR followed by electrospray ionization mass spectrometry. J Clin Microbiol. 2011;49: 908–17.

Mdivani N, Li H, Akhalaia M, et al. Monitoring therapeutic efficacy by real-time detection of *Mycobacterium tuberculosis* mRNA in sputum. Clin Chem. 2009;55: 1694–700.

Miller MB, Tang YW. Basic concepts of microarrays and potential applications in clinical microbiology. Clin Microbiol Rev. 2009;22:611–33.

Moller JK, Pedersen LN, Persson K. Comparison of the Abbott Realtime Ct new formulation assay with two other commercial assays for detection of wild-type and new variant strains of *Chlamydia trachomatis*. J Clin Microbiol. 2010;48:440–3.

Murdoch DR. Diagnosis of *Legionella* infection. Clin Infect Dis. 2003;36:64–9.

Neonakis IK, Gitti Z, Krambovitis E, et al. Molecular diagnostic tools in mycobacteriology. J Microbiol Methods. 2008;75:1–11.

Oliveira K, Procop GW, Wilson D, et al. Rapid identification of *Staphylococcus aureus* directly from blood cultures by fluorescence in situ hybridization with peptide nucleic acid probes. J Clin Microbiol. 2002;40:247–51.

Olshen E, Shrier LA. Diagnostic tests for chlamydial and gonorrheal infections. Semin Pediatr Infect Dis. 2005;16:192–8.

Parsons LM, Somoskovi A, Gutierrez C, et al. Laboratory diagnosis of tuberculosis in resource-poor countries: challenges and opportunities. Clin Microbiol Rev. 2011;24:314–50.

Persing DH, Tenover FC, Tang YW, et al., editors. Molecular microbiology: diagnostic principles and applications. Washington, DC: American Society for Microbiology Press; 2011.

Peterson LR. Molecular laboratory tests for the diagnosis of respiratory tract infection due to *Staphylococcus aureus*. Clin Infect Dis. 2011;52 Suppl 4: S361–6.

Picard FJ, Bergeron MG. Laboratory detection of group B *Streptococcus* for prevention of perinatal disease. Eur J Clin Microbiol Infect Dis. 2004;23: 665–71.

Qian Q, Venkataraman L, Kirby JE, et al. Direct detection of methicillin resistance in Staphylococcus aureus in blood culture broth by use of a penicillin binding protein 2a latex agglutination test. J Clin Microbiol. 2010;48:1420–1.

Rupnik M, Wilcox MH, Gerding DN. *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol. 2009;7:526–36.

Schmidt BL. PCR in laboratory diagnosis of human *Borrelia burgdorferi* infections. Clin Microbiol Rev. 1997;10:185–201.

Shamputa IC, Rigouts L, Portaels F. Molecular genetic methods for diagnosis and antibiotic resistance detection of *Mycobacteria* from clinical specimens. APMIS. 2004;112:728–52.

Soini H, Musser JM. Molecular diagnosis of mycobacteria. Clin Chem. 2001;47:809–14.

Stamper PD, Cai M, Howard T, et al. Clinical validation of the molecular Bd Geneohm StaphSR Assay for direct detection of *Staphylococcus aureus* and methicillin-resistant *Staphylococcus aureus* in positive blood cultures. J Clin Microbiol. 2007; 45:2191–6.

Tang YW, Ellis NM, Hopkins MK, et al. Comparison of phenotypic and genotypic techniques for identification

of unusual aerobic pathogenic gram-negative bacilli. J Clin Microbiol. 1998;36:3674–9.

Tortoli E. Clinical manifestations of nontuberculous *Mycobacteria* infections. Clin Microbiol Infect. 2009;15:906–10.

Underwood A, Green J. Call for a quality standard for sequence-based assays in clinical microbiology: necessity for quality assessment of sequences used in microbial identification and typing. J Clin Microbiol. 2011;49:23–6.

Wang F, Massire C, Li H, et al. Molecular characterization of drug-resistant *Mycobacterium tuberculosis* isolates circulating in China by multilocus PCR and electrospray ionization mass spectrometry. J Clin Microbiol. 2011;49:2719–21.

Whiley DM, Tapsall JW, Sloots TP. Nucleic acid amplification testing for *Neisseria gonorrhoeae*: an ongoing challenge. J Mol Diagn. 2006;8:3–15.

Wilson ML. Recent advances in the laboratory detection of *Mycobacterium tuberculosis* complex and drug resistance. Clin Infect Dis. 2011;52:1350–5.

Xu HB, Jiang RH, Sha W, et al. PCR-single-strand conformational polymorphism method for rapid detection of rifampin-resistant *Mycobacterium tuberculosis*: systematic review and meta-analysis. J Clin Microbiol. 2010;48:3635–40.

Yang S, Rothman RE. PCR-based diagnostics for infectious diseases: uses, limitations, and future applications in acute-care settings. Lancet Infect Dis. 2004;4:337–48.

Zhang Y, Yew WW. Mechanisms of drug resistance in *Mycobacterium tuberculosis*. Int J Tuberc Lung Dis. 2009;13:1320–30.

Zirakzadeh A, Patel R. Vancomycin-resistant enterococci: colonization, infection, detection, and treatment. Mayo Clin Proc. 2006;81:529–36.

Molecular Parasitology and Mycology 27

Vincent J. LaBombardi and Wei Wang

Contents

27.1 Parasitology 781
27.1.1 Malaria 781
27.1.2 Leishmaniasis 783
27.1.3 Toxoplasmosis 784
27.1.4 Trichomoniasis 785

27.2 Mycology 786
27.2.1 Candidiasis 786
27.2.2 Cryptococcus 788
27.2.3 Aspergillosis 789
27.2.4 Zygomycosis (Mucormycosis) 790
27.2.5 Pneumocystis 791

Further Reading 792

V.J. LaBombardi, PhD (✉)
Department of Microbiology, New York Hospital-Queens, Flushing, NY, USA

W. Wang, PhD
Department of Microbiology, Mount Sinai Medical Center, New York, NY, USA

27.1 Parasitology

27.1.1 Malaria

- General characteristics
  - Infections are caused by a protozoan of the genus *Plasmodium*. There are five species that are recognized as causing infections in humans, P. *falciparum, P. vivax, P. malariae, P. ovale, and P. knowlesi*
  - It is estimated that there are up to 500 million cases of malaria yearly with 1–3 million deaths
  - In the United States, cases of malaria are associated with travel to, or immigration from, endemic areas especially sub-Saharan Africa, Southeast Asia, and India
  - The life cycle of malaria includes two hosts. When an infected Anopheles mosquito takes a blood meal, it will inoculate the parasite into the human host. The malaria sporozoites will infect liver cells where they multiply. When released, they will invade erythrocytes where they undergo asexual reproduction. Two species, *P. vivax* and *P. ovale*, can become dormant in liver cells and can initiate infections even years after the initial infection
  - The periodicity of the fever varies by species. *P. knowlesi* exhibits a 24-h periodicity, whereas with *P. falciparum, P. vivax, and P. ovale* it is 48 h and with *P. malariae* it is 72 h

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_27,

- Clinical presentation
  - The incubation period of the disease varies by Plasmodium species. *P. falciparum* has an incubation period that ranges from 8 to12 days; *P. vivax and P. ovale*, 7–10 days; and *P. malariae*, 22–40 days. The incubation period for infections with *P. knowlesi* is uncertain at this time
  - Symptoms include fever, chills, headache, and malaise
  - Gastrointestinal symptoms including nausea, vomiting, and diarrhea (sometimes bloody) are also seen
  - Infection with either *P. falciparum* or *P. knowlesi* can result in high levels of parasitemia. This can result in the clogging of capillaries which causes severe and life-threatening disease. Exchange blood transfusions may be required as an adjunct to drug therapy
  - Coma and generalized convulsions can be seen in severe malaria
- Conventional diagnostic methods
  - Specimens: Peripheral blood or blood obtained by finger stick
  - Conventional microscopy
    - Bright field microscopy remains the gold standard for the diagnosis of malaria. Thick and thin smears are prepared, with the thin smears being stained immediately after drying with a Giemsa stain (Figs. 27.1 and 27.2)
    - Thick smears are allowed to dry overnight, dehemoglobinized, and then stained.
    - Thick smears can increase the sensitivity of the test, but species identification may be difficult
    - *P. knowlesi* can be confused with *P. malariae* in smears except that with *P. knowlesi* the rate of parasitemia can exceed 5% which would not be seen with *P. malariae*
  - Antigen detection
    - There are immunochromatographic assays for the detection of malarial parasites from whole blood. These utilize antigens that will be present in all five species such as lactate dehydrogenase or adolase or an antigen, histidine-rich protein2 (HRP2), that is specific for *P. falciparum*
    - This test cannot distinguish among the four species of Plasmodium other than *P. falciparum*. False positive reactions due to cross-reactions with rheumatoid factor are possible
    - Can be used to monitor therapy based on the disappearance of the reaction
    - Cannot distinguish mixed infections

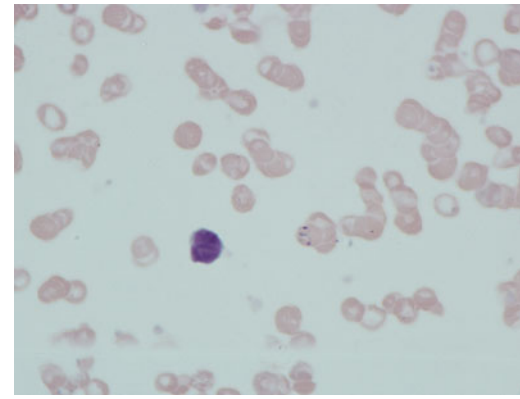


Fig. 27.1 A Giemsa-stained blood film showing a doubly infected RBC with ring forms of *Plasmodium falciparum*

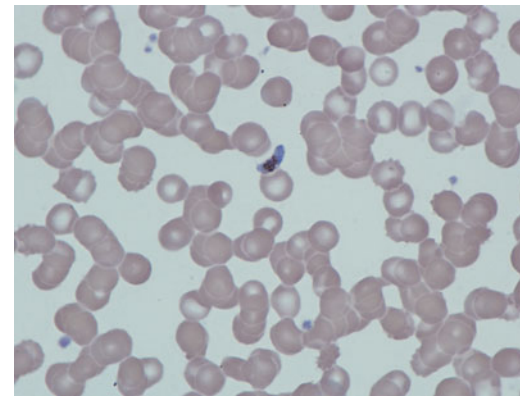


Fig. 27.2 A Giemsa-stained blood film showing the gametocyte form of *Plasmodium falciparum* obtained from a patient with multiple prior episodes of infection with this organism

  - Antibody detection
    - Antibodies directed against the malarial parasites are produced by virtually all individuals infected. There are various methods available to test
    - Positive serology is not always indicative of active clinical disease
    - Cross-reactions among the different species that infect humans can be observed
- Molecular diagnostic methods
  - Qualitative methods such as those amplifying Plasmodium-specific small subunit ribosomal RNA (SSU rRNA) have largely been replaced by real-time (RT) PCR assays
  - Real-time assays
    - *P. falciparum* primers amplify a region of the cytochrome oxidase subunit 1 (COX1) mitochondrial gene that is unique to this organism. Detected using probes containing molecular beacons
    - Plasmodium-specific primers amplify a conserved region of the 18s rRNA gene that would be present in all species of Plasmodium that infects humans
    - Primers that are capable of identifying the organism to the species level have been validated for the species that can cause disease in humans except for *P. knowlesi*.
    - Commercial assays are now available
  - Clinical utility
    - RT PCR assays are more sensitive than microscopy at diagnosing malaria
    - Species-specific primers can correct misidentifications made by microscopy
  - Pitfalls
    - Genus-specific primers cannot distinguish mixed infections
    - The assays lack standardization

27.1.2 Leishmaniasis

- General characteristics
  - Leishmania are protozoan parasites; they have a life cycle that includes an infectious phase with promastigotes which multiply in the salivary glands of the sand fly
  - Promastigotes are phagocytized by macrophages where they are transformed into amastigotes
  - Amastigotes that multiply within histiocytes are released and infect other cells
- Clinical presentation
  - Cutaneous leishmaniasis
    - Papule at the site of insect bite
    - Overtime usually within weeks develops into an ulcer; satellite lesions occur frequently
    - A delayed hypersensitivity reaction can occur and results in chronic mucosal leishmania which can result in major tissue destruction
    - Leishmania recidivans, characterized by red painless satellite lesions, are seen mainly in the Middle East caused by *L. tropica*
  - Visceral leishmaniasis (VL)
    - Found in parts of Europe, Asia, East Africa, and Brazil with approximately 500,000 cases a year with significant mortality
    - Transmission is through the bite of an infected sand fly of the genus Phlebotomus
    - Transmission from person to person or via infected dogs or rodents
    - Three species are seen to cause visceral leishmaniasis; *L. donovani*, *L. infantum*, or *L. chagasi*. The species vary by geographic region
    - The amastigotes in histiocytes travel to the lymph nodes and spread to infect the liver, spleen, and bone marrow
    - Symptoms include fever, malaise, weight loss, diarrhea, nonproductive cough, and hepatosplenomegaly
- Diagnostic methods
  - Giemsa or H&E stain of biopsy specimens. Amastigotes are ovoid 1–5 μm long by 1–2 μm wide with a large nucleus and a rod-shaped kinetoplast (Fig. 27.3)
  - Culture

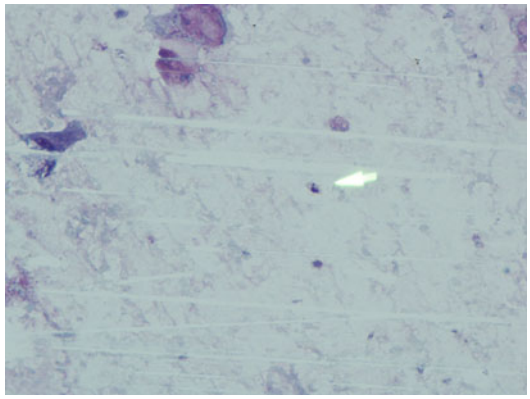


Fig. 27.3 A Giemsa-stained touch preparation from a biopsy specimen showing the amastigote form in a case of visceral leishmaniasis. Note the large nucleus and prominent kinetoplast

- The organisms can be cultured in Novy-MacNeal-Nicolle (NNN) media

– Serology
  - Methods available include the following: direct agglutination test (DAT) or fast agglutination screening test (FAST), uses freeze-dried promastigote antigen
  - Indirect fluorescent method (IFA) prepared from an antigen suspension derived from the species of interest
  - ELISA test using the H39 and other species-specific antigens
  - Urinary antigen latex test (LAtex) detects a nonprotein parasite-specific antigen in voided urine of patients suspected of having VL sensitivity and specificity of approximately 95% in patients with acute disease
– PCR: utilize blood, bone marrow, spleen, urine, and skin biopsy specimens
  - Primers specific for regions of the small subunit RNA gene (18s rRNA gene) sensitivity of 86–97% in VL
  - Primers derived from 18s rRNA gene sequence of the species responsible for VL used to test whole blood
  - Testing of filter paper impressions using primers specific for kinetoplast DNA (kDNA) followed by subspecies-specific primer using the cysteine protease B (CPB)
  - Heat shock protein 70 (Hsp70) has been used in VL
  - Urine testing for VL using primers designed from Leishmania minicircle DNA
  - Oligo C-Test and NASBA oligochromatography that target 18s ribosomal DNA and RNA, respectively, are commercially available as per United States Food and Drug Administration (FDA)
– Pitfalls
  - Sensitivity of PCR and other molecular-based assays is in the 70–95% range. Negative results for a blood sample do not rule out disease. Splenic biopsies have a higher yield in VL but are not without risk
– Clinical utility
  - Molecular methods offer higher sensitivity and specificity than traditional methods but are expensive and not widely used in resource-poor settings
  - Molecular methods do allow for the assessment of therapy in VL

27.1.3 Toxoplasmosis

- General characteristics
  – Disease caused by the coccidian protozoan *Toxoplasma gondii*
  – Infection is seen in birds and mammals
  – Cats, especially kittens, are a source for infection
  – Found in the soil. Raw or undercooked meat is also a source of infection
  – The organism exhibits two stages: tachyzoites associated with acute infection and bradyzoites seen in chronic infection
- Clinical presentation
  – Infections with *T. gondii* are largely asymptomatic
  – Maternal transmission to the fetus can occur even in asymptomatic infections resulting in congenital infection with manifestations

including encephalitis, jaundice, and growth retardation
  - Acute febrile disease with pneumonia, myocarditis, and hepatitis
  - Chorioretinitis
  - Acute or chronic encephalitis in an immunocompromised patient
- Diagnostic methods
  - Specimen
    - Stained smears of tissue biopsies
    - In vitro culture: blood, CSF, tissue, and amniotic fluid
    - PCR: blood, tissue, amniotic fluid, urine, or other body fluids
    - Serology: serum
  - Conventional tests
    - Microscopy
      - Tachyzoites can be visualized in Giemsa-, H&E-, and Papanicolaou-stained tissues
      - Cysts containing bradyzoites can be seen in tissues stained with H&E
    - Culture
      - Inoculation into Swiss mice
      - Follow with serology
      - Brain biopsy of those mice that seroconvert
    - Serology
      - IgG for immune status should be obtained in pregnant women
      - IgG can be useful to determine if certain patients, either HIV+ or those undergoing transplantation, are at risk for reactivation
      - IgM utilized to diagnose congenital and acute infections
  - Sensitivity and specificity
    - Smears
      - Can be confused with Leishmania
      - Requires an experienced microscopist
  - Serology
    - Interpretation of serological results can be difficult
    - One positive IgM result does not always correlate with active disease
    - IgG antibody levels are important to assess the risk of reactivation in transplant patients
  - Molecular methods
    - Real-time PCR
      - Primers target the B1 gene with 35 copies per organism
      - Primers also target the highly repetitive sequence REP-529 with 200–300 copies for cells
    - Loop-mediated isothermal amplification
      - Primer targets REP-529 sequence and the SAG1 and SAG2 genes
      - Performed at 65 °C
  - Clinical utility
    - Sensitivity and specificity varies, depending on primers used and the choice of specimen
    - PCR testing of blood specimens is a good predictor of invasive disease in high-risk individuals
    - Sensitivity on testing placenta or amniotic fluid to predict congenital disease ranges from 71% to 92%
    - Serology: useful in case of congenital toxoplasmosis
    - PCR: useful in diagnosing ocular, cerebral, and congenital toxoplasmosis

27.1.4 Trichomoniasis

- General characteristics
  - Caused by the protozoan *Trichomonas vaginalis*
  - Is a sexually transmitted infection (STI)
  - The relative prevalence of infections caused by this organism is unknown, but it is estimated to surpass that of chlamydia and gonorrhea
- Clinical presentation
  - Infections often asymptomatic
  - Can result in vaginitis, cervicitis, and urethritis
  - Can cause prostatitis in men
  - Complications include preterm labor and pelvic inflammatory disease (PID)
  - Can facilitate HIV transmission in both males and females

- Conventional diagnostic methods
  - Wet mount of vaginal secretions
  - Culture of vaginal secretions and urine in the IM Pouch™
  - "Whiff" test of vaginal secretions
  - Rapid antigen test, OSOM Trichomonas rapid test (Genzyme Diagnostics), is a point of care test that can be used on vaginal specimens
- Molecular methods
  - Affirm assay (BD Diagnostics) utilized probes directed against three causes of vaginitis/vaginosis, *Candida albicans, Gardnerella vaginalis*, and *T. vaginalis*
  - Aptima ATV assay uses transcription-mediated amplification (TMA) directed against 16s RNA specific for *T. vaginalis* and utilizes female vaginal or endocervical swabs and urine or male urine or urethral swabs
  - Clinical utility
    - Wet mounts are approximately 50% sensitive
    - Cultures are specific but can take up to 5 days to be positive
    - The Affirm assay had a reported sensitivity of only 63% in one study
    - The ATV assay uses multiple specimen types and has reported sensitivities ranging from 92% to100%

27.2 Mycology

27.2.1 Candidiasis

- General characteristics
  - Etiologic agent *Candida* sp.
  - There are over 200 species in the genus Candida, but only a small fraction of this number is implicated in causing human disease
  - Found in the environment and on mucocutaneous membrane
  - These yeasts are ovoid often with buds
  - *C. albicans* is the species most often isolated from clinical specimens
  - Candida sp. including *C. albicans*, *C. glabrata*, *C. parapsilosis*, and *C. tropicalis* are responsible for causing the majority of candidemias and are usually the fourth and fifth most prevalent isolates from blood cultures obtained from patients in the ICU
  - Isolation from clinical specimen does not always correlate with clinical disease
  - Can form biofilms or indwelling medical devices
- Clinical presentation
  - Mucocutaneous disease
    - Thrush in newborns owing to the low pH in the mouths. Characterized by a pseudomembrane which covers the tongue
    - Oropharyngeal candidiasis in the adult white patchy plaques evident
    - Esophagitis
    - Vaginitis
    - Balanitis
    - Gastritis
  - Cutaneous
    - Paronychia and onychomycosis
    - Diaper rash
    - Candida granuloma
  - Systemic disease
    - Pneumonia – rare, although Candida species are often isolated from respiratory specimens obtained from hospitalized patients
    - Urinary tract infection, often a colonizer but can cause infection
    - Fungemia and endocarditis
    - Meningitis – very rare
- Therapy
  - Fluconazole: *C. albicans* usually susceptible. *C. dubliniensis* usually susceptible but can develop resistance during therapy. *C. krusei* are intrinsically resistant. *C. glabrata* usually susceptible or susceptible dose-dependent
  - Amphotericin B: *C. lusitaniae* – resistant
  - Echinocandins: *C. parapsilosis* may have MICs in the intermediate range
- Diagnostic methods
  - Specimens

- Blood – collected into blood culture bottles
- Swab specimens obtained from the throat
- Tissue samples
- Urine
- Bronchial lavage
- Body fluids

– Conventional testing
  - Microscopy: smears made from clinical specimen may be Gram-stained to visualize the typical oval budding yeast
  - Wet mounts may be examined with KOH or by calcofluor white staining
  - Tissue sections in addition may be stained with H&E or GMS
– Culture
  - *Candida* sp. will grow on routine bacteriological culture media such as blood and chocolate agar plates incubated at 35–37 °C
  - *Candida* sp. will grow on commonly used fungal isolation media such as Sabouraud dextrose agar
  - Primary plating of specimen can also be performed on chromogenic media such as Candida ID agar and CHROMagar which can aid in the differentiation of some *Candida* sp.
– Identification to species
  - Germ tube tests positive for *C. albicans* and *C. stellatoidea* which is considered, possibly incorrectly, as a subspecies of *C. albicans*
  - Identification of non-*albican* species can be performed in automated systems (Vitek 2) or in test strips that couple biochemical tests with carbohydrate assimilations
– Serologic assays
  - Antibody studies have little utility as *Candida* sp. often can be found colonizing various areas of the body
  - (1–3) β-D-glucan assays can be used on blood specimens as a marker of invasive fungal disease. This assay is not specific for any one genus as this compound is found in the cell wall of many fungi, excluding *Cryptococcus* and the zygomycetes
– Molecular methods
  - Many amplification methods target the rDNA internal transcribed spacer region (ITS) because the variable nature of this region allows it to be used for species identification and there are multiple copies in each cell
  - Seminested PCR has been used to identify *Candida* sp. from blood culture bottles after an overnight incubation
  - High-resolution derivative melt analysis targeting the ITS regions of rDNA has been shown to be capable of distinguishing clinically relevant species of Candida
  - A LightCycler-based assay utilizing species-specific primers with melting temperature analysis can also be used to identify clinically important species of Candida
  - xMAP technology for Luminex and a PCR electrospray ionization mass spectrometry by Ibis Technology are commercially available as an aid to the identification of Candida species
  - PNA FISH which is a fluorescence in situ hybridization assay is commercially available. This test can be used directly from positive blood cultures and identifies the isolates as either *C. albicans*/*C. parapsilosis*, *C. tropicalis*, or *C. glabrata*/*C. krusei*. Cross-reactions with other lesser known *Candida* sp. can occur
– Clinical utility
  - Candida species are often found in clinical specimens, and care must be taken to differentiate a colonizer from a true pathogen
  - Candida grows on a variety of common laboratory media and can be readily identified using conventional methodologies
  - Amplification methods hold promise especially for detection of *Candida* sp. directly from clinical specimens
  - Extracting nucleic acids from yeasts and fungi is more difficult than from bacteria
  - Commercial PCR systems to detect and identify *Candida* sp. are beginning to become available

27.2.2 Cryptococcus

- General characteristics
  - Etiologic agents include *Cryptococcus neoformans* and *C. gattii*
  - Round to ovoid yeast with a diameter ranging from 3.5 to 8 μm
  - Yeast reproduce by budding with the daughter cell connected via a thin stalk
  - Yeasts are encapsulated by a capsule containing galactoxylomannan
  - In vivo the capsule size can vary greatly depending upon the immune status of the host
  - This organism can be found in pigeon droppings and is transmitted to humans viainhalation
- Clinical presentation
  - *C. neoformans* typically infects immunosuppressed patients especially AIDS patients and elderly debilitated patients
  - *C. gattii* infects both immune competent and immunocompromised individuals. Originally described as causing infection in tropical and subtropical areas, but more recently, cases have been described in the Pacific Northwest between the United States and Canada. The source of this species is the Eucalyptus and Douglas fir trees
  - Infection is acquired through inhalation of the organism where they can colonize the nasal cavity before spreading to other sites
  - Patients can experience a flu-like illness characterized by a productive cough, fever, and malaise. *C. neoformans* pulmonary disease will typically form pulmonary infiltrates. *C. gattii* pulmonary disease usually results in the formation of lung nodules
  - Disseminated disease occurs in immunocompromised patients and can involve many body sites
  - *Cryptococcal meningitis* is characterized by headache, fever, and stiffness of the neck
  - Infections with *C. gattii* are more likely to result in the formation of brain lesions than will infections with *C. neoformans*
- Therapy
  - Cryptococci are generally susceptible to fluconazole. They are resistant to echinocandins. *Cryptococcal meningitis* is usually treated with combination therapy with amphotericin B and 5-flucytosine and less commonly with fluconazole and 5-flucytosine
- Diagnostic methods
  - Specimens
    - CSF, tissue specimens, blood inoculated into blood culture bottles, sputum, and urine
  - Microscopy
    - Gram-stained CSF or tissue smears can be used to visualize the yeasts
    - India ink wet mount with CSF specimen will demonstrate capsules enclosing yeast cells. Typical budding forms may be present
    - Visualizing encapsulated budding yeasts in CSF specimens is diagnostic for *C. neoformans*
    - Mucicarmine-stained tissue sections are used to detect the polysaccharide capsule in tissue sections
  - Culture and identification
    - Cryptococci will grow on routine bacteriological media including blood and chocolate agar plates as well as on routine fungal media such as Sabouraud dextrose agar or brain-heart infusion agar with blood
    - Growth should be evident in 48–72 h in cultures incubated at 30–37 °C
    - Rapid presumptive identification can be accomplished by inoculating suspected *C. neoformans* colonies onto a caffeic acid disk. *C. neoformans* will turn the disk brown after a brief incubation period
    - *C. neoformans* will also form a brown pigment when grown on birdseed agar
    - The organisms can be identified using the yeast ID cards on the Vitek 2 and in test strips such as the API 20C
    - Matrix-assisted laser desorption/ionization time-of-flight mass

spectrometry (MALDI-TOF) can also be utilized to identify cryptococci
- Molecular testing
  - Amplification of a region of the CAP10 capsular gene mRNA was developed using LightCycler technology. This can be utilized for isolate identification or identification directly from a clinical specimen
  - PCR targeting the sequence variable region of the internal transcribed spacer 1 (ITS 1) of ribosomal DNA has been used as part of a multiplex PCR test to detect *C. neoformans* in blood cultures
  - Taqman PCR amplifying a 59-bp region in the ITS 2 rDNA gene has been used to detect isolates belonging to the *C. neoformans* group from blood, bone marrow, CSF, and BAL specimens
  - Amplified products of the ITS 2 region were hybridized to microarrays which contained species-specific probes bound to the surface. The signals were detected using gold nanoparticles and silver deposition
- Clinical utility
  - Direct staining of specimens can reliably presumptively identify *C. neoformans* group from clinical specimens
  - *C. neoformans* group can be readily identified using routine methods
  - *C. neoformans* cannot be easily distinguished from *C. gattii* using routine conventional methods. *C. gattii* can be identified by its ability to grow on canavanine glycine bromothymol blue (CGB) agar turning the medium from green to blue
  - PCR techniques have been developed to identify *Cryptococcus* directly from clinical specimens
  - PCR is expensive and not readily available in all clinical laboratories
  - To identify yeasts from clinical specimens, multiplex PCR would be more clinically useful

27.2.3 Aspergillosis

- General characteristics
  - The genus *Aspergillus* contains over 175 different species. These are found widely in the environment in soil, water, and agricultural products. The most predominant species isolated from clinical specimens includes *A. fumigatus*, *A niger*, *A. flavus*, *A. nidulans*, and several others
- Clinical presentation
  - *Aspergillus* sp. causes a wide range of diseases
  - In immunocompetent individuals, illness arises from the ingestion of mycotoxins, through allergic or hypersensitivity reactions, sinusitis, other noninvasive infections as well as invasive *aspergillosis*
  - Immunocompromised individuals are at major risk for invasive disease, particularly those patients with neutropenia following transplantation, both solid organ and bone marrow
- Diagnostic methods
  - Specimens
    - Blood, tissue biopsies, bronchial alveolar lavage (BAL), and sterile body fluids
  - Microscopy
    - Gram-stained smears can be examined for the presence of hyphal elements
    - Smears can be examined with calcofluor white to demonstrate septate hyphal elements with dichotomous branching at a 45° angle
    - H&E-stained tissue biopsy sections for invasive fungal elements showing typical morphology
  - Culture and identification
    - Specimens can be cultured using standard fungal media including Sabouraud dextrose agar (SDA), Mycosel, brain-heart infusion, and potato dextrose agar. *Aspergillus* will also grow well on routine bacteriological media such as blood agar or chocolate agar plates
    - Incubate the plates at 30 °C in ambient air
    - Growth can be evident in 48–72 h

- Colonies can start out white then become yellow, green, brown, or black. The reverse of the colony can be white, tan, grey, yellow, orange, or red
- Identification to species will require microscopic evaluation of lactophenol cotton blue-stained tease preparations or slide cultures (Fig. 27.4)

– Antigen tests
- Galactomannan assay detects *Aspergillus*-specific antigens in blood. Reported sensitivity varies greatly (30% to >95%) with specificities ranging from <40% to >95%. The cutoff values for a positive in blood are 0.5 ug/ml in the United States and 1.5 ug/ml in Europe
- The galactomannan assay may be useful in detecting invasive aspergillosis using other body fluids such as BAL fluid, but a reliable cutoff value has yet to be determined
- Latex tests and ELISA assays are commercially available, but the sensitivity is poor and not clinically useful

• Molecular assay
– PCR
- Extraction methods must be sufficient to release nucleic acid from the cells.

– The following have been used on blood and other body fluids as an aid to diagnosing invasive *aspergillosis*:
- Nested PCR which specifically amplifies the 18s rRNA gene
- Multiplex PCR that targets β-tubulin (βtub) and rodlet A (rodA) partial gene sequences
- Real-time PCR that targets a conserved region of the 18s rRNA gene. This assay amplifies genes from *Aspergillus* species that cause human disease, but cross-reactions with *Penicillium* species have been observed
- 18s rRNA gene sequencing is also a tool that can be used for species identification

– Therapy
- The choice of therapeutic agent varies as to the type of aspergillosis being treated. Caspofungin, voriconazole, itraconazole, posaconazole, and amphotericin B all have activity against Aspergillus and can be found in the IDSA guidelines

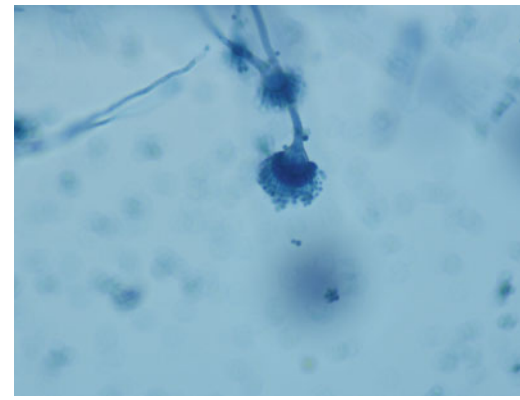


Fig. 27.4 A lactophenol blue-stained tease preparation from a culture of *Aspergillus fumigatus*

– Clinical utility
- In cases of invasive aspergillosis, culture is only approximately 50% sensitive
- Stained smears and tissue sections showing hyphal elements may not be *Aspergillus*
- Low positive results with the galactomannan assay must be interpreted with caution
- PCR testing of clinical specimens affords better sensitivity than culture but must be correlated with other clinical findings
- Gene sequencing to speciate *Aspergillus* requires the use of a vetted database

27.2.4 Zygomycosis (Mucormycosis)

• General characteristics
– The fungi responsible for causing disease in humans belong to the order *Mucorales*. They are further divided into six families, but the family most often involved in disease is the *Mucoraceae*. The predominant genera

involved are *Rhizopus, Mucor, Rhizomucor*, and *Absidia*. These fungi consist of filamentous hyphae that are rarely separate and which may or may not contain rhizoids. These fungi are ubiquitous in the environment and have a worldwide distribution

- Clinical presentation
  - Rhinocerebral mucormycosis
    - Seen in immunocompromised patients, in uncontrolled diabetics, in patients with leukemia, and in transplant recipients
    - Infection begins in the nasal cavity or the paranasal sinus
    - The infection spreads to the orbit and travels through the cribriform plates to infect the brain. This will lead to frontal lobe involvement with necrosis and abscess formation
  - Pulmonary mucormycosis
    - Seen most often in patients with diabetes, leukemia, or lymphoma
    - A primary infection which is acquired through the inhalation of spores
    - Signs include bronchitis and pneumonia with extensive cavitation
  - Cutaneous mucormycosis
    - Seen in burn patients, also in diabetics, and trauma victims
    - Can be primary or the result of dissemination from other sites
    - Characterized by painful lesions that develop into ulcers
  - Disseminated mucormycosis
    - Organisms spread via the blood stream
    - Found in the brain, heart, and spleen with other organ involvement possible
- Therapy
  - Infections caused by these organisms have a high mortality rate, owing to underlying disease
  - Posaconazole and amphotericin B are the most active agents against the zygomycetes
  - Case reports describing antifungal therapy with these organisms are anecdotal
- Diagnostic methods
  - Specimens
    - Tissue biopsy material, sinus scrapings, and nasal discharge secretions
  - Microscopy
    - H&E-stained tissue sections
    - Specimens stain with calcofluor white to demonstrate ribbon-like hyphal elements that are normally nonseptate
  - Culture and identification
    - Grows on routine fungal culture media incubated at 30 °C in ambient air
    - Potato dextrose agar used to induce sporulation
    - These species are characterized by rapid growth usually within 4 days
    - Identification accomplished using macroscopic and microscopic morphology
  - Molecular methods
    - Nested PCR targets 18s ribosomal DNA
    - DNA extracted from paraffin blocks can be utilized in this testing
    - Heterogeneous nature of the target gene segments allows for species identification
    - Real-time PCR with high resolution melt analysis used on fungal colonies and on paraffin tissue samples to identify the species
  - Clinical utility
    - The zygomycetes may be difficult to isolate from clinical specimens
    - Identification of fungal isolates is time consuming and requires experienced individuals
    - PCR is more sensitive than are fungal cultures for diagnosing disease
    - PCR cannot distinguish between colonizing and infecting fungi

27.2.5 Pneumocystis

- General characteristics
  - Originally described by Chagas in 1909
  - Thought to be a protozoan
  - DNA analysis characterized this as a fungus
  - Lacks ergosterol in the cell wall

  - Does not respond to antifungal agents
  - Original name was *Pneumocystis carinii*
  - Human pathogens were renamed *Pneumocystis jirovecii* in 1999
- Clinical presentation
  - Children usually show seroconversion by 6 years of age
  - Can cause a primary infection in immunocompetent infants
  - Prior to the institution of highly active antiretroviral therapy (HAART), it was the defining illness in approximately 40% of US AIDS patients
  - Clinical presentation as well as the organism load will vary depending on the degree of immunosuppression in the patient
  - Organism resides in the alveoli
  - Typical clinical presentation includes dyspnea, fever, and nonproductive cough.
  - Signs include tachypnea, tachycardia, and cyanosis
  - Chest X-ray typically shows diffuse infiltrates involving the perihilar regions
- Therapy
  - Two double-strength trimethoprim/sulfamethoxazole tablets three times a day
  - Alternative agents include pentamidine or clindamycin plus primaquine
- Conventional diagnostic methods
  - Giemsa, calcofluor white, or Gomori methenamine silver (GMS) stain of induced sputa, bronchoalveolar lavage (BAL) fluid, or lung biopsy touch preps. Giemsa-stained smears will not stain cyst walls however
  - Direct fluorescence antibody test to detect Pneumocystis cysts and trophozoites contained in respiratory tract specimens. Induced sputa and BAL specimens are preferred over sputum specimens
  - (1–3) β-D glucan assay can be utilized on blood specimens, but it is not specific for Pneumocystis as this compound is found in most fungi. It can be used to compliment other diagnostic methods
- Molecular diagnostic methods
  - Single-round PCR based on the major surface glycoprotein gene has been employed
  - Nested PCR based on the mitochondrial large subunit rRNA and on the internal transcribed spacer (ITS) region has also been developed. Amplified products are visualized by electrophoresis in agarose gels
  - A commercial real-time PCR assay targeting the mitochondrial ribosomal large subunit using molecular beacons to detect the amplified product is available (MycAssay Pneumocystis kit, Myconostica Ltd, Manchester, UK)
- Clinical utility
  - The home-brew PCR assays have not been standardized
  - The MycAssay Pneumocystis kit is available outside of the US only
  - The PCR assays may be positive in patients who are colonized but not infected with *Pneumocystis jirovecii*

Further Reading

Alfarazdeg AS, Ahmed NG, Osman OS, et al. Diagnostic accuracy of the *Leishmania* oligoc-TesT and NASBA-oligochromatography for diagnosis of leishmaniasis in Sudan. PLoS Negl Trop Dis. 2010;4:e776. doi:10.1371/journal.pntd.0000776.

Amjad M, Kfoury N, Cha R, Mobarak R. Quantification and assessment of viability of *Cryptococcus neoformans* by LightCycler amplification of capsule gene mRNA. J Med Microbiol. 2004;53:1201–6.

Andrea SB, Chapin KC. Comparison of Aptima *Trichomonas vaginalis* transcription-mediated amplification assay and BD affirm VPIII for detection of *T. vaginalis* in symptomatic women: performance parameters and epidemiological implications. J Clin Microbiol. 2011;49:866–9.

Bialek R, Konrad F, Kern J, et al. PCR based identification and discrimination of agents of mucomycosis and aspergillosis in paraffin wax embedded tissue. J Clin Path. 2005;58:1180–4.

Boggild AK, Valencia MV, Espinosa D, et al. Detection and species identification of *Leishmania* DNA from filter paper lesion impressions for patients with American cutaneous leishmaniasis. Clin Infect Dis. 2010;50:e1–6.

Cox-Singh J, Davis TME, Lee KS, et al. *Plasmodium knowlesi* Malaria in humans is widely distributed and potentially life threatening. Clin Infect Dis. 2008;46:165–71.

Damiani C, Le Gal S, Lejeune D, et al. Serum (1–3)-β-D-glucan levels in primary infection and pulmonary

colonization with *Pneumocystis jirovecii*. J Clin Microbiol. 2011;49:2000–2.

Farrugia C, Cabaret O, Botterel F, et al. Cytochrome *b* gene quantitative PCR for diagnosing *Plasmodium falciparum* infection in travelers. J Clin Microbiol. 2011;49:2191–5.

Fricker-Hidalgo H, Bulabois CE, Brenier-Pinchart MP, et al. Diagnosis of toxoplasmosis after allogeneic stem cell transplantation: results of DNA detection and serological techniques. Clin Inf Dis. 2009;48:e9–15.

Grupta R, Mirdha BJ, Guleria R, et al. Diagnostic significance of nested polymerase chain reaction for sensitive detection of *Pneumocystis jirovecii* in respiratory clinical specimens. Diag Microbiol Infect Dis. 2009;64:381–8.

Hauser PM, Bille J, Lass-Florl C, et al. Multicenter, prospective clinical evaluation of respiratory samples from subjects at risk for *Pneumocystis jirovecii* infection by use of a commercial real-time PCR assay. J Clin Microbiol. 2011;49:1872–8.

Hrncirova K, Lengerova M, Kocmanova I, et al. Rapid detection and identification of mucormycetes from culture and tissue samples by use of high-resolution melt analysis. J Clin Microbiol. 2010;48:3392–4.

Jarboui MA, Sellami A, Sellami H, et al. Molecular diagnosis of *Pneumocystis jirovecii* pneumonia in immunocompromised patients. Mycoses. 2009;53:329–33.

Lau A, Sorrell TC, Chen S, et al. Multiplex tandem PCR: a novel platform for rapid detection and identification of fungal pathogens from blood culture specimens. J Clin Microbiol. 2008;46:3021–7.

Lemrani M, Hamdi S, Laamrani A, Hassar M. PCR detection of Leishmania in skin biopsies. J Infect Developing Countries. 2009;3:115–22.

Mandviwala T, Shinde R, Kalra A, et al. High-Throughput identification and quantification of *Candida* species using high resolution derivative melt analysis of panfungal amplicons. J Mol Diagn. 2010;12:91–101.

Meyers WM, Firpo A, Wear DJ, editors. Topics on the pathology of protozoan and invasive arthropod diseases. Washington, DC: Armed Forces Institute of Pathology; 2010.

Nye MB, Schwebke JR, Body BA. Comparison of APTIMA *Trichomonas vaginalis* transcription-mediated amplification to wet mount microscopy, culture, and polymerase chain reaction for diagnosis of trichomoniasis in men and women. Am J Obstet Gynecol. 2009;200:188e1–e7.

Robert-Gangneux F, Dupretz P, Yvenou C, et al. Clinical relevance of placenta examination for the diagnosis of congenital toxoplasmosis. Pediatr Infect Dis J. 2010;29:33–50.

Srivastava P, Mehrotra S, Tiwary P, et al. Diagnosis of Indian visceral leishmaniasis by nucleic acid detection using PCR. PLoS One. 2011;6:e19304. doi:10.1371/journal.pone 0019304.

Veron V, Simon S, Blanchet D, Aznar C. Real-time polymerase chain reaction detection of *Cryptococcus neoformans* and *Cryptococcus gattii* in human samples. Diagn Microbiol Infect Dis. 2009;65:69–72.

Wallon M, Franck J, Thulliez P, et al. Accuracy of real-time polymerase chain reaction for *Toxoplasma gondii* in amniotic fluid. Obstet Gynecol. 2010;115:727–33.

Molecular Microbiology in Transplantation

28

Adnan A. Alatoom and Robin Patel

Contents

28.1 Common Infections in Transplant Recipients 795

28.2 Cytomegalovirus 797

28.3 Epstein–Barr Virus (Posttransplantation Lymphoproliferative Disease) 801

28.4 BK Virus 802

28.5 JC Virus 804

28.6 Hepatitis C Virus 804

28.7 Herpes Simplex Virus 807

28.8 Varicella Zoster Virus 807

28.9 Human Herpes Virus 6 808

28.10 Parvovirus (Erythrovirus) B19 808

28.11 Adenovirus 809

28.12 Hepatitis B Virus 809

28.13 *Mycobacterium tuberculosis* 809

28.14 *Toxoplasma gondii* 809

Further Reading 810

A.A. Alatoom, MD, PhD, ABMM, FCAP (✉)
Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX, USA

R. Patel, MD(CM), FRCP(C), D(ABMM), FIDSA, FACP, F(AAM)
Divisions of Clinical Microbiology and Infectious Diseases, Mayo Clinic, Rochester, MN, USA

28.1 Common Infections in Transplant Recipients

- Solid organ and stem cell transplantations are standard therapeutic options for selected diseases
- Allograft rejection and opportunistic infection are major causes of morbidity and mortality in transplant recipients
- Infection in organ transplant recipients occurs in three phases posttransplantation, related to surgical factors, level of immunosuppression, and environmental exposures (Fig. 28.1)
  - First month posttransplantation
    - Infection relating to surgical complications (e.g., wound infection, pneumonia, urinary tract infection, line sepsis, infection of drainage catheters)
    - Reactivated herpes simplex virus (HSV) infection (individuals seropositive for HSV pretransplantation not receiving antivirals active against HSV)
  - 1–6 months posttransplantation
    - Opportunistic pathogens
      - Cytomegalovirus (CMV), *Pneumocystis jiroveci*, *Aspergillus* species, *Nocardia* species, and Epstein–Barr virus (EBV)-related posttransplantation lymphoproliferative disease (PTLD)
    - Reactivation of infection present in recipient before transplantation

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_28,

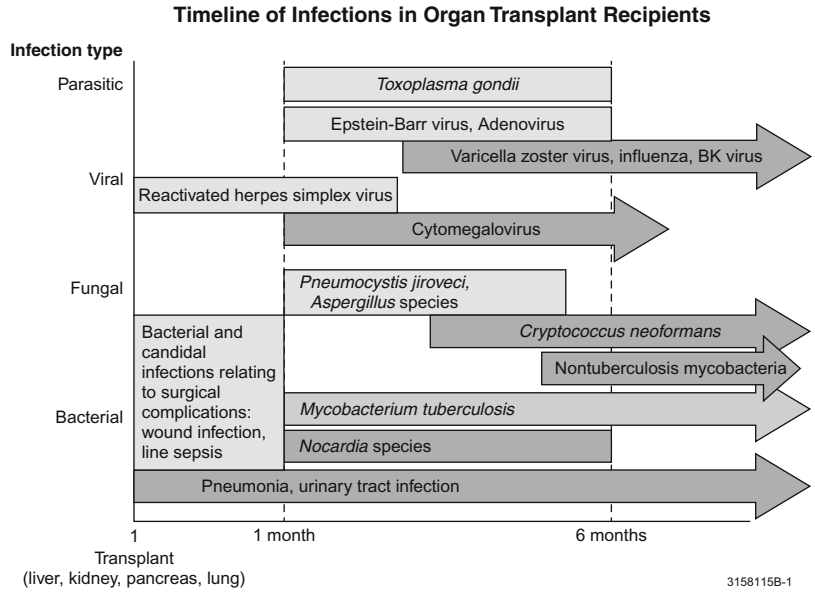


Fig. 28.1 Timeline of infections in organ transplant recipients

(e.g., *Mycobacterium tuberculosis*, viral hepatitis, *Histoplasma capsulatum*, *Coccidioides immitis*)
    - Presentation of donor-transmitted infection [e.g., human immunodeficiency virus, hepatitis B virus (HBV), hepatitis C virus (HCV), fungal and mycobacterial infection]
  - >6 months posttransplantation
    - Infection types typical of general population (e.g., influenza, urinary tract infection, pneumococcal pneumonia)
    - Reactivated varicella zoster virus (VZV) infection (herpes zoster)
- In stem cell transplant (SCT) recipients, infection also occurs in three phases related to the level and type of immunosuppression (i.e., neutropenia, cell-mediated and humoral immune deficiency; Fig. 28.2)
  - Preengraftment (typically first 30 days after SCT, infection associated with neutropenia, and breaks in mucosal barriers – autologous and allogeneic SCT recipients)
    - Oral, gastrointestinal, and skin flora (including *Candida* species)
    - *Aspergillus* species
    - Reactivated HSV infection (individuals seropositive for HSV pretransplantation not receiving antivirals active against HSV)
  - Postengraftment (30–100 days after SCT – allogeneic SCT recipients)
    - Opportunistic pathogens
      - CMV pneumonia, hepatitis, and colitis
      - *P. jiroveci* and *Aspergillus* species
      - EBV-related PTLD
  - Late (>100 days after SCT – allogeneic SCT recipients, especially with graft versus host disease)
    - CMV
    - Infection types typical of general population (e.g., influenza, community-acquired respiratory viruses)
    - Infection with encapsulated bacteria (e.g., *Streptococcus pneumoniae, Haemophilus influenzae*)
    - Reactivated VZV infection (herpes zoster)
- Molecular methods used in diagnosis and management of infections in transplant populations.
  - Application of molecular methods in common posttransplant infections is presented below.

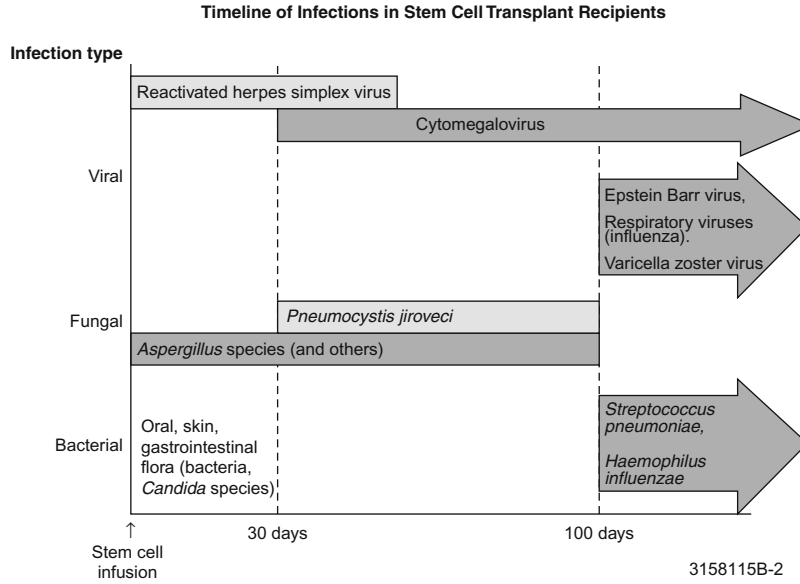


Fig. 28.2 Timeline of infections in stem cell transplant recipients

28.2 Cytomegalovirus

- Enveloped, double-stranded DNA virus, *Herpesviridae* family
- ~50–80% US adults infected
- Asymptomatically, intermittently shed (e.g., urine, saliva, semen, tears)
- Transmission
  - Primary infection, secondary infection, superinfection
    - Primary infection develops when CMV-seronegative transplant recipient becomes infected
      - Most often transmitted from donor
    - Secondary infection or reactivation infection develops when endogenous latent virus is reactivated in CMV-seropositive transplant recipient
    - Superinfection or reinfection occurs when seropositive recipient receives latently infected cells from seropositive donor and virus that reactivates posttransplantation is of donor origin
  - Occurrence, severity of disease in transplant recipients related to
    - CMV serostatus of recipient and donor
      - Organ transplant recipients – CMV-seronegative recipients of organs from CMV-seropositive donors highest risk
      - SCT recipients – allogeneic CMV-seropositive recipients of stem cells from CMV-seronegative donors highest risk
    - Type of organ transplanted
    - Immunosuppression
      - Degree, type (e.g., antilymphocyte antibodies, mycophenolate mofetil)
  - Other means of transmission – close contact with individual excreting CMV in saliva or other body fluids (especially young children) or blood transfusion
- Clinical presentation
  - Wide range of clinical manifestations – asymptomatic infection to severe, lethal, CMV disease
  - Mostly mild to moderate severity and rarely fatal
  - Fever and malaise alone is common; leukopenia with or without thrombocytopenia may be present. Myalgias, arthralgias, and occasionally frank arthritis may occur

- Solid organ transplant recipients
  - Organ involvement correlates with organ transplanted
    - Hepatitis most frequent in liver transplant recipients
    - Pancreatitis most frequent in pancreas transplant recipients
    - Pneumonia most frequent in lung transplant recipients
- Gastroenteritis
  - Affects any segment of gastrointestinal tract, including esophagus, stomach, and small and large intestines
  - Symptoms – dysphagia, odynophagia, nausea, vomiting, abdominal discomfort, gastrointestinal hemorrhage, and/or diarrhea
  - Potential intestinal perforation
  - Endoscopic findings
    - Erythema and diffuse, shallow erosions or localized ulcerations (nonspecific)
    - Biopsy – definitive diagnosis
  - Particularly common among CMV-seronegative recipients of allografts from CMV-positive organ donors following completion of prophylaxis – delayed primary CMV disease
- CMV pneumonia
  - Fever, dyspnea, cough, and/or hypoxemia
  - Radiography – bilateral interstitial, unilateral lobar, nodular infiltrates
- CMV retinitis
  - Asymptomatic or blurred vision, scotomata, and/or decreased visual acuity
  - Diagnosis by fundoscopy
  - Usually presents >6 months posttransplantation
- Other sites of involvement
  - Liver, gallbladder, pancreas, epididymis, biliary tract, skin, endometrium, central nervous system, etc.
- CMV infection/disease in organ transplant recipients associations
  - Acute rejection
  - Graft failure [vasculopathy (cardiac transplant), bronchiolitis obliterans (lung transplant), cirrhosis (liver transplantation)]
  - More aggressive relapse of HCV
  - Immunosuppression (other opportunistic, especially fungal, infections)
  - Other viral infections (EBV PTLD, HHV6)

- Prevention
  - CMV-seronegative, filtered or leukoreduced blood products
  - Selection of CMV-seronegative donors for CMV-seronegative recipients (not typically done)
  - Passive immunoprophylaxis (CMV immunoglobulin)
  - Pharmacologic prophylaxis (e.g., oral valganciclovir, oral ganciclovir)
  - Preemptive therapy
    - Asymptomatic at-risk patients tested for CMV DNA (or antigen) in blood, and if positive preemptive therapy given *before* symptoms develop
      - May be preferred over prophylaxis for patients at low or intermediate risk for CMV disease
      - Common in allogeneic SCT recipients due to risk of ganciclovir- or valganciclovir-associated neutropenia
- Treatment
  - Intravenous ganciclovir (oral valganciclovir)
    - Most widely used drug
    - Undergoes initial phosphorylation by thymidine kinase encoded by CMV *UL97* (phosphotransferase). Cellular enzymes add two more phosphates to make active form (analog of guanosine triphosphate) which is incorporated into replicating DNA by CMV DNA polymerase (encoded by *UL54* [DNA polymerase]) resulting in termination of replication
    - Side effect: leukopenia
  - Foscarnet
    - Inhibits DNA replication
    - Usually administered with intolerance of or failure to respond to ganciclovir
    - Side effects: nephrotoxicity, electrolyte disturbances

Table 28.1 Nonmolecular tests for cytomegalovirus

Assay	Principle	Advantages	Disadvantages	Comments
Histology, cytology	Intranuclear inclusions in tissue, immunohistochemistry, in situ hybridization	Specific and definitive diagnosis	Insensitive and requires invasive procedure	Useful for gastrointestinal disease which may occur without detectable DNAemia
Culture	Focal cytopathic effect in cell lines [e.g., human embryo lung fibroblasts (MRC5)]		Low sensitivity for CMV viremia, prolonged turnaround time	Gold standard for CMV detection
Shell vial assay	Immunofluorescent-labeled antibodies against CMV antigens	More sensitive and faster than conventional culture (1–2 days)	Low sensitivity for CMV viremia	
Serology	IgG produced during primary infection, persists lifelong		IgM not useful in transplant recipients	Used to assess immune status prior to transplantation and for donor testing
CMV antigenemia	Detects structural protein pp65 on surface of infected lymphocytes	Rapid turnaround time	Labor-intensive; requires processing within 6–8 h of specimen collection: compromised by low leukocyte count	

  - Cidofovir
    - Inhibits DNA replication
  - CMV immunoglobulin
- Diagnosis
  - Nonmolecular assays for CMV diagnosis – Table 28.1
  - Clinical utility of molecular CMV testing in transplant recipients
    - Diagnosis of active CMV disease
      - CMV causes latent infection and may be detected without symptoms; latent infection distinguished from active disease by
        - Cutoff values – CMV disease associated with higher viral loads asymptomatic infection (cutoffs may vary between assays)
        - Targeting mRNA (present in active infection) instead of DNA (present in active *and* latent infection)
        - Measurement of viral load in plasma versus whole blood or leukocytes
    - Monitoring response to antiviral therapy
      - Treatment of clinical CMV disease usually requires $\geq$2 weeks full-dose intravenous ganciclovir
        - Document clearance of viremia before intravenous therapy is stopped
      - Clearance time inversely correlates with pretreatment viral load
      - Persistent on therapy elevated viral load suggests ganciclovir resistance
- Molecular assays for detection and quantification of CMV nucleic acid
  - Conventional or real-time PCR [qualitative, quantitative (viral load)]
    - Performance varies due to differences in acceptable specimen type, nucleic acid extraction, target genes, primer sequences, quantitation standards, amplified product detection methods, etc.
  - CMV viral load by PCR – plasma or whole blood
    - Indications
      - Preemptive therapy marker (thresholds vary between SCT and solid organ transplant recipients)

  - Diagnosis of CMV-associated signs and symptoms
  - Monitoring response to antiviral therapy
- World Health Organization International Standard for CMV available from National Institute for Biological Standards and Control (NIBSC product code: 09/162) for standardization
  - Without this standard, results not always comparable between assays use same assay for monitoring individual patients
- PCR advantages versus pp65 antigenemia
  - CMV DNA stable in whole blood/plasma for prolonged periods
  - PCR may be performed with low leukocyte count
- Qualitative PCR
  - Aid to diagnose CMV disease
    - Various clinical specimens (e.g., cerebrospinal fluid, urine, various tissues, respiratory specimens, body fluids)
- Formalin-fixed, paraffin-embedded tissues
  - CMV antigens demonstrated by immunohistochemistry
  - DNA demonstrated by in situ hybridization
- FDA-cleared non-PCR molecular diagnostic test example
  - CMV pp67 mRNA (bioMérieux, Durham, NC)
    - Nucleic acid sequence-based amplification (NASBA)
    - Qualitative
    - Detects pp67 mRNA in whole blood
    - High levels pp67 expressed in active disease (not latent infection)

- Testing for antiviral resistance
  - De novo resistance rare – resistance usually occurs after prolonged therapy
  - Risk factors
    - CMV donor seropositive/recipient seronegative organ transplant
    - High viral load
    - Profound immunosuppression
    - Suboptimal antiviral therapy
    - Lung or kidney–pancreas transplant recipient
  - Indication for antiviral resistance testing
    - Inadequate antiviral treatment response
      - Stable or rising viral load or persistence of symptoms after 2–3 weeks full-dose intravenous therapy
  - Assay types for detection of antiviral resistance
    - Phenotypic detection of CMV antiviral resistance
      - Growth of CMV in presence of varying concentrations of antivirals
      - Plaque reduction assay – concentration that reduces number of plaques by 50% relative to control wells without drug is 50% inhibitory concentration (IC_{50})
      - Time-consuming (4–6 weeks), labor-intensive, subjective, not standardized
  - Genotypic
    - Genotypic detection of CMV antiviral resistance – *UL97* and/or *UL54* amplified using PCR and PCR products sequenced
    - From cultured virus or directly from clinical specimen
    - Ganciclovir resistance most commonly due to point mutations/deletions in *UL97* (foscarnet and cidofovir not affected)
      - Decreased levels of ganciclovir triphosphate in CMV-infected cells
      - Mutations at codons 460, 594, and 595 most common
    - *UL54* point mutations/deletions
      - Less frequent than *UL97* mutations
      - If selected by ganciclovir or cidofovir, confers cross-resistance to one another but usually not foscarnet
      - If selected by foscarnet, usually does not confer cross-resistance to ganciclovir or cidofovir

Table 28.2 Risk factors and poor prognostic factors of posttransplant lymphoproliferative disease (PTLD)

Risk factors	Poor prognostic factors
Early PTLD	Poor performance status Multisite disease Central nervous system disease Monoclonal disease T cell or NK cell PTLD EBV-negative PTLD Hepatitis B or C virus coinfection Recipient (versus donor) origin disease Proto-oncogene or tumor suppressor gene mutation
Primary EBV infection Young recipient age Type of organ transplanted CMV disease Antilymphocyte antibody receipt	
Late PTLD	
Duration/amount of immunosuppression Type of organ transplanted Older recipient age	

28.3 Epstein–Barr Virus (Posttransplantation Lymphoproliferative Disease)

- Enveloped, double-stranded DNA virus, *Herpesviridae* family
- >90% adults infected
- Transmission
  - Donor-transmitted infection (organ donor seropositive/recipient seronegative)
  - Blood transfusion, exposure to saliva of infected asymptomatic person
  - Reactivation
- Disease associations
  - Posttransplantation lymphoproliferative disease (PTLD)
    - Rarely non-PTLD disease
- Infectious mononucleosis
  - Sore throat, malaise, fever, headache
- Hemophagocytic syndrome
- Posttransplantation Lymphoproliferative Disease Pathogenesis
  - EBV replication, without lymphocytes that normally control expression of EBV-infected, transformed B cells (due to antilymphocyte therapy)
    - Allogeneic SCT recipient, EBV-infected B cells usually donor-derived
    - Organ transplant recipient, EBV is typically released from transplanted organ and infects recipient B cells
- Risk factors
  - EBV seronegativity prior to organ transplantation
  - CMV disease
  - Type and intensity of immunosuppression (especially administration of antilymphocyte therapy for rejection)
  - High EBV viral load
- Clinical presentation
  - Median time to onset
    - Organ transplant recipients – 6 months
    - SCT recipients – 3 months
  - SCT recipients – more severe, disseminated versus organ transplant recipients
  - 1–15% of transplant patients
    - Small intestinal transplant recipients, up to 20%
    - Pancreas, heart, lung, and liver transplants recipients, 3–12%
    - Renal transplant and SCT recipients, 1–2%
  - Mortality, 40–70%
  - More frequent in children than adults (children more likely seronegative)
  - Risk factors and prognostic factors; see Table 28.2
  - EBV genome in majority (>90%) of early (within first year after organ transplant) B cell PTLD
  - 21–38% of late PTLD EBV-negative, non-B cell

  - Clinical presentations include fever (including "fever of unknown origin"), weight loss, adenopathy, abdominal pain, anorexia, jaundice, gastrointestinal bleeding, intestinal perforation, renal dysfunction, liver dysfunction, pneumothorax, and pulmonary infiltrates or nodules
  - Often multicentric, may involve central nervous system, eyes, gastrointestinal tract, liver, spleen, lymph nodes, lungs, allograft, oropharynx, and other organs
- Diagnosis
  - Serology
    - Determines pretransplant donor and recipient EBV serostatus (assess PTLD risk)
    - Unreliable for diagnosis of PTLD
  - PTLD diagnosis
    - Definitive diagnosis – biopsy
    - Monoclonal, oligoclonal, and polyclonal
    - B cell or T cell lymphoma
    - Diagnosis of EBV-associated PTLD – EBV DNA, RNA, or protein in biopsy tissue
      - Gold standard: in situ hybridization targeting *EBER1* and/or *EBER2*
        - More sensitive than targeting viral DNA because EBER is expressed at high levels in infected cells
        - More sensitive than immunohistochemistry
        - Example commercial systems (Ventana [Tucson, AZ], Leica [Bannockburn, IL], Dako [Glostrup, Denmark], Invitrogen [Carlsbad, CA], Biogenex [San Ramon, CA])
  - Conventional or real-time PCR [qualitative or quantitative (viral load)]
    - Workup of transplant recipient with symptoms suggestive of PTLD
      - High EBV load triggers search for mass lesions with computerized tomography of chest, abdomen, and pelvis, or organ dysfunction, pinpointing potential site(s) for biopsy
    - Increase in EBV viral load in peripheral blood (whole EDTA blood, peripheral blood lymphocytes, plasma) may be detected *before* development of EBV-associated PTLD; levels typically decrease with effective therapy
      - Levels of EBV DNA above threshold trigger evaluation for PTLD (see above); even if PTLD not found may trigger preemptive lessening of immunosuppression, administration of murine humanized chimeric anti-CD20 monoclonal antibody (Rituximab), or adoptive immunotherapy
      - Cutoff values to trigger evaluation unclear
        - Some pediatric liver and heart transplant recipients have chronic high viral loads
        - Low level elevated EBV viral load frequent; may resolve without intervention
    - No standardization of optimal assay technique, sample type (i.e., whole blood, lymphocytes, plasma), or sampling schedule
- Treatment
  - Early detection; reduction in immunosuppression – response high in low risk patients
  - Murine humanized chimeric anti-CD20 monoclonal antibody (Rituximab)
  - Adoptive immunotherapy using donor-derived cloned EBV-specific cytotoxic T cells
  - Chemotherapy
  - Surgical resection
  - Radiotherapy
  - Antiviral drugs (limited role)

28.4 BK Virus

- Nonenveloped, double-stranded DNA virus, *Polyomaviridae* family
  - 75% sequence homology with JC virus
  - Four genotypes (1–4)
- Primary infection in childhood; 60–90% adults infected

  - Following primary infection, remains latent in urogenital tract, B lymphocytes, or other tissues (e.g., spleen, brain)
  - May reactivate with urinary shedding of infected urothelial cells following immunosuppression
  - Up to 5% of normal hosts – asymptomatic viruria
- Viruria after transplantation
  - Kidney transplantation, 27%
  - Heart transplantation, 26%
  - Liver transplantation, 8%
- Viremia after transplantation
  - Kidney transplantation, 12%
  - Heart transplantation, 7%
- Clinical presentation
  - Polyomavirus-associated nephropathy (PVAN) of allograft – kidney transplant recipients
    - Incidence varies, up to 8%
    - Histologic progression – initial cytopathic stage, followed by cytopathic inflammatory stage and late stage tubular atrophy and fibrosis
    - Mostly in first posttransplant year
    - Graft loss 15–80%
    - Definitive diagnosis requires renal allograft biopsy with in situ hybridization for BK virus
  - Hemorrhagic cystitis (HC) – SCT recipients
    - Risk greater for allogeneic than autologous SCT recipients (rare in organ transplant recipients)
    - Typically 1–18 weeks after transplantation
    - Symptoms similar to bacterial urinary tract infection
    - Resolves spontaneously
- Diagnosis and monitoring
  - Urine cytology
    - BK virus-infected – "decoy" cells – enlarged round nuclei with ground glass appearance and marginal chromatin
      - Useful for early diagnosis PVAN with 100% sensitivity but low positive predictive value (<30%)
      - Not specific for BK virus – may be found with JCV, adenovirus
  - Renal biopsy
    - Viral cytopathic changes in renal tubular epithelium
    - Early stage PVAN may be focal (false negative biopsy)
  - Molecular assays
    - Active replication of BK virus in urothelial cells, with tissue damage, releases BK virus into urine and blood
    - Early diagnosis of BK virus replication before renal function deteriorates
    - Guide management of immunosuppressive therapy
    - Monitor response to intervention
    - Real-time PCR – method of choice for BK virus viral load
    - Assays differ in equitable detection of genotypes, limit of quantitation and dynamic ranges, sample type (urine vs. urine sediment vs. unextracted urine; plasma vs. serum vs. whole blood), DNA extraction and purification method, primer and probe sequences, and PCR amplification conditions
      - Plasma preferred over serum, whole blood
      - Whole urine preferred over urine sediment
- Clinical utility of molecular testing in kidney transplant recipients
  - Diagnosis of PVAN with allograft dysfunction or preemptively (Fig. 28.3)
  - Urine testing
    - Cytology may be used as screening test, followed by BK viral load testing, if positive
    - PCR more sensitive than cytology; BK virus DNA may be present in urine earlier than decoy cells
  - Plasma testing
    - BK viremia occurs only when virus is detected in urine and is prerequisite for progression to PVAN
      - Rare transient BK viremia without urinary excretion may reflect reactivation of BK virus in

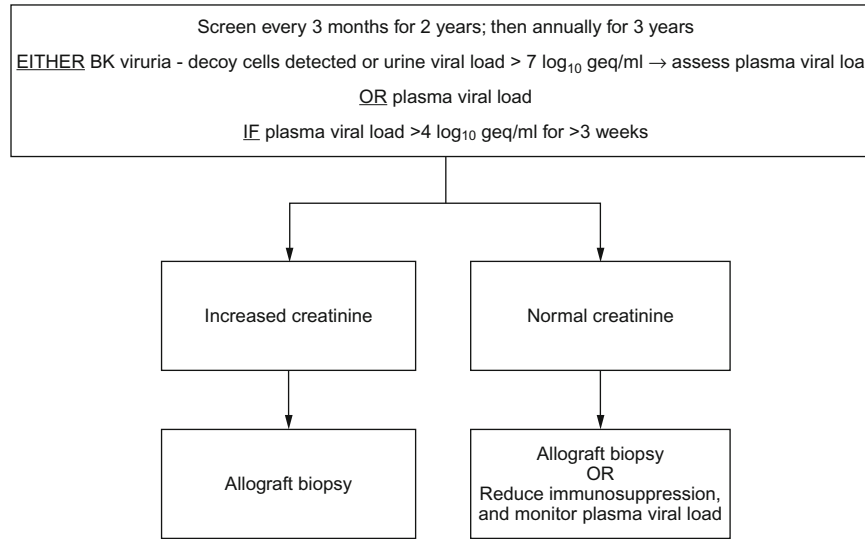


Fig. 28.3 BK virus monitoring algorithm example

circulating leukocytes, PCR inhibitors in urine, or sequence differences between urine and plasma BK virus
    - Median interval from onset of viruria to onset of viremia – 1–3 months
    - Viremia precedes PVAN – median 1–12 weeks
    - PVAN higher BK viral load than no PVAN
    - Plasma viral load $>10^4$ genome equivalents/ml >3 weeks – presumptive PVAN
    - Monitoring of therapy
- Treatment
  - Reduction in immunosuppression
    - Enables immune system to recover and control infection
    - Risk of acute rejection
  - Cidofovir, leflunomide, or intravenous immunoglobulin
  - Viremia resolves before viruria
    - Clearance $t_{1/2}$ 6 h to 17 days
      - Monitor viral load every 2–4 weeks

28.5 JC Virus

- Nonenveloped, double-stranded DNA virus, *Polyomaviridae* family
  - 75% sequence homology with BK virus
- Etiologic agent of progressive multifocal leukoencephalopathy (PML)
  - Fatal central nervous system demyelinating disease
  - Brain biopsy
    - Characteristic pathologic changes localized mainly in oligodendrocytes and astrocytes
    - In situ hybridization for JC virus on brain tissue
  - Detection of JC virus DNA by PCR in cerebrospinal fluid has largely replaced tissue biopsy for diagnosis

28.6 Hepatitis C Virus

- Enveloped, positive sense, single-stranded RNA virus, *Flaviviridae* family
  - Genome encodes single open reading frame coding structural proteins (one core and two envelope proteins) and nonstructural proteins (NS1–NS5)
    - Envelope region hypervariable – high mutation rate
  - Six genotypes based on sequence homology of 5′ NTR
    - Genotypes 1–6
      - <70% homology in nucleotide sequence

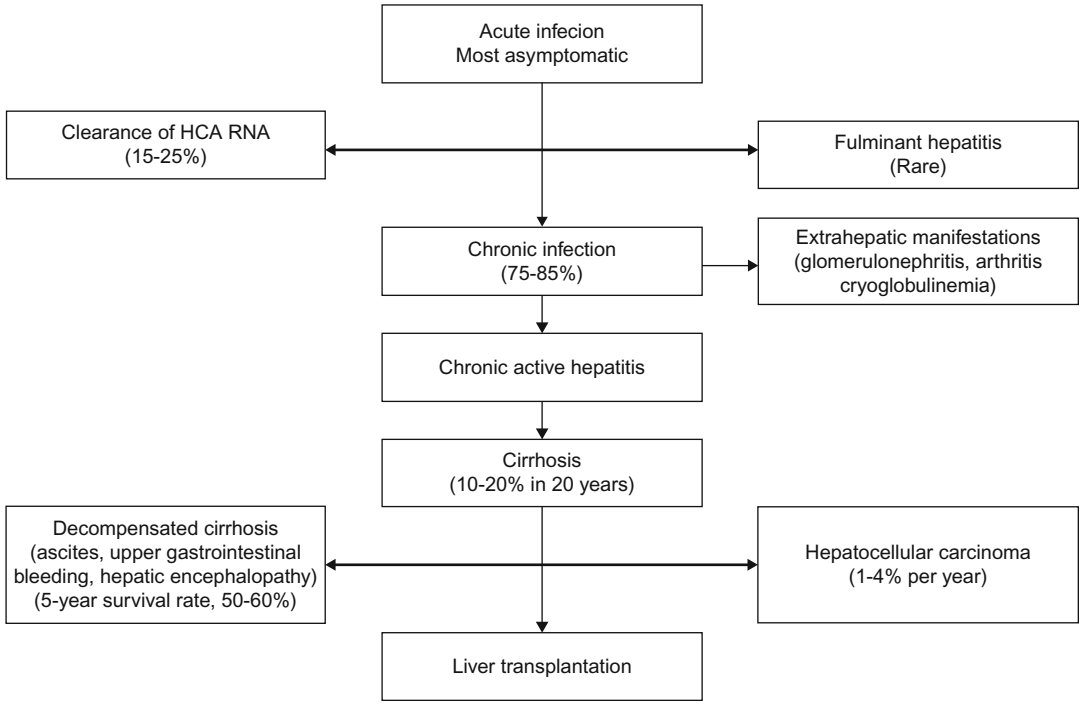


Fig. 28.4 Clinical presentation of hepatitis C virus infection

  - Types la and lb ~60% of infections
  - United States – ~70–75% genotype 1, most remaining types 2 or 3
- ~1.6% adults infected; most common indication for liver transplantation
- Transmission
  - Injection drug use, transfusion of blood products (before 1992), transplantation, dialysis, needle stick injury, vertical (infected mother to child), sexual
- Clinical presentation
  - Clinical presentation of HCV presented in Fig. 28.4
- Diagnosis of HCV infection
  - Serologic assays
    - Immunoassay
      - Enzyme-linked immunosorbent assay (ELISA)
      - Chemiluminescent assay
    - Supplemental or reflex testing distinguishes new from past infection
    - Antibody production absent
      - Early infection (window period, 7–8 weeks)
      - Some immunocompromised patients (e.g., dialysis, agammaglobulinemia)
      - Delayed in some intravenous drug users
    - False positive (up to 30% in low prevalence populations)
    - Confirmatory test – recombinant immunoblot assay (RIBA)
      - Most helpful in low signal-to-cutoff ratio samples
      - Samples with high signal-to-cutoff ratio are rarely RIBA negative
  - Molecular assays
    - Qualitative assays
      - Uses
        - Confirm infection
        - Document viral clearance during therapy, end of therapy, and following therapy completion

  - Technologies
    - Reverse transcription–polymerase chain reaction
    - Transcription-mediated amplification (TMA)
  - Ideally detect 50 HCV RNA IU/ml or less and have equitable sensitivity for detection of genotypes
- Quantitative assays
  - Uses
    - Quantify baseline RNA
    - Measure viral decline during therapy
  - Technologies
    - Reverse transcription–polymerase chain reaction
    - Branched DNA
    - TMA
  - HCV RNA levels above upper limit of quantification of assay underestimated
    - Samples retested after 1/10 to 1/100 dilution for accurate quantification
- Genotype determination
  - Cure rates with therapy higher with genotypes 2 and 3
    - Allows shorter duration of therapy (compared to genotype 1)
- Interleukin – 28B genotype
  - Single nucleotide polymorphisms on chromosome 19 in or near the interleukin-28B gene predict successful antiviral treatment response
    - C (vs. T) allele advantages for single nucleotide polymorphism rs129789860
    - T (vs. G) allele advantages for single nucleotide polymorphism rs8099917
- Clinical utility of molecular testing in transplant recipients
  - Organs from donors with HCV antibody and detectable HCV RNA more likely to transmit HCV infection than those from donors with HCV antibody but without detectable HCV RNA
  - HCV infection independent risk factor for mortality following liver transplantation
  - Liver transplantation for HCV cirrhosis
    - Recurrent HCV disease in most with viremia at transplant
    - Spontaneous viral clearance rare
    - May lead to graft failure, cirrhosis
    - Histologic progression accelerated compared to general population
  - Acute hepatitis associated with recurrent HCV infection >50%, typically in first 6 months following liver transplant – increased viral load, rise in serum aminotransferase levels, and/or histologic evidence of acute HCV infection
  - Risk factors for severe recurrent HCV and poor allograft survival after liver transplantation
    - Fibrosing cholestatic HCV syndrome (infrequent severe disease with high viral load occurring in first 6 months after transplant)
    - Early recurrence or severe histologic inflammatory activity within 1 year
    - Increased donor age (>50 years)
    - High serum HCV RNA levels before or after transplantation
    - Certain immunosuppressive agents (e.g., high dose steroids)
    - Other infections (CMV, human immunodeficiency virus)
  - Serum viral RNA
    - May reach pretransplant levels within first few days postoperatively
    - Serum RNA peaks 1–3 months posttransplant, reaching levels many fold higher than in pretransplant period
  - Liver biopsy for evaluation of posttransplant recurrent HCV infection
  - Treatment
    - Preemptive approach (treat immediately following transplantation)
    - Recurrence-based approach (patients with histologic liver disease selected for treatment)
    - Recurrent HCV infection
      - Liver transplant recipients usually treated
      - Other types of organ transplant recipients usually not treated (graft rejection/loss secondary to interferon's immunomodulatory effects)

  - Kidney transplantation of HCV positive recipients – no increased risk of graft loss or death following transplantation
    - RNA levels and liver biopsy assist in management of HCV infection before transplantation
    - Treat before transplantation
  - HCV infection associated with transient hepatitis and veno-occlusive disease in posttransplant period in SCT recipients
    - Persistent hepatitis associated with increased risk of earlier cirrhosis compared with HCV-infected nontransplanted patients

28.7 Herpes Simplex Virus

- Enveloped, double-stranded DNA virus, *Herpesviridae* family
- Following primary infection, virus remains latent in sensory nerve ganglia
- Anti-HSV IgG antibodies ~75% adults
- Most commonly presents as reactivation infection
  - Usually in first month after transplantation
  - Oral or genital mucocutaneous lesions
  - Most orolabial infections are mild, although severe ulceration and discomfort, complicated by bacterial superinfection or esophageal involvement, are possible
  - Anogenital infection usually presents as ulceration and may or may not have typical vesicular appearance
  - Rare zosteriform lesions on buttocks, or mucocutaneous nodules or plaques
  - Prevented by acyclovir, valacyclovir, ganciclovir, or valganciclovir
- Occasional primary infection, transmitted by contact from person to person or via allograft
- Reactivation or primary HSV infection occasionally causes:
  - Pneumonitis, tracheobronchitis
    - HSV pneumonitis usually secondary process in intubated patients with pneumonia of other etiology
    - Due to oropharyngeal shedding, detection of HSV in respiratory secretions does not definitively imply HSV pneumonitis
      - Lung biopsy needed for definitive diagnosis of pneumonitis
  - Esophagitis
    - Dysphagia
    - Mimics candidal and CMV esophagitis
  - Hepatitis
  - Disseminated infection
  - Ocular infection
  - Central nervous system infection
- Diagnosis
  - PCR more sensitive and faster than culture, more sensitive and specific than Tzanck, and more sensitive than direct fluorescent antibody testing
    - Mucosal or skin lesion – vigorously rub culture transport swab over suspect lesion
      - Assay should detect and differentiate HSV types 1 and 2

28.8 Varicella Zoster Virus

- Enveloped, double-stranded DNA virus, *Herpesviridae* family
- ~90% of adult organ transplant recipients seropositive before transplantation – at risk for reactivation (zoster) typically after first 6 months of transplantation
  - 11% renal transplant recipients within 4 years of transplant
  - 37% SCT recipients within 3 years of transplant
  - Dermatomal distribution (may involve multiple adjoining dermatomes), occasionally with a few or many sites of cutaneous dissemination at distant sites, or more widespread dissemination with visceral and/or central nervous system involvement
- 10% adult solid organ transplant recipients seronegative
  - Risk primary infection after transplantation
    - Occurs at any time
    - Acquired via contact with infected individual via respiratory route

  - Primary infection
    - Chickenpox syndrome, hepatitis, or fatal disseminated infection
  - Screen transplant candidates for antibody to VZV prior to transplantation
    - Vaccinate organ transplant candidates before transplantation if no history of prior VZV disease and VZV seronegative
    - Postexposure prophylaxis in nonimmune transplant recipients
- Diagnosis
  - PCR more sensitive and faster than culture
    - Culture transport swab vigorously rubbed over suspect lesion
    - Cerebrospinal fluid tested as an aid to diagnosis of VZV central nervous system infection
- Treatment
  - Localized dermatomal zoster – acyclovir, famciclovir, or valacyclovir
  - Primary infection – intravenous acyclovir +/− varicella zoster immune globulin

28.9 Human Herpes Virus 6

- Enveloped, double-stranded DNA virus, *Herpesviridae* family
- Primary infection
  - Transmitted via infected donor organs, cells, or cellular blood products
  - Nontransplant childhood disease – roseola infantum, asymptomatic infection
- Reactivation infection (common)
- Clinical presentation
  - Asymptomatic
  - Symptomatic
    - Febrile illness
    - Rash
    - Pneumonitis
    - Myelosuppression
    - Hepatitis
    - Encephalitis
  - CMV reactivation
- Diagnosis
  - PCR
    - Plasma, serum, whole blood, peripheral blood mononuclear cells, whole blood, biopsy, and tissue specimens
    - Ideally detects and differentiates variants A and B
    - As with CMV, does not differentiate replicating from latent virus
      - Differentiate latent from active infection
        - High/increasing viral load or reverse transcription PCR to detect RNA
        - Serum rather than whole blood
  - In situ hybridization of formalin-fixed, paraffin-embedded tissues
  - Culture from peripheral blood mononuclear cells (other clinical specimens)
    - Labor-intensive
    - Takes up to 21 days (1–3 days with shell vial assay)
- Treatment – ganciclovir, foscarnet

28.10 Parvovirus (Erythrovirus) B19

- Nonenveloped, single-stranded DNA virus, *Parvoviridae* family
- 60–90% adults seropositive
- Infects erythroid progenitor cells by binding P antigen
- Clinical presentation
  - Nontransplant patients
    - Children – erythema infectiosum
    - Adults – occasional arthropathy
    - Pregnancy – hydrops fetalis
    - Hemolytic disorders – severe anemia
  - Transplant patients
    - Usually presents in first 3 months after transplantation
      - Anemia is most common (weakness, dyspnea, orthostasis)
      - Rare – hepatitis, myocarditis, pneumonitis, glomerulopathy, graft dysfunction, leukopenia, thrombocytopenia, fever and flulike manifestations, rash, arthralgia, carditis, or hepatitis

- Diagnosis
  - PCR (plasma or serum) most sensitive non-invasive technique for diagnosis of parvovirus B19-related anemia
    - May remain positive for extended periods
  - Bone marrow examination
    - Giant pronormoblasts (not always detected) suggest parvovirus B19 infection
  - Serology – IgM lacks sensitivity
- Treatment
  - Reduction in immunosuppression
  - Intravenous immune globulin (contains parvovirus B19-specific antibodies)

28.11 Adenovirus

- Nonenveloped, double-stranded DNA virus, *Adenoviridae* family – 52 serotypes
- Respiratory transmission (also water, fomites)
- Normal host – respiratory tract infection, gastroenteritis, or conjunctivitis
- Transplant patients – respiratory tract infection (pneumonia, upper respiratory tract infection), disseminated infection, gastroenteritis, hemorrhagic cystitis, meningoencephalitis, or hepatitis
- Diagnosis – viral culture (long turnaround time), PCR (qualitative, quantitative), or histopathology
- Treatment – symptomatic, cidofovir, or ribavirin
- Preemptive therapy – based on viral load, used in some pediatric SCT populations

28.12 Hepatitis B Virus

- Enveloped, DNA virus, *Hepadnaviridae* family
- Recurrent HBV 80–90% patients following liver transplantation
  - Detected by appearance of hepatitis B surface antigen 2–6 months after transplantation, followed by hepatocellular injury
- Molecular methods used for diagnosis and resistance testing
- Primary HBV, acquired from donor liver or by blood product transfusion, usually mild
- Cirrhosis of liver allograft uncommon in patients transplanted for fulminant HBV

28.13 *Mycobacterium tuberculosis*

- Transplant recipients increased risk for primary and reactivation infection
- Disseminated disease more common in transplant recipients than in other populations
- Rarely transmitted by allograft
- Clinical presentation
  - Cavitary and noncavitary pulmonary disease
  - Extrapulmonary – intestinal, skeletal, cutaneous, central nervous system, or disseminated disease
- Molecular methods used to diagnose *M. tuberculosis* from positive cultures, and directly from clinical specimens (e.g., sputum, tissues), and for drug resistance testing
  - FDA-approved test example
    - AMPLIFIED™ *Mycobacterium tuberculosis* Direct Test (MTD) [Gen-Probe, San Diego, California] – target-amplified nucleic acid probe test for detection of *M. tuberculosis* complex ribosomal RNA in sputum, bronchoalveolar lavage fluid, or bronchial or tracheal aspirates

28.14 *Toxoplasma gondii*

- Toxoplasmosis generally develops within the first six months after transplantation with the highest incidence in the second and third months
- Usually results from reactivation of latent disease in the seropositive donor heart when transplanted into seronegative recipient but may occur in any type of organ transplant recipient

- Prophylaxis given to seronegative heart transplant recipients who receive allografts from seropositive donor
- Meningoencephalitis, brain abscess, pneumonia, myocarditis, pericarditis, hepatitis, or retinochoroiditis
- PCR can be performed on blood, cerebrospinal fluid, respiratory secretions, or tissues

Further Reading

Abdel Massih RC, Razonable RR. Human herpesvirus 6 infections after liver transplantation. World J Gastroenterol. 2009;15:2561–9.

Aitken C, Barrett-Muir W, Millar C, et al. Use of molecular assays in diagnosis and monitoring of cytomegalovirus disease following renal transplantation. J Clin Microbiol. 1999;37:2804–7.

Allen U, Preiksaitis J, AST Infectious Diseases Community of Practice. Epstein-Barr virus and posttransplant lymphoproliferative disorder in solid organ transplant recipients. Am J Transplant. 2009;9:S87–96.

Bechert CJ, Schnadig VJ, Payne DA, et al. Monitoring of BK viral load in renal allograft recipients by real-time PCR assays. Am J Clin Pathol. 2010;133:242–50.

Blanckaert K, De Vriese AS. Current recommendations for diagnosis and management of polyoma BK virus nephropathy in renal transplant recipients. Nephrol Dial Transplant. 2006;21:3364–7.

Caliendo AM, Schuurman R, Yen-Lieberman B. Comparison of quantitative and qualitative PCR assays for cytomegalovirus DNA in plasma. J Clin Microbiol. 2001;39:1334–8.

Caliendo AM, St. George K, Allega J, et al. Distinguishing cytomegalovirus (CMV) infection and disease with CMV nucleic acid assays. J Clin Microbiol. 2002;40:1581–6.

Campe H, Jaeger G, Abou-Ajram C, et al. Serial detection of Epstein-Barr virus DNA in sera and peripheral blood leukocyte samples of pediatric renal allograft recipients with persistent mononucleosis-like symptoms defines patients at risk to develop posttransplant lymphoproliferative disease. Pediatr Transplant. 2003;7:46–52.

Charlton M, Seaberg E, Wiesner R, et al. Predictors of patient and graft survival following liver transplantation for hepatitis C. Hepatology. 1998;28:823–30.

Chevaliez S, Pawlotsky JM. Hepatitis C virus serologic and virologic tests and clinical diagnosis of HCV-related liver disease. Int J Med Sci. 2006;3:35–40.

Costa C, Elia M, Astegiano S, et al. Quantitative detection of Epstein-Barr virus in bronchoalveolar lavage from transplant and nontransplant patients. Transplantation. 2008;86:1389–94.

Diaz-Mitoma F, Leger C, Miller H. Comparison of DNA amplification, mRNA amplification, and DNA hybridization techniques for detection of cytomegalovirus in bone marrow transplant recipients. J Clin Microbiol. 2003;41:5159–66.

Drew WL. Cytomegalovirus resistance testing: pitfalls and problems for the clinician. Clin Infect Dis. 2010;50:733–6.

Eid AJ, Brown RA, Patel R, Razonable RR. Parvovirus B19 infection after transplantation: a review of 98 cases. Clin Infect Dis. 2006;43:40–8.

Fishman JA. Infection in solid-organ transplant recipients. N Engl J Med. 2007;357:2601–14.

Gartner BC, Schafer H, Marggraff K. Evaluation of use of Epstein-Barr viral load in patients after allogeneic stem cell transplantation to diagnose and monitor posttransplant lymphoproliferative disease. J Clin Microbiol. 2002;40:351–8.

Gilbert C, Boivin G. Human cytomegalovirus resistance to antiviral drugs. Antimicrob Agents Chemother. 2005;49:873–83.

Gonzalez SA. Management of recurrent hepatitis C following liver transplantation. Gastroenterol Hepatol. 2010;6:637–45.

Hirsch HH, Randhawa P, AST Infectious Diseases Community of Practice. BK virus in solid organ transplant recipients. Am J Transplant. 2009;9:S136–46.

Humar A, Gregson D, Caliendo AM, et al. Clinical utility of quantitative cytomegalovirus viral load determination for predicting cytomegalovirus disease in liver transplant recipients. Transplantation. 1999;68:1305–11.

Humar A, Kumar D, Boivin G, et al. Cytomegalovirus (CMV) virus load kinetics to predict recurrent disease in solid-organ transplant patients with CMV disease. J Infect Dis. 2002;186:829–33.

Ikewaki J, Ohtsuka E, Satou T, et al. Real-time PCR assays based on distinct genomic regions for cytomegalovirus reactivation following hematopoietic stem cell transplantation. Bone Marrow Transplant. 2005;35:403–10.

Kogan-Liberman D, Burroughs M, Emre S, et al. The role of quantitative Epstein-Barr virus polymerase chain reaction and preemptive immunosuppression reduction in pediatric liver transplantation: a preliminary experience. J Pediatr Gastroenterol Nutr. 2001;33:445–9.

Kotton CN, Fishman JA. Viral infection in the renal transplant recipient. J Am Soc Nephrol. 2005;16:1758–74.

Lee TC, Savoldo B, Rooney CM, et al. Quantitative EBV viral loads and immunosuppression alterations can decrease PTLD incidence in pediatric liver transplant recipients. Am J Transplant. 2005;5:2222–8.

Munoz P, Fogeda M, Bouza E, et al. Prevalence of BK virus replication among recipients of solid organ transplants. Clin Infect Dis. 2005;41:1720–5.

Pang XL, Doucette K, LeBlanc B. Monitoring of polyomavirus BK virus viruria and viremia in renal allograft recipients by use of a quantitative real-time PCR

assay: one-year prospective study. J Clin Microbiol. 2007;45:3568–73.
Patel R, Paya CV. Infections in solid-organ transplant recipients. Clin Microbiol Rev. 1997;10:86–124.
Peffault de Latour R, Ribaud P, Robin M, et al. Allogeneic hematopoietic cell transplant in HCV-infected patients. J Hepatol. 2008;48:1008–17.
Randhawa P, Brennan DC. BK virus infection in transplant recipients: an overview and update. Am J Transplant. 2006;6:2000–5.
Randhawa P, Ho A, Shapiro R, et al. Correlates of quantitative measurement of BK Polyomavirus (BKV) DNA with clinical course of BKV infection in renal transplant patients. J Clin Microbiol. 2004;42:1176–80.
Schiano TD, Martin P. Management of HCV infection and liver transplantation. Int J Med Sci. 2006;3:79–83.
Sessa A, Esposito A, Giliberti A, et al. BKV reactivation in renal transplant recipients: diagnostic and therapeutic strategy – case reports. Transplant Proc. 2008;40:2055–8.
Sia IG, Wilson JA, Groettum CM, et al. Cytomegalovirus (CMV) DNA load predicts relapsing CMV infection after solid organ transplantation. J Infect Dis. 2000;181:717–20.
Smith TF, Espy MJ, Mandrekar J, et al. Quantitative real-time polymerase chain reaction for evaluating DNAemia due to cytomegalovirus, Epstein-Barr virus, and BK virus in solid-organ transplant recipients. Clin Infect Dis. 2007;45:1056–61.
Sriaroon C, Greene JN, Vincent AL, et al. BK virus: microbiology, epidemiology, pathogenesis, clinical manifestations and treatment. Asian Biomedicine. 2010;4:3–18.
Terrault NA, Adey DB. The kidney transplant recipient with hepatitis C infection: pre- and posttransplantation treatment. Clin J Am Soc Nephrol. 2007;2:563–75.
Tsai DE, Nearey M, Hardy CL, et al. Use of EBV PCR for the diagnosis and monitoring of posttransplant lymphoproliferative disorder in adult solid organ transplant patients. Am J Transplant. 2002;2:946–54.

Molecular Microbiology Epidemiology 29

Miao Cui, Fei Ye, and David Y. Zhang

Contents

29.1 General Information 813

29.2 Methods 814
29.2.1 Typing by Restriction Fragment Length Polymorphism 814
29.2.2 Ribotyping with Southern Blot Analysis 818
29.2.3 Typing by Polymerase Chain Reaction 818
29.2.4 Typing by Sequencing Analysis 819

Further Reading 822

M. Cui, MD (✉) • F. Ye, PhD • D.Y. Zhang, MD, PhD, MPH
Department of Pathology, Mount Sinai School of Medicine, New York, NY, USA

29.1 General Information

- Molecular epidemiology is the use of molecular methods to conduct epidemic investigation of bacterial outbreaks, including the recognition that a problem exists, establishment of a case–control definition, confirmation of cases, and completion of the case findings
  - The technologies most commonly used for analysis of bacterial DNA are restriction fragment length polymorphism (RFLP)-based assays, DNA repeat-based assays (ribotyping and VNTR), PCR-based assays (multiplex PCR and random-amplified polymorphic DNA analysis), and sequencing-based assays (multilocus sequence). RFLP–pulsed field gel electrophoresis (PFGE) used to be considered the current "gold standard" for typing bacteria is now replaced by sequencing-based assays due to more diversity in single sequence polymorphism (SNP) than in RFLP
  - Molecular epidemiology is used to investigate environmental and hospital-based (nosocomial) outbreaks. Establishing lineage or clonality of pathogens can aid in the identification of the source (environmental or personnel) of organisms, distinguish infectious from noninfectious strains, and distinguish relapse from reinfection. Information from molecular epidemiological studies facilitates the design of rational pathogen control methods
- Nosocomial infections are an important source of morbidity and mortality in hospital settings

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_29,

- An estimated 2 million patients in the United States are affected each year
- Accounts for approximately 5% of hospitalized patients
- Results in an estimated 88,000 deaths annually
- 4.5 billion dollars in excess healthcare costs
- Bacterial agents remain the most commonly recognized cause of hospital-acquired infections
- Multidrug-resistant pathogens represent a major problem, including
  - Gram-positive nosocomial pathogens: glycopeptide (vancomycin)-resistant *Enterococci*, methicillin-resistant *Staphylococcus aureus* (MRSA), and, more recently, glycopeptide-intermediate and glycopeptide-resistant *S. aureus*
  - Gram-negative bacilli: extended spectrum beta-lactamase-producing strains of *Escherichia coli* and *Klebsiella pneumoniae*, and fluoroquinolone-resistant strains of *Pseudomonas aeruginosa* and *E. coli*
- The foundation of molecular epidemiology is based on establishing clonality of microorganisms and understanding the distribution and relatedness of microorganisms
  - There are a number of key factors that are essential in an epidemic investigation, including the recognition that a problem exists, establishment of a case–control definition, confirmation of cases, and completion of the case findings
  - There are a number of important attributes for successful typing schemes: the methodologies should be standardized, sensitive, specific, and objective
  - All typing systems can be characterized in terms of typeability, reproducibility, discriminatory power, ease of performance and interpretation, and cost (in terms of time and money)
    - Typeability – the ability of a technique to assign an unambiguous result (type) to each isolate
    - Reproducibility – the ability to yield the same result upon repeat testing of a bacterial strain. Poor reproducibility may reflect technical variation in the method or biologic variation occurring during in vivo or in vitro passage of the organisms
    - Discriminatory power – the ability to differentiate among epidemiologically unrelated isolates
  - Most molecular methods require costly materials and equipment but are relatively easy to learn and are applicable to a variety of species
- The spatial distribution of bacterial pathogens can be assessed at a local level (e.g., streets, hospital wards), at regional level (cities, provinces), or even globally with the aid of internet-based tools such as multilocus sequence typing (MLST) maps (http://maps.mlst.net/view_maps.php) which allows the user to enter and integrate MLST data together with location as an additional variable (http://www.spatialepidemiology.net/)
  - International alliances have been organized to coordinate the effort in molecular epidemiology. CEM–NET (for Centro de Epidemiologia Molecular) combined the interests and scientific skills of two university-based centers: the Laboratory of Microbiology and Infectious Disease at Rockefeller University in New York, USA, and the Laboratory of Molecular Genetics at Instituto de Tecnologia Química e Biológica (ITQB), affiliated with the Universidade Nova de Lisboa in Portugal. CEM–NET tracks antibiotic-resistant staphylococci and pneumococci in hospitals and in the community

29.2 Methods

29.2.1 Typing by Restriction Fragment Length Polymorphism

- RFLP–PFGE: Considered as a gold standard method because of its good discrimination power and reproducibility

Enzyme	Sequences
Sma I	5′ -CCC|GGG-3′ 3′ -GGG|CCC-5′
Xba I	5′ -T|CTAGA-3′ 3′ -AGATC|T-5′

Fig. 29.1 Recognition sequences of restriction endonucleases *SmaI* and *Xba1*

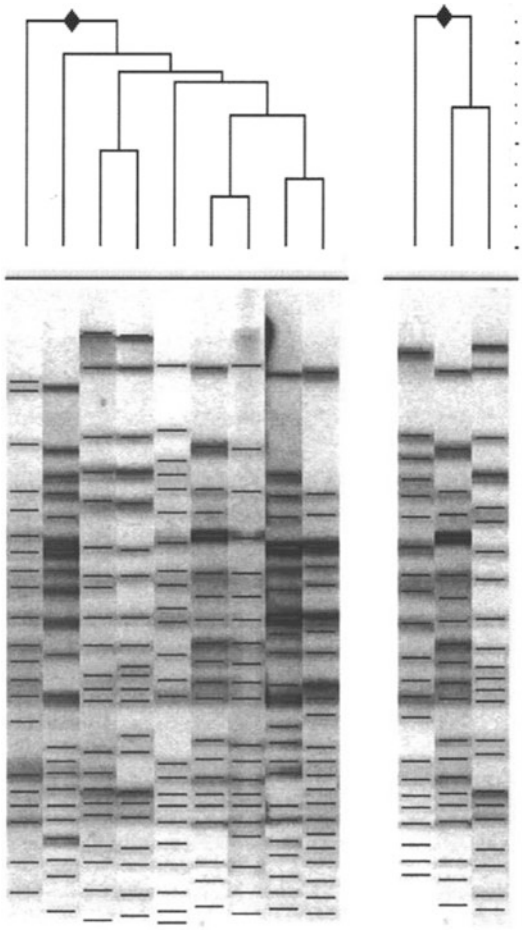


Fig. 29.2 Analysis of RFLP–PFGE results using a computer program (*bottom*) and grouping of isolates of *Acinetobacter* by dendrogram (*top*)

- Principle
  - Chromosomal DNA is digested with restriction enzymes, resulting in a series of fragments of different sizes that form different patterns (i.e., DNA fingerprinting) when analyzed by agarose gel electrophoresis
  - Differences in these patterns are referred to as RFLPs. Usually six nucleotide cutters (Fig. 29.1), such as *SmaI*, *XbaI*, and *SalI*, are used to digest DNA in order to generate relatively few DNA fragment for analysis
  - The product is usually analyzed by pulse field gel electrophoresis which allows the separation of DNA molecules of 20–1,000 kbp in length by periodically changing the direction of the electrical field. Field inversion gel electrophoresis utilizes a conventional electrophoresis chamber in which the orientation of the electric field is periodically inverted by 180°
    - CHEF (contour-clamped homogeneous electric field electrophoresis) uses a more complex electrophoresis chamber with multiple electrodes (24) to achieve highly efficient electric field conditions for separation; typically the electrophoresis apparatus reorients the DNA molecules by switching the electric fields at 120° angles
    - CHEF combined with a programmable autonomously controlled electrode gel electrophoresis (PACE) (Bio-Rad) is the most common pulsed field method used for DNA typing
  - Analysis of PFGE (pulse field gel electrophoresis) patterns is done using a software program such as BioNumerics (Applied Maths, Kortrijk, Belgium), Molecular Analyst Fingerprinting version 1.0 (Bio-Rad, Hercules, CA), or other programs that are available for the analysis of DNA fingerprint data
  - The typical phylogenic output is the dendrogram, which provides a visual representation of strain lineages and of genetic similarities and differences between groups (Fig. 29.2)

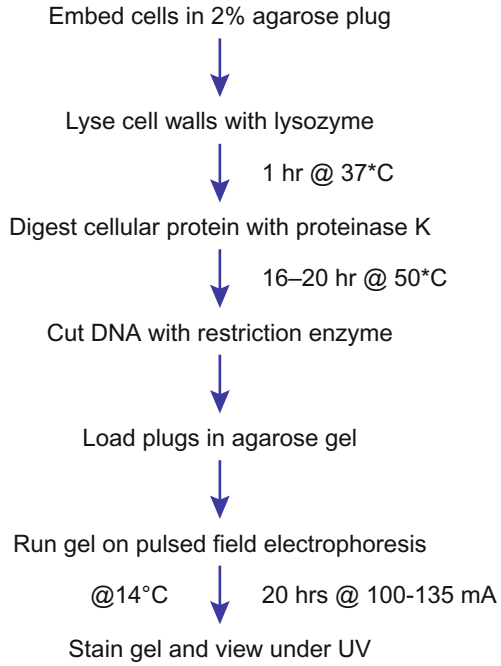


Fig. 29.3 Flow chart for bacterial genotyping using PFGP

- Procedure
  - Assay procedure (Fig. 29.3)
    - Bacterial cells are embedded in gel block
    - Cell lysis and release of intact chromosomal DNA by soaking the gel block in lysis solution, usually lysozyme, which digests cell wall
    - Restriction endonuclease digestion of chromosomal DNA in gel block
    - Gel block is mounted in agarose gel and DNA fragments are separated by PFGE at 14 °C for 22 h
    - Gel is stained by ethidium bromide
    - Analysis of DNA RFLP using a computer program
  - Genetic changes or events that alter RFLP patterns include point mutations, insertions/deletions, and rearrangements of DNA that change the number and/or locations of restriction sites (Figs. 29.4, 29.5, and 29.6)
    - Each bacterial clone will have a specific restriction profile, thus differentiate a particular clone from others. This correlation depends on the number of genetic events in bacterial DNA required to generate the observed pattern variation
    - However, random genetic events, such as point mutations or insertions and deletions of DNA that can alter the restriction profile obtained during the course of an outbreak can occur
    - Single genetic events, such as those that may alter or create a new restriction endonuclease site or change DNA fragment size by insertions/deletions, can occur unpredictably even within the time span of a well-defined outbreak (1–3 months)
  - The purpose of interpretative criteria is to establish a guide for distinguishing true differences in strains from random genetic polymorphisms that may occur over the time of a given nosocomial outbreak (Table 29.1). Appropriate interpretative criteria provide consistent, objective guidelines for correlating restriction pattern variations observed between individual isolates and the putative outbreak strain and provide an estimate of the likelihood that the isolate is part of the outbreak or "identical/related" to the outbreak strain
    - The isolates with the same banding patterns are considered as identical or clonally related
    - Three fragment differences in a band pattern could have occurred due to a single genetic event; thus these isolates are classified as closely related
    - Differences of four to five restriction fragments are likely due to two genetic events and are considered as possibly related subtypes of the same strain
    - Differences of greater than six restriction fragments are due to three or more genetic events and are considered as different or unrelated

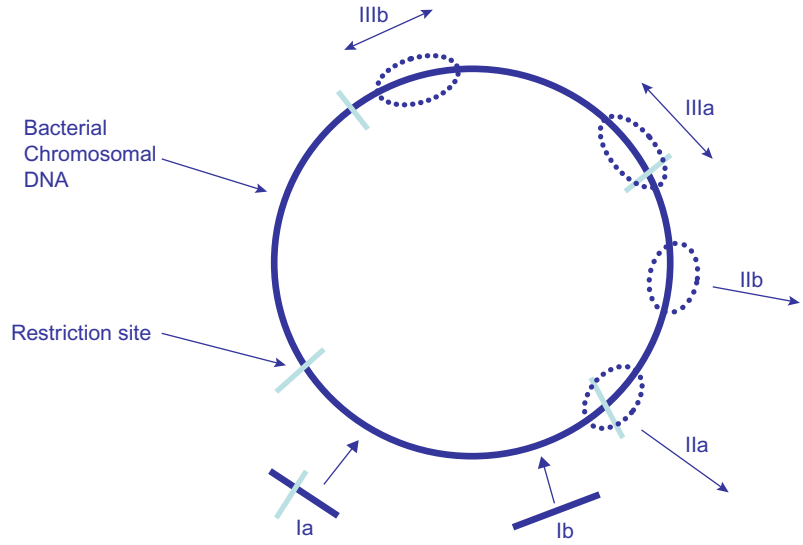


Fig. 29.4 Genetic alterations of bacterial genome which include insertions (Ia, Ib), deletions (IIa, IIb), and rearrangements (IIIa, IIIb). Some of alterations can also affect restriction sites (Ia, IIa, IIIa) that lead to the change of the length of restriction DNA fragments

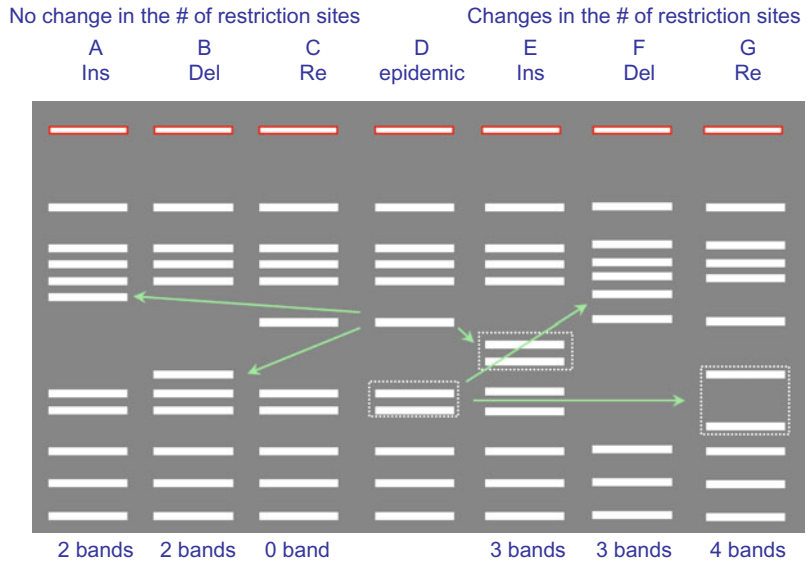


Fig. 29.5 Schematic representation of genetic alteration of restriction sites, number of DNA fragments, and DNA fragment mobility on electrophoresis

- Limitation
  - RFLP–PFGE requires large amounts of genomic DNA; the process is somewhat time-consuming and technically demanding, thus limiting the laboratory's ability to process large numbers of organisms simultaneously.

RFLP–PFGE analysis provides relatively global chromosomal overview, scanning more than 90% of the chromosome (the sum of the restriction fragment sizes), but it has only moderate sensitivity, since minor genetic changes may go undetected. The most common

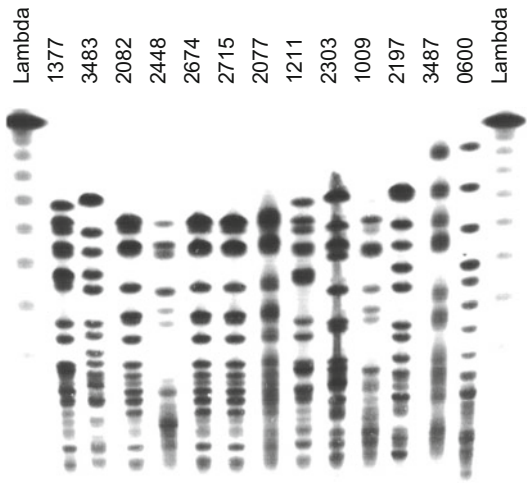


Fig. 29.6 RFLP–PFGE to detect vancomycin-resistant *Enterococcus*: Schematic representation of genetic alteration of restriction sites, number of DNA fragments, and DNA fragment mobility on electrophoresis

Table 29.1 Inherited disorders with increased risk of pancreatic neoplasia

Category	# Genetic diff	# Band diff	Interpretation
Identical	0	0	Part of the epidemic
Closely related	1	1–3	Probably part of the epidemic
Possibly related	2	4–5	Less likely to be part of the epidemic
Different	≥3	>6	Not part of the epidemic

Note: At least 10 fragments from each isolates are needed for adequate interpretation

problem associated with this assay is incomplete digestion, resulting in difficulty in the interpretation of band patterns. The complex profiles which consist of hundreds of bands that may be overlapping make it difficult to interpret

29.2.2 Ribotyping with Southern Blot Analysis

- Principle
  - Ribotyping is a method that can identify and classify microorganisms based upon differences in rRNA genes. Variations among bacteria in both the position and intensity of rRNA bands can be used for their classification and identification. Ribotyping generates a highly reproducible and precise fingerprint that can be used to classify bacteria from the genus through and beyond the species level
- Procedure
  - Bacterial DNA is digested using a frequent cutting restriction enzyme. The resulting DNA fragments are then separated by agarose gel electrophoresis and transferred (blotted) onto a nitrocellulose or nylon membrane
  - Next, a labeled (colorimetric or radioactive) oligonucleotide complementary to the target rRNA gene is used to probe the membrane
  - Under the appropriate conditions, the probe hybridizes to a complementary base pair, and the banding patterns are resolved through the detection of the probe label
  - The discriminatory power of this method is related to the copy numbers of the targeted genetic elements in the bacterial genome and their distribution among the restriction fragments following electrophoresis. Variations in the number and sizes of fragments detected are used to type the microorganisms
- Limitation
  - The limitation of ribotyping is that the discriminatory power is less than that of PFGE or some PCR-based methods. A potential benefit of ribotyping is that it can be automated, reducing the technologist time and limiting user variability

29.2.3 Typing by Polymerase Chain Reaction

- Principle
  - PCR is commonly used for typing organisms due to the ease of the assay and its good discriminatory power. Several PCR methods have been developed for this purpose
  - PCR methods can detect a difference of single nucleotide change (such as an addition or deletion) in a DNA fragment.

The variations at a single nucleotide position are far greater among different isolates; therefore, the discriminatory power is much higher
- In contrast, RFLP–PFGE analysis provides a relatively global chromosomal overview, scanning more than 90% of the chromosome (the sum of the restriction fragment sizes), but has only moderate sensitivity, since minor genetic changes may go undetected

29.2.3.1 Multiplex Polymerase Chain Reaction

- Multiple sets of primers are included in a single reaction tube to generate multiple fragments. Because the amplification products are noticeably different in their sizes, the products can be resolved on agarose gel. The band patterns can be used to discriminate the clones

29.2.3.2 Arbitrarily Primed Polymerase Chain Reaction

- Arbitrarily primed PCR (AP-PCR) and the randomly amplified polymorphic DNA assay are variations of the PCR technique in which a random primer, which is not targeted to amplify any specific bacterial DNA sequence, is used for amplification. The primers bind to target randomly, generating various length DNA fragments which are specific for the particular clone
  - Although the method is much faster than many of the other typing methods, it is much more susceptible to technical variation than most other methods. Slight variations in the reaction conditions or reagents can lead to difficulty in reproducibility of results and to differences in the band patterns generated. Therefore, trying to make comparisons among potential outbreak strains by interpretation of band patterns can be very problematic

29.2.3.3 Amplified Fragment Length Polymorphism

- Amplified fragment length polymorphism (AFLP) is a typing method that utilizes a combination of restriction enzyme digestion and PCR:
  - The DNA is digested with two different restriction endonucleases, usually chosen so that one cuts more frequently than the other does. This restriction strategy generates a large number of fragments
  - An adapter sequence where PCR primers bind is then linked to the ends of the restriction fragments. Following PCR, the reaction products are separated by gel electrophoresis, and their banding patterns can be resolved
  - The method utilizes the benefits of RFLP analysis with the increased sensitivity of PCR to generate profiles that are reproducible and relatively easy to interpret

29.2.3.4 Variable Number Tandem Repeat

- Variable number tandem repeat (VNTR) typing employs amplification of short, repetitive tandem sequences present in many bacterial genomes. The copy number of these VNTR sequences often varies among unrelated strains and can be used for genotyping. Often, fluorescently labeled PCR primers are designed to amplify the whole repeat region. Following amplification, the PCR products are separated by capillary electrophoresis and sized to determine the number of repeats present. Typically, multiple repeat regions are analyzed to determine the genotype

29.2.4 Typing by Sequencing Analysis

- Principle
  - Sequence-based molecular epidemiology is attractive in offering the promise of reproducible typing profiles that are highly amenable to standardization, uniform interpretation, and database cataloging, since they are based on simple quaternary data (A, T, G, and C)
  - Universal sequences can be used to genotype bacteria, including 16S rRNA genes, 16S–23S rRNA gene interspacer

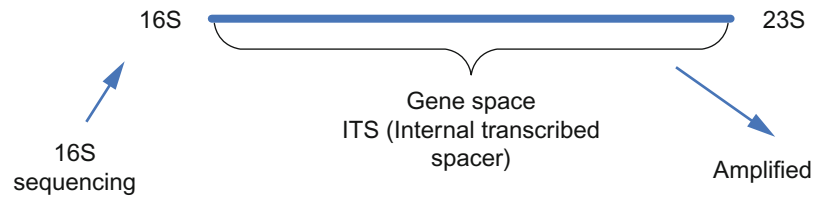


Fig. 29.7 16S–23S rDNA internal transcribed spacer

region (Fig. 29.7), and heat shock protein genes (i.e., hsp65)
- Sequence variation in a specific gene (i.e., gene for virulence, pathogenicity, drug resistance, etc.) at single nucleotide level or short repeats can be resolved by sequencing analysis. Therefore, the sequence data for specific loci/gene from different strains of the same species can be used for molecular epidemiologic application

29.2.4.1 Single Locus Sequence Typing

- The single locus sequence typing (SLST) involves analysis of a particular region of the staphylococcal protein A gene (*spa*) which is polymorphic due to 24-bp repeat sequences that may vary in both the number of repeats and the overall sequence contains in the polymorphic X or short sequence repeat region
  - *spa* typing appears to be very robust, with benefits in throughput, ease of use, and interpretation
  - It has a lower level of epidemiologic discrimination than that of established genotypic methods such as RFLP–PFGE

29.2.4.2 Multilocus Sequence Typing

- MLST utilizes a larger, and potentially more representative, portion of the genome
  - MLST compares the nucleotide sequences of internal 400–500-bp regions of a series of housekeeping genes (typically seven or more) which are present in all isolates of a particular species
  - For each gene fragment, genetic polymorphisms in sequences are considered distinct alleles. Each isolate is defined by the alleles at each of the sequenced housekeeping loci, which together comprise the allelic profile or sequence type
  - Because there are many potential alleles at each of the loci, it is unlikely that identical allelic profiles will occur by chance. Thus, isolates with the same allelic profile are assigned as members of the same clone
  - Databases containing MLST and associated data from hundreds or thousands of isolates can be accessed via the internet (http://www.mlst.net and http://pubmlst.org) (Fig. 29.8)
  - The potential limitation is its ability to track spatial spread over local scales due to the insufficient discriminatory power of the method

29.2.4.3 Whole Genome Sequencing

- Recent advancement of high-throughput deep sequencing technologies (such as Roche 454, Illumina Solexa, and ABI SOLiD) makes it possible to sequence entire microorganism genome
  - Tagged genomic libraries can be used to generate sequence data from multiple isolates in a single assay, providing sufficient information to discover single nucleotide polymorphisms (SNPs), small insertions or deletions (indels), and variation in gene content in multiple bacterial strains over a short time frame
  - The development and use of a custom SNP array (containing over 1,500 SNP loci) using the GoldenGate platform (Illumina) provided greater discriminatory power than other conventional methods
  - The sequencing and SNP typing of specific haplotypes is expected to provide additional discriminatory power of genotyping
- The advantage of sequencing-based approach is its ability of detecting phylogenetic

a

no. isolates	Sequence Type	adk	atpG	frdB	fucK	mdh	pgi	reCA
15	6	10	14	4	5	4	7	8
11	54	10	14	22	5	4	7	22
9	57	14	7	13	7	17	13	17
8	47	5	15	7	9	7	5	11
8	44	10	14	4	3	4	3	8
8	165	44	2	16	37	17	2	3
8	576	13	16	5	18	3	11	36
7	146	11	8	47	7	17	57	41
7	176	5	33	7	15	47	58	29
6	2	14	7	1	15	16	4	1
6	3	1	1	1	1	1	1	5
5	34	11	2	15	8	28	26	3
5	107	33	8	16	16	49	2	3
4	145	1	8	1	14	22	14	13
4	112	3	9	8	31	14	8	4
4	699	114	5	38	73	15	158	106
3	166	46	4	17	11	54	3	34

b

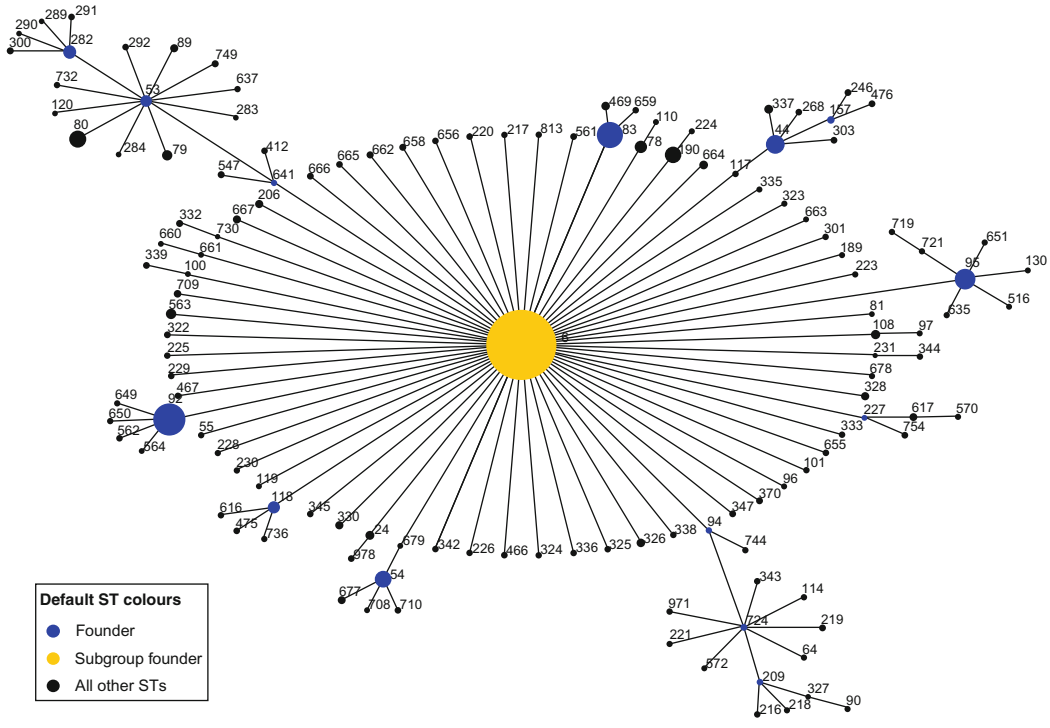


Fig. 29.8 Combining and exploring multilocus sequence typing (MLST) and geographical data for pathogen analysis with MLST maps. MLST maps (http://maps.mlst.net/view_maps.php) allow the user to enter and integrate MLST data together with location as an additional variable. (**a**) By clicking on a country of origin, the user can view the number of strains and the various sequences types (STs) in a text format. The image shows the locations in the US and the corresponding quantity of isolates of *Haemophilus influenza* with available MLST data. (**b**) MLST data for strains selected from within the MLST maps software can be sent directly to the server, allowing the user to identify groups of related genotypes in a single location or across multiple locations

informative genetic variations, therefore, offering broad coverage of sequence, enhanced sensitivity, specificity, reproducibility, and discrimination power. It also offers the better ability for special distribution analysis. Limitations include high cost for sequencing, limited throughput, requirement of extensive bioinformatics for data analysis and interpretation, and technical complexity

Further Reading

Baker S, Hanage WP, Holt KE, et al. Navigating the future of bacterial molecular epidemiology. Curr Opin Microbiol. 2010;13:640–5.

de Lencastre H, Tomasz A, et al. The CEM-NET initiative: molecular biology and epidemiology in alliance – tracking antibiotic-resistant staphylococci and pneumococci in hospitals and in the community. Int J Med Microbiol. 2011;301:623–9.

Kuroda M, Ohta T, Uchiyama I, et al. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. Lancet. 2001;357:1225–40.

MacLean D, Jones JD, Studholme DJ, et al. Application of 'next-generation' sequencing technologies to microbial genetics. Nat Rev Microbiol. 2009;7:287–96.

Nubel U, Strommenger B, Layer F, Witte W, et al. From types to trees: reconstructing the spatial spread of *Staphylococcus aureus* based on DNA variation. Int J Med Microbiol. 2011;301:614–8.

Singh A, Goering RV, Simjee S, Foley SL, et al. Application of molecular techniques to the study of hospital infection. Clin Microbiol Rev. 2006;19:512–30.

Section 5

Hematologic Malignancies

Immunoglobulin and T Cell Receptor Gene Rearrangement: Principle

30

Jeffrey E. Miller

Contents

30.1 Overview of Antigen Receptor Gene Rearrangement 825
30.1.1 Somatic Gene Rearrangement 826
30.1.2 Sequential Steps in Lymphoid Gene Rearrangement 827
30.1.3 Stages and Events During B Cell Development 828
30.1.4 Immunoglobulin Gene Rearrangement in B Cells 830
30.1.5 Stages and Events During T Cell Development 830
30.1.6 T Cell Receptor Gene Rearrangement in T Cells 831
30.1.7 Summary of Processes That Increase Diversity and Modify Effector Function 831

30.2 Clonality Testing 831
30.2.1 Specimen Types 832
30.2.2 Antibody-Based Clonality Testing 832
30.2.3 Nucleic Acid–Based Testing 833
30.2.4 Southern Blot Hybridization 833
30.2.5 Polymerase Chain Reaction 834
30.2.6 Detection Systems 836
30.2.7 Interpretative Guidelines 837

Further Reading 855

30.1 Overview of Antigen Receptor Gene Rearrangement

- Diversity of the immunoglobulin (Ig) and T cell receptor (TCR) molecules results from a series of sequential, developmentally regulated, lymphoid-specific, gene rearrangement processes
- The immune systems of jawed vertebrates can be separated into two divisions: the innate and adaptive systems
- The adaptive vertebrate immune system produces a repertoire of Ig and TCR molecules to protect the host from a constantly changing array of different pathogens using a relatively limited number of heritable germline gene segments
- Somatic gene rearrangement is the fundamental mechanism used to generate the vast repertoire of different immunoglobulin and TCR molecules, each with unique binding specificity
  - B cells produce both membrane-bound and secreted Ig or antibody molecules, which are heterodimers, composed of both heavy and light chains
  - Antibodies bind antigen independent of accessory molecules. The repertoire of antibody molecules constitutes the humoral immune response
  - T cells express only membrane bound antigen receptors, which are associated with a cytoplasmic membrane bound complex of proteins (CD3)

J.E. Miller, PhD
LabPMM, San Diego, CA, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_30,

- TCR are composed of two chains, either TCRαβ or TCRδγ heterodimers. TCRαβ molecules bind antigen that is processed and then presented by major histocompatibility complex (MHC) molecules. These antigen receptors mediate the cellular immune response
- The effector portions of both Ig and TCR molecules are the constant (C) regions, which are structurally conserved. Constant region genes are encoded in their entirety in the heritable genome
- The antigen-binding complementary determining region 3 (CDR3) of Ig and TCR must be assembled into contiguous coding units to produce a productive antigen receptor molecule. They are assembled by site-specific recombination from the library of available variable (V), diversity (D), and junction (J) gene segments that are widely separated in the germline
- These "somatic" V(D)J gene rearrangements markedly increase the number of different Ig and TCR molecules by randomly choosing and joining germline gene segments. This combinatorial diversity increases the repertoire of Ig and TCR molecules
- Enzymes expressed only during certain stages of development, and only in lymphoid cells are required to direct these somatic gene rearrangements
- Template independent, enzyme-mediated addition and/or excision of nucleotides at the junction of gene segments further increase the number of different molecules. This is termed junctional diversity
- Since Ig and TCR are heterodimers, the antigen-binding specificity is determined in part by proteins expressed from two genetic loci, which independently rearrange. This contributes additional diversity
- Somatic hypermutation (SHM) of the antigen binding regions of Ig heavy and light chain genes occurs in the germinal center. Mutation of CDR regions is coupled with a functional test of antigen binding, resulting in selection of lymphocytes encoding higher affinity receptors. This process is termed affinity maturation
- As many as 10^{12} antigen receptor molecules (Ig + TCR) can be generated starting with approximately 300 germline gene segments
- Lymphocytes undergo gene rearrangements to assemble *CDR3* coding regions that are unique in both size and DNA sequence. This is the basis for clonality testing

30.1.1 Somatic Gene Rearrangement

- Somatic gene rearrangement or recombination is a regulated stepwise process that occurs in lymphocytes in specific environments and at specific times during development. This coordinated multistep enzymatic recombination requires
  - Presence of specific DNA sequence motifs (RSSs)
  - Enzymes expressed only in lymphocytes and only during stages of antigen receptor gene rearrangement (recombination activating genes 1 and 2 [*RAG1* and *RAG2*], terminal deoxynucleotide transferase [TdT])
  - A number of ubiquitous cellular enzymes
- Recombination signal sequences (RSSs) are found immediately downstream of V segments, immediately upstream and downstream of D segments, and immediately upstream of J gene coding segments. They consist of
  - A conserved heptamer sequence 5′-CACAGTG-3′ that starts immediately after the coding gene segment
  - A spacer sequence of either 12 or 23 bases (one or two turns of DNA helix)
  - A conserved nonamer sequence 5′-ACAAAAACC-3′ at the end of the RSS
  - See Fig. 30.1
- VDJ recombination normally occurs only on the same chromosome and only between gene segments with 12 base pair and 23 base pair RSSs
  - This 12/23 rule helps prevent unproductive recombinations (e.g., V–V, D–D, J–J joining). Note: a small percentage of Ig have D–D rearrangements

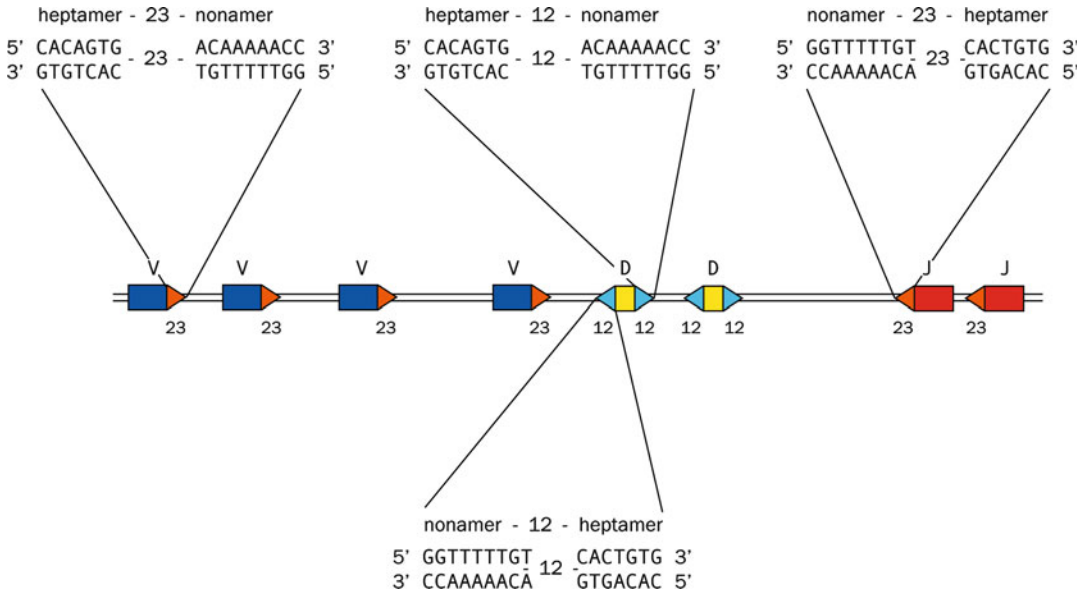


Fig. 30.1 A simplified schematic representation of the *IGH* locus showing RSSs flanking the *V*, *D*, and *J* gene segments. Note the position of the heptamer and nonamer and the length of the spacer segment flanking each gene segment. The 12/23 rule helps prevent unproductive recombinations

- *RAG1* and *RAG2* encode enzymes that are homologous to transposases. Genes encoding these enzymes lack introns and are only found in vertebrates
  - *RAG1* and *RAG2* are normally expressed only in lymphoid cells and only during periods of development when cells are rearranging the antigen receptor genes
  - They are members of the RAG complex that binds to and brings RSSs together in what is termed synapsis. This allows recombination to take place
- TdT
  - Catalyzes the addition of nucleotides at gene segment junctions. N region diversity
  - Expressed more in Pro-B cell stage when heavy chains are rearranged; expression declines in Pre-B stage during light chain rearrangement; hence, more N region diversity is seen in the heavy chain
- Ubiquitous cellular enzymes
  - Ku – Ku70:Ku80 heterodimer
  - DNA-dependent protein kinase catalytic subunit (DNA–PKcs)
  - Artemis nuclease
  - DNA ligase IV
- Enzymes correlated with primary immune deficiencies
  - RAG1, RAG2, DCLRE1C/Artemis, DNA-PKcs
  - Defects in RAG1 or RAG2 responsible for Omenn syndrome
  - Defects in Artemis produce a combined immunodeficiency, RS–SCID
  - TCR excision circle (TREC) assays can be used to screen for primary immunodeficiencies (PID)

30.1.2 Sequential Steps in Lymphoid Gene Rearrangement

- *RAG1*:*RAG2*, as part of the V(D)J recombinase complex, bind RSSs
- Following the 12:23 rule, the V(D)J recombinase complex brings RSSs together and aligns gene coding segments in what is termed "synapsis"

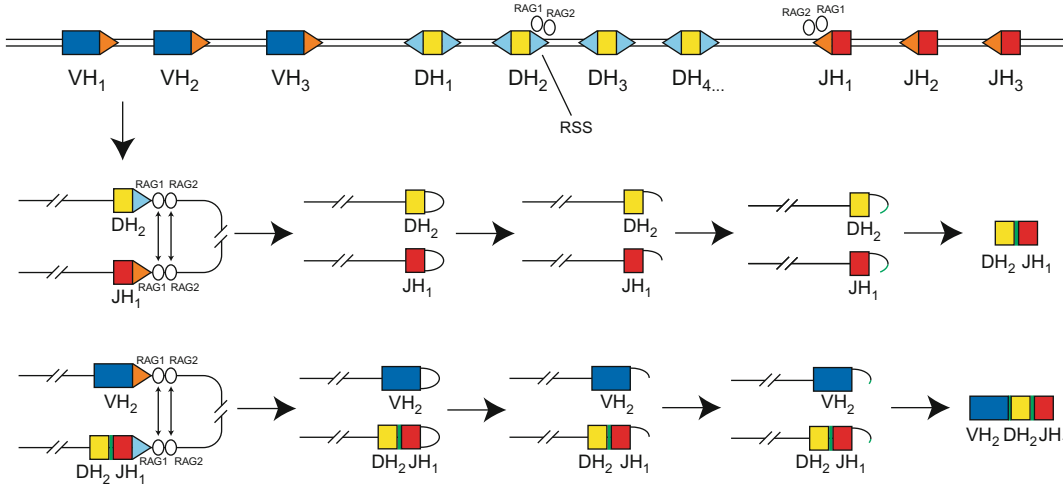


Fig. 30.2 The RAG1:RAG2 enzyme complex initiates V(D)J rearrangements binding to RSSs and bringing them together consistent with the 12/23 rule. The DNA is cleaved at the 5′ end, leaving the free 3′ end to create a hairpin loop. DNA-PK: Artemis reopens the hairpin loop, allowing TdT to add nucleotides and DNA ligase IV:XRCC4 to ligate the gene segments

- Nicking of DNA occurs at the 5′ end of each RSS heptamer leaving a free 3′-hydroxyl (−OH) at the end of each coding segment and a 5′ phosphate group at the end of the RSS. Free 3′ OH attacks opposite strand causing initial covalent closure of hairpin ends between what will become a contiguous gene coding units
- Ku, a heterodimer of Ku70:Ku80 binds to the ends
- DNA–PK: Artemis reopens the hairpin at a random site
- TdT and ubiquitous DNA repair enzymes process the ends, leading to N nucleotide additions and exonuclease subtractions
- DNA ligase IV: XRCC4 ligates ends to form a contiguous coding unit
- Next steps for the signal joints
  - Ku70:Ku80 binds ends
  - DNA ligase IV:XRCC4 ligates ends yielding precise signal joint in excision circle or integration in an inverted orientation in the chromosome
  - See Fig. 30.2

30.1.3 Stages and Events During B Cell Development

- The following antigen independent steps in B cell development occur in the bone marrow
  - Stem cell – heavy chain (*IGH*) and kappa and lambda light chain (*IGK* and *IGL*) genes are in germline configuration
  - Early pro-B cell – *IGH* undergoes D–J gene rearrangement with loss of DNA between the joined D and J segments
  - Late pro-B cell – *IGH* undergoes V–DJ rearrangement with loss of DNA between the joined V and D segments
  - Large pre-B cell – intracellular expression and transient surface expression of μ chain with invariant pseudo light chain (pre-B cell receptor). Successful cell surface expression of pre-B cell receptor triggers allelic exclusion to prevent rearrangement of second allele and also initiates pre-B cell proliferation, which results in different light chains matched with the same heavy chain rearrangement in different daughter cells

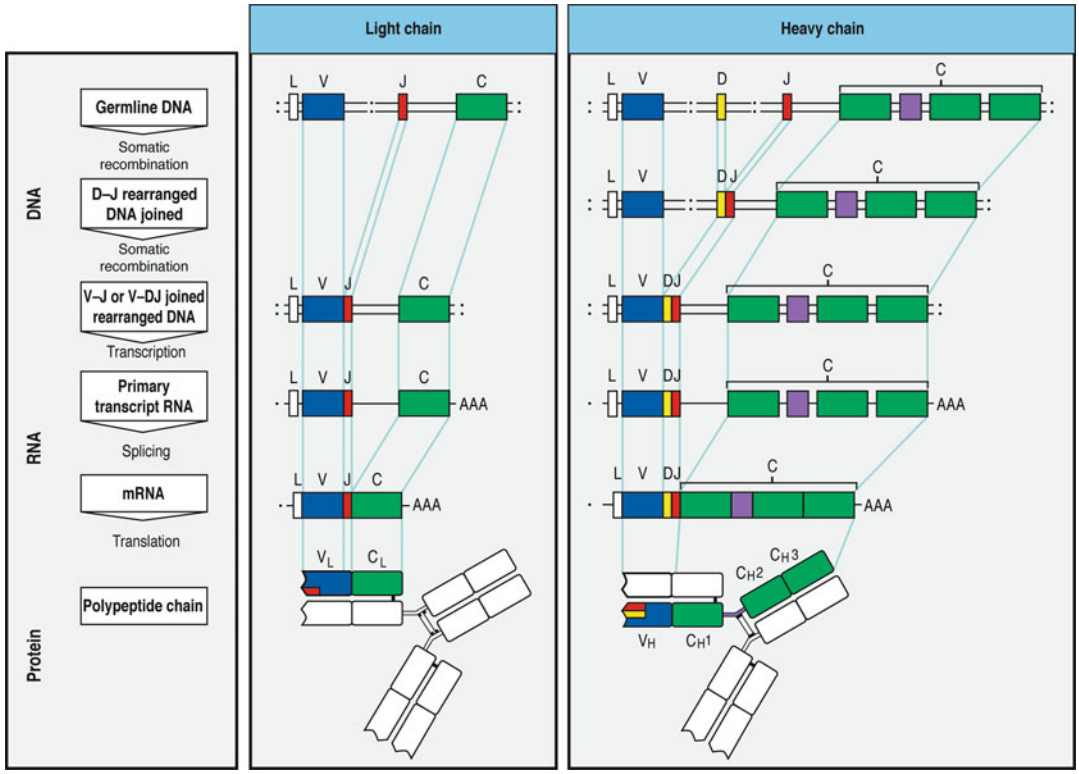


Fig. 30.3 During B cell maturation, rearrangement of the variable (V), diversity (D), and joining (J) regions in the Ig heavy chain, and V and J segments in Ig light chains, increases diversity of Ig molecules (Adapted from Murphy et al. 2008)

  - Small pre-B cell – V–J rearrangement of light chain gene(s). IgK is rearranged first then, if rearrangement of both *IGK* alleles is unsuccessful, IgL is rearranged
  - See Fig. 30.3
- The following antigen dependent steps in B cell development take place in the periphery
  - Immature B cell – IgM surface expression
  - Mature naïve B cell – IgD and IgM expressed on cell surface, made from alternatively spliced transcripts
  - Lymphoblast – alternative splicing results in secreted IgM
  - Memory B cell – isotype switch to IgG. SHM of IGH occurs in the germinal center of lymph nodes. Mutated Ig are selected for improved antigen binding in a process termed affinity maturation
  - Subset of B cell expressing surface IgD and IgM undergo a further DNA rearrangement, class switch, which results in expression of IgG, IgA, or IgE, changing the IgH constant region, and therefore effector functions associated with the Ig, without changing specificity
  - A subset of B cells survive as long-lived memory B cells, which respond rapidly to antigen and differentiate into antibody producing plasma cells
  - Terminally differentiated plasma cell – alternative splicing yields both membrane-bound and secreted Ig
- Note
  - SHM of IGH is a normal maturation process. Lack of somatic mutation in CLL is indicative of aggressive disease

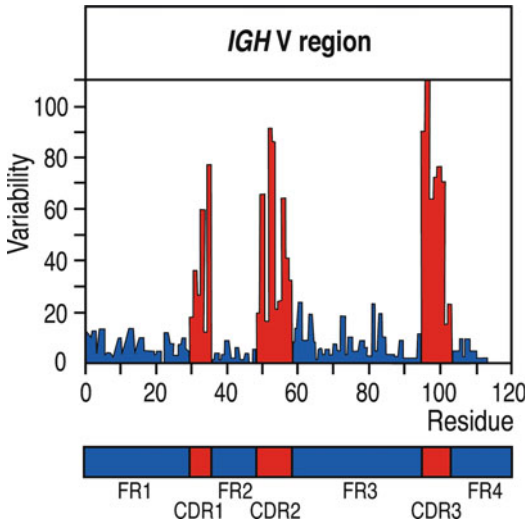


Fig. 30.4 Variability plot of amino acids encoded by the Framework (FR) and complementary determining regions (CDR) of the *IGH* V region. CDR regions exhibit greatest variability, with CDR1 and CDR2 encoded in their entirety in heritable V gene segments. The CDR3 is produced by combinatorial assembly from a library of individual V, D, and J segments (Adapted from Murphy et al. 2008)

  - V3–21 gene usage in CLL has poor prognosis, regardless of mutation status

30.1.4 Immunoglobulin Gene Rearrangement in B Cells

- The immunoglobulin heavy chain (*IGH*) locus at chromosome 14q32 contains a number of variable (V), diversity (D), joining (J) gene segments
- The *IGK* and *IGL* loci (chromosomes 2p11.2 and 22q11, respectively) are composed of V and J segments but lack D segments
- *IGH*, *IGK*, and *IGL* genes encode conserved framework regions (FR) interspersed with three hypervariable complementarity determining regions (CDR), which are involved in antigen binding
  - CDR1 and CDR2 are encoded within the germline V gene segments
  - CDR3 is the product of VDJ recombination
  - See Fig. 30.4
- *IGH* undergoes D–J joining. Intervening gene segments are spliced out and become extrachromosomal excision circles; exceptions occur when gene segments that are joined are in opposite transcriptional orientation in the germline. In these instances, intervening sequences are inverted and retained in the chromosome
- *IGH* undergoes V–DJ joining completion of recombination results in a contiguous VDJ encoding unit in genomic DNA
- Once *IGH* is rearranged, an enhancer, located in the J–C intron, is brought into proximity with the promoter elements to drive transcription
- Ig heavy chain is transcribed as L, V, D, J, intron with enhancer, Cμ, polyA
- *IGH* rearrangements are tested at the Pre-B stage to ensure they are productive (correct reading frame; no stop codons) prior to the commencement of light chain rearrangement. Functional testing on the cell surface is accomplished using an invariant surrogate light chain
- B cells go through cell division after *IGH* has rearranged but before light chains have rearranged. This further increases diversity as different light chains are paired with each functional heavy chain
- Once functionality of *IGH* product is confirmed, *IGK* rearranges (one allele, then the other, if necessary), then *IGL* rearranges (if neither *IGK* rearrangement is successful).
- A subset of B cells expressing surface IgD and IgM undergo a second round of DNA rearrangement termed class switch, which results in production of secreted IgG, IgA, and IgE

30.1.5 Stages and Events During T Cell Development

- Stem cells that give rise to T cells are present in the bone marrow
- The following antigen independent steps in T cell development occur in the thymus

  - Early double negative (CD4–, CD8–) T cells – D–J rearrangement of TCRB
  - Late double negative T cells – V–DJ rearrangement of TCRB
  - Early double positive T cells – V–J rearrangement of TCRA
- The following antigen dependent steps take place in the periphery
  - Naïve CD4 cells with proper specificity of TCR are activated by processed antigen bound to class II major histocompatibility complex (MHC) and presented by antigen presenting cells (APCs)
  - Cytotoxic CD8 cells recognize antigen presented in the context of class I MHC, present on the surface of virtually every cell in the body
  - TCRδγ cells are not MHC restricted and are found in highest abundance in the mucosa of the gut, where they contribute to a T cell population termed intraepithelial lymphocytes (IELs)
  - T cells that interact too strongly with antigen are eliminated through a process termed apoptosis, which helps protect the body against autoimmune diseases

30.1.6 T Cell Receptor Gene Rearrangement in T Cells

- Rearrangement of the TCR loci follows very similar processes and uses the same enzymes and cellular machinery used in Ig gene rearrangement
- Compared to Ig loci, the TCR loci have approximately the same number of V gene segments, but more J region gene segments
- *TCRB* and *TCRD*, similar to *IGH*, have V, D, and J gene segments
- *TCRA* and *TCRG*, similar to immunoglobulin light chain genes, have just V and J gene segments
- There are a limited number of TCR constant regions, which all encode membrane bound proteins. However, the TCR constant region genes lack diversity of effector function seen in immunoglobulin heavy chain constant region genes
- *TCRD* gene segments are located entirely within the *TCRA* locus and the orientation of *TCRA* genes ensures that rearrangement of *TCRA* results in loss of *TCRD* locus
- *TCRD* is generally rearranged first, followed by *TCRG* resulting in TCRδγ. Then *TCRB* with deletion of *TCRD*, leading to expression of *TCRA* (TCRαβ)
- *TCRab* cells do not always have rearranged *TCRG*, so *TCRD, TCRG, TCRB*, and *TCRA* rearrangement is not an obligate stepwise process

30.1.7 Summary of Processes That Increase Diversity and Modify Effector Function

- V, D, and J segments are independently joined to form the CDR3 (combinatorial diversity).
- P nucleotides and N nucleotides are added and removed randomly during D–J and V–DJ joining in a template independent manner. This imprecise joining results in junctional diversity
- In B cells, different light chain variable regions associate with each heavy chain variable region to produce different binding specificity
- Subsequent processes of selection and antigen stimulation result in either expansion or elimination of individual lymphocytes
- Activation-induced cytidine deaminase (AID) introduces C–U mutations only in germinal center activated B cells. These somatic mutations further increase diversity through sequential processes of mutation and antigen driven affinity maturation
- Class switch (isotype switching) in B cells is a mechanism that provides different effector functions to antibodies that have the same antigen specificity

30.2 Clonality Testing

- Principle: Since all leukemias and lymphomas originate from the malignant transformation of a single lymphoid cell, leukemias and

lymphomas should contain unique clonal rearrangements of one or more antigen receptor gene(s)
  - B cell malignancies should harbor clonal rearrangements of the immunoglobulin heavy and/or light chain genes (*IGH* + *IGK* and/or *IGL*)
  - T cell malignancies should harbor clonal rearrangements of two or more of the T cell receptor genes (*TCRA, TCRB, TCRD*, and *TCRG*)
  - Clonality is not equivalent to malignancy. Presence of clonality is not necessarily indicative of malignancy, as benign and reactive inflammatory conditions can generate clonal patterns due to antigen driven lymphocyte expansions
  - Diagnosis of leukemias and lymphomas must be done using a combination of all available molecular, clinical, histological, and immunological data
  - Clonal rearrangements are not lineage specific. B cell malignancies often have clonal rearrangements of *TCR* genes (clonal TCR rearrangements are often present in B-ALL). T cell malignancies can contain clonal rearrangements of the immunoglobulin genes

30.2.1 Specimen Types

- There are a number of clinical specimens collected for routine diagnostic assessment of leukemias and lymphomas
- Fresh and frozen tissue is a preferred specimen for clonality analysis
- Bone marrow aspirate – ACD, heparin, and EDTA are all acceptable anticoagulants
- Peripheral blood – ACD, heparin, and EDTA are all acceptable anticoagulants
- Fine needle aspirate biopsy
- Formalin-fixed paraffin-embedded (FFPE) tissue samples
  - Often a majority of specimens submitted for clonality analysis; especially *TCRG*
  - DNA can be fragmented, of low quality, and can present challenges for testing and interpretation
  - Factors that affect usefulness of FFPE specimens
    - Age of block material
    - Thickness
    - Fixative time
    - Fixative type (Bouin fixative is not recommended; neutral buffered fixative is recommended)
    - DNA extraction procedures
    - Presence of PCR inhibitors in extracted nucleic acid can make specimens untestable
- DNA
  - Advantages
    - Extremely stable and ubiquitous
    - Use circumvents concerns about stability
    - Allows testing of archived specimens
  - Disadvantages
    - Provides limited or no information on expression or active repertoire
- RNA
  - Advantages
    - Can identify expressed rearrangements and repertoire
  - Disadvantages
    - Extremely unstable
    - Difficult to extract from certain specimens (e.g., FFPE)
    - Testing can be more arduous and costly

30.2.2 Antibody-Based Clonality Testing

- Principle: Labeled monoclonal antibodies (MoAb) targeting Ig or TCR proteins on the cell surface are used to identify monoclonal lymphoid populations by flow cytometry or multiparameter flow cytometry
- Labeled MoAb-targeting cell surface immunoglobulin kappa and lambda light chains can identify clonality in suspected B cell lymphoproliferations. The ratio of kappa/lambda light chain in peripheral blood lies within a defined range (approximately 2:1). Ratios that lie outside the established range are indicative of clonality

- Panels of monoclonal antibodies targeting a variety of variable region TCR beta chains can identify clonality in suspected T cell lymphoproliferations, as disproportionate representation of Vβ usage is indicative of clonal cell population(s)
- Immunophenotyping studies are done first to establish lineage
- Advantages
  - Rapid and relatively inexpensive
  - Determines light chain use for subsequent monitoring
  - Identifies specific Vβ usage, which is useful for monitoring patients
- Disadvantages
  - Limited sensitivity
  - Limited specimen types amenable to flow analysis
  - Specimen stability requires timely analysis

30.2.3 Nucleic Acid–Based Testing

- Nucleic acid–based clonality tests can provide critical data necessary to discriminate between malignant and reactive lymphoproliferative processes cases where morphology, flow cytometric, and immunohistochemistry data are inconclusive
- Since leukemias and lymphomas all arise from a single malignantly transformed lymphoid cell, all leukemias and lymphomas should carry unique clonal rearrangements of one or more antigen receptor gene(s). These rearrangements are most readily identified using nucleic acid–based test methods
- In B cell malignancies clonal rearrangements should include the immunoglobulin heavy and/or light chain genes and in T cell malignancies clonal rearrangements of the *TCR gene*s
- In practice, there are inefficiencies, which reduce detection rates, as well as rearrangements that are not determinative of B cell or T cell lineage (e.g., clonal TCR rearrangements are often present in B-ALL) nucleic acid test methods are the current gold standard for detecting clonal cell populations
- Clonality does not equal malignancy, as benign and reactive inflammatory conditions can generate clonal patterns due to antigen-driven lymphocyte expansions
- Diagnosis of leukemias and lymphomas must be done using a combination of all available molecular, clinical, histological, and immunological data

30.2.4 Southern Blot Hybridization

- Principle: Gene rearrangements within the antigen receptor genes add, delete, or alter restriction enzyme sites, which change the size of the DNA fragments. Changes to the germline pattern can be identified using DNA probes to downstream gene segments that rearrange. Clonal rearrangements appear as new, nongermline bands. Value of test is related to the combinatorial repertoire of the locus
  - For many years the gold standard for clonality analysis
  - Digestion of genomic DNA with restriction enzymes followed by size fractionation and hybridization with DNA probes targeting conserved regions that rearrange. Generally J region probes are used
  - Individual gene rearrangements are not detected within the pattern unless they are overrepresented. These clonal rearrangements produce bands that differ from the germline DNA banding pattern
    - Stepwise standard protocol
    - Several aliquots of genomic DNA are digested with different restriction enzymes
    - Gel electrophoresis is used to separate digested fragments on the basis of size and the fractionated DNA is transferred to a membrane
    - Labeled DNA probes targeting joining regions hybridize to DNA on the membrane
    - The location of the probe is visualized to identify the location of germline and rearranged fragments

    - Prominent nongermline bands in two of three restriction patterns are generally required to yield a valid clonal result
    - J region probes pick up complete and incomplete rearrangements and can be used to analyze loci that have sufficient combinatorial diversity (*IGH, IGK*, and *TCRB*)
    - Sensitivity is approximately 5–10% (i.e., five to ten clonal cells in a background of 100 cells)
- Advantages
  - Comprehensive analysis is possible, including analysis of partial/incomplete rearrangements
  - Assay method and probe design allows testing of samples that have somatic mutations
- Disadvantages
  - Laborious, time consuming
  - Requires significant amount of high-quality DNA (10–20 μg), which limits testing of many specimens (e.g., FFPE tissue)
  - Requires skill to identify artifacts and technical problems (e.g., incomplete digestion)
  - Limited sensitivity (5–10% clonal population)

30.2.5 Polymerase Chain Reaction

- First described by Dr. Alexander Morley in 1990 (Trainor et al. 1990) as an improvement over Southern blot hybridization techniques
- PCR-based clonality assays are now routinely used to assist in differentiating between malignant and reactive processes in suspect lymphoproliferations
- PCR has replaced Southern blot hybridization assays as the "gold standard" for clonality assessment
- When used consistent with the appropriate guidelines and with a proven diagnostic testing strategy these standardized assays can identify clonality in approximately 99% of all B cell malignancies and 94% of all T cell malignancies (van Krieken et al. 2007)
- See Fig. 30.5
- Principle: V, D, and J segments of the antigen receptor genes are widely separated in the germline, preventing PCR amplification. Rearrangement in B cells and T cells brings V, D, and J gene coding segments together, producing a hypervariable CDR3 region that is unique in both length and sequence. This makes PCR-based clonality testing possible
- PCR tests amplify the DNA between primers targeting conserved regions of DNA sequence within the variable (V) and the joining (J) regions that lie on either side of the hypervariable antigen-binding CDR3 region, which is assembled from the multiple germline gene segments during lymphocyte development
- Since the antigen receptor genes are a population of heterogeneous DNA sequences, it is difficult to identify a set of DNA primer sequences that will target all of the conserved flanking regions around the V–J rearrangement. N region diversity and somatic mutation add further diversity to the DNA sequences in these regions. Therefore, master mixes contain a mixture of primers that target multiple V region families within framework regions (FR) to identify the majority of clonal rearrangements
- DNA from a normal population of lymphocytes generates a normal distribution of amplicon products within an expected size range. This Gaussian distribution reflects the heterogeneous population of V–J rearrangements resulting from both combinatorial and junctional diversity
- DNA containing a clonal population generates one or two prominent amplified products (amplicons) within a diminished polyclonal background
- PCR-based testing produces data that are specimen-type specific, so assay performance, interpretation, and validation should be independently determined for each specimen type. Test- and specimen-specific artifacts must be identified for each specimen type
- Optimally, primer pairs should generate products of less than 500 base pairs; less than 300 base pairs are preferred especially when testing FFPE specimens

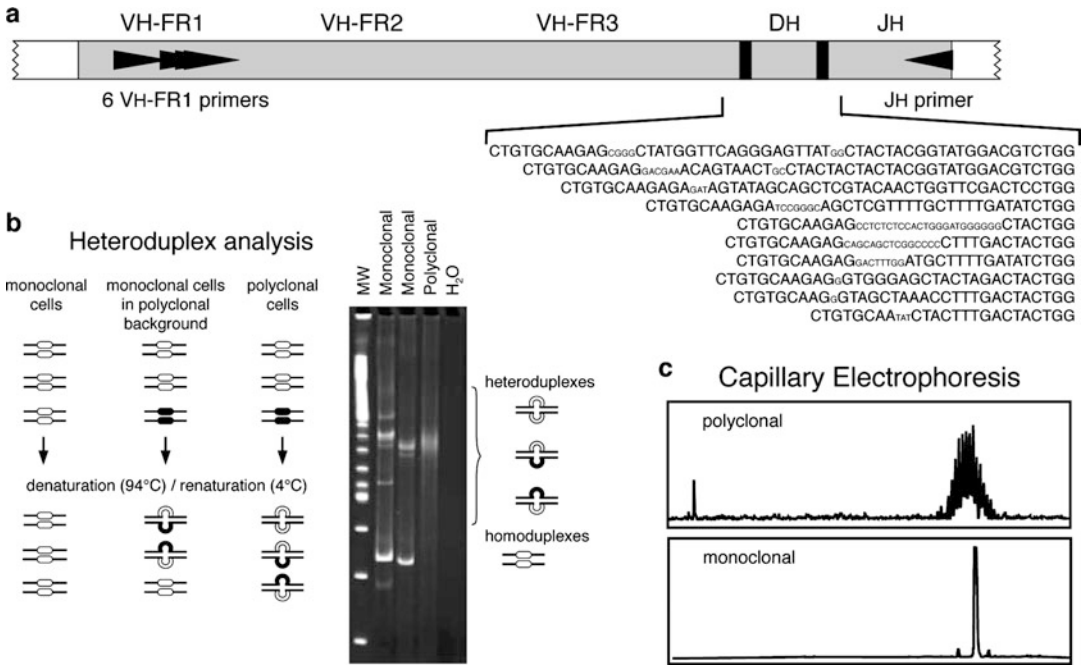


Fig. 30.5 Schematic diagram of PCR-based testing strategy and data generated testing the *IGH* locus on chromosome 14. (**a**) Primer targeting of conserved regions within the V region and J region segments showing the variability in length and sequence of amplified CDR3 region products. (**b**) Schematic representation of heteroduplex formation and data generated using heteroduplex analysis on a polyacrylamide gel. (**c**) Data generated using capillary electrophoresis (Adapted from the BIOMED-2 report; van Dongen et al. 2003)

- Capillary electrophoresis and heteroduplex analysis are preferred methods for fractionation prior to interpretation
- Recent advances include use of optimized consensus primers and standardized protocols for testing, detection, and interpretation
- Advantages
  - Comprehensive analysis is possible, including analysis of partial/incomplete rearrangements
  - Assays can be completed in a single day
  - Requires 50 ng to 1 μg of DNA, which facilitates testing of small samples, including FFPE specimens
  - Good sensitivity (1–5% clonal population)
- Disadvantages and limitations
  - False negatives due to improper or incomplete primer annealing
  - Multiplex reactions can result in cross-priming, inability to pick up specific gene family members, and generation of nonspecific products
  - False positives due to difficulty discriminating between polyclonal and monoclonal rearrangements
  - Requires skill to interpret data generated from some loci (e.g., IGK and TCRG) and some specimen types
  - Test data from several loci are needed for comprehensive clonality testing

30.2.5.1 General Rules for Polymerase Chain Reaction–Based Testing

- Tests should be conducted in a minimum of three separate work areas with control of personnel and workflow, equipment and supplies, and airflow
  - Master mix preparation and plate setup
  - Specimen extraction, preparation, and plate loading
  - Postamplification processing and data interpretation
- Test 50 ng to 1 μg of extracted DNA (cDNA)

- Quantity and volume of DNA should be validated
- Test in duplicate; interpret each result independently
- Use capillary electrophoresis and heteroduplex analysis whenever possible
- Confirm DNA (cDNA) quality and quantity was sufficient to yield a valid result by testing housekeeping genes (e.g., specimen control size ladder)

30.2.5.2 Quality Parameters for Polymerase Chain Reaction–Based Clonality Testing

- Standardized, clinically validated tests such as those developed by the BIOMED-2 Concerted Action helps ensure test reproducibility and assay performance. These tests are distributed worldwide by Invivoscribe Inc.
  - Consensus multiplex primers identified by 45 participating laboratories by iterative primer and master mix testing done over several years
  - Clinical test performance determined testing a range of well-characterized, WHO categorized, representative clinical samples
  - Defined master mix compositions, reaction volumes, thermocycler parameters, and post-PCR analysis steps
  - Clonal positive controls that generate size-specific products
  - Polyclonal controls (tonsil, peripheral blood) generate a Gaussian distribution of normal polyclonal products
  - Negative controls (HeLa or other non-lymphoid cell line)
  - Water blank control (test for contamination)
- Confirm quality of specimen DNA
  - Confirm DNA is of sufficient quality to yield a valid result
  - PCR targeting housekeeping genes generates a size ladder of amplicon products. Housekeeping products should exceed in size the upper valid size range of clonality assay
  - Housekeeping gene products confirm specimen DNA can be amplified (no inhibitors present)
- Confirm quantity of specimen DNA
  - Input quantity should be consistent with amounts used for assay validation
  - Spectrophotometric quantification of extracted DNA recommended. Pico green methods are favored over 260-nm/280-nm reading
- Test in duplicate
  - Each test result interpreted separately
- Confirm tissue representation
  - Low or absent polyclonal background is indicative of low number of total lymphocytes
  - Testing should be coupled with histological examination of FFPE and skin biopsy samples to ensure testing of adequate numbers of representative lymphocytes

30.2.6 Detection Systems

- Heteroduplex analysis
  - Limit of detection approximately 5%
  - Relies on separation based upon differences in mobility of homoduplex (clonal) versus heteroduplex (polyclonal) CDR3 regions. V–J regions containing the hypervariable CDR3 regions are amplified and the products denatured and then reannealed (gel electrophoresis/silver staining)
  - Amplified double-stranded DNA is heated to separate strands, followed by cooling to promote annealing of similar PCR products
  - Heteroduplexes generally predominate, as more dissimilar than similar complementary partners are present in the mixture
  - Homopolymers and heteroduplexes move differently through a gel, and they can be separated using gel electrophoresis
- Heteroduplex protocol
  - 10–20 μl of PCR product is denatured by heating for 5 min at 95 °C
  - PCR products are then reannealed for 60 min at 4 °C
  - 5 μl of ice-cold nondenaturing bromophenol blue loading buffer is added
  - 10–20 μl of reannealed product is loaded on precast ready to use polyacrylamide gels

(e.g., GeneGel Exel 12,5/24 Kit, Cat#: 17-6000-14, GE Healthcare)
  - Gels are run on a GenePhor (Pharmacia Biotech) electrophoresis subunit
  - Duplexes are visualized by silver staining, using the DNA silver staining kit (PlusOne™, Cat#: 17-6000-30, GE Healthcare)
- Capillary electrophoresis detects fluorescently labeled amplicon products
  - PCR products are end-labeled with fluorescent dyes during amplification
  - PCR products are denatured to produce single-stranded products; formamide and heat are used to prevent reannealing
  - Products travel along a polymer filled heated capillary past a laser that excites the fluorescent tag; emission spectra and run times are recorded, which are converted to peak amplitude/area and size
  - Instruments include: ABI 3130xl, ABI 3500 running GeneScan or GeneMapper software
  - Limit of detection approximately 1–5% clonal populations
- Denaturing gradient gel electrophoresis (DGGE)
  - Based upon differential melting of double-stranded DNA products as they migrate through a gel matrix with increasing gradient of temperature or other denaturing conditions. Sequence specific melting occurs resulting in retardation of denatured or partially denatured DNA, resulting in discrete bands that represent clonal product
- Single strand conformation polymorphism (SSCP)
  - Based upon differential migration of single-stranded DNA on a nondenaturing polyacrylamide gel. Following denaturation, amplicon products undergo intrastrand base pairing, which creates three-dimensional conformations that migrate at different rates through a gel matrix. Clonal products appear as specific bands
- Cloning and sequencing amplified products
  - Labor intensive but yields an unambiguous result and can identify patient- and tumor-specific V–J sequence for MRD testing

30.2.7 Interpretative Guidelines

- Interpretative guidelines for clonality tests are assay- and specimen-specific
- Most investigators do not support interpretation using purely objective criteria
- The following approaches have been proposed for *TCRG* assays
  - Relative peak height (RPH)
    - Lee et al. (2000)
    - 1.5–3 warrant further evaluation; >3 is clonal
  - Relative peak ratio (RPR)
    - Greiner and Rubocki (2002)
    - 2x highest peak/highest peak in polyclonal background; complicated by biallelic patterns that are often present
  - Height ratio (HR)
    - Sprouse et al. (2000)
  - Peak height ratio (PHR, or Rn)
    - Luo et al. (2001)
  - Normal distribution (ND)
    - Kuo et al. (2007)
    - Computer-aided method to evaluate significance using best fit to normal distribution and significance of peaks that occur outside background pattern
  - TCRG Algorithm (Invivoscribe)
    - Ringler et al. (2012)
    - Computer-aided method to evaluate significance using best fit to normal distribution and significance of peaks that occur outside background pattern
- Troubleshooting PCR-based clonality results
  - Bands or peaks that occur just outside the "valid size range" – accept as true rearrangement product; sequence to confirm
- Spectratyping
  - Used to monitor and measure antigen receptor repertoire diversity post lymphoid transplantation or in response to infection or immunization. Length heterogeneity is thought to be indicative of overall repertoire diversity
- Immunoscope analysis
  - Can be used to examine the antigen-specific immune response to identify both

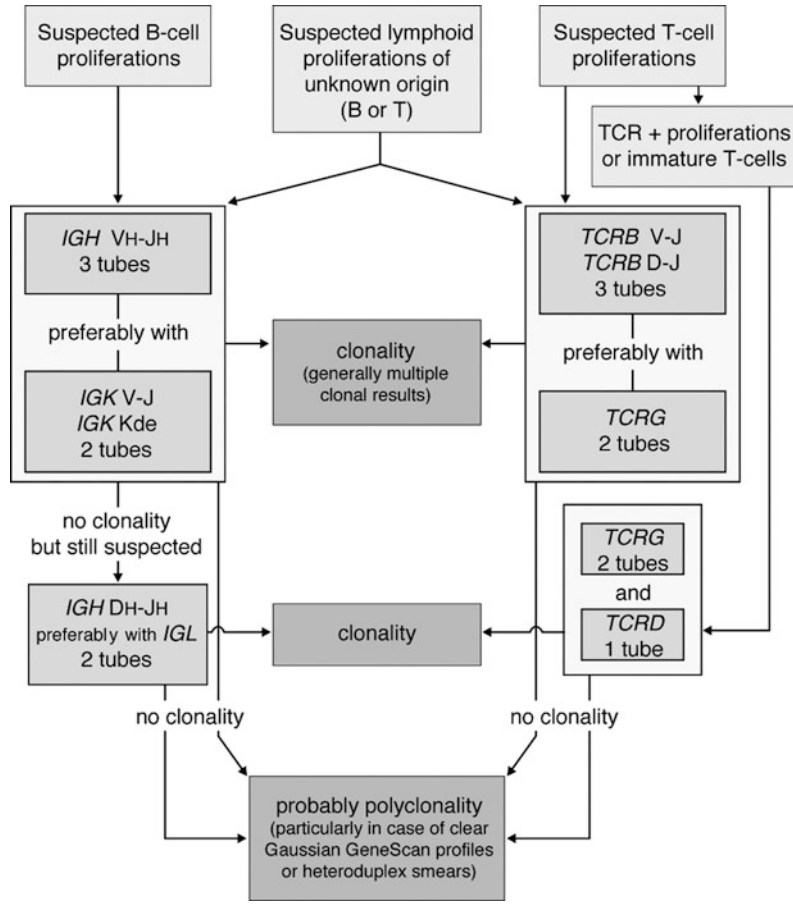


Fig. 30.6 BIOMED-2/EuroClonality testing algorithm. Strategy for PCR-based clonality diagnostics for suspect lymphoproliferations using the BIOMED-2 primer sets (Adapted from van Krieken et al. 2007). Invivoscribe, Inc., provides commercial versions of these reagents

private and public responses to an antigenic stimulus. For example, public T cells respond to an antigenic determinant in every individual and serve to drive the immune response

- Interpretation (*IGK* and *TCRG* can be challenging). Heteroduplex + capillary analyses recommended
- See Fig. 30.6

30.2.7.1 Testing Algorithm to Assist in Diagnostic Assessment

- Algorithm developed following extensive testing of a variety of clinical samples
- Primers target conserved V framework and joining regions
- Secondary targets include D and J segments to pick up incomplete rearrangements
- Reactions are composed of multiple consensus primers, each targeting a number of gene family members
- In-frame amplification products generate Gaussian distribution to assist in interpretation

30.2.7.2 Immunoglobulin Heavy Chain Gene Locus

- The immunoglobulin heavy chain (*IGH*) locus is located on chromosome 14q32.3
- Forty-six to 52 functional V segments (depending on haplotype) that comprise seven homologous gene families
- Twenty-seven functional D segments
- Six functional J segments
- Large number of gene segments resulting in extensive combinatorial diversity
- Additional diversity at both the VD and DJ junctions (junctional diversity)

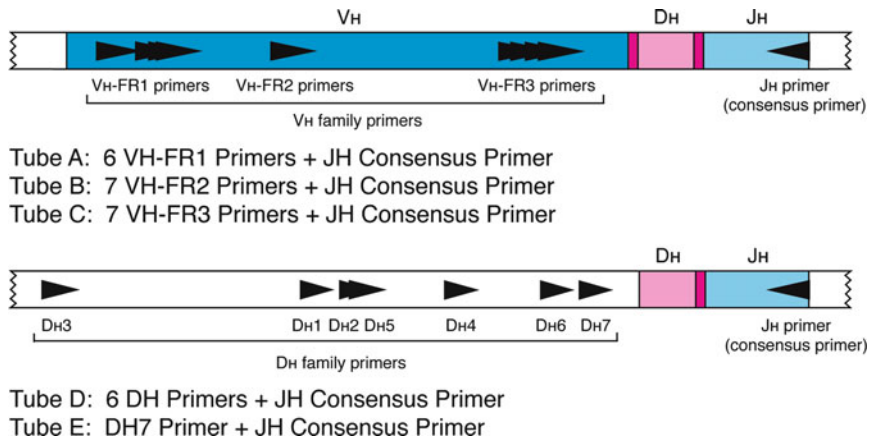


Fig. 30.7 Schematic diagram of the rearranged *IGH* gene showing the testing strategy employed using BIOMED-2 primer sets. Tubes A, B, and C target framework regions FR1, FR2, and FR3, respectively. Tubes D and E target diversity regions to identify incomplete clonal rearrangements. A consensus J region primer is used as the downstream primer in all tubes. *Black arrows* show the approximate placement of the upstream and downstream DNA primers. The numbers of primers and their specificity are listed for master mix tubes A, B, C, D, and E

- VH1, 3, 4, and 6 (in B-ALL) families constitute the vast majority of VH usage
- Antigen-binding complementary determining regions (CDR) are highly variable. These regions preferentially undergo SHM during B cell maturation in the germinal centers
- See Fig. 30.7

30.2.7.3 Immunoglobulin Heavy Chain Gene Rearrangement Assays

- Principle: PCR assays targeting the *IGH* gene are routinely used for the identification of clonal B cell populations. Each B cell has a single productive *IGH* V–J rearrangement that is unique in both length and sequence
  - IGH is a primary target for clonality assessment of suspect B cell lymphoproliferations.
  - To maximize sensitivity it is recommended testing for both *IGH* and *IGK* gene rearrangements
- Advantages
  - Well-characterized genetic locus
  - Regions of sequence homology in gene family framework regions (FR1, FR2, FR3) provide good targets for PCR-based methods
  - Six functional J segments with sufficient sequence homology to design a consensus J_H primer
  - Approximate size range of 60 nucleotides produces a characteristic Gaussian distribution of polyclonal products from peripheral blood and lymphoid tissues that aids in interpretation
- Disadvantages
  - SHM can interfere with primer targeting, especially in postgerminal center malignancies (FCL and large B cell lymphomas)
- See Fig. 30.8a, b
- *IGH* gene clonality assays amplify the DNA between conserved sequences in the variable (V) and joining (J) regions
- Primers targeting the D and J regions are used to pick up incomplete clonal rearrangements
- Normal, polyclonal lymphoid populations generate a normal distribution around a statistically favored, average-sized rearrangement. This Gaussian distribution reflects the heterogeneous population of V–J rearrangements
- One or two prominent amplified products within a diminished polyclonal background are apparent testing samples containing a clonal population

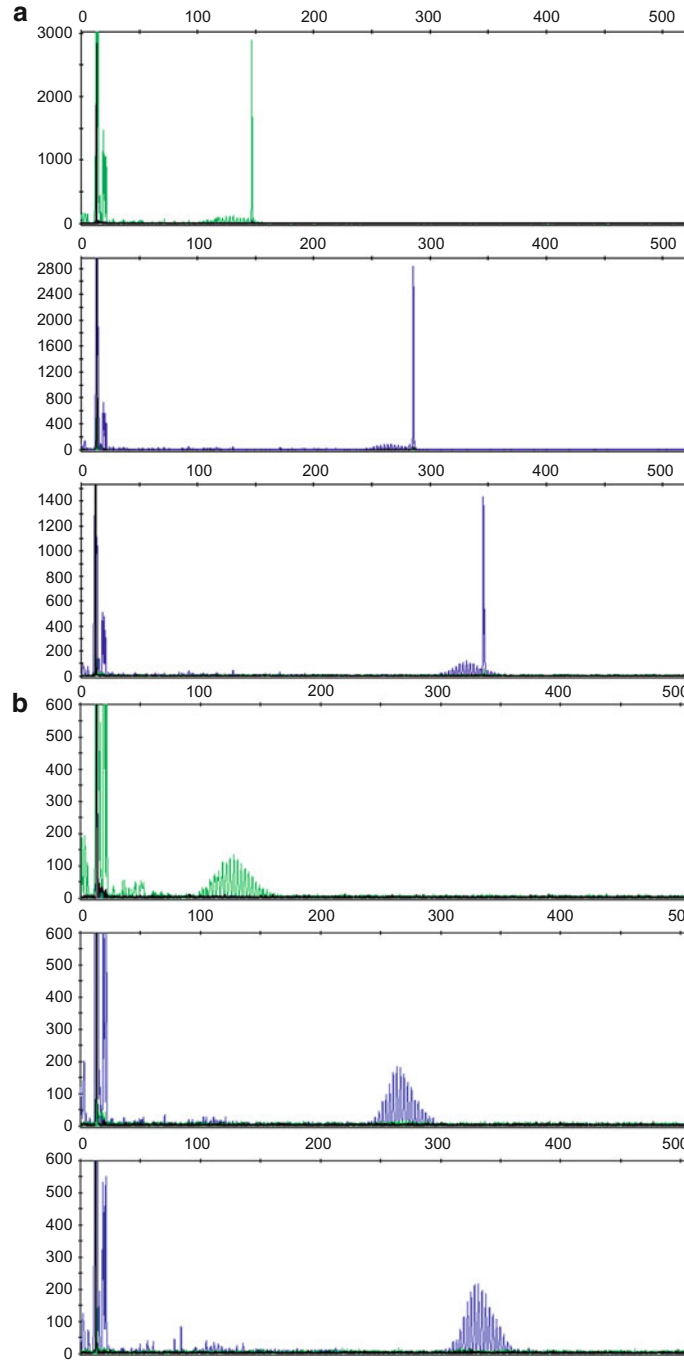


Fig. 30.8 (**a**, **b**.) Data generated using tubes A, B, and C of the BIOMED-2 *IGH* Gene Clonality Assay (IVS #1-101-0061) formatted for the ABI capillary electrophoresis instrument. (*Upper panel*) Data generated testing a clonal cell line DNA diluted to 5% into polyclonal tonsil DNA. Panel displays data generated between the FR1 and J region (*top*), FR2 and J region (*middle*), and FR3 and J region (*lower*) regions. Note that the clonal peak occurs in the same area of the Gaussian distribution when testing each framework region. (*Lower panel*) Data generated testing polyclonal tonsil DNA using the same tests described in the *Upper panel*

- Since the antigen receptor genes are polymorphic, it is difficult to employ a single set of DNA primer sequences that will target all of the conserved flanking regions around V–J rearrangement. Junctional diversity and somatic mutation add further DNA sequence diversity to these regions. So master mixes composed of primers that target conserved

sequences within the FR region of multiple related gene family members are used to identify the majority of clonal rearrangements
- To maximize detection primer sets individually targeting each of the three distinct FR regions are used to amplify the different FR regions; each produces a correspondingly different size-range of V–J products
- BIOMED-2 assay
  - *IGH* tube A – primers target conserved sequences in the FR1 region of IGH V_H1–V_H7 + J_H consensus primer
  - *IGH* tube B – primers target conserved sequences in the FR2 region of IGH V_H1–V_H7 + J_H consensus primer
  - *IGH* tube C – primers target conserved sequences in the FR3 region of IGH V_H1–V_H7 + J_H consensus primer
  - *IGH* tube D – primers target conserved sequences in the D_H region of IGH D_H1–D_H6 + J_H consensus primer (identifies incomplete clonal rearrangements)
  - *IGH* tube E – primers target conserved sequences in the D_H region of IGH D_H7 + J_H consensus primer (identifies incomplete clonal rearrangements)

30.2.7.4 Performance Characteristics

- The *IGH* PCR test is a rapid and reliable procedure that is far more sensitive than Southern blot (SB) analysis in detecting clonality in suspect lymphoproliferations
- The clinical and histopathological diagnosis correlates well with PCR results in a higher number of patients in comparison with SB results, as determined by two papers
  - van Dongen et al. (2003)
  - Sandberg et al. (2005)
    - PCR/SB concordance: IGH: 93% sensitivity/92% specificity
    - PCR/SB concordance: IGH + IGK: 85% sensitivity
- PCR versus SB analysis relative to histopathology and final diagnosis
  - PCR/SB concordance: IGH + IGK: 85%
  - PCR sensitivity: 98%
  - SB sensitivity: 39%

30.2.7.5 Minimal Residual Disease Testing

- Principle: Minimal residual disease (MRD) testing is used to detect residual disease
- Clinical outcome correlates with the level of MRD at the end of induction therapy
- Clinical value depends on technique and ability to detect the clone(s) or subclone(s) that reoccur
- Tests target patient-specific and tumor-specific DNA sequences, providing quantitative data with better sensitivity than assays used to detect clonality during initial diagnosis
- The limit of detection of IGH MRD assays is generally in the range of 1 clonal cell in 10,000 normal cells. This superior sensitivity is the result of primers that target the hypervariable DNA sequence of the CDR3 region as the tumor- and patient-specific DNA sequence of the *CDR3* region is unique for each clonal B cell population

30.2.7.6 Standard Test Development Protocol

- Clonality is confirmed testing diagnostic specimens
- Patient-specific clonal IGH rearrangements are identified, gel extracted, and sequenced to identify tumor- and patient-specific CDR3 region sequences
- CDR3 sequence information is used to design primer and probe sets
- Diagnostic samples are used to confirm performance of patient-specific tests
- See Fig. 30.9

30.2.7.7 Immunoglobulin Heavy Somatic Hypermutation Analysis

- Principle: SHM coupled with affinity maturation are normal steps in B cell development
  - Absence of SHM reflects a disruption in the regulated maturation of B cells
  - Mutations are quantified by sequencing the V region of clonal rearrangements of IGH and comparing the sequence to the corresponding germline encoded gene segments

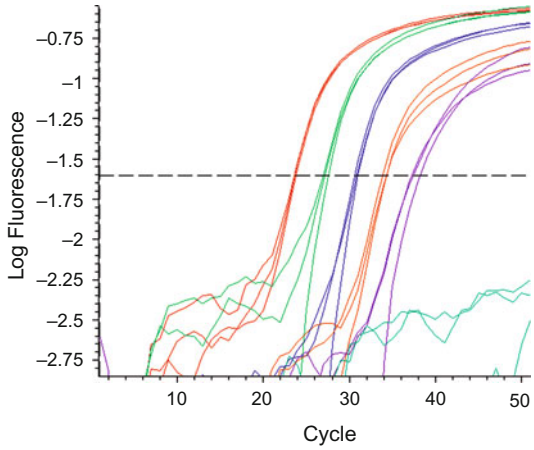


Fig. 30.9 Real-time PCR data generated using a patient-specific minimal residual disease (MRD) assay. Test was run using a patient-specific master mix with primers designed to target the unique CDR3 region *IGH* gene rearrangement identified from the initial diagnostic patient specimen. Extracted DNA from the diagnostic specimen of a leukemia patient was serially diluted into normal tonsil DNA. DNA extracted from the original diagnostic specimen and serial dilutions of the original DNA were made into normal polyclonal tonsil DNA and run in triplicate (*red*: 1x original diagnostic; *green*: 10(−1); *blue*: 10(−2); *yellow*: 10(−3); *magenta*: 10(−4)). The normal control DNA has a Ct >40 cycles. No amplification was seen testing the negative water blank (*cyan*)

  - The degree of somatic mutation in the immunoglobulin heavy (IGH) chain variable (V) genes is one of the best prognostic tools in the treatment of patients with chronic lymphocytic leukemia (CLL) and small lymphocytic lymphoma (SLL)
  - Clonal PCR products are identified, and they are gel extracted and sequenced. For SHM analysis, the full variable region (FR1–FR3) or a partial variable region (CDR1–FR3) are sequenced to determine mutational status
  - Mutational status is determined by comparing the sequence of the IGH V region of the patient sample to the most homologous germline V sequence
  - Sequences that differ by more than 2% from their corresponding germline sequences are considered highly mutated, whereas sequences that differ by less than 2% are considered unmutated
  - *IGH* V3–21 gene usage in CLL has poor prognosis, regardless of mutation status
  - See Fig. 30.10
- Listed below are several websites available to aid in IGH SHM analyses
  - IMGT – The International ImMunoGeneTics information system (initiator and coordinator: Marie-Paule Lefranc, Montpellier, France)
  - http://imgt.cines.fr
- Analysis tools: IMGT/V-QUEST and IMGT/Junction Analysis
  - V BASE – The MRC Centre for protein engineering's database of human antibody genes
  - http://vbase.mrc-cpe.cam.ac.uk/ and http://www.vbase2.org
- Analysis tools: DNAPLOT
  - NCBI – National Center for Biotechnology Information
  - http://www.ncbi.nlm.nih.gov/igblast/
- Analysis tools: IgBLAST (basic local alignment search tool)
  - For the latest recommendations for CLL hypermutation testing visit:
  - ERIC – European Research Initiative on CLL
  - http://ericll.org

30.2.7.8 Hypermutation Analysis and Reporting

- For sequences obtained by cloning, ignore vector sequences
- Align multiple sequence data obtained from a given sample and check for concordance; the majority of the sequences should be identical
- For sequences obtained by direct sequencing, adjust sequences so that they are displayed in the V to J orientation
- Check sequences to ensure they represent real V–D–J product(s) (Figs. 30.11, 30.12, and 30.13). This can be done with the aid of one or more of the websites listed below
  - IMGT – The international immunogenetics information system (initiator and coordinator Marie-Paule Lefranc, Montpellier, France); http://imgt.cines.fr

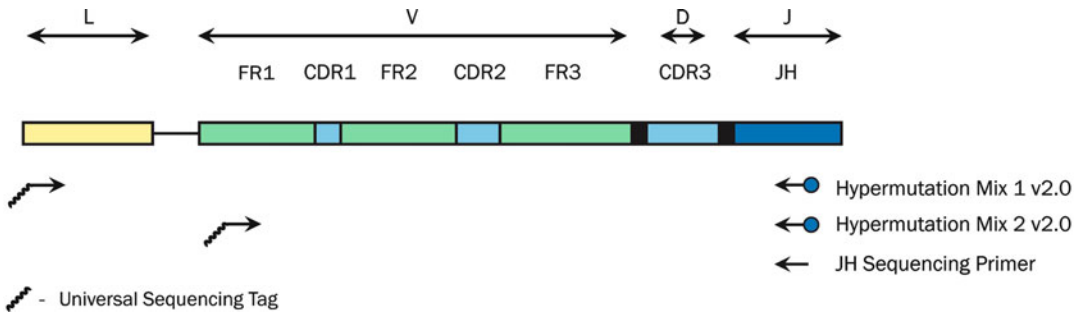


Fig. 30.10 Somatic hypermutation assay (Invivoscribe, Inc.). Schematic diagram of a rearranged *IGH* gene showing the location of primers targeting the leader, framework 1 (FR1), and joining (J) regions of the variable segment. Universal sequencing tag allows sequencing from the variable region with a single primer. The downstream primer targets the joining region consensus sequence, so a single primer for sequencing can be used. Bidirectional sequencing is required to ensure reliability of sequencing data

  - Analysis tools: IMGT/V-QUEST and IMGT/junction analysis
    - V BASE – The MRC Centre for protein engineering's database of human antibody genes; http://vbase.mrc-cpe.cam.ac.uk/ and http://www.vbase2.org
  - Analysis tools: DNAPLOT
    - NCBI – National Center for Biotechnology Information; http://www.ncbi.nlm.nih.gov/igblast/
- Analysis tools: IgBLAST (basic local alignment search tool)
  - Different databases may produce varying results and may have different amino acid numbering definitions for FR and CDR regions
  - After identifying a valid sequence, IGH SHM analysis can be done using one of the listed websites. Alternatively this can be done manually
  - If the sequence was obtained using primers that target the Leader to the J region (Hypermutation Mix 1), analysis of the full FR1–FR3 region can be completed
  - If the sequence was obtained using primers that target FR1 to the J region (Hypermutation Mix 2), analysis of only the CDR1–FR3 region can be completed
    - Find the germline V region sequence that best corresponds to the sample sequence
    - Align the germline V region sequence to the V region sequence of the sample
    - Determine the number of mismatched bases and the total number of bases that are being compared
    - The % divergence = number of mismatched bases ÷ total number of bases compared, and the % homology = 100% – % divergence
    - Complete these steps for both the forward and reverse sequences
    - If the forward and reverse sequences are discordant for mutational status, repeat the assay

30.2.7.9 Immunoglobulin Kappa Gene Locus

- The human *IGK* locus is composed of two large duplicated clusters on chromosome 2p11.2
- The upstream cluster is inverted with respect to downstream Jk segments, so inversion rearrangements are required for these upstream Vk to combine with Jk
- Seven gene families and five J regions upstream of constant region
- Other elements in the locus are involved in recombination
- Kde, located downstream of Jk-Ck can rearrange with Vk segments or with an RSS located just downstream of Jk gene segments in the Jk-Ck intron (intronRSS-Kde)

```
                                 1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21
EMBL    Locus     Name           ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Z14071  1-02      DP-75/VI-2...+ CAG GTG CAG CTG GTG CAG TCT GGG GCT GAG GTG AAG AAG CCT GGG GCC TCA GTG AAG GTC TCC
X92208  1-02      1-1+           CAG GTG CAG CTG GTG CAG TCT GGG GCT GAG GTG AAG AAG CTT GGG GCC TCA GTG AAG GTC TCC
Z12310  1-02      DP-8+          CAG GTG CAG CTG GTG CAG TCT GGG GCT GAG GTG AAG AAG CCT GGG GCC TCA GTG AAG GTC TCC
Z14213  -         VHGL1.2        CAG GTG CAG CTG GTG CAG TCT GGG GCT GAG GTG AAG AAG CCT GGG GCC TCA GTG AAG GTT TCC
X07448  1-02      V35/VI-2b+     CAG GTG CAG CTG GTG CAG TCT GGG GCT GAG GTG AAG AAG CCT GGG GCC TCA GTG AAG GTC TCC

                                                    H1
                                                                        CDR1
                                 22  23  24  25  26  27  28  29  30  31 31a 31b 32  33  34  35  36  37  38  39  40
EMBL    Locus     Name           ... ... ... ... ... ... ... ... ... GAC ... ... TAC TAT TTG CAC TGG GTG CGA CAG GCC
Z14071  1-02      DP-75/VI-2...+ TGC AAG GCT TCT GGA TAC ACC TTC ACC -G- ... ... --- --- A-- --- --- --- --- --- ---
X92208  1-02      1-1+           TGC AAG GCT TCT GGA TAC ACC TTC ACC -G- ... ... --- --- A-- --- --- --- -X- --- ---
Z12310  1-02      DP-8+          TGC AAG GCT TCT GGA TAC ACC TTC ACC -G- ... ... --- --- A-- --- --- --- --- --- ---
Z14213  -         VHGL1.2        TGC AAG GCA TCT GGA TAC ACC TTC ACC AG- ... ... --- --- A-- --- --- --- --- --- ---
X07448  1-02      V35/VI-2b+     TGC AAG GCT TCT GGA TAC ACC TTC ACC -G- ... ... --- --- A-- --- --- --- --- --- ---

                                                                                    H2
                                                                                                        CDR2
                                 41  42  43  44  45  46  47  48  49  50  51  52  52a 52b 52c 53  54  55  56  57  58
EMBL    Locus     Name           CCT GGA CAA GGG TTT GAG TGG ATG GGA TGG ATC AAC CTT ... ... AAG AAT GGT GCC ATC AGG
Z14071  1-02      DP-75/VI-2...+ --- --- --- --- C-- --- --- --- --- --- --- --- -C- ... ... --C -G- --- -G- -CA -AC
X92208  1-02      1-1+           --- --- --- --- C-- --- --- --- --- --- --- --- -C- ... ... --C -G- --- -G- -CA -AC
Z12310  1-02      DP-8+          --- --- --- --- C-- --- --- --- --- --- --- --- -C- ... ... --C -G- --- -G- -CA -AC
Z14213  -         VHGL1.2        --- --- --- --- C-- --- --- --- --- --- --- --- -C- ... ... --C -G- --- -G- -CA -AC
X07448  1-02      V35/VI-2b+     --- --- --- --- C-- --- --- --- C-- --- --- --- -C- ... ... --C -G- --- -G- -CA -AC

                                 59  60  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79
EMBL    Locus     Name           TAT GCA CAG AGG TTT GAG GAC AGG GTC ACC ATG ACC AGT GAC ACG TCC ATT AGC ACA GCC TAC
Z14071  1-02      DP-75/VI-2...+ --- --- --- -A- --- C-- -G- --- --- --- --- --- --G --- --- --- --C --- --- --- ---
X92208  1-02      1-1+           --- --- --- -A- --- C-- -G- --- --- --- --- --- --G --- --- --- --C --- --- --- ---
Z12310  1-02      DP-8+          --- --- --- -A- --- C-- -G- T-- --- --- --- --- --G --- --- --- --C --- --- --- ---
Z14213  -         VHGL1.2        --- --- --- -A- --- C-- -G- --- --- --- --- --- --G --- --- --- --C --- --- --- ---
X07448  1-02      V35/VI-2b+     --- --- --- -A- --- C-- -G- --- --- --- -GT --- --G --- --- --- --C --- --- --- ---

                                 80  81  82  82a 82b 82c 83  84  85  86  87  88  89  90  91  92  93  94  95
EMBL    Locus     Name           ATG GAG CTG GGC AGG CTG AAA TCA GAC GAC ACG GCT GTA TAT TAC TGT GCG AGA GAG
Z14071  1-02      DP-75/VI-2...+ --- --- --- A-- --- --- -G- --T --- --- --- --C --G --- --- --- --- --- --
X92208  1-02      1-1+           --- --- --- A-- --- --- -G- --T --- --- --- --C --G --- --- --- --- --- --
Z12310  1-02      DP-8+          --- --- --- A-- --- --- -G- --T --- --- --- --C --G --- --- --- --- --- --
Z14213  -         VHGL1.2        G-- --- --- A-- --- --- -G- --T --- --- --- --C --G --- --- --- --- --- --
X07448  1-02      V35/VI-2b+     --- --- --- A-- --- --- -G- --T --- --- --- -TC --G --- --- --- --- --- --
```

Fig. 30.11 The following is an example of a sequence alignment using V BASE DNAPLOT analysis tool and the sequence data of a patient sample. The patient sample was sequenced using the Hypermutation Mix 2, therefore only data from CDR1 to FR3 (amino acids 31–95) was used. DNAPLOT gives five germline sequences with the highest degree of homology. In this case the sequence has the highest degree of homology with the germline sequence from DP-75, which corresponds to VH1-02. The % divergence of this sample = 21 ÷ 206 = 10.2%

```
ID%                      <-----CDR1----->  <--------------------FWR2------------------>  <----------------------
        tmpseq_0   3     --CTACTATTTGCAC  TGGGTGCGACAGGCCCCTGGACAAGGGTTTGAGTGGATGGGA  TGGATCA-ACCTTAAGAATGGTG  79
90      VH1-2      91    GG.......A.....  ............................C..............  .......-...C...C.G.....  169
86      VH1-8      104   -------------..  ................A...........C..............  .....G.-...C...C.G....A  169
84      VH1-46     93    --.......A.....  ............................C..............  ATA....-...C..GTGG....A  169

ID%                      --CDR2----------------------->  <-------------------------------------------FWR
        tmpseq_0   80    CCATCAGG-TATGCACAGAGGTTTGAGGAC  AGGGTCACCATGACCAGTGACACGTCCATTAGCACAGCCTACATGGAGCT  158
90      VH1-2      170   G..CA.AC-..........A....C...G.  ..................G...........C.....................  248
86      VH1-8      170   A..-....C..........A...CC...G.  ..A...............GA....C.....A.....................  248
84      VH1-46     170   G..CA..C-..C.......A...CC...G.  ..A...............G..........CG.......T............  248

ID%                      3------------------------------------------>
        tmpseq_0   159   GGGCAGGCTGAAATCAGACGACACGGCTGTATATTACTGTGCGAGA   204
90      VH1-2      249   .A.........G...T...........C..G................   294
86      VH1-8      249   .A....C....G...T..G........C..G................   294
84      VH1-46     249   .A....C....G...T..G........C..G................   294
```

Fig. 30.12 The following is the same sequence analyzed in with NCBI IgBlast tool analyzing data from CDR1 to FR3 (amino acids 31–95). The sequence was found to have the highest degree of homology with VH1-02. IgBlast provides the percentage of identity (or homology) given by (ID%). For this sample, it is 90% homologous to the germline VH1-02, therefore the % divergence = 100% − 90% = 10%

```
                        <---------------------  F   R   1   -   I   M   G   T  ----------------

                        1               5                   10                  15
input                   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
X62106 IGHV1-2*02       CAG GTG CAG CTG GTG CAG TCT GGG GCT ... GAG GTG AAG AAG CCT GGG GCC TCA
X92208 IGHV1-2*03       CAG GTG CAG CTG GTG CAG TCT GGG GCT ... GAG GTG AAG AAG CTT GGG GCC TCA
Z12310 IGHV1-2*04       CAG GTG CAG CTG GTG CAG TCT GGG GCT ... GAG GTG AAG AAG CCT GGG GCC TCA
X07448 IGHV1-2*01       CAG GTG CAG CTG GTG CAG TCT GGG GCT ... GAG GTG AAG AAG CCT GGG GCC TCA
L06612 IGHV1-46*03      CAG GTG CAG CTG GTG CAG TCT GGG GCT ... GAG GTG AAG AAG CCT GGG GCC TCA

                        ------------------------------>
                                                                  CDR1 - IMGT
                            20                  25                  30                  35
input                   ... ... ... ... ... ... ... ... AGA AAC ACA TTC AGC GAC TAC TAT ... ...
X62106 IGHV1-2*02       GTG AAG GTC TCC TGC AAG GCT TCT G-- T-- --C --- -C- -G- --- --- ... ...
X92208 IGHV1-2*03       GTG AAG GTC TCC TGC AAG GCT TCT G-- T-- --C --- -C- -G- --- --- ... ...
Z12310 IGHV1-2*04       GTG AAG GTC TCC TGC AAG GCT TCT G-- T-- --C --- -C- -G- --- --- ... ...
X07448 IGHV1-2*01       GTG AAG GTC TCC TGC AAG GCT TCT G-- T-- --C --- -C- -G- --- --- ... ...
L06612 IGHV1-46*03      GTG AAG GTT TCC TGC AAG GCA TCT G-- T-- --C --- -C- AG- --- --- ... ...

                                <-------------------    F   R   2   -   I   M   G   T  --------
                        ________
                                    40                  45                  50
input                   ... ... TTG CAC TGG GTG CGA CAG GCC CCT GGA CAA GGG TTT GAG TGG ATG GGA
X62106 IGHV1-2*02       ... ... A-- --- --- --- --- --- --- --- --- --- --- C-- --- --- --- ---
X92208 IGHV1-2*03       ... ... A-- --- --- --- -X- --- --- --- --- --- --- C-- --- --- --- ---
Z12310 IGHV1-2*04       ... ... A-- --- --- --- --- --- --- --- --- --- --- C-- --- --- --- ---
X07448 IGHV1-2*01       ... ... A-- --- --- --- --- --- --- --- --- --- --- C-- --- --- --- ---
L06612 IGHV1-46*03      ... ... A-- --- --- --- --- --- --- --- --- --- --- C-- --- --- --- ---

                        -->                                                 <--------------------------
                            _____________ CDR2 - IMGT _______________
                        55                  60                  65                  70
input                   TGG ATC AAC CTT AAG AAT GGT GCC ATC ... ... AGG TAT GCA CAG AGG TTT GAG
X62106 IGHV1-2*02       --- --- --- -C- --C -G- --- -G- -CA ... ... -AC --- --- --- -A- --- C--
X92208 IGHV1-2*03       --- --- --- -C- --C -G- --- -G- -CA ... ... -AC --- --- --- -A- --- C--
Z12310 IGHV1-2*04       --- --- --- -C- --C -G- --- -G- -CA ... ... -AC --- --- --- -A- --- C--
X07448 IGHV1-2*01       C-- --- --- -C- --C -G- --- -G- -CA ... ... -AC --- --- --- -A- --- C--
L06612 IGHV1-46*03      ATA --- --- -C- -GT GG- --- AG- -CA ... ... --C --C --- --- -A- --C C-

                        ------------------------    F   R   3   -   I   M   G   T  ------------
                                75                  80                  85                  90
input                   ... GAC AGG GTC ACC ATG ACC AGT GAC ACG TCC ATT AGC ACA GCC TAC ATG GAG
X62106 IGHV1-2*02       ... -G- --- --- --- --- --- --G --- --- --- --C --- --- --- --- --- ---
X92208 IGHV1-2*03       ... -G- --- --- --- --- --- --G --- --- --- --C --- --- --- --- --- ---
Z12310 IGHV1-2*04       ... -G- T-- --- --- --- --- --G --- --- --- --C --- --- --- --- --- ---
X07448 IGHV1-2*01       ... -G- --- --- --- -GT --- --G --- --- --- --C --- --- --- --- --- ---
L06612 IGHV1-46*03      ... -G- --A --- --- --- --- --G --- --- --- -CG --- --- -T- --- --- ---

                        ------------------------------------------------------>
                                                                                        _______
                                    95                  100                 105
input                   CTG GGC AGG CTG AAA TCA GAC GAC ACG GCT GTA TAT TAC TGT GCG AGA
X62106 IGHV1-2*02       --- A-- --- --- -G- --T --- --- --- --C --G --- --- --- --- ---
X92208 IGHV1-2*03       --- A-- --- --- -G- --T --- --- --- --C --G --- --- --- --- ---
Z12310 IGHV1-2*04       --- A-- --- --- -G- --T --- --- --- --C --G --- --- --- --- ---
X07448 IGHV1-2*01       --- A-- --- --- -G- --T --- --- --- -TC --G --- --- --- --- ---
L06612 IGHV1-46*03      --- A-- --C --- -G- --T --G --- --- --C --G --- --- --- --T ---
```

Fig. 30.13 The following is the same sequence analyzed with IMGT V-QUEST tool. Only data from CDR1 to FR3 was used, however, IMGT has a different delineation for FR and CDR regions. Therefore the CDR1-FR3 region used for this analysis corresponded to amino acids 27 through 104. The sequence was found to have the highest degree of homology with VH1-02 (additional sequences are listed). The % divergence = 25 ÷ 231 = 10.8%. Data from VBase2 DNAPLOT (not shown) is similar to VBase; however, Vbase2 uses the IMGT delineations for FR and CDR regions

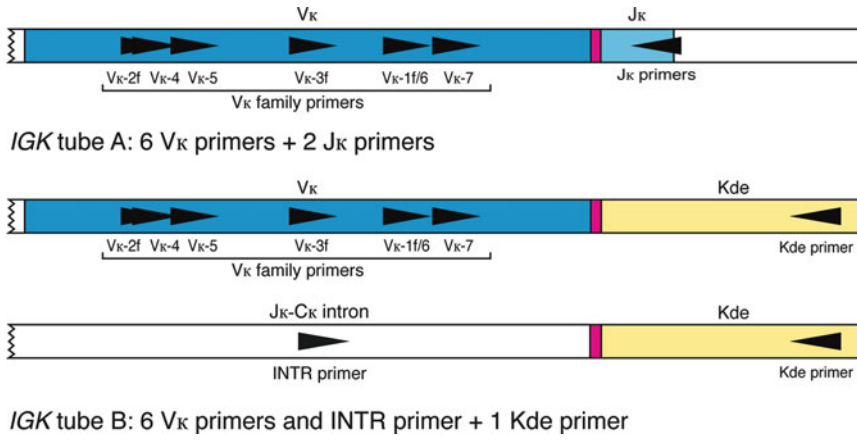


Fig. 30.14 Schematic diagram of rearranged *IGK* variable region, Vκde and Intron-κde on chromosome 2p11.2. Two multiplex master mixes target rearrangements of the kappa gene and rearrangements of the kappa deleting element. *Black arrows* show the approximate location of upstream and downstream DNA primers. The number and primer specificities are listed for *IGK* tubes A and B

- Functional inactivation of an *IGK* allele can happen by intronRSS-Kde or Vk-Kde rearrangements
- See Fig. 30.14

30.2.7.10 Immunoglobulin Kappa Gene Rearrangement Assays

- Advantages
  - *IGK* provides a clear additional value for clonality testing and complements *IGH* testing, particularly in cases where SHM may interfere with *IGH* testing
  - To maximize sensitivity, it is recommended testing for both *IGK* and *IGH* gene rearrangements
  - Rearrangements involving Kde are thought to be free of SHM of *IGH*, as it removes *IGK* enhancer that is required for hypermutation process
  - Virtually all lambda positive cells have Kde rearrangements
- Disadvantages
  - Clonality can be mistaken for oligoclonality as multiple rearrangements are possible on an allele
  - Heteroduplex analysis is slightly preferred over capillary electrophoresis for analysis
- BIOMED-2 assay
  - *IGK* tube A – primers target Vκ1f-7 + Jκ1-5
  - *IGK* tube B – primers target Vκ1f-7 + Kde and intron + Kde

30.2.7.11 Immunoglobulin Lambda Gene Locus

- The human *IGL* locus is located on chromosome 22q11.2
- IGL is composed of eight gene families
- Limited junctional diversity makes interpretation difficult

30.2.7.12 Immunoglobulin Lambda Gene Rearrangement Assays

- Testing this locus adds little to overall ability to detect clonal rearrangements

30.2.7.13 T Cell Receptor Gamma Gene Locus

- The human *TCRG* locus is located on chromosome 7p14
- Fourteen variable (V) segments + five J segments. The J segments are arranged in a duplicated cluster of joining (J) and constant region (C) segments (J1.1, J1.2, J1.3 + C1, and J2.1 + J2.3 + C2)
- Gene family plus two sets of accepted nomenclature (table; gene family, IMGT nomenclature; alternative names). IMGT nomenclature is the current accepted standard
- See Fig. 30.15

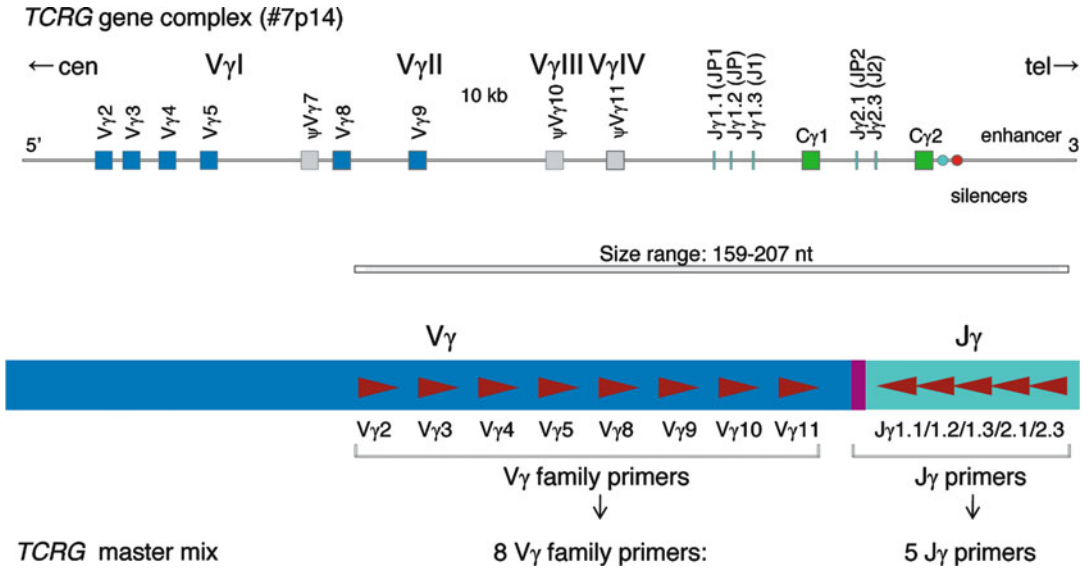


Fig. 30.15 (**a**) Schematic diagram of the human *TCRG* locus on chromosome band 7p14 showing functional V region gene segments (*blue*) and nonfunctional V region genes (*gray*). (**b**) Schematic diagram of *TCRG* rearrangement showing the eight Vγ and five Jγ primers included in the *TCRG* Gene Rearrangement Assay 2.0 (IVS# 1-207-0101), which generates a single product Gaussian distribution of 159–207 nucleotides (Adapted and modified from Groenen et al. 2008)

30.2.7.14 T Cell Receptor Gamma Gene Rearrangement Assays

- Primary target for clonality assessment of suspect T cell lymphoproliferations (see section on "Diagnostic Algorithm")
- Recommended use with *TCRB* clonality test. Detection of clonal *TCRB* and clonal *TCRG* rearrangements can confirm clonality independent of other tests
- Advantages
  - Genetic locus is well characterized and has a restricted germline repertoire, which facilitates design of PCR-based tests
  - Good target for clonality testing as *TCRG* is rearranged early, just after *TCRD*, and *TCRG* is not deleted in TCRαβ cells
  - PCR-based *TCRG* tests have been used for clonality assessment for more than a decade
  - Six V segments are expressed and nine are rearranged at a frequency that they are viable targets for PCR-based assays
  - V gene segments f1, 2, 3, 4, 5, 7, and 8 are homologous and can be grouped in one gene family (VγI). Gene segments V9, 10, and 11 each represent a single member gene family (VγII, VγIII, VγIV, respectively)
  - Assays that target all rearranging gene segments and generate products that fall under a single normal Gaussian distribution facilitate interpretation and allow testing of more standard specimen types, such as FFPE specimens
  - See Fig. 30.16
- Disadvantages
  - No D segments and limited junctional diversity results in V–J junctional length limited to 20–30 bases. This complicates interpretation
  - Minor or canonical rearrangements, most notably V9–J1.2 (AKA V9–JP) are present at a frequency of about 1% in peripheral blood and increase with age. They generate peaks that can appear clonal
  - V segments may not amplify if they are mutated or deleted (e.g., in cases of T-ALL)
  - Not appropriate for alpha/beta versus delta/gamma or B versus T cell lineage determinations. *TCRG* is not deleted in alpha/beta

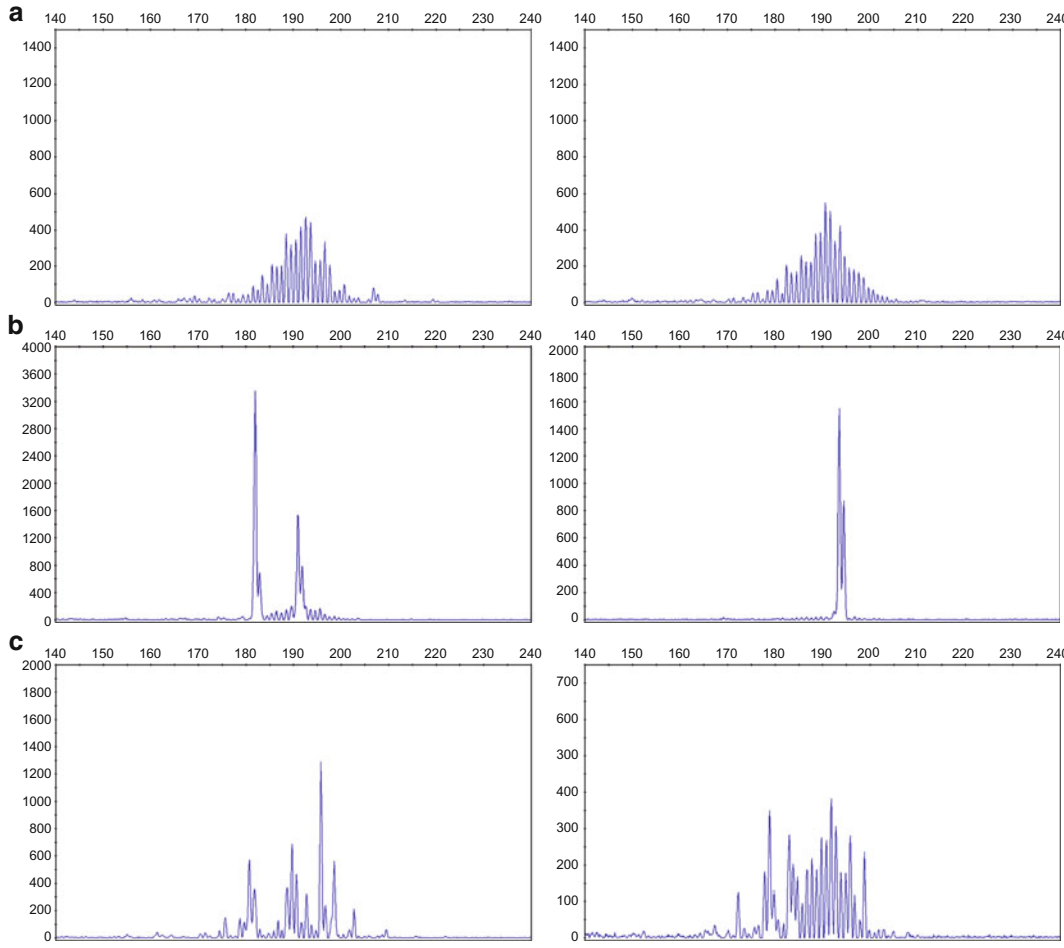


Fig. 30.16 (**a–c**) Data generated using the *TCRG* Gene Rearrangement Assay 2.0 (IVS# 1-207-0101). (*Upper panels*) These Gaussian data generated testing FFPE samples are consistent with normal lymphocyte populations. (*Middle panels*) These data generated testing FFPE samples confirm presence of clonal cell population(s). (*Lower panels*) These data generated testing FFPE samples are consistent with presence of oligoclonal cell populations

T cells, and clonal *TCRG* rearrangements are present in a majority of B-ALL
  - Interpretative criteria must be specimen specific
- BIOMED-2 *TCRG* clonality assay
  - *TCRG* tube A contains primers that target the Vγ1-8 + Vγ10 genes and Jγ1.1, Jγ1.3, Jγ2.1, and Jγ2.3 gene segments (alternative nomenclature: JγP1, Jγ1, JγP2, and Jγ2, respectively)
  - *TCRG* tube B contains primers that target the Vγ9 + Vγ11 genes and Jγ1.1, Jγ1.3, Jγ2.1, and Jγ2.3 genes
  - Jγ1.2 was not targeted due to concerns about generating peaks that can appear clonal but represent canonical rearrangements (e.g., Vγ9-Jγ1.2 seen routinely in peripheral blood)
- *TCRG* gene rearrangement assay 2.0 (IVS# 1-207-0101)
  - Single master mix contains V region primers and J region primers that target every gene segment that undergoes rearrangement
  - Products fall under a single Gaussian distribution in a smaller product size range

consistent with testing DNA extracted from FFPE specimens
  - Biallelic rearrangements are most often identified in clonal populations

30.2.7.15 T Cell Receptor Beta Gene Locus

- The human *TCRB* locus is located on chromosome 7q34
- Sixty-five V segments, and separate groupings with six and seven J segments, each group has an upstream D segment and constant region. Though the Db1 can rearrange with either Jβ1 or Jβ2 segments, the Db2 segment generally rearranges with Jβ2 and Jβ2 rearrangements are far more common than Jβ1 rearrangements
- Two different nomenclatures are used. The nomenclature described by Arden groups Vβ into 34 families. The alternative nomenclature of Rowen was adopted and is used by IMGT
- Two rearrangements on an allele are possible. Incomplete Jβ2–Dβ2 can accompany a productive or incomplete Jβ1Dβ1 rearrangement
- The diversity of rearrangements is enhanced by the addition and subtraction of an average of three to five nucleotides from gene segments during D–J and V–DJ rearrangement processes
- See Fig. 30.17

30.2.7.16 T Cell Receptor Beta Gene Rearrangement Assays

- Primary test used for purposes of testing suspected T cell proliferations
- It is recommended *TCRB* testing be done in combination with the *TCRG* assay to maximize detection of clonal rearrangements
- The complexity of the V region genes as well as consecutive D–J gene clusters long delayed design of a simple, comprehensive PCR-based approach for *TCRB* testing
- The BIOMED-2 Group was responsible for developing the first practical PCR-based approach for *TCRB* clonality testing (van Dongen et al. 2003)
- The range of Gaussian distributions is approximately 40–45 nucleotides. Although the Gaussian distributions are well beyond that detected in *TCRG* or *TCRD* rearrangements. they are less than in *IGH* CDR3 as the majority of diversity is primarily determined by a region encoding nine to ten amino acids
- Due to the extensive combinatorial repertoire and limited homology in the V segments, any manageable PCR-based test design employs multiplex primers in several independent reactions; each with multiple V and J segment or D and J segment primers
- The alternative more time-consuming approach is to use primers that target family and subfamily members, but this is not a viable approach for routine diagnostic testing
- Capillary electrophoresis based methods are not typically designed to pick up D–D rearrangements and heteroduplex analysis complements analysis using capillary electrophoresis methods
- *TCRB* rearrangements are present in the majority of mature T cell malignancies as well as a majority of T-ALLs and about a third of precursor B-ALLs
- Use of Vβ gene family members is nonrandom as certain Vβ families predominate in the peripheral blood and others in the thymus. In fact, seven Vβ gene segments account for half of the entire *TCRB* repertoire. Jβ2 segments are also represented more than twice as often as Jβ1
- *TCRB* rearrangement patterns differ between T cell malignancies
- CDR3 length as assessed by Spectratype analysis of expressed Vβ family and Vβ subfamily members is used as a surrogate for determining T cell repertoire diversity (Kepler et al. 2005)
- BIOMED-2 array
  - *TCRB* tube A identifies complete rearrangements of the TCRB locus using 23 Vβ primers and nine Jβ primers (Jβ1.1–1.6, 2.2, 2.6, and 2.7)

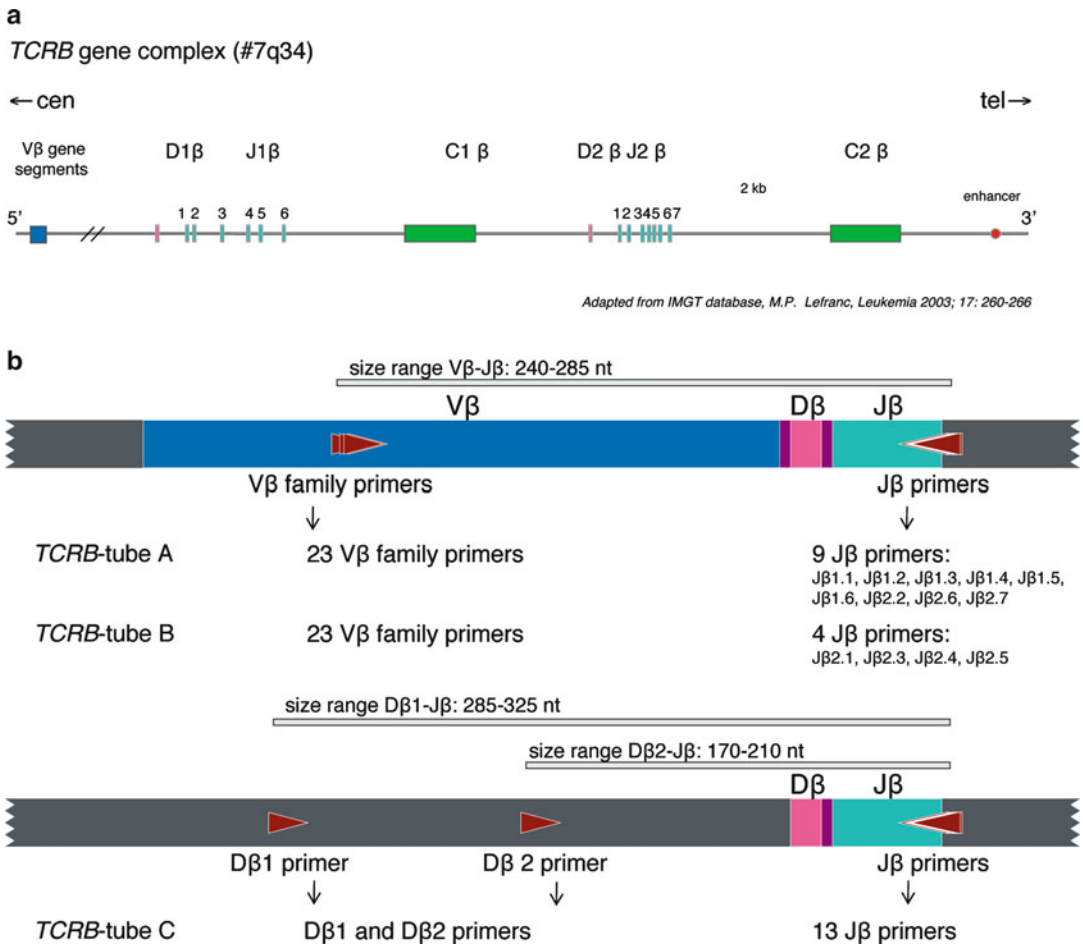


Fig. 30.17 (**a**) Schematic diagram of the human *TCRB* locus on chromosome band 7q35 showing the D1β, J1β, C1β and D2β, J2β, and C2β organization. (**b**) *Red arrows* show the approximate placement of the upstream and downstream DNA primers included in the *TCRB* Gene Clonality Assay (IVS# 1-205-0011). The numbers of Vβ, Dβ and Jβ primers as well as the expected size range of products are listed for master mix tubes A, B, and C. Tube C targets incomplete TCRβ D–J gene rearrangements (Adapted from Groenen et al. 2008)

- *TCRB* tube B identifies complete rearrangements of TCRB employing the identical 23 Vβ primers and an alternative set of four Jβ primers (Jβ2.1, 2.3, 2.4, and 2.5)
- *TCRB* tube C contains Dβ1 and Dβ2 primers with all 13 Jβ primers
- Optimal PCR assessment is obtained by combined use of capillary electrophoresis fragment analysis and heteroduplex analysis
- *TCRG* gene rearrangement assays pick up additional rearrangements that may be missed by testing TCRB alone

30.2.7.17 T Cell Receptor Delta Gene Locus

- Eight Vδ gene segments are interspersed with Vα gene segments
- There are three D segments and four J segments

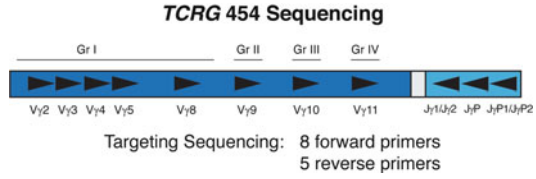


Fig. 30.18 Targeted deep sequencing. Schematic representation of the *TCRG* locus on chromosome band 7p14 showing PCR amplification primers used for generating amplicon targets for sequencing. Primers target eight V regions and five J regions that undergo rearrangement

Fig. 30.19 Primers incorporating midi sequence tags. Schematic diagram of upstream and downstream PCR amplification primers (*blue arrows*) incorporating the midi sequences (*red*) required for sample tracking and data analysis

- Intermediate *TCRD* rearrangements can be deleted during *TCRA* rearrangement processes
- Five of eight Vδ segments can rearrange with Jα genes

30.2.7.18 T Cell Receptor Delta Gene Rearrangement Assays

- Not routinely used in T cell clonality diagnostics
- Recommended only when
  - There is evidence of TCRγδ proliferation
  - There is suspicion of immature T cell neoplasm
- Advantages
  - First TCR locus to rearrange during T cell development
  - Good marker for both B cell and T cell ALL
  - Hierarchical rearrangements allow differentiation of subtypes of ALL
- Disadvantages
  - Age-related restriction of Vδ1 and Vδ2 repertoires leads to appearance of oligoclonality in blood and intestines of adults
  - Though limited number of gene segments extensive rearrangement of the locus during D–J and V–DJ joining processes result in a large size range of PCR products. This disrupts a normal Gaussian distribution and complicates interpretation using capillary electrophoresis methods. Accordingly, heteroduplex is preferable method for analysis of products
  - Existing test methods generate nonspecific bands that can appear clonal
  - This is at best a second tier assay for clonality testing

30.2.7.19 Clonality Testing Using Targeted Deep Sequencing

- Principle: The consensus primers used in clonality testing can be modified for targeted deep sequencing. This approach generates a wealth of DNA sequence data around Ig and *TCR* gene rearrangements associated with leukemias and lymphomas (see Fig. 30.18)

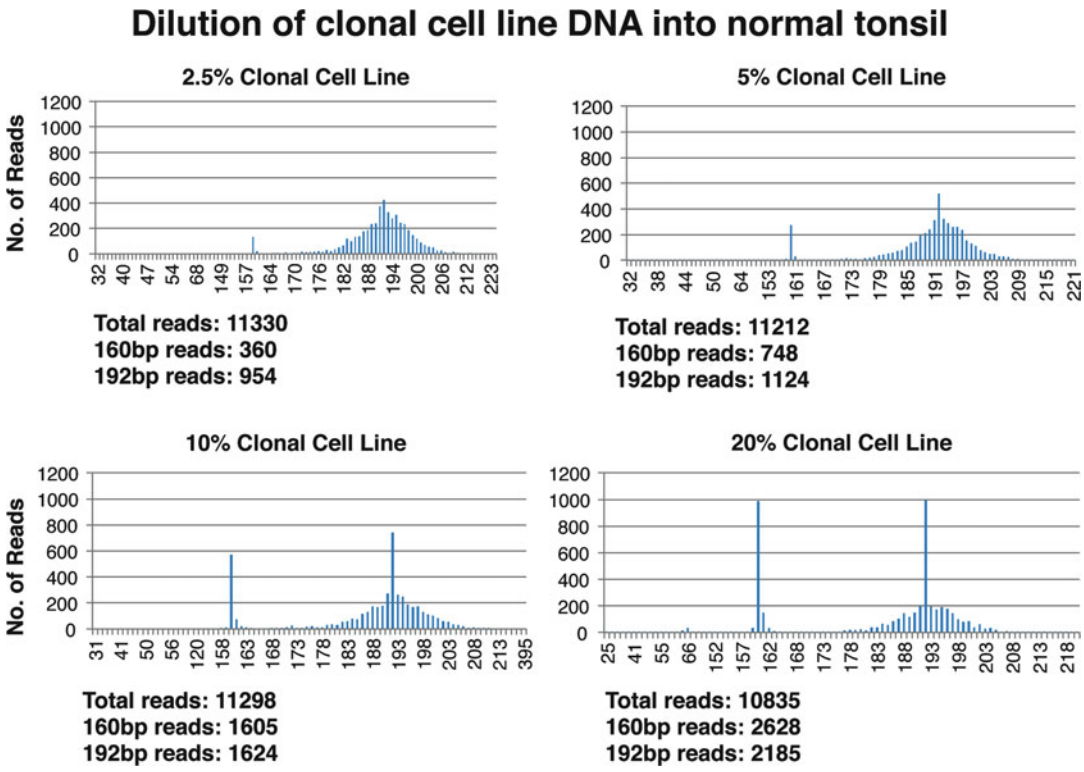


Fig. 30.20 Dilutions of a clonal positive into a polyclonal background. Roche 454 sequencing data generated testing dilutions of DNA extracted from a clonal positive cell line diluted into normal tonsil DNA. Panels indicate the number of reads versus size distribution of data generated from 2.5%, 5%, 10%, and 20% (as indicated) amplified using *TCRG* gene amplification/deep sequencing primers. Note the presence of two clonal rearrangements at approximately 160 and 192 base pairs. The number of reads for clonal peaks at 160 and 192 base pairs appears proportional to the dilution of the clonal population into the normal polyclonal DNA background

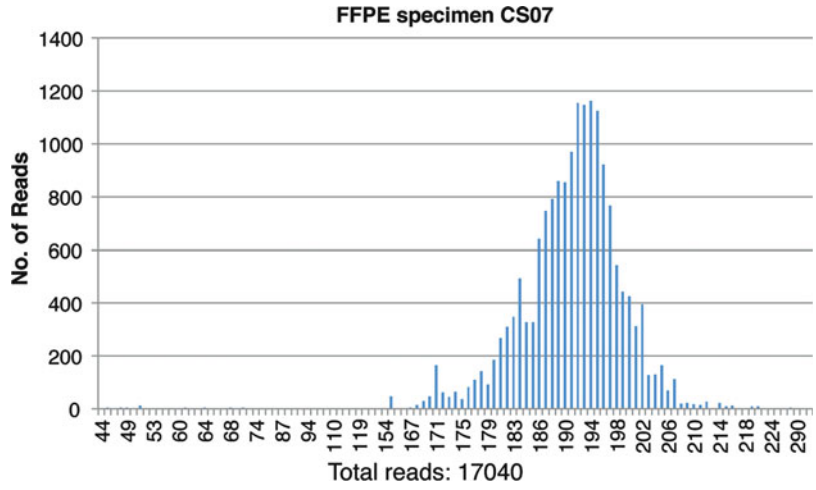


Fig. 30.21 Normal polyclonal FFPE specimen. Roche 454 sequencing data generated from DNA extracted from a FFPE specimen CS07 amplified using *TCRG* gene amplification/deep sequencing primers. Note the normal or Gaussian distribution as displayed as number of rearrangements identified when sequencing data was sorted by size. These data are consistent with the presence of a normal polyclonal population of lymphocytes

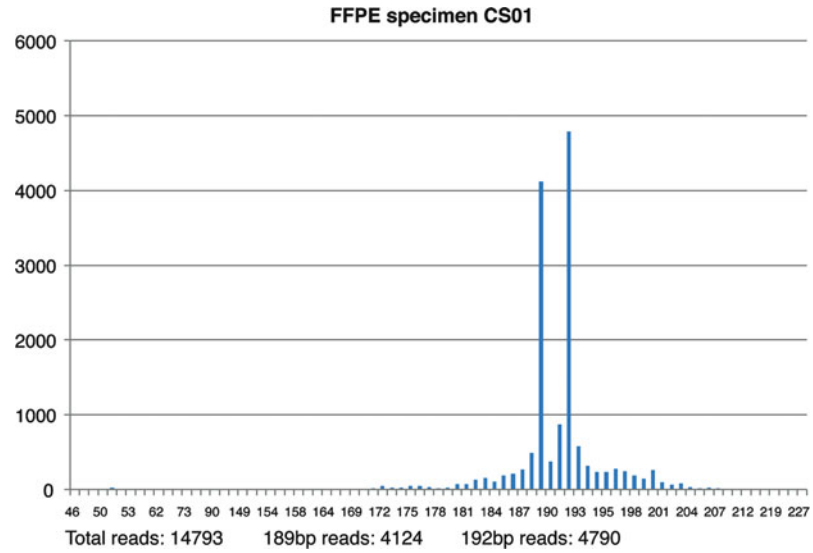


Fig. 30.22 Clonal FFPE specimen. Roche 454 sequencing data generated from DNA extracted from a FFPE specimen CS01 amplified using *TCRG* gene amplification/deep sequencing primers. Note the presence of two clonal rearrangements at approximately 189 and 192 base pairs identified when sequencing data were sorted by size. These data are consistent with the presence of a clonal population of lymphocytes

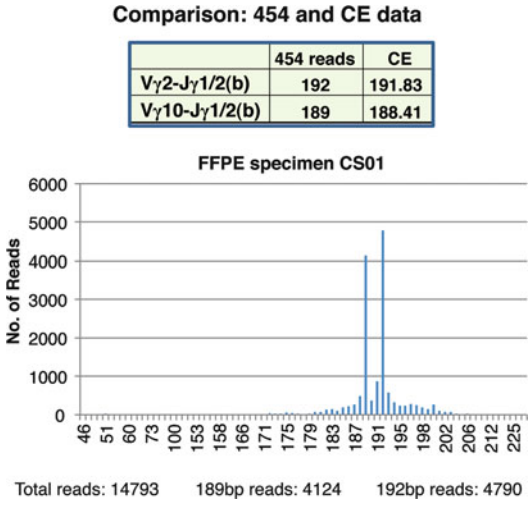
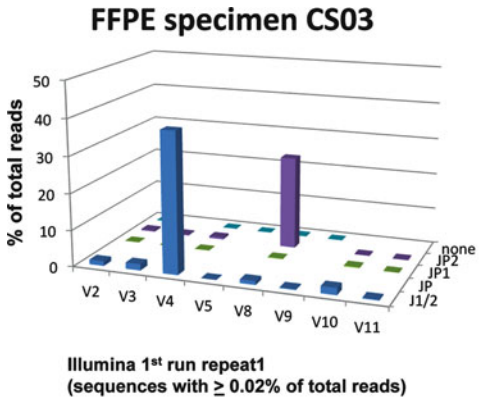


Fig. 30.23 Size comparison. 454 sequencing versus fragment analysis data testing clonal FFPE specimen. Roche 454 sequencing data generated testing DNA extracted from a FFPE specimen CS01 was compared with data generated testing the same FFPE specimen with primers from the *TCRG* gene rearrangement assay 2.0 (1-207-0101) using fragment analysis GeneMapper software and running amplicon product on the ABI 3130xl instrument. Note that both capillary electrophoresis and targeted 454 sequencing identified clonal amplicon product of 189 and 192 base pairs

30.2.7.20 Roche 454 Platform

- Approximately 10^6 DNA sequence reads can be generated
- PCR using consensus primers to conserved regions allows diagnostic testing
- Flexible plate configuration can be adjusted for depth of sequence coverage sufficient for clonality testing, repertoire analysis, or minimal residual disease testing
- Sequence tags incorporated into the sequencing primers track samples; allow filtering of data
- Patient- and tumor-specific sequences can be readily identified from the initial diagnostic sample
- Advantages
  - Sequencing identifies both tumor-specific sequences and V and J segment usage
  - Diagnostic clone and subclone sequences are readily identified and are useful for residual disease testing
- Disadvantages
  - Costly
  - Data analysis requires custom software (reagents and software available from Invivoscribe)
- Workflow
  - Patient samples are PCR amplified using consensus or gene-specific primers
  - Amplicon products are quantified and pooled
  - Library is prepared and sequenced on Roche 454 Genome Sequencer FLX System
  - Data are analyzed



	454 1st run	454 2nd run	Illumina 1st run repeat 1	Illumina 1st run repeat 2
top 10 sequences (%)	32.43	31.93	30.95	32.81
	25.05	29.45	22.2	20.81
	2.36	1.24	1.49	0.92
	0.68	0.92	0.97	0.63
	0.51	0.49	0.76	0.57
	0.37	0.48	0.55	0.55
	0.36	0.43	0.55	0.55
	0.32	0.38	0.51	0.54
	0.29	0.37	0.49	0.47
	0.27	0.36	0.49	0.38
total reads	16050	10933	62032	66371

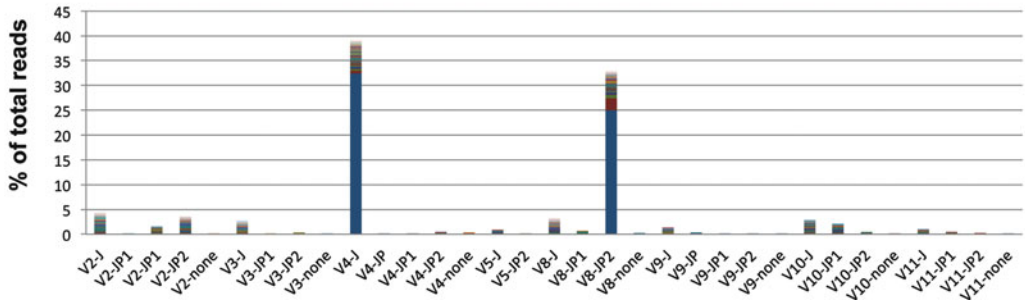


Fig. 30.24 V and J gene usage identified testing clonal FFPE specimen CS03. Roche 454 and Illumina MiSeq sequencing data generated testing DNA extracted from a FFPE specimen using *TCRG* gene amplification/deep sequencing primers. Panels show the representation of V and J specific gene segments in the sequencing data, as well as top 10 sequence percentages. Each of the solid colors in the V-J specific columns in the lower panel represent clonal populations. Note the overrepresentation of Vγ4 and Vγ8 gene members. These data are consistent with presence of a clonal population. Reagents and Bioinformatics from Invivoscribe Technologies

- Reads are identified for each sample
- Parse reads by MID sequences
- See Fig. 30.19
- Sequence amplicon size is consistent with sizing data generated by capillary electrophoresis
- Number of clonal reads is proportional to the representation of clonal cells in the total population
- See Figs. 30.20, 30.21, 30.22, and 30.23

– Establish V and J gene usage
- Align reads to reference sequence
- Reads are sorted by V and J alignment
- Number of clonal reads is proportional to the representation of clonal cells in the total population
- See Fig. 30.24

30.2.7.21 Illumina MiSeq Platform

- Approximately ten million DNA sequence reads can be generated (single reads: 12–15 million; paired-end reads: 24–30 million)
- PCR using consensus primers to conserved regions allows diagnostic testing
- Up to 96 samples can be placed in one run for high sequence coverage sufficient for clonality testing, repertoire analysis, or minimal residual disease testing
- Sequence tags incorporated into the sequencing primers track samples; allow filtering of data.
- Patient- and tumor-specific sequences can be readily identified from the initial diagnostic sample
- Advantages
  – Sequencing identifies both tumor-specific sequences and V and J segment usage

  - Diagnostic clone and subclone sequences are readily identified and are useful for residual disease testing
- Disadvantages
  - Costly
  - Data analysis requires custom software (reagents and software available from Invivoscribe)
- Workflow
  - Patient samples are PCR amplified using consensus or gene-specific primers
  - Patient-specific indexes are incorporated
  - Amplicon products are quantified and pooled
  - Library is prepared and sequenced on Illumina MiSeq System
  - Data are analyzed
    - Reads are identified for each sample
    - Parse reads by index sequences
    - Read 1 and Read 2 from paired-end reads are combined
    - See Fig. 30.19
    - Sequence amplicon size is consistent with sizing data generated by capillary electrophoresis
    - Number of clonal reads is proportional to the representation of clonal cells in the total population
    - See Figs. 30.20, 30.21, 30.22, and 30.23
  - Establish V and J gene usage
    - Align reads to reference sequence
    - Reads are sorted by V and J alignment
    - Number of clonal reads is proportional to the representation of clonal cells in the total population
    - See Fig. 30.24

Acknowledgements I would like to thank Shaun Lau for preparation of the figures, Huihong You for preparation of the 454 and Illumina sequencing data, and Mitchell Kronenberg, Jordan Thornes, and Konstantin Sidelnikov for their assistance reviewing, suggesting changes, and editing this chapter.

Further Reading

Bottaro M, Berti E, Biondi A, et al. Heteroduplex analysis of T-cell receptor gamma gene rearrangements for diagnosis and monitoring of cutaneous T-cell lymphomas. Blood. 1994;83:3271–8.

Boyd SD, Marshall EL, Merker JD, et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. Sci Transl Med. 2009;1:12ra23.

Boyd SD, Gaëta BA, Jackson KJ, et al. Individual variation in the germline Ig gene repertoire inferred from variable region gene rearrangements. J Immunol. 2010;184:6986–92.

Brüggemann M, White H, Gaulard P, Garcia-Sanz R, et al. Powerful strategy for polymerase chain reaction-based clonality assessment in T-cell malignancies. Report of the BIOMED-2 Concerted Action BHM4 CT98-3936. Leukemia. 2007;21:215–21.

Davi F, Rosenquist R, Ghia P, Belessi C, Stamatopoulos K. Determination of IGHV gene mutational status in chronic lymphocytic leukemia: bioinformatics advances meet clinical needs. Leukemia. 2008;22:212–14.

Fais F, Ghiotto F, Hashimoto S, et al. Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. J Clin Invest. 1998;102:1515–25.

Ghia P, Stamatopoulos K, Belessi C, et al. Geographic patterns and pathogenetic implications of IGHV gene usage in chronic lymphocytic leukemia: the lesson of the IGHV3-21 gene. Blood. 2005;105:1678–85.

Ghia P, Stamatopoulos K, Belessi C, et al. ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. Leukemia. 2007;21:1–3.

Goeldel AL, Cornillet-Lefebvre P, Durlach A, et al. T-cell receptor gamma gene rearrangement in cutaneous T-cell lymphoma: comparative study of polymerase chain reaction with denaturing gradient gel electrophoresis and GeneScan analysis. Br J Dermatol. 2010;162:822–9.

Groenen PJ, Langerak AW, van Dongen JJ, van Krieken JH. Pitfalls in *TCR* gene clonality testing: teaching cases. J Hematop. 2008;1(2):97–109.

Greiner TC, Rubocki RJ. Effectiveness of capillary electrophoresis using fluorescent-labeled primers in detecting T-cell receptor gamma gene rearrangements. J Mol Diagn. 2002;4:137–43.

Hodges E, Krishna MT, Pickard C, Smith JL. Diagnostic role of tests for T cell receptor (TCR) genes. J Clin Pathol. 2003;56:1–11.

Kepler TB, He M, Tomfohr JK, et al. Statistical analysis of antigen receptor spectratype data. Bioinformatics. 2005;21:3394–400.

Kneba M, Bolz I, Linke B, Bertram J, Rothaupt D, Hiddemann W. Characterization of clone-specific rearrangement T-cell receptor gamma-chain genes in lymphomas and leukemias by the polymerase chain reaction (PCR) DNA sequencing and automated high resolution PCR fragment analysis. Blood. 1994;84:574–81.

Kneba M, Bolz I, Linke B, Hiddemann W. Characterization of rearranged T-cell receptor beta-chain genes by the polymerase chain reaction and DNA sequencing. Blood. 1995;86:3930–7.

Kuo FC, Hall D, Longtine JA. A novel method for interpretation of T-cell receptor γ gene rearrangement

assay by capillary gel electrophoresis based on normal distribution. J Mol Diagn. 2007;9:12–19.

Langerak AW, Szczepanski T, Van Der Burg M, Wolvers-Tettero IL, Van Dongen JJ. Heteroduplex PCR analysis of rearranged T cell receptor genes for clonality assessment in suspect T cell proliferations. Leukemia. 1997;11:2192–9.

Langerak AW, Molina TJ, Lavender FL, et al. Polymerase chain reaction-based clonality testing in tissue samples with reactive lymphoproliferations: usefulness and pitfalls. A report of the BIOMED-2 concerted action BMH4-CT98-3936. Leukemia. 2007;21:222–9.

Lee SC, Berg KD, Racke FK, et al. Pseudo-spikes are common in histologically benign lymphoid tissues. J Mol Diagn. 2000;2:145–52.

Luo V, Lessin SR, Wilson RB, et al. Detection of clonal T-cell receptor gamma gene rearrangements using fluorescent-based PCR and automated high-resolution capillary electrophoresis. Mol Diagn. 2001;6:169–79.

Miller JE, Wilson SS, Jaye DL, Kronenberg M. An automated semi-quantitative B and T cell clonality assay. Mol Diagn. 1999;4:101–17.

Murphy K, Travers P, Walport M, et al., editors. Janeway's immunobiology. 7th ed. New York: Garland Science/Taylor and Francis; 2008.

Ria F, van den Elzen P, Madakamutil LT, Miller JE, Maverakis E, Sercarz EE. Molecular characterization of the T cell repertoire using immunoscope analysis and its possible implementation in clinical practice. Curr Mol Med. 2001;1:297–304.

Ringler RL, Shaw J, Miller JE. A comprehensive T cell receptor gamma gene rearrangement assay and interpretative algorithm for data analysis. Poster, AMP 2012 Annual Meeting on Genomic Medicine. Long Beach, CA. USA

Robins HS, Srivastava SK, Campregher PV, et al. Overlap and effective size of the human CD8+ T cell receptor repertoire. Immunology. 2010;2:47ra64.

Robins HS, Campregher PV, Srivastava SK, et al. Comprehensive assessment of T-cell α-chain diversity in αβ T cells. Blood. 2009;114:4099–107.

Rock EP, Sibbald PR, Davis MM, et al. CDR3 length in antigen-specific immune receptors. J Exp Med. 1994;179:323–8.

Sandberg Y, van Gastel-Mol EJ, Verhaaf B, Lam KH, van Dongen JJM, Langerak AW. BIOMED-2 multiplex immunoglobulin/T-cell receptor polymerase chain reaction protocols can reliably replace Southern Blot analysis in routine clonality diagnostics. J Mol Diagn. 2005;7:495–503.

Sherwood AM, Desmarais C, Livingston RJ, et al. Deep sequencing of the human TCRγ and TCRβ repertoire suggests that TCRβ rearranges after αβ and γδ T cell commitment. Sci Trans Med.. 2011;3:90ra61.

Sprouse JT, Werling R, Hanke D, et al. T-cell clonality determination using polymerase chain reaction (PCR) amplification of the T-cell receptor gamma-chain gene and capillary electrophoresis of fluorescently labeled PCR products. Am J Clin Pathol. 2000;113:838–50.

Tonegawa S. Somatic generation of antibody diversity. Nature. 1983;302:575–81.

Trainor KJ, Brisco MJ, Story CJ, Morley AA. Monoclonality in B-lymphoproliferative disorders detected at the DNA level. Blood. 1990;75:2220–2.

van Dongen JJM, Langerak AW, Bruggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene rearrangements in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia. 2003;17:2257–317.

van Krieken JH, Langerak AW, Macintyre EA, et al. Improved reliability of lymphoma diagnostics via PCR-based clonality testing: report of the BIOMED-2 concerted action BHM4-CT98-3936. Leukemia. 2007;21:201–6.

Wang C, Sanders CM, Yang Q, et al. High throughput sequencing reveals a complex pattern of dynamic interrelationships among human t-cell subsets. Proc Natl Acad Sci USA. 2010;107:1518–23.

Warren RL, Freeman JD, Zeng T, et al. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveal signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. Genome Res. 2011;21:790–7.

Lymphoblastic Leukemia and Lymphoma: Molecular Diagnostics

31

Dan Jones and Zeqiu J. Han

Contents

31.1 Stages of Lymphocyte Development 857
31.1.1 Bone Marrow Precursor B cell Maturation 857
31.1.2 Thymic Precursor T Cell Maturation 857

31.2 Clinical Features of Lymphoblastic Leukemia and Lymphoma 858
31.2.1 Disease Presentation 858
31.2.2 Patterns of Relapse and Disease Progression 858

31.3 Transforming Mechanism in Lymphoblastic Leukemia and Lymphoma 858
31.3.1 B-LBL/ALL 858
31.3.2 T-LBL/ALL 859

31.4 Diagnostic Assays 861
31.4.1 Morphologic Features 861
31.4.2 Immunologic Features 861
31.4.3 Cytogenetic and Molecular Features 862

31.5 Therapy and Response Predictors 862
31.5.1 Standard and Experimental Therapies 862
31.5.2 Prognostic Factors 863
31.5.3 Pharmacogenomics for Therapy Personalization 863
31.5.4 Monitoring After Therapy 863

31.6 Summary 864

Further Reading 864

D. Jones, MD, PhD (✉)
Medical Director, Cancer Diagnostics Services, Quest Diagnostics Nichols Institute, Chantilly, VA, USA

Z.J. Han, MD
Department of Pathology, Quest Diagnostics Nichols Institute, Chantilly, VA, USA

31.1 Stages of Lymphocyte Development

31.1.1 Bone Marrow Precursor B cell Maturation

- Normal B cells arise as precursors in the bone marrow called lymphoblasts that migrate through a number of phenotypically distinguishable stages that have been historically described as pre–pro-, pro-, and pre-B cell stages but are now more precisely defined by sequential expression of lineage-specific commitment factors
- A critical event in shift from precursor to mature B cells is surface expression of the B cell receptor (BCR) comprised of immunoglobulin heavy chain and one of two types of immunoglobulin light chain (IGK or IGL)
- B cells with intact/functional BCRs then migrate into the peripheral blood as long-lived mature forms, a process that is largely completed in childhood

31.1.2 Thymic Precursor T Cell Maturation

- The thymus has several different microanatomic compartments where immunophenotypic shifts reflecting functional changes occur on immature T cells ("thymocytes")

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_31,

- The most immature CD2+ CD3− CD4− CD7 + CD8− ("double-negative") thymocytes are derived from precursors that migrate from the bone marrow to the corticomedullary junction
- Upon migration into the thymic cortex, CD1a+ CD3+ CD4+ CD8+ ("double-positive") T cells undergo sequential and programmed rearrangement of their TCR genes
  - Rearrangement of the TCR gamma chain (TCRG) locus is followed by TCR delta, then TCR beta, and then TCR alpha
  - This DNA rearrangement process ceases when a functional and properly folded TCR of one particular antigen specificity is expressed on the surface of each thymocyte
- In the cortex and subsequently upon migration to the thymic medulla, the TCR on each thymocyte is tested for the ability to signal properly (termed "positive selection")
- T cells are then eliminated if their TCR binds to endogenous antigens too avidly or signals without antigen binding ("negative selection")
- CD4+ CD8− or CD4− CD8+ "single-positive" thymocytes exit into the peripheral circulation

31.2 Clinical Features of Lymphoblastic Leukemia and Lymphoma

31.2.1 Disease Presentation

- Acute lymphoblastic leukemia (ALL) and its lymph node/tissue-based counterpart lymphoblastic lymphoma (LBL) are among the most common malignancies of childhood but are rare in adulthood
- ALL/LBL is distinguished from mature lymphoid malignancies based on its morphologic appearance (blastic chromatin), immunophenotype (expression of TdT), and the genetic alterations described below
- Leukemias with a precursor B cell phenotype comprise 80% of ALL
  - In children, B-ALL/LBL arise de novo, with most cases presenting before age 10
  - In adults, lymphoid blast crisis due to transformation of an underlying myeloproliferative neoplasm, particularly chronic myelogenous leukemia (CML), is common
  - A localized cutaneous variant of B-LBL occurs in adults
  - Burkitt lymphoma, previously grouped with precursor ALL, is now regarded as a mature B cell neoplasm and is covered in Chap. 44
- Leukemias with precursor T cell phenotype (15–20% of ALL) more frequently occur in older children and adolescents
  - Most present with a mediastinal mass with variable marrow involvement

31.2.2 Patterns of Relapse and Disease Progression

- Current therapies for pediatric ALL/LBL are quite effective in inducing remission
- In contrast, remissions in most cases of adult ALL/LBL are of short duration with overt relapses in blood, bone marrow, lymph nodes, and tissues common
- In children, relapses often occur in immune-privileged sites such as the central nervous system and testes where optimal chemotherapy levels often cannot be achieved

31.3 Transforming Mechanism in Lymphoblastic Leukemia and Lymphoma

31.3.1 B-LBL/ALL

- The initial stage of transformation in B-ALL/LBL involves lymphoid maturation arrest
- In most cases, this occurs through chromosomal fusions or overexpression of B-cell transcriptional factor such as RUNX1 (Fig. 31.1) that act on many B cell-related differentiation genes simultaneously
  - Most common characteristic chromosomal translocations in B-ALL/LBL, except for t(9;22), drive this maturation arrest (Table 31.1)

Table 31.1 Common cytogenetic alterations in B-ALL/LBL

Chromosomal abnormality	Genetic alteration	Frequency, percent of total B-ALL	Phenotypic correlations	Prognosis
t(12;21) (p13;q22)	ETV6-RUNX1 (TEL-AML1), often cryptic	C: 20–25% A: Rare	CD25+, often CD13+	Good
t(9;22)(q34;q 11.2)	BCR-ABL1	C: 2–4% A: 25%	Often CD13+, CD33+	Poor
t(v; 11q23), commonly t(4;11)(q21;q23)	MLL gene fusions	Infants <1 year old, older adults	Often CD10-, CD24-, CD15+	Poor
t(5;14)(q31;q32)	IL3-IGH	Rare	Peripheral eosinophilia	Poor
t(l;19)(q23;p13.3)	TCF3-PBX1 (E2A-PBX1), often add1q	C: 5–6% A: 3%	mu-chain+, CD34–	Intermediate
t(7;9)(q11.2;p13.2)	PAX5/AUTS2	C: 2%,		Intermediate
t(17;19)(q22;p13)	TCF3/HLF	Uncommon		Poor
IGH enhancer-driven t(14;19)(q32;q13) t(14;19)(q32;q13) t(8;14)(q11;q32) t(14;14)(q11;q32) t(14;20)(q32;p13) t(6;14)(p22;q32) t(X;14)(p22;q32) t(8;14)(q34;q32)	IGH/CEBPA IGH/EPOR CEBPD/IGH CEBPE/IGH IGH/CEBPB ID4/IGH CRLF2/IGH MYC/IGH	C:5–7% A: 8%	Variable	Intermediate, except MYC which is poor
Hyperdiploid	>50 chromosomes	C: 25–30% A: rare	None	Good Excellent if "triple trisomy" (i.e., +4, +10, +17)
Hypodiploid	<46 chromosomes	C: 5%	None	Intermediate for 45 chromosomes; worse prognosis if <45 chr

Abbreviations: *C* children, *A* adults

  - MLL gene fusions (v;11q23) appear to transform by altering the epigenetic state of many different genes and thus regulate proliferation and cell differentiation simultaneously (Fig. 31.2)
  - In t(9;22)+, hyperdiploid, hypodiploid, and normal diploid karyotype B-ALL/LBL, somatic alterations in lineage commitment genes such as IKZF1 likely substitute for the effects of oncogenic fusion proteins (Table 31.2)
- Subsequent stages of transformation in B-ALL/LBL including mutations in growth promoting oncogenes such as KRAS, JAK1, JAK2 (Table 31.2) and/or loss of cell cycle regulators such as CDKN2A/p16 (Fig. 31.1c)

31.3.2 T-LBL/ALL

- The initial stage of transformation in T-ALL/LBL also appears to be maturation arrest
  - Chromosomal translocations that produce overexpression of T cell regulatory factors LYL1 (early thymocytes), HOX11 (early cortical thymocyte), and TAL1 (late cortical thymocyte) juxtapose these genes to the TCRA/TCRD or TCRG T cell-specific transcriptional enhancers
  - Duplication/amplification of the global transcription regulator MYB occurs in 5–10%
- NOTCH1 growth regulatory mutations are seen in over 50% of T-ALL/LBL and occur largely independent of cytogenetic changes

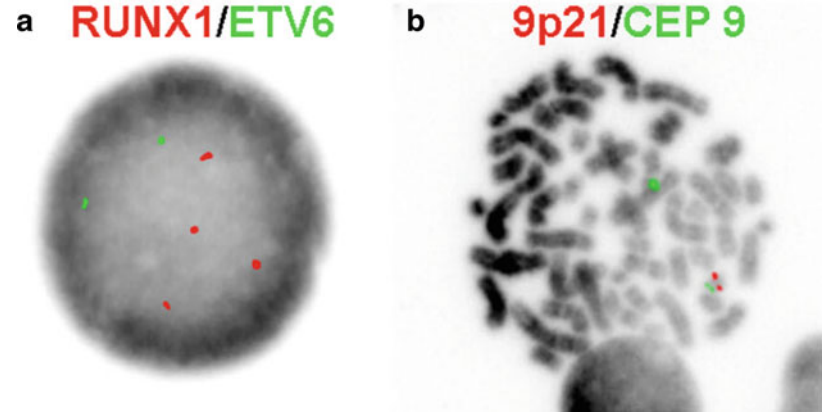


Fig. 31.1 RUNX1 amplification in B-lymphoblastic leukemia. Flow cytometry revealed 96% low side scatter/CD45-dim blasts that also coexpressed B cell markers CD10, CD19, and CD79a as well as TdT consistent with a typical pre-B cell phenotype (not shown). (**a**) Cytogenetic analysis revealed a hyperdiploid 48,XX,+10,–20,+21,+21 karotype with amplification of RUNX1 at chr 21q22 confirmed by FISH. (**b**) FISH analysis additionally revealed deletion of one copy of the 9p21/CDKN2A probe

Table 31.2 Gene alterations commonly seen in ALL

Gene implicated	Type of alteration	Frequency	Prognosis/Clinical context
Primarily mediating maturation arrest			
IKZF1	Deletions and dominant-negative mutations	C: 15-30% of high risk, 80% BCR–ABL1+ ALL, 66% of CML in lymphoid blast crisis	Poor
PAX5	Deletions, point mutations fusion transcripts	Pre-B-ALL: 30-40%, TCF3–PBX1, BCR–ABL1	Associated with progression
RUNX1	Amplification fusion transcripts	2% of B-ALL; older children or adolescents	Poor
PHF6	Deletions and dominant-negative mutations	T-ALL/LBL only C: 16%, A: 38%	Associated with TLX-translocated cases
MYB	Amplification as dup6q23	T-ALL/LBL only 5–10%	Not clearly established
Primarily mediating growth and cell cycle progression			
CDKN2A/p16	Loss of one or both alleles (del 9p21)	C: 30–60% A: 15-30% Both T-ALL/B-ALL	Similar outcome to nondeleted cases in same risk group
JAK1, JAK2	Activating point mutations	10–20% high-risk B-ALL, 20% of T-ALL	Associated with high-risk groups
KRAS/NRAS	Activating point mutations	15%	High WBC count, outcome dependent on translocation type
CRLF2	Translocation to IGH enhancer	C: 5–16% A: >50% high-risk B-ALL	Associated with JAK1/2-mutated and IKZF1 groups
NOTCH1 FBW7	Biallelic alterations affecting NOTCH dimerization, signaling	T-ALL/LBL only, 50–80%	High WBC count, outcome dependent on translocation type
PTEN	Biallelic deletion, inactivating mutations	T-ALL/LBL mostly	Associated with progression
FLT3	Activating mutation	CD117+ T-ALL/LBL	Poor
LEF1	Monoallelic or biallelic microdeletion, mutation	T-ALL/LBL only, 10%	Dependent on translocation type

Abbreviations: *C* children, *A* adults

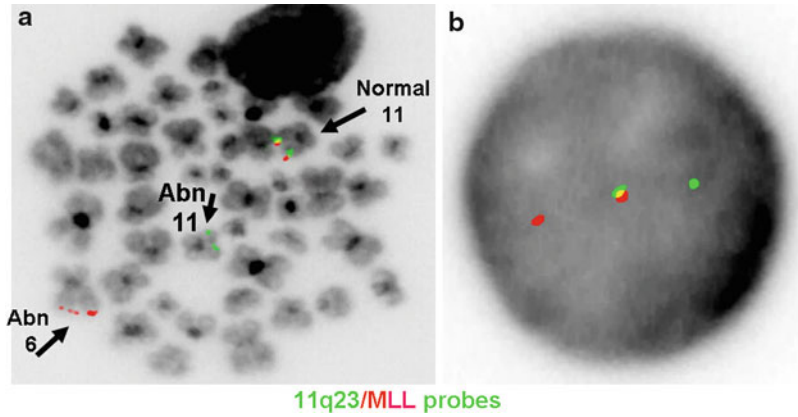


Fig. 31.2 MLL translocation in T-lymphoblastic leukemia. (**a**) Karyotype shows 46,XX,t(6;11)(q27;q23), with FISH analysis on a metaphase spread using a break-apart MLL/11q23 probe set showing a normal cell with two *red/green* fusion signals and a leukemia cell with splitting of one signal into red and green on different chromosomes. (**b**) High magnification of one cell with MLL/11q23 rearrangement

- Other common growth regulatory alterations in T-ALL/LBL include deletion of the kinase negative regulator PTEN and overexpression of CRLF2

31.4 Diagnostic Assays

31.4.1 Morphologic Features

- Most cases of ALL are diagnosed based on blood/marrow smear blast morphology with supporting immunophenotype confirming the diagnosis
  - Cytoplasmic vacuoles can be seen in precursor ALL but are more numerous in Burkitt lymphoma
- High mitotic rate, with tumor cell infiltration of tissues in a single-file fashion, is characteristic of LBL

31.4.2 Immunologic Features

- Regardless of lineage, ALL/LBL variably express markers of immaturity including TdT (90% of cases), CD34 (50% of cases), and CD99
  - Some immunophenotypic features are associated with certain cytogenetic abnormalities (Table 31.1)
- Lineage markers commonly expressed in B-ALL/LBL include CD10, CD19, cytoplasmic (c)CD22 and cCD79a, as well as HLA-DR
  - Pre-B type express CD10 (CALLA) and usually lack CD20
  - More immature pro-B cell subtypes lack CD10 and CD20
  - An uncommon more mature CD20+ IgM/IgD+ subtype overlaps immunophenotypically with Burkitt lymphoma
- Most cases of T cell ALL/LBL are positive for CD2, cCD3, and CD7 but usually lack HLA-DR expression unlike mature T cell neoplasms
  - Cases often show expression of other T cell markers in a pattern mimicking normal thymocyte maturation
    - The majority are surface CD3+, CD4+, CD5+, and CD8+ resembling the common "double-positive" thymocytes
    - Immature "double-negative" cases lacking CD1a, sCD3, CD4, CD5, and CD8
    - More mature CD4+ or CD8+ "single-positive" tumors that lack CD1a

Table 31.3 Common cytogenetic alterations in T-ALL/LBL

(Cyto)genetic alterations	Genes	Frequency	Prognosis
T cell receptor enhancer-driven translocations (TCRA/D at chr 14Q11 or TCRB at chr 7q34-35)			
t(10;14)(q24;q11) or t(7;10)(q35:q24)	TLX1-TCRA/D TCRB-TLX1 (HOX11)	C: 4–7% A: 30%	Favorable
t(11;14)(p15;q11)	LMO1-TCR	C: <1%	
t(11;14)(p13;q11)	LMO2-TCR	C: 5–10%	
t(1;14)(p32;q11)	TAL1-TCR	C: 3%	
inv(7)(p15q32)	HOXA-TCRB	C: Rare	
t(7;19)(q34;p13)	LYL1-TCRB		
t(5;14)(q35;q32)	TLX3 (HOX11L2)-BCL11B	C: 20%, A: 10–15%	
Gene fusions			
del1(p32) (cryptic)	SIL-TAL1	C: 9–30%	Poor
9(q34) episomes	NUP214-ABL1	10%	Poor
t(10;11)(p13;q14)	CALM-AF10	8–10%, TCR-γ/δ+	Variable, by immunophenotype
t(11;19)(q23;p13)	MLL-ENL	Uncommon	Variable

Abbreviations: C: children, A: adults

31.4.3 Cytogenetic and Molecular Features

- B-ALL/LBL is subtype in the 2008 World Health Organization (WHO) schema based on cytogenetic findings, with subgroups for specific chromosomal translocations (Table 31.1)
  - Given their prognostic significance, hyperdiploid, hypodiploid, and normal karyotype groups are also usually recognized
  - Additional genetic alterations with well-established clinicopathologic correlates occur but are uncommon
- T-ALL/LBL is not currently subclassified by cytogenetics on a routine basis given its low incidence, frequent tissue presentation, and typically less informative karyotype
  - 35–70% of T-ALL/LBL have detectable translocations by FISH, but many are cryptic on routine karyotype (Table 31.3)
- Detection of most ALL/LBL translocations can be accomplished by dual-probe/fusion FISH
  - Break-apart FISH probes have also been used for loci with multiple translocation partners, such as the MLL gene at chromosome 11q23 (Fig. 31.2) or the TCRA/D locus in T-ALL
  - Several multiplex PCR assays have also been developed that detect up to ten common ALL translocations in a single amplification reaction
- Expression microarrays and genomic analysis by comparative genomic hybridization (CGH) or single nucleotide polymorphism (SNP) arrays have been proposed as tools to replace routine karyotyping or FISH but are not yet standard diagnostics

31.5 Therapy and Response Predictors

31.5.1 Standard and Experimental Therapies

- Therapy is multiagent induction chemotherapy followed by consolidation/maintenance chemotherapy and early stem cell transplantation in appropriate patient groups
- Tyrosine kinase inhibitors (TKIs) are added for t(9;22)-bearing ALL
  - The TKIs imatinib and dasatinib have resulted in improved outcomes even without stem cell transplant in pediatric populations

- Rituximab (anti-CD20 therapeutic antibody) can be added sometimes for CD20+ ALL
- BRD4 inhibitors may preferentially target MLL-driven ALL/LBL

31.5.2 Prognostic Factors

- The Children's Cancer Group (CCG) and Pediatric Oncology Group (POG) have defined low-, standard-/intermediate-, high-, and very-high-risk groups in pediatric ALL/LBL based on age, presenting white blood cell count and the type of cytogenetic abnormality
  - Cytogenetic features have independent prognostic value (Tables 31.1, 31.3)
- Molecular predictors such as mutation or deletion of IKZF1 in t(9;22)+B-ALL have prognostic significance but are not often routinely tested currently

31.5.3 Pharmacogenomics for Therapy Personalization

- Detection of inherited/germline differences in drug-metabolizing genes can influence the levels and efficacy of some of the standard drugs in ALL regimens and is used at some pediatric centers to tailor therapy
- Variants in the drug-conjugating enzymes glutathione *S*-transferase (GST) isoforms and thiopurine S-methyltransferase (TPMT) that influence handling of mercaptopurine and thioguanine levels can be detected by PCR
- Methotrexate activity can be influenced by PCR-detectable SNPs in two of its targets: dihydrofolate reductase (DHFR) and thymidylate synthase (TYMS)

31.5.4 Monitoring After Therapy

- In childhood ALL/LBL, presence of minimal residual disease (MRD) following therapy for ALL has shown to be the most important prognostic marker for both relapse and outcome
  - Both morphology and conventional cytogenetics have a similar detection sensitivity of 5%
  - FISH for a tumor-specific aberration increases the sensitivity to 1–2% but is not sufficiently sensitive for outcome monitoring
- Three widely used more sensitive methods for MRD assessment include
  - Multiparameter flow cytometry tracking the blast-specific immunophenotype having a sensitivity $\geq$0.01%
    - Detecting MRD in B-ALL can be challenging due to the interference of hematogones which are normal immature B cells that often increase following chemotherapy
  - Reverse transcription quantitative PCR, for translocation-specific fusion transcripts, sensitivity $\geq$0.01%
    - Limited to those cases with specific gene fusions for which PCR primers can be designed
  - Clone-specific quantitative PCR for ALL-associated IGH and/or TCR gene rearrangements, with a potential sensitivity of $\geq$0.001%
    - Expensive and laborious, but is more widely applicable since the vast majority of ALL have at least one clonal IGH, TCRG, or TCRD gene rearrangement that can be used to generate tumor-specific primer sets
- MRD methods have been generally correlated with each other, with MRD flow cytometry predominating
- A Children's Oncology Group (COG) study showed that the presence of even low levels of MRD detected by flow cytometry in day 8 peripheral blood (1% or greater) and day 29 marrow (0.01–0.1% or greater) was associated with shorter event-free survival in all risk groups
- Using clone-specific PCR monitoring at days 33 and 78 postinduction, the AIEOP-BFM ALL 2000 study of B-ALL showed that absence of residual disease at the end of induction was the strongest predictor of favorable outcome (5-year event-free survival >90%)

31.6 Summary

- Nearly all cases of T-ALL/LBL and B-ALL/LBL have detectable maturation arrest and growth regulatory molecular alterations that correlate with outcome
- The large number of prognostic subgroups requires extensive testing with karyotyping and FISH panels for risk stratification; mutation testing is not yet standard
- MRD analysis by flow cytometry or PCR methods is an important predictor of long-term outcome, particularly in pediatric cases
- Improved outcomes are seen with addition of adjuvant targeted therapies in selected subgroups

Further Reading

Cheok MH, Pottier N, Kager L, et al. Pharmacogenetics in acute lymphoblastic leukemia. Semin Hematol. 2009;46:39–51.

Ferrando AA, Look AT. DNA microarrays in the diagnosis and management of acute lymphoblastic leukemia. Int J Hematol. 2004;80:395–400.

Ferrando AA, Neuberg DS, Staunton J, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell. 2002;1:75–87.

McCarthy N. Leukaemia: MLL makes friends and influences. Nat Rev Cancer. 2010;10:529.

Mullighan CG, Su X, Zhang J, et al. Deletion of *IKZF1* and prognosis in acute lymphoblastic leukemia. N Engl J Med. 2009;360:470–80.

Roberts KG, Millighan CG. How new advances in genetic analysis are influencing the understanding and treatment of childhood acute leukemia. Curr Opin Pediatr. 2011;23:34–40.

Myeloid Neoplasms: Molecular Diagnostics

32

C. Cameron Yin and Dan Jones

Contents

32.1 Overview of the Molecular Biology of Hematopoiesis 865
32.1.1 Overview of Normal Hematopoiesis 865
32.1.2 Important Molecules Involved in Myeloid Differentiation and Maturation 866

32.2 Practical Molecular Diagnostics of Myeloid Neoplasms 867
32.2.1 How Leukemia Specimens Are Handled 867
32.2.2 Core Molecular Technologies Used in Leukemia Diagnostics 867

32.3 Clinical and Molecular Genetic Features of Specific Myeloid Neoplasms 869
32.3.1 General Principles of Leukemia Diagnostics 869
32.3.2 Myelodysplastic Syndrome 869
32.3.3 Acute Myeloid Leukemia 872
32.3.4 Chronic Myelogenous Leukemia 880
32.3.5 JAK2-Mutated Myeloproliferative Neoplasms 881
32.3.6 Mastocytosis 882
32.3.7 Hypereosinophilic Syndromes (HES) 883
32.3.8 Newer Therapeutic Agents for AML, MDS, and MPNs 884

32.4 Summary 885

Further Reading 886

C.C. Yin, MD, PhD (✉)
Department of Hematopathology, UT MD Anderson Cancer Center, Houston, TX, USA

D. Jones, MD, PhD
Cancer Diagnostics Services, Quest Diagnostics Nichols Institute, Chantilly, VA, USA

32.1 Overview of the Molecular Biology of Hematopoiesis

32.1.1 Overview of Normal Hematopoiesis

- Hematopoiesis is a highly regulated process in which hematopoietic stem cells give rise to granulocytic, erythroid, megakaryocytic, monocytic, lymphoid, and dendritic lineages
- The stepwise differentiation and maturation of granulocytic, erythroid, and megakaryocytic lineages are defined based on morphological features, which roughly correlate with acquisition of surface markers and functional capacity (Fig. 32.1)
- Hematopoietic stem cells represent a poorly defined set of precursors with multilineage potential for bone marrow (BM) repopulation
  - The term "blast" is a morphologic description of an immature hematopoietic precursor with typically fine nuclear chromatin and often expression of surface CD34
- Maturation of all hematopoietic lineages occurs in the marrow microenvironment except the later stages of T cell maturation, which occur in the thymus
- Later stages of myeloid maturation involve the acquisition of cytoplasmic granule-associated enzymes and immune modulators required for the antibacterial activity of neutrophils and the proinflammatory functions of eosinophils and basophils

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_32,

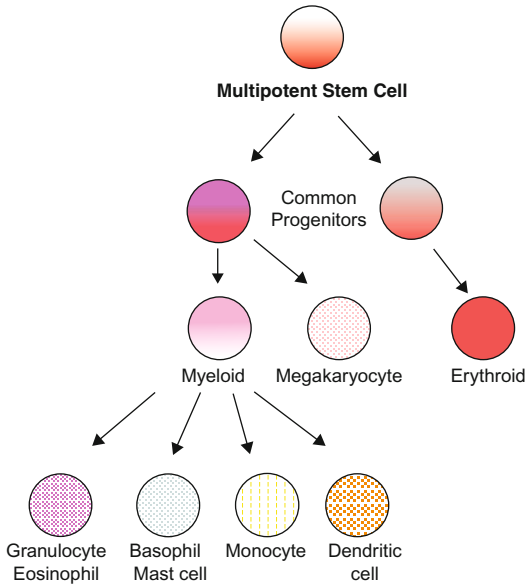


Fig. 32.1 Role of stem cells in hematopoiesis. Subsequent roles of maturation produce progressively more committed precursors with some plasticity maintained depending on microenvironmental influences

- With maturation, monocytes and dendritic cells acquire surface immune-associated molecules involved in antigen presentation
- Migration of hematopoietic elements out of the bone marrow (BM) environment into blood and sites of inflammation is mediated in part by chemotactic chemokines and cytokines

32.1.2 Important Molecules Involved in Myeloid Differentiation and Maturation

- Core-binding factors (CBFs) are a family of transcription factors required for normal hematopoiesis
  - Heterodimeric factors including RUNX1 (AML1, CBFA) and CBFB
  - Bind to enhancers in genes involved in myeloid differentiation, including *IL3*, *CSF1R*, *TRB@*, and *MPO*
  - CBFs are frequent targets for chromosomal translocations and inversions in AML including
    - t(3;21)(q26;q22)/EVI1-RUNX1
    - t(8;21)(q22;q22)/RUNX1-RUNX1T1
    - t(12;21)(p13;q22)/TEL-RUNX1
    - inv(16)(p13.1q22)/CBFB-MYH11
    - t(16;21)(q24;q22)/MTG16-RUNX1
    - t(X;21)(p22;q22)/PRDX4-RUNX1
  - The fusion proteins generally act as dominant negative regulators and inhibit normal CBF function, thus blocking myeloid maturation at specific stages
- *Cytokines* signaling through JAK/STAT-linked receptors are a family of small soluble cytokines essential for production of hematopoietic elements
  - Examples include granulocyte colony-stimulating factor (G-CSF), erythropoietin (EPO), and thrombopoietin (TPO) that regulate granulocyte, red blood cell, and platelet production, respectively
  - Activated cytokine receptors bind JAK tyrosine kinases, which phosphorylate STAT proteins that then translocate to the cell nucleus to act as transcriptional factors affecting growth and differentiation
- Other myeloid growth factors signal through receptor tyrosine kinases (RTKs)
  - Examples include FLT3 ligand, and its receptor FLT3, and stem cell factor (SCF) and its receptor CD117/KIT
  - RTK autophosphorylation drives formation of multimolecular complexes through SH2 and SH3 association domains on the receptors leading to further phosphorylation of a range of substrates
  - Signaling downstream of RTKs involves ubiquitous Ser/Thr kinases and typically drives proliferation through the action of mitogen-activated protein kinases (MAPK)
- Cytoplasmic granule products are functional molecules that define and differentiate myeloid cell subtypes and thus represent excellent lineage markers
  - Myeloperoxidase (MPO) is a core enzyme of granulocytes and is expressed at the beginning of the progranulocytic stage but is abnormally expressed in acute myeloid leukemia (AML)

– Nonspecific esterase (e.g., butyrate esterase) activity is characteristic of the monocytic lineage

32.2 Practical Molecular Diagnostics of Myeloid Neoplasms

32.2.1 How Leukemia Specimens Are Handled

32.2.1.1 Isolating Cells

- Myeloid cells can be isolated by direct centrifugation (i.e., buffy coat preparation) followed by lysis of residual red blood cells
- Myeloid and lymphoid blasts can also be partially purified by density centrifugation using carbohydrate polymers (e.g., Histopaque or Ficoll)
- More specific cell separation methodologies include
  - Flow cytometric methods: high specificity but requires expensive sorters
  - Magnetic bead affinity purification
  - Cell depletion/negative selection methods that remove all but the relevant cell population

32.2.1.2 Uses of DNA

- Used for most lymphoma translocation assays (i.e., detection of translocation breakpoints) and for mutation detection
- DNA extraction methods for bone marrow aspirate and peripheral blood can be easily automated
- DNA is chemically stable and can be stored at 4 °C for weeks to months. Longer term storage is at −20 °C

32.2.1.3 Uses of RNA

- Used for most leukemia translocation assays (i.e., detection of fusion transcripts)
- In most assays, RNA requires conversion into complementary DNA (cDNA) using the enzyme reverse transcriptase (derived from retroviruses)
- RNA should be stored at −80°or in liquid nitrogen

32.2.1.4 Analyzing Protein from Cells

- Leukemia-associated oncoproteins may be detected and quantified by immunofluorescence of fixed cells on smears or by immunohistochemistry on tissue sections using appropriate antibodies
- Protein levels can be assessed in the entire tumor population by lysis of purified cells directly into solubilization buffers containing detergents and protease inhibitors
  - Subcellular fractionation of cytoplasmic and nuclear proteins is done using differential solubilization strategies or density gradient centrifugation
- Protein lysates are analyzed by Western blot following separation by electrophoresis, transfer to solid support (e.g., nitrocellulose), and probing with a tagged antibody or primary/secondary antibody combinations
  - Enzyme-linked immunosorbent assay (ELISA) is a highly sensitive protein analysis technique using primary and capture antibodies in a plate format

32.2.2 Core Molecular Technologies Used in Leukemia Diagnostics

- Karyotypic (cytogenetic) analysis is performed by Giemsa (G) staining of disaggregated chromosomes from tumor cells dropped onto glass slides
  - Karyotyping requires short-term tissue culture followed by trapping of cells in mitosis by colchicine before dropping chromosomes
  - Experienced personnel then to examine the G-banding pattern for each chromosome to identify missing, amplified, and translocated material
  - Given the inability of banding to detect alterations below 100–1,000 megabases, conventional metaphase analysis of chromosome can miss some leukemia-associated translocations
- Fluorescence in situ hybridization (FISH): fluorescently labeled complementary nucleic acid probes are bound to the chromosomes in

a tumor sample and then visualized by microscopy
  - FISH done on metaphase spreads is optimal for correlation with karyotypic banding
  - Interphase FISH, done on nondividing cells, is becoming the dominant technique given its applicability to both fresh BM smears and fixed material
    - Paraffin-embedded archival material can be used for FISH, but decalcification of BM biopsies usually precludes its use in core trephines (BM clot sections are alternative samples for FISH)
  - Dual fusion probes (i.e., one color probe for each partner chromosomal region) are widely applicable to leukemias due to high frequency of reciprocal translocations
    - False-positive results due to cell overlap limit sensitivity of detection, typically 1–5% tumor cells, depending on the number of cells examined
    - Breakapart probes (i.e., a single probe whose signal splits to appear on two chromosomes due to the rearrangement) are especially useful for leukemia-associated genes with multiple fusion partners such as the mixed lineage leukemia, or myeloid lymphoid leukemia [*MLL*] gene
    - Automated methods of FISH counting are becoming available; however, manual scoring by eye remains the dominant technique
- Conventional polymerase chain reaction (PCR)
  - Most useful for the qualitative detection of fusion transcripts
  - Detection methods for PCR products include
    - Gel electrophoresis, with or without probe hybridization
    - Capillary electrophoresis, when one primer is labeled with a fluorochrome
    - Bead array assays (e.g., Luminex methodology) that have different capture probes bound to different beads that are detected by flow cytometry
- Quantitative reverse transcription (qRT) PCR: most leukemia-associated translocation produces *fusion transcripts*; so cDNA can be efficiently used for diagnosis and detection of minimal residual disease (MRD) on posttreatment samples
  - Reverse transcription is an enzymatic process, typically using a modified viral reverse transcriptase, such as AMV-RT, that converts RNA into DNA template for PCR
  - qRT-PCR method detects amplification product when PCR reaction is still in exponential phase (i.e., product is doubling with every cycle)
  - One-step RT-PCR techniques allow efficient PCR but may be more limited in sensitivity
  - TaqMan (Applied Biosystems, Foster City, CA) and LightCycler (Roche, Pleasanton, CA) are the most common methodologies (described in Chap. 25)
- Mutation detection methodologies used in leukemias
  - Direct sequencing using the dideoxy chain termination (Sanger) method is the gold standard but has a usual sensitivity of 15% mutated alleles
  - Pyrosequencing is sequencing by synthesis method that determines DNA sequence by adding individual nucleotides in a predetermined order and detecting subsequent light emission catalyzed by a series of enzyme reactions
    - Sensitivity is about 1 in 10 mutation-bearing cells
    - Can be quantitative based on the peak height on program
    - Good for short-length DNA with well-defined "hot spot" point mutation
  - Allele-specific oligonucleotide (ASO) PCR relies on primers and/or probes that separately bind the unmutated and mutated sequence
    - Maximal sensitivity is 1:10,000 mutation-bearing cells
    - Difficult to optimize given the limited sequence difference in single base pair in point mutations

– Advanced sequencing technologies are beginning to emerge as routine diagnostic methodologies (see Chap. 9)

32.3 Clinical and Molecular Genetic Features of Specific Myeloid Neoplasms

32.3.1 General Principles of Leukemia Diagnostics

- Myeloid and lymphoid leukemia arising in younger patients (children and young adults) usually have defining reciprocal chromosomal translocations that are generally predictive of clinical outcome
- Myeloid leukemias arising in the elderly population typically lack reciprocal translocation, show complex chromosomal gains and losses, and may be preceded by preleukemic myelodysplastic syndrome (MDS)
- Leukemogenesis can be modeled as a multistep process involving several broadly defined transforming events (Fig. 32.2)
  - Mutations and cytogenetic changes producing maturation arrest and usually affect transcription factors or components of the transcriptional coactivation complex that lead to impaired myeloid differentiation (Fig. 32.3, left side). When these changes occur first, AML usually results
  - Mutations affecting growth regulators usually involve genes encoding RTK (e.g., *FLT3*, *KIT*) or downstream effectors (e.g., *JAK2*, *RAS*) that activate signal transduction pathways, providing a proliferative and/or survival advantage to the tumor cells (Fig. 32.3, right side). If growth regulatory mutations occur first, myeloproliferative neoplasms (MPNs) usually arise
  - Mutations within each of these two complementary groups occur infrequently in the same tumor, whereas mutations between the two complementary groups often occur together in the same AML patient

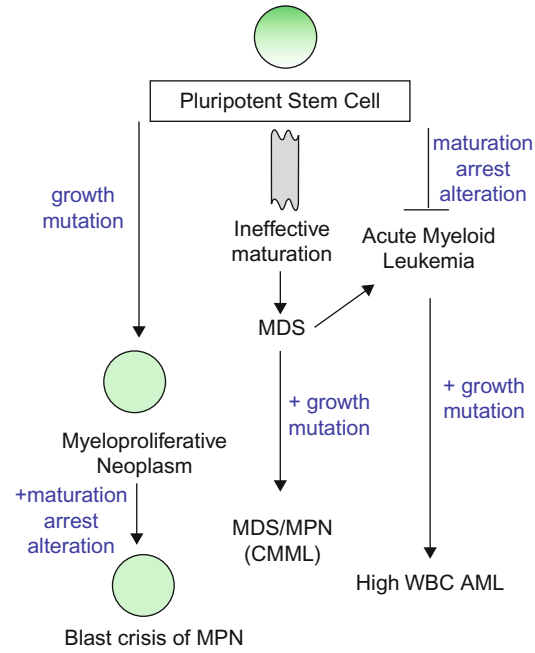


Fig. 32.2 An integrated model for pathogenesis of myeloid neoplasms. Depending on the type of mutation in the hematopoietic stem cell, a myeloproliferative neoplasm (MPN), myelodysplastic syndrome (MDS), or acute myeloid leukemia (AML) may result. Subsequent mutations can shift the presenting neoplasm such as blast transformation of MPN, MDS evolving to AML, or AML relapsing with a very high white cell count

  - Epigenetic shifts and microenvironmental influences predominate in the early stages of MDS (Fig. 32.3, middle)

32.3.2 Myelodysplastic Syndrome

- MDS is a "preleukemic" clonal disorder characterized by ineffective hematopoiesis and dysplastic hematopoiesis associated with cytopenias and progression to AML (Fig. 32.3)
- Clinical features: In 1982, the French–American–British (FAB) Cooperative Group classified MDS into five major categories. In 2001, classification was revised by the World Health Organization (WHO) schema, which was further revised in 2008 (Table 32.1)

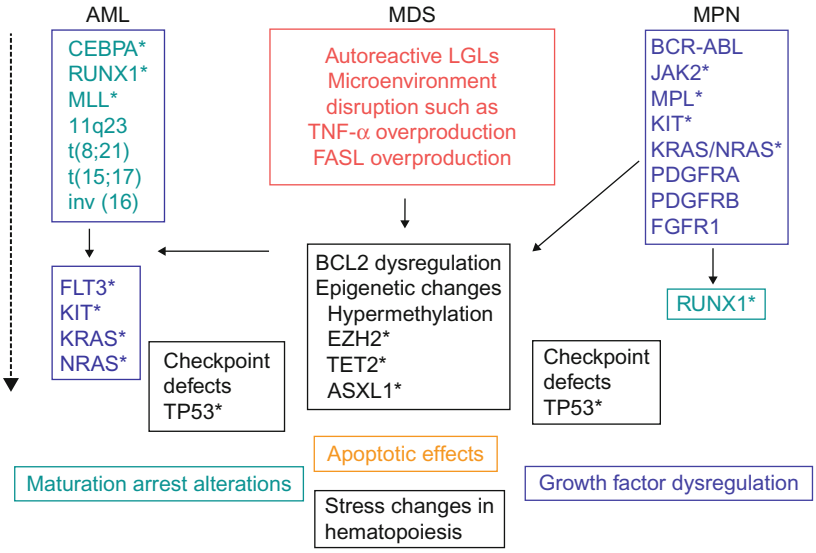


Fig. 32.3 Genes involved in myeloid neoplasms. In AML, mutations (indicated by *asterisk*) or gene rearrangements that produce maturation arrest are often followed by growth regulatory mutations that drive blast expansion. In MDS, the initial stage is often characterized by excessive apoptosis of progenitor cells due to immune attack or epigenetic or microenvironment changes leading to ineffective hematopoiesis. As MDS progresses, this is counterbalanced by mutations which produce an increased proliferation of dysplastic hematopoietic elements. MPNs usually arise as a result of growth regulatory mutations

Table 32.1 Comparison of FAB and WHO classifications of MDS

FAB classification	WHO classification
Refractory anemia (RA)	Refractory cytopenias with unilineage dysplasia (RCUD)
Refractory anemia with ring sideroblasts (RARS)	RARS
	Refractory cytopenia with multilineage dysplasia (RCMD)
Refractory anemia with excess of blasts (RAEB)	RAEB-1 and RAEB-2
Refractory anemia with excess of blasts in transformation (RAEB-t)	AML
Chronic myelomonocytic leukemia (CMML)	Chronic myelomonocytic leukemia (CMML)
NA	5q- syndrome

Table 32.2 International prognostic scoring system for MDS

Score	0	0.5	1.0	1.5	2.0
Blasts %	<5	5–10		11–19	20–30
Karyotype[a]	Good	Intermediate	Poor		
Cytopenia[b]	0–1	2–3			

[a]Classification of karyotype type as in Table 32.3
[b]Cytopenia defined as hemoglobin <10 mg/dL, neutrophils <1,800/uL, platelets <100 k/uL

- The principal feature of classification in both systems is the number of myeloblasts and monocytes and the pattern of dysplasia (i.e., the extent of abnormal maturation in three hematopoietic lineages)
- Isolated 5q deletion syndrome is listed as a separate entity in the WHO scheme due to its better prognosis
- MDS can arise de novo or subsequent to chemotherapy with alkylating agents or topoisomerase II inhibitors, or radiation therapy
- Risk prediction: Cytogenetic aberrations, percentage of bone marrow blasts, and number of cytopenias have been combined in the International Prognostic Scoring System (IPSS) (Table 32.2)
  - Four risk groups are recognized: low, 0; intermediate-1, 0.5–1.0; intermediate-2, 1.5–2.0; and high, ≥2.5

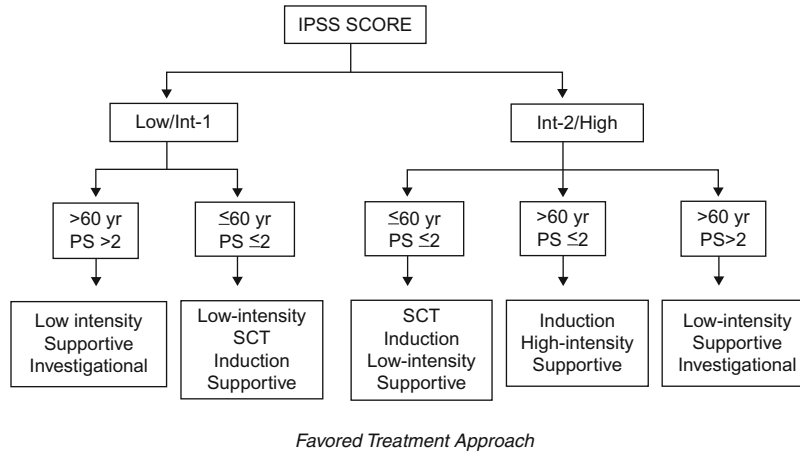


Fig. 32.4 Treatment strategies for MDS. Low-intensity therapy: supportive care, erythropoietin/G-CSF, immunosuppressive therapy; High-intensity therapy: intensive chemotherapy +/− autologous transplantation (Modified from Hofmann et al. Myelodysplastic syndrome. Ann Rev Med 2005;56:1–16)

    - The median survival of MDS patients according to this classification ranges from 6 years for low-risk patients to 6 months for high-risk patients
  - Treatment decisions are often based on the IPSS, age, and performance status (Fig. 32.4)
- Pathologic features: Patients with MDS usually have hypercellular bone marrow often with proportional increase in erythroid elements due to ineffective erythropoiesis
  - Dysplastic erythroid precursors can be multinucleated or show nuclear irregularities or dysynchrony between nuclear and cytoplasmic maturation. Ring sideroblasts represent an abnormal pattern of iron accumulation in erythroid precursors that is common in MDS
  - Dysplastic myeloid elements can be hypergranulated or hypogranulated or show abnormal nuclear maturation
  - Dysplastic megakaryocytic forms show abnormal nuclear lobation and abnormal cell size
  - In approximately 10% of MDS, bone marrow is hypocellular and the distinction from aplastic anemia remains problematic
  - Fibrotic MDS is often associated with increased blasts (best identified by immunostain for CD34 on the core biopsy) and an aggressive clinical course
- Immunophenotypic features: Most cases of MDS show immunophenotypic aberrancies that correspond to the dysplastic morphological features. Multicolor flow cytometry to assess the maturation pattern of the myeloid cell population and to determine the percentage and immunophenotype of the blasts may be required to definitively identify abnormal populations
  - Low-grade MDS can be associated with autoimmune attack on the marrow. In these cases there may be an increase in large granulocyte lymphocytes (LGLs) or either CD8+ T cell or NK cell type
  - Immunohistochemistry has a limited role in the diagnosis of MDS
- Cytogenetic findings: Cytogenetic studies have a major role in the evaluation of MDS in regard to diagnosis and risk stratification
  - De novo MDS shows cytogenetic aberrations in 30–50% of cases
  - Treatment-related MDS has detectable cytogenetic aberrations by conventional karyotyping in 80% of cases
  - Some clonal cytogenetic abnormalities, in the presence of a refractory cytopenia without morphologic evidence of dysplasia, are considered presumptive evidence for MDS (Table 32.3)
    - −Y, +8, or del(20q) as the sole abnormality is not definitive for a diagnosis of MDS in the absence of morphological evidence of dysplasia
  - Three major risk categories of cytogenetic abnormalities have been recognized as

Table 32.3 Recurring cytogenetic abnormalities listed in the WHO classification as presumptive evidence for myelodysplastic syndrome

Type	Cytogenetic abnormalities
Unbalanced	−7 or del(7q), −5 or del(5q),i(17q) or t (17p), −13 or del(13q), del(11q), del(12p) or t(12p), del(9q), idic(X)(q13)
Balanced	t(11;16)(q23;p13.3), t(3;21)(q26.2;q22.1), t(1;3)(p36;q21.2), t(2;11)(p21;q23), inv(3) (q21q26.2), t(6;9)(p23;q34)

Table 32.4 Risk categories of cytogenetic findings in MDS

Group	Cytogenetic abnormalities
Good risk	Normal, isolated del(5q), isolated del (20q), −Y
Poor risk	Complex abnormalities (>= 3), chromosome 7 abnormalities
Intermediate risk	All other abnormalities

Table 32.5 WHO classification of AML

AML with recurrent genetic abnormalities
t(8;21)(q22;q22)(*RUNX1*/*RUNX1T1*)
inv(16)(p13.1q22) or t(16;16)(p13.1;q22)(*CBFB*/*MYH11*)
t(15;17)(q24;q21)(*PML*/*RARA*)
t(9;11)(p22;q23)(*MLLT3*/*MLL*)
t(6;9)(p23;q34)(*DEK*/*NUP214*)
inv(3)(q21q26.2) or t(3;3)(q21;q26.2)(*RPN1*/*EVI1*)
t(1;22)(p13;q13)(*RBM15*/*MKL1*)
Mutated *NPM1*
Mutated *CEBPA*
AML with myelodysplasia-related changes
Therapy-related myeloid neoplasms
AML not otherwise categorized
Myeloid sarcoma
Myeloid proliferations related to Down syndrome
Blastic plasmacytoid dendritic cell neoplasm

a prognostic indicator for MDS (Table 32.4)

- Genetic events in disease progression: As in AML, activating mutations in RTKs (e.g., *FLT3*) and transcriptional regulatory factors (e.g., *CEBPA*) have a role in MDS progression but are not often determined clinically given the difficulty of separating the minor mutation-bearing blast population from the rest of the hematopoietic cells

32.3.3 Acute Myeloid Leukemia

- AML is characterized by the clonal growth of immature blasts of myeloid lineage due to disturbed differentiation and proliferation
- Two general classes of AML are recognized
  - AML with myelodysplasia-related changes
  - AML with simple recurrent cytogenetic abnormalities
- Pathological features
  - Diagnosis of AML by WHO criteria requires 20% or more blasts in the bone marrow aspirate or peripheral blood, except as below (Table 32.5)
    - WHO classification requires both immunophenotypic and cytogenetic and/or molecular studies for classification
    - The older FAB classification defined AML as 30% or more blasts and was based on morphologic and cytochemical findings only
  - AML with t(8;21)(q22;q22), inv(16)(p13.1q22) or t(16;16)(p13.1;q22), and acute promyelocytic leukemia (APL) with t(15;17)(q24;q21) do not require a blast count of $\geq$20% in bone marrow aspirate or peripheral blood
  - AML with 11q23/MLL rearrangements t(6;9)(p23;q34), inv(3)(q21q26.2) or t(3;3)(q21;q26.2), and t(1;22)(p13;q13) do require a blast count of $\geq$20%
  - Cases of that show markedly hypocellular marrow with clusters of blasts have been termed "smoldering AML," regardless of blast count
- Immunophenotypic features: Multiparameter flow cytometry has replaced cytochemistry for phenotyping of AML blasts
  - Typical flow cytometry panels include panmyeloid markers (CD13, CD15, CD33, CD117, myeloperoxidase),

Table 32.6 Risk stratification of AML based on cytogenetic abnormalities

Prognosis	Cytogenetic abnormalities	Comment
Favorable	Balanced chromosomal aberrations (translocations/inversions)	Includes t(8;21), inv(16)/t(16;16), t(15;17)
		Deregulate genes encoding transcription factors important in hematopoiesis
Poor	Unbalanced chromosomal aberrations (losses, deletions)	Includes abn(3q), −5/del(5q), t(6;9), −7/del(7q), t(9;22), abn(9q), abn(11q), abn(17p), abn(20q), abn(21q), >= 3 abn
		Oncogenes are involved in the pathogenesis
Intermediate	None or other abnormalities	Largest cytogenetic group (~45% of de novo AML have normal karyotype)
		Novel molecular markers with prognostic and therapeutic implications

Based on Southwest Oncology Group/Eastern Cooperative Oncology Group studies

monocytic markers (CD14, CD64), megakaryocytic markers (CD41, CD61), and sometimes erythroid markers (glycophorin A, hemoglobin)
  - TdT and lymphoid markers are also added to exclude lymphoblastic leukemia or mixed phenotype acute leukemia
- Cytogenetic and molecular findings: Required for accurate WHO classification and prognostication
  - Pretreatment karyotype has been recognized as the most important independent predictor of clinical outcome in AML
    - Karyotypic analysis identifies three prognostic groups of AML (Table 32.6)
  - Mutations in myeloid transcriptional regulatory factors are common in AML, particularly in cases with normal diploid karyotype
    - Two of these mutations have been separated out to define new provisional entities
      - AML with mutated *NPM1*
      - AML with mutated *CEBPA*

32.3.3.1 AML with Myelodysplasia-Related Changes

- Included are AML arising from a prior MDS, AML with morphologic evidence of myelodysplasia, and AML with myelodysplasia-related cytogenetic abnormalities
- Typically have complex cytogenetic abnormalities: −5/del(5q), −7/del(7q), +8, del(9q), del(11q), del(12p), −13/del(13q), i(17q), del(20q), +21, idic(X)(q13)
- Much more common in the elderly with poor prognosis

32.3.3.2 AML with Recurrent Genetic Abnormalities

- Typically arise de novo without preceding myeloid disorders
- More common in children and young adults
- Often associated with favorable prognosis
- AML with t(8;21)(q22;q22)/*RUNX1-RUNX1T1*
  - Constitutes 5–10% of AML and predominantly affects young patients
  - Clinical features: Good prognosis with conventional chemotherapy
  - Pathological features: Usually classified as AML with maturation (FAB M2) but can also present as AML M1 or rarely other types
  - Immunophenotypic features: Typically expresses CD19 with myeloid markers
  - Molecular genetic findings: The t(8;21)(q22;q22) fuses the *RUNX1*/*AML1*/*CBFA* gene to the *RUNX1T1*/*ETO* gene
  - RUNX1 normally enhances gene transcription by interacting with transcriptional coactivators such as p300 and CREB-binding protein or suppresses gene transcription by interacting with transcriptional corepressors
  - The *RUNX1-RUNX1T1* likely acts as a dominant negative regulator of *RUNX1* function

  - The presence of t(8;21)(q22;q22) is most efficiently detected by FISH but cytogenetic analysis, Southern blot, and RT-PCR can all be used
  - The RUNX1-RUNX1T1 transcript is expressed at very high levels which persists following therapy making qualitative qRT-PCR of limited utility in MRD monitoring
  - qRT-PCR permits serial measurement of tumor burden over time and may be useful in predicting impending relapse
- AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22)/*CBFB-MYH11*
  - Constitutes 5–10% of AML and occurs predominantly in young patients
  - Clinical features: Good response to high-dose cytarabine, high rate of complete remission, and favorable prognosis
  - Pathological features: Mixed monocytic and granulocytic differentiation and the presence of abnormal eosinophils (FAB M4Eo)
    - The diagnosis may not require 20% blasts in the bone marrow or peripheral blood if molecular confirmation is obtained
  - Immunophenotypic features: Mixed population of immature and maturing myeloid and monocytic elements are present which all bear the fusion protein
  - Molecular genetic features: Both chromosome 16 abnormalities result in the fusion of *CBFB* gene at 16q22 (a transcription factor) to *MYH11* at 16p13.1 (smooth muscle myosin heavy chain)
    - The type A break point is seen in 85–95% of AML
      - Nine other fusion transcripts have been reported, mostly related to the number of MYH11 exons included
    - The fusion protein likely acts as a dominant negative regulator of CBF function
    - This fusion is best detected by FISH, Southern blot, or RT-PCR
      - The inv16 or t(16;16) can be subtle and missed by conventional karyotyping
    - Quantitative RT-PCR permits serial measurement of tumor burden over time and may be useful in predicting impending relapse
- APL with t(15;17)(q24;q21)/*PML-RARA* and variants
  - Constitutes 5–8% of AML and affects predominantly middle-aged adults
  - Clinical features: High rate of complete remission and favorable prognosis
    - Frequently associated with disseminated intravascular coagulation likely due to the procoagulant effects of tumor cell granules
    - One of the first examples of effective targeted therapy with use of all-*trans*-retinoic acid (ATRA) to target the abnormal retinoic acid receptor alpha (RARA) protein produced by the chromosomal fusion
    - Timely confirmation of the diagnosis is critical so ATRA therapy can be instituted
  - Pathological features: In the classical form, large promyelocytes with heavily granulated cytoplasm, Auer rods, and bilobed nuclei outnumber myeloblasts
    - The microgranular type has indistinct granules and may be misdiagnosed
    - The diagnosis of APL does not require 20% blasts in the bone marrow or peripheral blood if t(15;17) is present
  - Molecular genetic findings: 95–99% of APL have t(15;17)(q24;q21), resulting in fusion of the *RARA* gene at 17q21 with the promyelocytic leukemia (*PML*) gene at 15q24
    - The breakpoints in the *RARA* gene almost always occur in intron 2, but there are two major breakpoints in the *PML* gene producing long-form and short-form fusion transcripts
    - In both fusion proteins, the normal function of RARA is disrupted
    - t(15;17) is best detected by FISH although Southern blot and RT-PCR can also be used. Conventional cytogenetic analysis may have up to 15% false-negative rate

Table 32.7 Translocations occurring in APL

Translocations	Genes involved
t(15;17)(q24;q21)	*PML-RARA*
t(5;17)(q35;q21)	*NPM1-RARA*
t(11;17)(q13;q21)	*NUMA1-RARA*
t(11;17)(q23;q21)	*ZBTB16-RARA*[a]
der(17)(interstitial deletion)	*STAT5-RARA*

[a]*ZBTB16* formally known as *PLZF*

Table 32.8 Chromosomal abnormalities involving MLL at 11q23

Translocations	Genes involved	Types of leukemia
t(4;11)(q21;q23)	*AF4-MLL*	B-ALL or biphenotypic
t(6;11)(q27;q23)	*AF6-MLL*	AML-M4/M5
t(9;11)(q34;q23)	*AF9-MLL*	AML-M5
t(10;11)(p12;q23)	*AF10-MLL*	AML-M4/M5
t(11;19)(q23;p13.1-13.3)	*ENL-MLL*	AML-M4/M5
del(11q23)	*MLL*	Various

- Quantitative RT-PCR permits serial measurement of tumor burden over time and has been shown to predict overt relapse
- Rare cases of APL have RARA fused to variant translocation partners (Table 32.7)
- All molecular types of APL except the t(11;17)(q23;q21) variant can be effectively treated with ATRA

- AML with t(9;11)(p22;q23)/*MLLT3-MLL*
  - Constitutes approximately 10% of pediatric AML and 2% of adult AML
  - Clinical features
    - May present with disseminated intravascular coagulation
    - Has a tendency to develop extramedullary tumor
    - Overall intermediate to poor prognosis
      - An aggressive subtype occurs following chemotherapy with topoisomerase II inhibitors
  - Pathologic features
    - Usually associated with monocytic differentiation (FAB M4 or M5)
    - The diagnosis requires ≥20% of blasts in the bone marrow or peripheral blood
      - Myeloblasts, monoblasts, and promonocytes should all be considered as blasts
  - Immunophenotypic features
    - Usually express monocytic markers (i.e., CD11c, CD14, and CD64)
  - Molecular genetic findings
    - *MLL* stands for mixed lineage leukemia or myeloid lymphoid leukemia that encodes a histone methyltransferase that regulates gene transcription via chromatin remodeling
    - In addition to *MLLT3*, *MLL* translocations involve over 80 different partner genes (Table 32.8). Many of the fusion partner genes are putative transcription factors
    - Fusion of *MLL* and its partner genes leads to a gain of function of the *MLL* gene that affects differentiation of hematopoietic stem cells by deregulating homeobox (*HOX*) gene expression patterns
    - *MLL* abnormalities are best detected by breakapart FISH probes. Southern blot is also useful to point MLL breakpoints in an 8.3 kb region of 11q23
    - Genetic alterations of MLL through partial tandem duplication also occur in AML
- AML with t(6;9)(p23;q34)/*DEK-NUP214*
  - Constitutes approximately 1% of AML and is more often seen in FAB M2, M4, and M5
  - Clinical features
    - Often preceded by MDS
    - Associated with a poor prognosis
  - Pathological features
    - Associated with basophilia and myelodysplasia
    - The diagnosis requires ≥20% of blasts in the bone marrow or peripheral blood
  - Molecular genetic findings
    - Chimeric fusion transcripts between *DEK* (DNA-binding protein, transcriptional regulator and signal transducer) at

6q23 and *NUP214*/*CAN* (a putative oncogene) at 9q34
    - The resulting nucleoporin fusion protein acts as an aberrant transcription factor and alters nuclear transport
    - High frequency of *FLT3* internal tandem duplications
- AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2)/*RPN1-EVI1*
  - Constitutes approximately 1% of AML and also can be seen in MDS and blast phase of chronic myelogenous leukemia (CML)
  - Clinical features: Associated with an aggressive clinical course
  - Pathological features
    - Multilineage dysplasia, especially hypolobated micromegakaryocytes
    - The diagnosis requires ≥20% of blasts in the bone marrow or peripheral blood
  - Molecular genetic findings
    - Both cytogenetic abnormalities result in fusion of *EVI1* at 3q26.2 and *RPN1* at 3q21
    - The fusion leads to overexpression of *EVI1* that encodes a zinc-finger-containing transcription factor and causes increased proliferation and impaired differentiation
    - Other cytogenetic aberrations involving *EVI1* includes t(3;21)(q26.2;q22)/*RUNX1-EVI1*
- AML with t(1;22)(p13;q13)/*RBM15-MKL1*
  - Occurs in <1% of AML (megakaryocytic)
  - Clinical features
    - Most often seen in infants without Down syndrome
    - No preceding myelodysplastic syndrome
    - Prominent organomegaly
    - Poor prognosis suggested by early reports, but good response to intensive chemotherapy by recent studies
  - Pathological features
    - Increased megakaryoblasts and micromegakaryocytes
    - Often reticulin and collagen fibrosis
    - The diagnosis requires ≥20% of blasts in the bone marrow or peripheral blood
  - Immunophenotypic features
    - The megakaryoblasts express CD41, CD61, CD36, and less frequently CD42
    - Usually negative for CD34, HLA-DR, and myeloperoxidase
  - Molecular genetic findings
    - The translocation results in fusion between *RBM15*/*OTT* (RNA-binding motif protein-15, encodes RNA recognition motifs) at 1p13 and *MKL1*/*MAL* (megakaryocyte leukemia-1, encodes a DNA-binding motif involved in chromatin organization) at 22q13
    - The fusion protein may modulate chromatin organization, HOX-induced differentiation, and extracellular signaling pathways

32.3.3.3 Therapy-Related Myeloid Neoplasms

- Encompass therapy-related AML, therapy-related MDS, and therapy-related myelodysplastic/myeloproliferative neoplasms
- Occur as late complications of a range of chemotherapeutic agents and radiation therapy
- Myelodysplasia is commonly seen
- Cytogenetic abnormalities occur in over 90% of cases and typically are similar to those seen in AML with myelodysplasia-related changes
  - AML following topoisomerase II inhibitor therapy can show balanced translocations involving 11q23 (*MLL* gene)

32.3.3.4 Genes Mutated in AML with Prognostic Significance

- *FLT3* mutation
  - Fms-like tyrosine kinase 3 (*FLT3*), located at 13q12, encodes a receptor tyrosine kinase that is expressed by hematopoietic progenitor cells and downregulated during myeloid differentiation
  - *FLT3* gene alterations are one of the most common known abnormalities in AML with a diploid karyotype (CN-AML) as well as in MDS

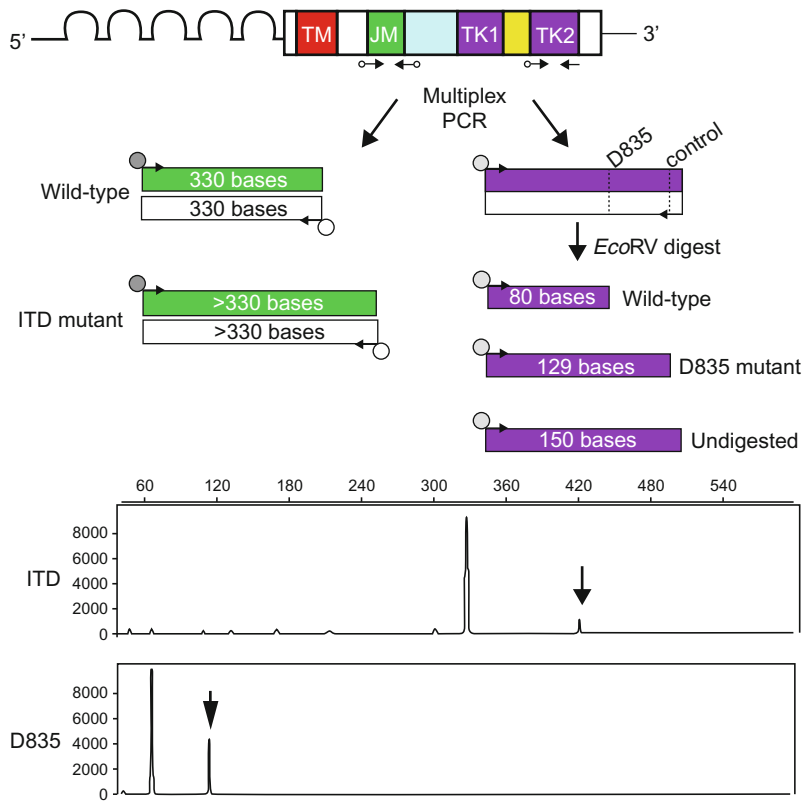


Fig. 32.5 Mutational analysis of *FLT3*. Lower panel shows capillary electrophoretic trace of PCR products with unmutated *FLT3* and ITD (*arrow*, *top*) and the larger-sized PCR product with a D835 *FLT3* point mutation (*arrowhead*, *bottom*) that prevents the complete cutting of the PCR product

  - *FLT3* is an adverse prognostic factor in all AML subtypes and is strongly associated with progression in MDS
  - Most common alterations are variably sized internal tandem duplications (ITD) in the juxtamembrane region of *FLT3* that lead to ligand-independent signaling. *FLT3* ITD occurs in 20% of AML
  - Point mutation in codon 835 (or less commonly codon 836) of the activation loop of the kinase occurs in about 5–7% of AML. An EcoRV restriction site that spans codons 835 and 836 can be used to detect these mutations (Fig. 32.5)
  - *FLT3* ITD may be lost at relapse in a small subset of AML patients carrying *FLT3*-ITD at initial diagnosis. However, when used with caution, assessment of the level of FLT3 mutation in a bone marrow or blood sample also can be used as a MRD assay for *FLT3*-mutated AML and MDS
- *KIT* mutation
  - *KIT*, located at 4q11–12, encodes a receptor tyrosine kinase
  - Activating mutations confer constitutive KIT phosphorylation and activation of downstream effectors including STAT3, MAPK, and PI3K
  - Seen in 5–20% of core-binding factor AML and associated with a poor prognosis
    - Although blasts in 80% of AML cases express CD117/KIT, mutations are infrequent
  - Most common mutations in AML occur in exon 17, especially codon 816, identical to those seen in mastocytosis
  - Can be detected by PCR followed by DNA sequencing
- *RAS* mutation
  - *KRAS* and *NRAS* mutations are detected in 5–15% of AML and MDS cases
    - More commonly seen in chronic myelomonocytic leukemia and AML with monocytic differentiation (FAB M4 and M5)

  - *KRAS*, *NRAS*, and *HRAS* constitute a family of guanine nucleotide-binding proteins which are activated through growth factor signaling and mediate response to the BRAF and MAPK
  - Mutations almost always involve a single amino acid substitution at codons 12, 13, or 61 that alter intrinsic GTPase activity and lead to constitutive *RAS* activation
  - In the absence of *RAS* mutations, the Ras signaling pathway may be activated through other mechanisms, such as gain of function mutations in the upstream tyrosine kinases *KIT* and *FLT3*
  - Can be detected by pyrosequencing, or PCR followed by DNA sequencing
- *NPM1* mutation
  - *NPM1*, located at 5q35, is found to be mutated and mislocalized to the cytoplasm of AML cells in approximately 35% of patients
  - NPM1 is normally a nucleolar protein that interacts with ARF (alternative reading frame) and targets it to the nucleus (ARF-MDM2-p53 pathway)
  - Mutations consist typically of insertions/deletions in exon 12 that lead to frameshift and affect intracellular NPM1 trafficking
    - A duplication of TCTG at position 956–959 is seen in 75–80% of cases, and insertion of CATG or CGTG is detected in 10–15% of cases
  - *NPM1* mutations in AML cosegregate with normal karyotype, CD34 negativity, high frequency of *FLT3-ITD* mutations and good response to induction therapy
  - Can be detected by PCR/capillary electrophoresis or PCR/DNA sequencing
    - Immunohistochemistry to detect abnormally localized NPM1 may be an efficient tool to screen for NPM1 mutations
  - *NPM1* mutation is found to be very stable at relapse, and has become one of the best molecular markers for MRD monitoring in AML patients
- *CEBPA* mutation
  - *CEBPA*, located at 19q13.1, encodes a DNA-binding protein that belongs to basic region leucine zipper (bZIP) family that is essential for granulocyte differentiation
  - Mutations seen in approximately 7% of AML, most commonly FAB M2
  - Mutations span a large number of sites within the protein contributing to a differentiation block specific to AML (dominant-negative effect)
    - N terminal mutations are usually nonsense mutations leading to expression of truncated CEBPA
    - C terminal mutations are usually in-frame mutations (del, ins, dup) resulting in CEBPA mutants with decreased DNA-binding potential
  - Mutations predict favorable prognosis in AML with normal cytogenetics (intermediate risk)
  - CEBPA may also be dysregulated by post-transcriptional mechanisms in core-binding factor AML
  - Best be detected by PCR/DNA sequencing
- *WT1* mutation
  - *WT1*, located at 11p13, encodes a bifunctional protein that consists of
    - N terminal transcriptional regulatory domain (exons 1–6) that acts as a transcription factor
    - C terminal domain that contains 4 zinc finger domains (exons 7–10) that is involved in RNA and protein interaction
  - It functions as a potent transcriptional regulator for genes involved in cellular growth and metabolism
  - Its expression is restricted to primary cells of developing genitourinary system and hematopoietic progenitors, but can be seen in 75–100% AML
  - Germline mutations identified in WAGR syndrome, Denys–Drash syndrome, etc.
  - Mutations seen in approximately 10% of CN-AML
    - Clustered in exons 7 and 9, also seen in exons 1, 2, 3, and 8
  - Mutations associated with younger age, higher serum LDH, higher blast count, and *FLT3-ITD*

  - Impact on survival is variable in different subtypes
- *IDH1 and IDH2* mutations
  - NADP+-dependent isocitrate dehydrogenases that catalyze the oxidative decarboxylation of isocitrate to alpha-ketoglutarate and reduce NADP+ to NADPH
  - Mutations occur in the majority of malignant gliomas and glioblastomas
    - Favorable prognosis
  - Mutations also seen in CN-AML
    - Reported frequency: 15–30%
    - Poor prognosis in *FLT3* wild-type group
    - Mutations also seen in MDS and MPN and have been associated with poor prognosis
  - Mutations identified to date affect a single amino acid within the isocitrate-binding site (exon 4), most often *IDH1* R132 (usually R132H) and *IDH2* R140 and R172
  - Can be detected by Sanger sequencing
- *TET2* mutation
  - *TET2*, located at 4q24, is epigenetic regulator of corepressor/transcription complexes
  - Loss of function mutation seen in 5–20% MPNs
    - Also seen in AML, MDS, CMML, and mastocytosis
    - Mainly in primary myelofibrosis or AML evolved from MPN
  - Associated with monocytosis
  - Limited prognostic relevance
    - Better responses to hypomethylating agents
    - No significant correlation with overall survival
- *CBL* mutation
  - *CBL*, located at 11q23.3, encodes several proteins including E3 ubiquitin-protein ligase CBL that is involved in cell signaling and protein ubiquitination
  - Seen in ~5% MPN and AML
    - Abrogated CBL ubiquitin ligase activity
    - Conferred a proliferative advantage
  - No significant prognostic value on overall survival
- *ASXL1* mutation
  - *ASXL1*, located at 20q11.1, encodes a protein that functions as a coactivator for retinoic acid receptor and is also involved in the regulation of histone methylation
  - Mutations seen in various types of myeloid neoplasms, including MDS, MPN, MDS/MPN, and AML
  - Mutations occur in exon 12 and include frameshift and nonsense mutations
  - Mutations are associated with old age, male gender, trisomy 8, *RUNX1* mutation, higher WBC, monocytosis, lower hemoglobin, progression to AML, and shorter overall survival
- *DNMT3A* mutation
  - *DNMT3A*, located at 2p23, encodes a DNA methyltransferase that catalyzes the addition of a methyl group to the cytosine residue of CpG dinucleotides
    - Increased methylation of CpG islands in the promoter region of a gene is associated with reduced gene expression
    - Hypermethylation of CpG islands in the promoters of tumor-suppressor genes is common in many tumors
  - Mutations seen in ~20% of AML, especially CN-AML
    - A variety of mutations (missense, frameshift, nonsense, splice-site mutation, deletion)
    - Hot spot R882
  - Independently associate with adverse outcomes
- *BAALC* overexpression
  - *BAALC*, located at 8q22.3, encodes a protein of yet unknown function and may represent a marker of early hematopoietic precursors
  - Expressed in a subset of AML cases, especially core-binding factor AML
  - High expression has been associated with *FLT3-ITD*, wild-type *NPM1*, mutated *CEBPA*, *MLL-PTD*, and high *ERG* expression
  - Level of expression has been identified as an independent prognostic marker in CN-AML

    - High BAALC expression correlates with poor prognosis
- *ERG* overexpression
  - *ERG*, located at 21q22, encodes a downstream effector of cellular signaling that regulates proliferation, apoptosis, and differentiation
  - Overexpression is seen in both CN-AML and AML with abnormal karyotype
  - High *ERG* expression has been reported as an independent poor prognostic marker in patients with CN-AML
- *MN1* overexpression
  - *MN1*, located at 22q12, encodes a protein that participates in a gene transcription regulator complex with the nuclear receptor RAR-RXR or vitamin D receptor
  - Acts as a coactivator of transcription
  - Overexpression more commonly seen in AML with inv(16)
  - Higher *MN1* expression has been reported to be independently associated with wild-type *NPM1*, higher *BAALC* expression, and poor clinical outcome in CN-AML

32.3.3.5 Epigenetic Changes in AML

- Many genes have CpG islands in their promoter region that can be methylated at the 5′ position of cytosine, which silences expression of these genes
- For example, P15/INK4b, a cyclin-dependent kinase inhibitor, is frequently methylated in AML and MDS
- Methylation-based PCR or sequencing is done following exposure of tumor DNA to bisulfate, which converts unmethylated C to T, to determine the levels of methylation on each promoter CpG
- Promoter methylation analysis may also be useful in monitoring response to demethylating agents such as decitabine (Dacogen) and azacytidine (Vidaza), which are becoming commonly used in AML and MDS

32.3.4 Chronic Myelogenous Leukemia

- Clinical features: A common leukemia that affects a wide age range and presents with marked leukocytosis, prominent basophilia, and splenomegaly
  - As currently defined, all cases of CML must have the *BCR-ABL1* gene fusion, usually through the t(9;22)(q34;q11.2) chromosomal translocation known as the Philadelphia chromosome (Ph)
  - Cases of *BCR-ABL1*-negative/Ph-negative myeloproliferative neoplasms are no longer classified as CML
  - Usually consists of an initial chronic phase, an ill-defined accelerated phase, and a terminal blast phase
  - Standard therapy for CML is continuous daily oral imatinib, a small molecule inhibitor, which is selective for the ABL1, PDGFR, and KIT tyrosine kinases by competitively inhibit ATP binding
    - Resistance to imatinib is often due to emergence of CML clones with point mutations in the *ABL1* kinase domain, *BCR-ABL1* gene amplification (extra Ph copies), and/or clonal evolution
    - Second-generation tyrosine kinase inhibitors (TKIs, e.g., dasatinib, nilotinib) overcome most forms of imatinib resistance but not T315I and may select new mutation-bearing clones
- Pathological features
  - Peripheral blood shows left-shifted leukocytosis and basophilia
  - Bone marrow is markedly hypercellular with left-shifted granulocytic hyperplasia, erythroid hypoplasia, and small hypolobated megakaryocytic hyperplasia
  - Increased basophils ($\geq$20%) or blasts (10–19%) are features of accelerated phase
  - Blasts $\geq$20% constitute blast phase
- Immunophenotypic features: Routine phenotyping is not done in CML
  - In blast phase, tumor cells are usually immature myeloid (70%) but can be lymphoid (20%) or biphenotypic (10%)
- Molecular genetic findings: Detection of *BCR-ABL1* gene fusion at time of diagnosis can be accomplished by karyotypic analysis, fusion FISH (Fig. 32.6a), or qRT-PCR for the BCR-ABL1 fusion transcript

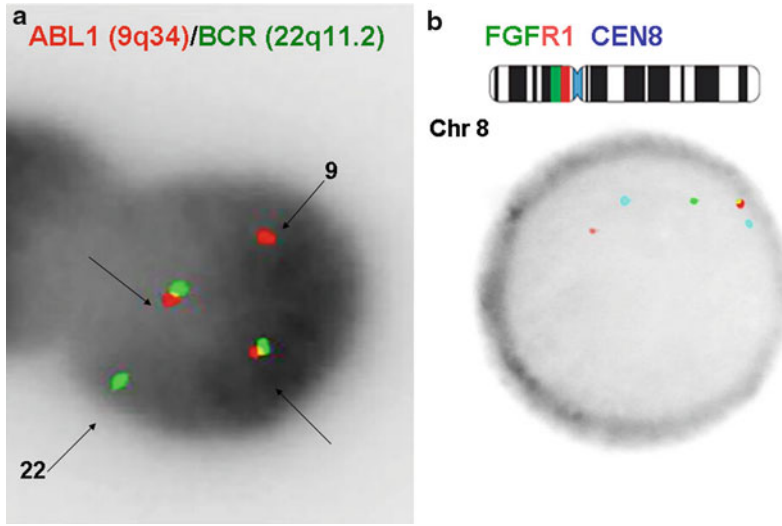


Fig. 32.6 Myeloproliferative neoplasms detected by FISH analysis. (**a**) The red–green (RG) fusion from red (ABL1 on 9q34) and green (BCR on 22q11) locus-specific probes associated with the presence of the t(9;22)/BCR-ABL chromosomal fusion. (**b**) FGFR1 translocation detected using a RG breakapart FGFR1 probe along with a chromosome 8 centromere probe (CEN8) included to assess for copy number

- Standard monitoring for therapy effectiveness in CML is BCR-ABL1 qRT-PCR every 3–6 months while on imatinib therapy. Effective therapy is associated with a three-log reduction in BCR-ABL1 transcript levels from baseline values
  - Complete disappearance of the BCR-ABL1 transcript is unusual with imatinib therapy but is the goal following stem cell transplantation
- Karyotypic analysis of bone marrow aspirate is typically done at the time of disease progression
  - Cytogenetic changes associated with clonal evolution include extra cope of Ph (der22q), isochromosome 17q (TP53 gene deletion), trisomy 8 and trisomy 19
  - Acquisition of AML-type translocations involving core-binding factors [e.g., t(8;21), t(3;21), inv(16)] is a feature associated with sudden blast transformation
- FISH can be performed when bone marrow aspirate or sufficient metaphases are not available
- ABL1 mutational analysis by Sanger sequencing or pyrosequencing is indicated at the time of TKI resistance or disease progression

32.3.5 JAK2-Mutated Myeloproliferative Neoplasms

- Three MPNs with overlapping clinical features, essential thrombocythemia (ET), polycythemia vera (PV), and primary myelofibrosis (PMF), have been shown in the majority of cases to share a common molecular pathogenesis, namely, *JAK2* V617F point mutation
- Cases of MPNs that lack *JAK2* mutations often have mutations in *ASXL1*, *TET2*, and *EZH1* and the other genes discussed above (Fig. 32.3). Diagnostic assays for the mutations are beginning to be used for diagnosis
- Clinical features: Shared clinical features of MPNs include an indolent clinical course, roughly equal gender distribution, a wide range of age at presentation, and stepwise progression from an initial indolent phase

with elevated white blood cells, red blood cells, and/or platelets to a late stage with myelofibrosis, ineffective hematopoiesis, increased blasts and organomegaly
  - JAK2 kinase inhibitors, such as ruxolitinib, are emerging as therapies for MPNs, particularly PMF
- Pathological features: The features of each of these neoplasms are overlapping, and both PV and ET can progress to myelofibrosis
  - PMF is characterized by a hypercellular marrow with thickened bony trabeculae and marrow fibrosis with patent marrow sinuses containing abnormal megakaryocytes. In the spent phase, the marrow is hypocellular and the bony trabeculae markedly thickened
  - All three MPNs progress through an accelerated phase to a blast crisis, analogous to CML, with progression to AML most frequent in PMF
- Molecular findings: A single point mutation (GTC to TTC) in codon 617 of *JAK2* changing Val to Phe is seen in 95% of PV, 50% of ET, and 50% of PMF
  - This mutation results in autoactivation of the JAK2 kinase and hypersensitivity to a group of JAK/STAT-linked cytokine receptors, including granulocytic-monocyte colony-stimulating factor, erythropoietin, and thrombopoietin
  - Three patterns of *JAK2* point mutation are seen: mutation of one *JAK2* allele, mutation of both alleles, or mutation at one allele and gene deletion at the other allele
    - Homozygous mutation of *JAK2* is more common in PV and may contribute to higher hemoglobin levels and a more aggressive disease course
  - Mutations in exon 12 of the *JAK2* gene seen in 5% of PV
  - A point mutation at codon 515 of the *MPL* gene, W515L/K, has been reported in approximately 1% of ET and 1% of PMF
  - The molecular events mediating transformation of the PV, ET, and PMF lacking *JAK2* and *MPL* mutations are not yet known
  - Although most cases are diploid by conventional karyotyping, del 13q is the most commonly detected cytogenetic change seen in both *JAK2*-mutated and unmutated cases of PV, ET, and IMF

32.3.6 Mastocytosis

- Clinical features: A relatively common disorder with a cutaneous type (urticaria pigmentosa) and systemic form with both skin and bone marrow involvement
  - Systemic mastocytosis with associated clonal hematological non-mast cell lineage disease is common in systemic cases
- Pathological features: Neoplastic mast cells commonly have a spindled appearance and form aggregates surrounded by lymphocytes and with prominent fibrosis. Such cases may be confused with lymphoma
  - Diffuse patterns of mast cell infiltration and a leukemic form also occur
- Immunophenotypic features: Neoplastic mast cells can have a variety of immunophenotypic features, with bright CD117/KIT and CD2 coexpression commonly used to gate on mast cells for flow cytometric analysis
  - CD25 expression is seen in neoplastic mast cells but not in normal mast cells
- Molecular generic findings: The majority cases of systemic mastocytosis have an activation point mutation in the *KIT* tyrosine kinase, which is the receptor for stem cell factor
  - Almost always occur at codon 816 in exon 17 (usually D816V)
    - Other rare sites of *KIT* mutation include D820G, E839K, F522C, and V560G
    - The D816V confers resistance to imatinib
  - Occurs in 50–95% of adult patients with systemic cases and 30–50% of pediatric patients with cutaneous lesions, depending on the diagnostic criteria and technology used for mutation detection
  - Given the focal nature of mast cell aggregates and their poor sampling in marrow

aspirates due to myelofibrosis, a sensitive approach to mutation detection is required
- Pyrosequencing (lower limit of sensitivity of 1%) or mutation-specific quantitative PCR (lower limit of sensitivity of 0.01%) on DNA extracted from grossly microdissected bone marrow biopsy is favored
- Sanger sequencing or nonquantitative PCR is not recommended

32.3.7 Hypereosinophilic Syndromes (HES)

- Reactive causes of eosinophilia, such as drug reaction, allergy, parasitic infection, and T-cell lymphoma or Hodgkin lymphoma are far more common than the primary hematopoietic disorders
- Clinical features: Patients usually have pruritus, splenomegaly, and respiratory, cardiac, or gastrointestinal symptoms due to eosinophilic infiltrate. Diagnostic criteria for HES include
  - Persistent eosinophilia greater than 1.5×10^9/L on two occasions at least 6 months apart
  - Evidence of end organ damage, including histologic evidence of tissue infiltration by eosinophils
  - Secondary or other underlying causes need to be excluded: Common myeloid neoplasms associated with eosinophilia include acute myelomonocytic leukemia, chronic myelomonocytic leukemia, and CML
- Pathologic features: Hypercellular bone marrow with eosinophilia including hyposegmented and abnormally granulated dysplastic forms
- Immunophenotypic features: Immunophenotyping is rarely used for characterization of eosinophils, given their characteristic morphologic appearance but a CD10–CD16–CD52+ myeloid immunophenotype is characteristic
- Molecular genetic findings: A subset of HES shows activation of the RTKs PDGFRA, PDGFRB, or FGFR1, by gene fusions through chromosome rearrangements
  - Multiple fusion gene partners have been reported, with the most common being *FIP1L1*, *ETV6/TEL*, and *ZNF198* for *PDGFRA*, *PDGFRB*, and *FGFR1*, respectively (Table 32.9)

Table 32.9 Partner genes involved in the rearrangements of *PDGFRA*, *PDGFRB*, and *FGFR1*

Kinase	Chromosome location	Partner gene (most common partner first)	Chromosome location
PDGFRA	4q12	*FIP1L1*	4q12
		STRN	2p24
		CDK5RAP2	9q33
		KIF5B	10p11
		ETV6	12p13
		BCR	22q11
PDGFRB	5q31–33	*ETV6/TEL*	12p13
		TPM3	1q21
		PDE4D1P	1q22
		WDR48	3p22
		GOLGA4	3p22
		PRKG2	4q21
		HIP1	7q11
		CCDC6/H4	10q21
		GPIAP1	11p13
		GIT2	12q24
		NIN	14q24
		KIAA1509	14q32
		TRIP11/ CEV14	14q32
		TP53BP1	15q22
		NDE1	16p13
		SPECC1	17p11
		RABEP1	17p13
FRF1	8p11	*ZNF198/ RAMP/FIM*	13q12
		FGFR1OP/ FOP	6q27
		TIF1	7q34
		CEP110	9q33
		FGFR1OP2	12p11
		MYO18A	17q11
		LOC113386/ HERV-K	19q13
		BCR	22q11

Table 32.10 Targets and investigational therapy agents in MDS and AML

Class of drugs	Genes targeted	Examples
Angiogenesis inhibitors	*VEGF*	Thalidomide, revlimid, bevacizumab, sunitinib, vatalanib, As2O3
	VEGFR	
Anticytokine agents	*TNFR*	Etanercept
	TNF	Infliximab
Farnesyl transferase inhibitors	*RAS*	Tipifarnib, lonafarnib
Small molecule kinase inhibitors	*FLT3*, *KIT*	Sorafenib, midostaurin, tandutinib sunitinib
Hypomethylating agents	Epigenetic program including *P16/INK4b*	Decitabine, 5-azacytidine
Histone deacetylase inhibitors	*MLL*, *CEBPA*, *AML1-ETO*, *CBFB-MYH11*	Vorinostat/SAHA, MS-275
Proteasome inhibitors	*NFKB*	Bortezomib
Nuclear receptor ligands	*PML-RARA*	ATRA
mTOR inhibitors	mTOR pathway	Everolimus

  - The most common molecular abnormality in this group is the cytogenetically occult *FIP1L1-PEGFRA* gene fusion that results from an 800-bp interstitial deletion at 4q12
    - Fusion produces constitutive activation of the PDGFRA tyrosine kinase that can be effectively blocked with imatinib
    - Best detected by FISH for loss of the *CHIC2* gene in the intervening deleted chromosomal segment or by RT-PCR
  - Karyotyping and FISH can be used to detect PDGFRB- or FGFR1-translocated cases (Fig. 32.6b)

32.3.8 Newer Therapeutic Agents for AML, MDS, and MPNs

- Insights into the molecular pathogenesis of MDS and AML have led to the development of new classes of therapeutic agents (Table 32.10). To date, most of these therapies have had limited efficacy as single agent but are now being tested in combination with the standard cytotoxic chemotherapy regimens. Classes of new agents include the following
- Angiogenesis inhibitors: Dysregulation of angiogenesis by abnormal secretion of angiogenic cytokines and growth factors is essential for apoptosis of marrow progenitor cells
  - Small molecule inhibitors of angiogenic agents and receptors, such as vascular endothelial growth factor (VEGF) and its receptors have been developed
  - Thalidomide and the related lenalidomide and pomalidomide are anti-VEGF agents that also have immunomodulatory and antitumor necrosis factor alpha (TNFα) effects
  - Bevacizumab is a recombinant anti-VEGF monoclonal antibody and also inhibits bone marrow production of TNFα
- Farnesyl transferase inhibitors: Activating point mutations in the farnesylated proteins NRAS and KRAS are detected in 5–15% of AML and 50% of CMML
  - Farnesylation of C terminal consensus sequences by farnesyl transferase is the rate-limiting posttranslational modification of RAS proteins
  - Farnesyl transferase inhibitors, including tipifarnib and lonafarnib, represent a novel class of potent inhibitors of RAS activation that are able to modulate multiple signaling pathways implicated in the pathogenesis of MDS and CMML
- FLT3 inhibitors: FLT3 internal tandem duplications and activating point mutations cause constitutive activation of the receptor tyrosine kinase and lead to signaling through the RAS, MAPK, and STAT5 pathways, contributing to the development of leukemias in animal models
  - PKC412, CEP701, MLN518, and SU11248 are FLT3 inhibitors currently in clinical trials

  - Given the wide expression of FLT3, it remains unclear whether AML with *FLT3* genetic alterations are the appropriate cases for treatment with these kinase inhibitors
- DNA methyltransferase inhibitors: Abnormalities of cytosine methylation constitute the most common epigenetic changes in AML and MDS and represent a potentially reversible method lead to altered gene expression
- DNA methyltransferase inhibitors decitabine and azacytidine are both approved by Food and Drug Administration (FDA) for the treatment of AML/MDS
- Histone deacetylase (HDAC) inhibitors and proteosome inhibitors: Posttranslational modification of histones by dynamic acetylation and deacetylation is mediated by histone acetyltransferase (a transcriptional coactivator) and HDACs, a transcriptional corepressor
- HDACs are associated with transcriptionally inactive chromatin (heterochromatin)
- HDAC inhibitors (HDACi) modulate chromatin structure and gene expression by inducing histone hyperacetylation. They also induce growth arrest, cell differentiation, and apoptosis of tumor cells
- Certain leukemia-associated fusion genes, such as *RUNX1-RUNX1T1* and *PML-RARA*, specifically recruit nuclear corepression complexes including HDACs and silence groups of differentiation-related genes
  - HDACi may be utilized to specifically reverse the transcriptional repression induced by the fusion proteins
- The proteosome inhibitor bortezomib has demonstrated to have preclinical synergistic activity with HDACi, as well as potential single-agent activity in AML and MDS

32.4 Summary

- The current leukemia classification combines morphologic, immunophenotypic, and molecular genetic data
  - With the rapidly expanding list of novel molecular abnormalities identified in myeloid neoplasms, new disease entities may emerge quickly and revision of the current classification scheme is therefore anticipated
  - Molecular diagnosis has a major role in MRD detection in leukemias with defining fusion transcripts [e.g., CML, APL, AML with inv(16)], as well as those with gene mutations (e.g., *FLT3, NPM1, JAK2*)
  - Identification and detection of molecular abnormalities also play a role in risk stratification (e.g., *FLT3* mutation, *NPM1* mutation) and development of target therapy (e.g., TKIs, FLT3 inhibitors, DNA methyltransferase inhibitors)
- Development of high-throughput molecular diagnostic tests is mandated in the era of molecular diagnosis and target therapy
  - Conventional karyotypic analysis is still routinely performed on all new cases of myeloid neoplasms
    - Some common leukemia-associated cytogenetic abnormalities are relatively commonly missed by conventional karyotyping, including inv(16)(p13q22), t(15;17)(q24;q21) and 11q23 abnormalities
  - FISH is a robust method for initial diagnosis of nearly all leukemias with characteristic chromosomal translocations/inversions since false-negative results can occur by PCR due to the heterogeneity of chromosomal breakpoints
  - Qualitative or quantitative PCR is the mainstay for detection of chromosomal translocations and inversions
    - Qualitative PCR may detect very low levels of leukemia-associated translocations in normal individuals
    - Quantitative PCR may be too sensitive for some types of leukemia, particularly t(8;21), to provide meaningful predictive value in MRD
  - Gene mutational analysis is best performed by Sanger sequencing and pyrosequencing (for cases with well-defined "hot spot" point mutation)

– Gene expression profiling by DNA microarray technology, high-throughput mutational analyses, and proteomic approaches represent alternative strategies for leukemia classification that may emerge in the future

Further Reading

Arber DA, Brunning RD, Le Beau MM, et al. Acute myeloid leukemia with recurrent genetic abnormalities. In: Swerdlow SH, Campo E, Harris NL, et al., editors. WHO classification of tumours of haematopoietic and lymphoid tissues. Lyon: IARC; 2008. p. 110–23.

Brunning RD, Orazi A, Germing U, et al. Myelodysplastic syndromes. In: Swerdlow SH, Campo E, Harris NL, et al., editors. WHO classification of tumours of haematopoietic and lymphoid tissues. Lyon: IARC; 2008. p. 87–108.

Burnett A, Wetzler M, Lowenberg B. Therapeutic advances in acute myeloid leukemia. J Clin Oncol. 2011;29:487–94.

Tefferi A, Vainchenker W. Myeloproliferative neoplasms: molecular pathophysiology, essential clinical understanding, and treatment strategies. J Clin Oncol. 2011;29:573–82.

Theilgaard-Monch K, Boultwood J, Ferrari S, et al. Gene expression profiling in MDS and AML: potential and future avenues. Leukemia. 2011;25:909–20.

Vainchenker W, Delhommeau F, Constantinescu SN, et al. New mutations and pathogenesis of myeloproliferative neoplasms. Blood. 2011;118:1723–35.

Walker A, Marcucci G. Impact of molecular prognostic factors in cytogenetically normal acute myeloid leukemia at diagnosis and relapse. Haematologica. 2011;96:640–3.

Yin CC, Jones D. Myeloproliferative Neoplasms. In: Jones D, editor. Neoplastic hematopathology – experimental and clinical approaches. New York: Springer/Humana Press; 2010. p. 177–92.

Lymphoid Malignancies: Molecular Diagnostics

33

Vasiliki Leventaki and Francisco Vega

Contents

33.1 Overview of the Molecular Biology of Lymphocytes 887
33.1.1 Differentiation and Maturation of B Cells 887
33.1.2 Differentiation and Maturation of T Cells 890
33.1.3 Natural Killer Cells 891

33.2 Practical Molecular Diagnostics of Lymphoid Malignancies 892
33.2.1 General Principles 892
33.2.2 How Lymphoma Specimens Are Handled 892
33.2.3 The Core Technologies Used in Lymphoma Diagnostics 892

33.3 Clinical and Molecular Genetic Features of Specific Lymphoid Malignancies 898
33.3.1 Immature B Cell and T Cell Leukemia/Lymphomas 898
33.3.2 Mature B Cell Lymphomas 902
33.3.3 Follicular Lymphoma 906
33.3.4 B Cell Lymphomas, Unclassifiable ("Gray-Zone" Lymphomas) 913
33.3.5 Mature T Cell Lymphoma 916
33.3.6 Classical Hodgkin Lymphoma 923
33.3.7 Primary Cutaneous B Cell Lymphomas 923
33.3.8 Cutaneous T Cell Lymphomas 926

Further Reading 928

V. Leventaki, MD • F. Vega, MD, PhD (✉)
Department of Hematopathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

33.1 Overview of the Molecular Biology of Lymphocytes

33.1.1 Differentiation and Maturation of B Cells

- The stepwise differentiation and maturation of B cells include
  - Commitment to the B cell lineage: multistep process that occurs in the bone marrow
  - V(D)J immunoglobulin (Ig) gene rearrangement: DNA recombination event occurring in bone marrow that ultimately generates a unique antibody molecule in each precursor B cell and its progeny
  - Progressive phenotypic maturation: stepwise maturation of pro-B, pre-B, and mature, naïve B cells with different surface markers lost and gained at each stage
  - Antigen recognition: through surface Ig (aka B cell receptor) and processed antigen presented on the surface of an antigen-presenting cell (APC) in association with class II major histocompatibility antigens (MHC)
  - Antigen selection: triggered by antigen binding to the surface Ig
    - Antigen-stimulated B cells move into the germinal center (GC) where a process of targeted mutagenesis (*somatic hypermutation*) introduces changes into the V_H gene, particularly in complementarity-determining regions (CDR)-1 and -2

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_33,

    - The mutated Ig is functionally tested for improved antigen binding in the GC, a process termed *affinity maturation*
    - A subset of B cells undergo shift from sIgD and sIgM expression to expression of secreted IgG, IgA, or IgE by a second DNA recombination event known as *class switch*
  - Terminal B cell differentiation and antibody production
    - B cells leave the GC following termination of antigen selection and mature into long-lived antibody-producing plasma cells, which reside in the medullary areas of lymph node and the bone marrow
- The terminally differentiated plasma cells lack most B cell markers including CD20 and CD22 and express CD38 and CD138
- A subset of B cells survive as long-lived memory B cells (antigen-experienced B cells) that provide immunologic memory quickly differentiating into plasma cells in antigen recall responses
- Some important molecules involved in B cell differentiation and maturation
  - Interleukin (IL)-7: cytokine that is important for early B cell growth and proliferation
  - Rag-1/2 endonucleases and terminal deoxynucleotidyl transferase polymerase: these proteins are required for the recombination/ligation and N base addition during the VDJ recombination process
  - Paired box protein-5 (PAX5/BSAP): transcription factor required for rearrangement of V_H gene segments and for commitment to and maintenance of the B cell differentiation pathway. Target genes of PAX5 include the pan-B cell marker CD19
  - OCT2 and BOB.1 transcription factors: control the expansion and/or maintenance of mature B cells
  - Polydomain transcription factors: transcription factors required for in vitro proliferation of B cells following stimulation with T cell-independent antigens
  - B cell lymphoma 6 (BCL6): multifunctional transcription factor required for activation of the GC transcription program
  - Activation-induced cytosine deaminase: enzyme required for somatic hypermutation and Ig class switch in the GC reaction
  - B lymphocyte-induced maturation protein 1 (BLIMP1): transcriptional repressor required for plasma cell differentiation. BLIMP1 represses PAX5
  - Interferon regulatory factor 4 (IRF4) is upregulated in a subset of centrocytes and expressed at high levels in plasma cells. IRF4 represses BCL6, prerequisite for plasma cell differentiation
  - Nuclear factor kappa B (NFκB) subunits are expressed in a subset of centrocytes and provide survival signals necessary for cell exit from the GC or their recirculation within the GC

33.1.1.1 Ig Gene Rearrangement in B Cells

- Functional Ig and T cell receptor (TCR) genes are assembled from separate germline coding sequences by a site-specific V(D)J DNA recombination
- V(D)J recombination is the main mechanism for generating antigen receptor diversity in mammals, with generation of an almost unlimited repertoire of different antigen receptors with unique specificity
- These genes initially reside in germline DNA as a large number of discrete and discontinuous segments located at chromosomes 14q32 (*IGH*), 2p11.2 (*IGK*), and 22q11 (*IGL*) (Fig. 33.1)
  - *IGH* is composed of number of variable (V), diversity (D), and joining (J) gene segments. Initially, the D_H segment is joined with a J_H segment with deletion or insertion of a variable number of nucleotides (N) between D_H and J_H. This is followed by a V_H to D_HNJ_H joining with N basepair (bp) insertion or deletions between V_H and D_H (Fig. 33.2)
- The VDJ junctional sequence produced, termed the CDR3, is essentially unique to any precursor B cell and its progeny.

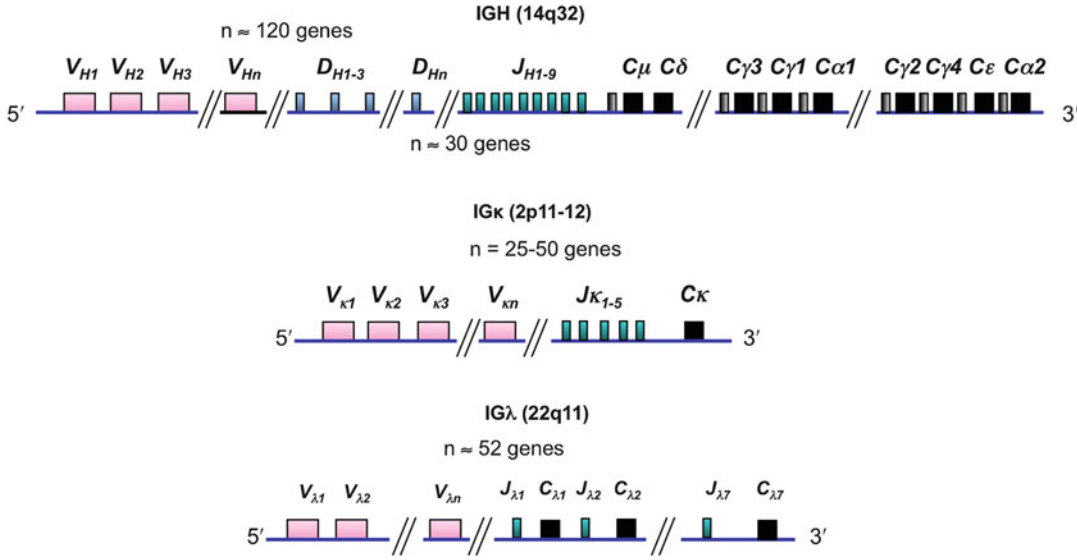


Fig. 33.1 Structure of the *IGH*, *IGK*, and *IGL* genes. The Ig genetic loci are composed of a number of variable (*V*), joining (*J*), and constant (*C*) regions. The *IGH* loci also have diversity (*D*) regions

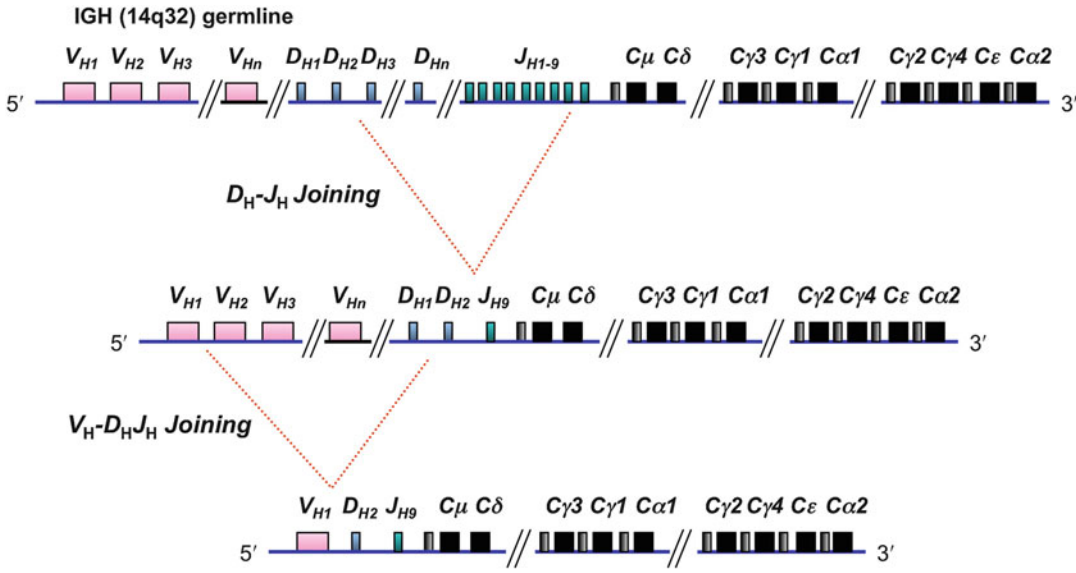


Fig. 33.2 Rearrangement of the Ig heavy chain locus. This *DNA recombination* process occurs in pre-B cells in the bone marrow in a stepwise fashion generating an IGH chain with a particular V_H, D_H, and J_H segment and a unique VDJ junctional sequence termed the CDR3. All B cells that are produced from this precursor cell have an identical VDJ IGH molecule

The unique size of this region for each B cell clone allows the detection of monoclonality in B cell lymphomas

– The VDJ segment joins with a C_H constant region segment, which can either be inframe (correct for encoding antibody sequences) or out of frame

- If there is an out of frame (nonfunctional) VDJC joining, a second attempt at VDJC joining occurs on the other *IGH* allele

- When the VDJC joining produces an inframe functional protein, the precursor B cell proceeds to rearrange the light chain gene(s)
  - The Ig light chains are composed of one of two genes kappa (κ) and lambda (λ), each containing multiple V and J segments (no D), which rearrange sequentially
- This order of recombination gives rise to a predominance of Igκ-expression over Igλ-expression among mature B cells
- V(D)J recombination occurs only in lymphocytes and has the following features
  - Lineage specificity: Ig gene assembly in B cells but not T cells and *TCR* gene assembly in T cells but not B cells
  - Developmental stage specificity: assembly of *IGH* genes occurs before Ig light genes. Class switch producing IgG, IgA, and IgE follows antigen selection
  - Allelic exclusion: only one IGH and one Ig light chain protein is expressed in a given B cell. The silencing of out-of-frame Ig transcripts is mediated in part by nonsense-mediated degradation at the time of translation

33.1.2 Differentiation and Maturation of T Cells

- The stepwise differentiation and maturation of T cells include
  - Commitment to the T cell lineage: a multistep process that occurs in the bone marrow and in the thymus
  - Progressive phenotypic maturation: stepwise maturation of pre-T cells, early double-positive cells (CD4+, CD8α+, CD8β−), double-positive cells (CD4+, CD8α+, CD8β+), and finally mature α/β- and γ/δ-T cell subsets
  - V(D)J *TCR* gene rearrangement: DNA recombination ultimately generates a dimeric surface *TCR* composed of either γ/δ-chains or α/β-chains
  - α/β-T cells: specialized functional subsets include
    - Cytotoxic T cells (mostly CD8+) that recognize foreign antigens through MHC class I binding
    - Helper T cells (mostly CD4+) that stimulate B cell activation and recognize processed antigen and class II MHC on the surface of APC. A subset produces interleukin (IL)-17 (T_H 17) and other proinflammatory cytokines, IL-17A, IL-17F, IL-22, TNF, IL-26, and IL-6
    - Regulatory T cells (mostly CD25-bright CD4+ and FOXP3+) that downmodulate T cell activation
  - γ/δ-T cells: these lymphocytes have APC roles and function predominantly in the innate immune system, which recognizes altered antigen patterns, particularly following infection
  - T cell activation following antigen binding: binding of antigen to surface TCR of either type triggers signaling that results in cell proliferation and cytokine secretion
- Some important molecules involved in T cell differentiation and maturation
  - NOTCH: a family of signal transduction proteins that regulate production and differentiation of precursor and mature T cells
  - Hedgehog (Hh) ligands (sonic Hh, Indian Hh, and dessert Hh) are produced by stroma thymic epithelial and dendritic cells and regulate proliferation and differentiation of thymocytes. Indian Hh is also produced by thymocytes in double-positive cells. As T cell differentiation progresses, immature thymocytes move away from the corticomedullary junction (which has high expression of Hh proteins), resulting in a gradual loss of Hh stimulation and induction of TCRβ rearrangements
  - Glioma-associated oncogene homologue-3 (GLI-3) is necessary for double-negative to double-positive differentiation, especially after pre-TCR signaling

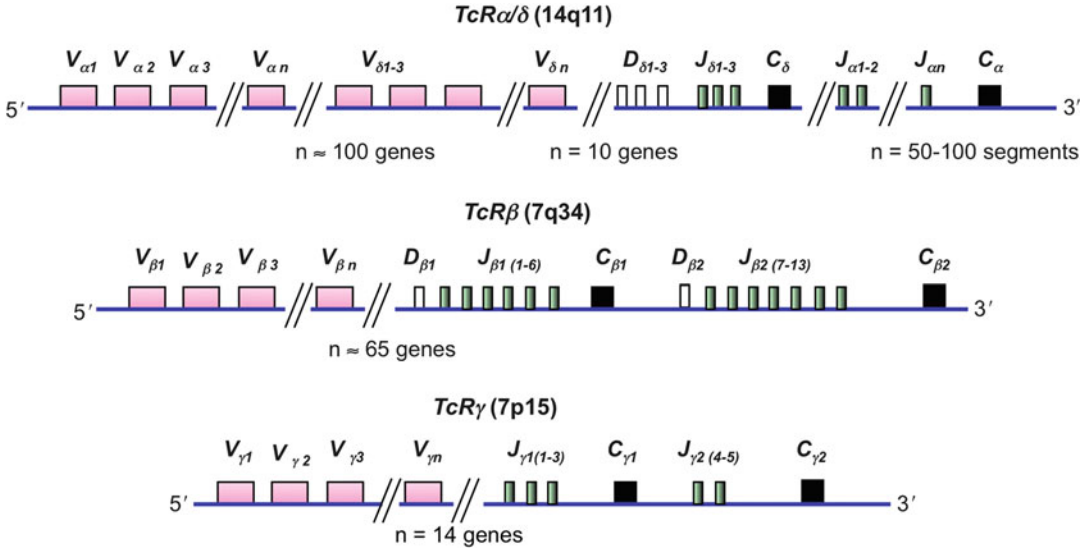


Fig. 33.3 Structure of the *TCRA*, *TCRD*, *TCRB*, and *TCRG* genes. The *TCR* loci, similar to the *Ig* genes, are composed of a number of variable (*V*), joining (*J*), and one or more constant (*C*) regions. The *TCRB* and *TCRD* genes also have diversity (*D*) regions. The *TCRD* is contained within the *TCRA* locus and is usually deleted following *TCRA* rearrangement

- IL-2: the most important growth regulatory cytokine for mature T cells
- CD25: the high-affinity subunit of the IL-2 receptor. High CD25 expression is a marker of regulatory (suppressor) T cells
- IL-12: cytokine that promotes the production of the Th1 subset of T cells that produce interferon-γ
- IL-4 and IL-10: both cytokines promote the production of the Th2 subset of T cells that stimulate antibody production by B cells

33.1.2.1 TCR Gene Rearrangement in T Cells

- *TCR* genes reside on chromosomes 14q11 (*TCRA, TCRD)*, 7q34 (*TCRB*), and 7p15 (*TCRG*) and utilize the same recombination machinery (Rag1/2, TdT) as B cells (Fig. 33.3)
- *TCR* expression also exhibits lineage restriction, developmental specificity, and allelic exclusion, but *TCR* genes do not undergo somatic hypermutation
- The *TCR* γ-chain has a limited number of germline V and J segments and rearranges first
- The *TCR* δ-chain has a small number of V, D, and J gene segments and often shows minimal N base addition contributing to the limited diversity of the *TCR*-γ/δ system
- If rearrangement produces an inframe functional *TCR*-γ/δ, a γ/δ-T cell is produced and recombination ceases
- If no functional TCR is produced, *TCRB* rearrangement proceeds in the thymus followed by rearrangement of the *TCRA* gene
  - The large number of *TCRB* V, D, and J and *TCRA* V and J segments results in a highly unique CDR3 that mediates the highly specific antigen binding of *TCR*-α/β system
  - The *TCRA* gene straddles the *TCRD* gene on chromosome 14q11 so α/β-T cells usually have deletion of one or both of the *TCRD* genes
- Because of the above developmental sequence, all mature T cells have genomic rearrangement of the *TCRG* locus, but only α/β-T cells have rearranged *TCRB* genes

33.1.3 Natural Killer Cells

- NK cells are involved in innate immunity, which is the rapid activation component of

the immune response that does not require immunologic memory
- NK cells lack antigen-specific antigen receptors but recognize general classes of foreign antigens (e.g., bacterial glycolipids) through toll-like receptors
- NK cells interact with nonclassical MHC-like molecules like CD1d and HLA-E on lymphocytes and APCs through growth stimulatory and inhibitory NK receptors (NKR)
  - NKRs are polymorphic and thus mediate additional immunologic diversity between individuals

33.2 Practical Molecular Diagnostics of Lymphoid Malignancies

33.2.1 General Principles

- Molecular diagnostics and cytogenetic analysis have complementary roles with other analysis techniques in lymphoma diagnostics including morphological evaluation, flow cytometry, and immunohistochemistry
- Detection of an expanded (clonal) population bearing a specific *IGH, IGK, TCRG*, or *TCRB* gene rearrangement is the basic technique for establishing clonality in B cell and T cell tumors
  - Except in early lymphoblastic tumors, the process of gene rearrangement precedes neoplastic clonal expansion so can be used as a clonal tumor marker
  - Benign reactive lymphoid proliferations usually give a polyclonal pattern of amplification due to the variable usage of V, D, and J segments and the variable-sized CDR3 produced during the recombination process described earlier
  - Pseudoclonal or oligoclonal proliferation of T cell and B cells due to infection or inflammatory conditions can occasionally give a "false"-positive clonality test
- NK cell lymphomas do not harbor monoclonal antigen receptor gene rearrangements

33.2.2 How Lymphoma Specimens Are Handled

- Frozen sections: snap-frozen tissues prior to sectioning provide an ideal source for long-term storage at −20 °C for later DNA, RNA, or protein analysis
- Cytology preparations: DNA and RNA molecular analysis can be performed on freshly isolated lymphocytes or ethanol-fixed cell pellets from fine needle aspirate samples. Air-dried slides can be used for FISH analysis
- Paraffin-embedded material: tissue blocks can be used for PCR and FISH but not for Southern blot analysis. Transcript analysis may be limited in older material due to RNA degradation
  - Fixatives, particularly formalin, produce breaks in DNA and RNA that decrease efficiency of PCR amplification of long amplicons
- Uses for DNA: chemically stable and can be stored at 4 °C for weeks to months. Longer-term storage at −20 °C is recommended
  - Detection of lymphoma translocations
  - Detection of point mutations
- Uses for RNA: sensitive to degradation, should be stored at −70 °C or in liquid nitrogen
  - Expression microarray studies
  - PCR to detect fusion transcript (e.g., API2 mucosa-associated lymphoid tissue-1 [MALT1]) requires conversion of RNA into complementary DNA (cDNA) using the enzyme *reverse transcriptase* (derived from retroviruses) before PCR

33.2.3 The Core Technologies Used in Lymphoma Diagnostics

- Use of Southern blot analysis
  - Requires high molecular weight DNA usually obtained from fresh or frozen tissue
  - Useful for B cell and T cell clonality assays and for detecting rearrangement of

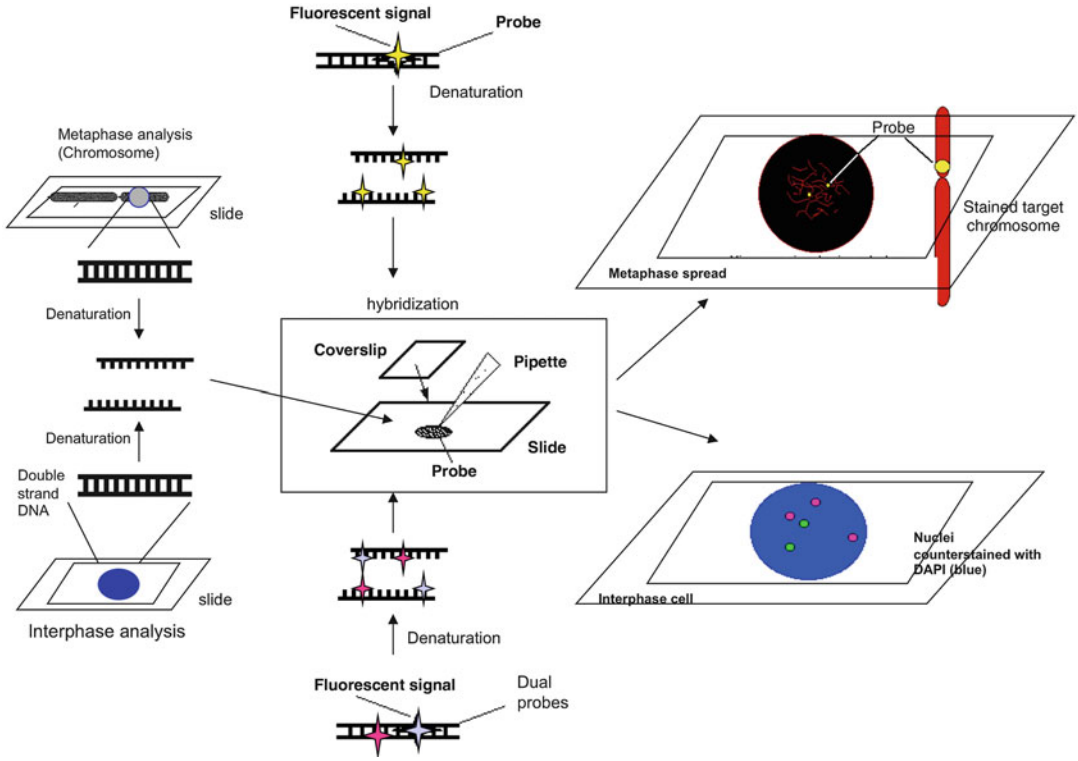


Fig. 33.4 Schematic representation of the FISH technique (interphase and metaphase analysis)

genes with multiple fusion partners (e.g., *MLL* at chromosome 11q23) or altered genes that have widely scattered break points (e.g., *MYC* in Burkitt lymphoma [BL])

- Conventional PCR detection methods include
  - Agarose/polyacrylamide gel electrophoresis, with or without probe hybridization
  - Capillary electrophoresis, where one primer is labeled with a fluorochrome
- Quantitative real-time PCR (qPCR)
  - The two most commonly used methodologies, TaqMan (Applied Biosystems, Foster City, CA) and LightCycler (Roche Molecular Systems, Pleasanton, CA), have wide dynamic ranges making them useful for minimal residual disease (MRD) assessment
  - MRD assessment of *BCL2*/*IGH* and *CCND1*/*IGH* rearrangements can use qPCR
- Mutation detection methodologies used in lymphomas
  - Direct sequencing using dideoxy chain termination methods
    - Sensitivity is approximately 1 in 5 cells bearing the mutation
  - Pyrosequencing: sequencing by synthesis method
    - Sensitivity is approximately 1 in 10 cells bearing the mutation
- Fluorescence in situ hybridization (FISH)
  - This technique involves the hybridization of DNA-specific probes onto interphase and metaphase chromosomal DNA (Fig. 33.4)
- Comparative genomic hybridization (CGH): CGH is a double-color hybridization procedure that provides a genomic-wide view of DNA copy number changes
  - In the classical CGH approach, total genomic DNA is isolated from test and reference cell populations, labeled with

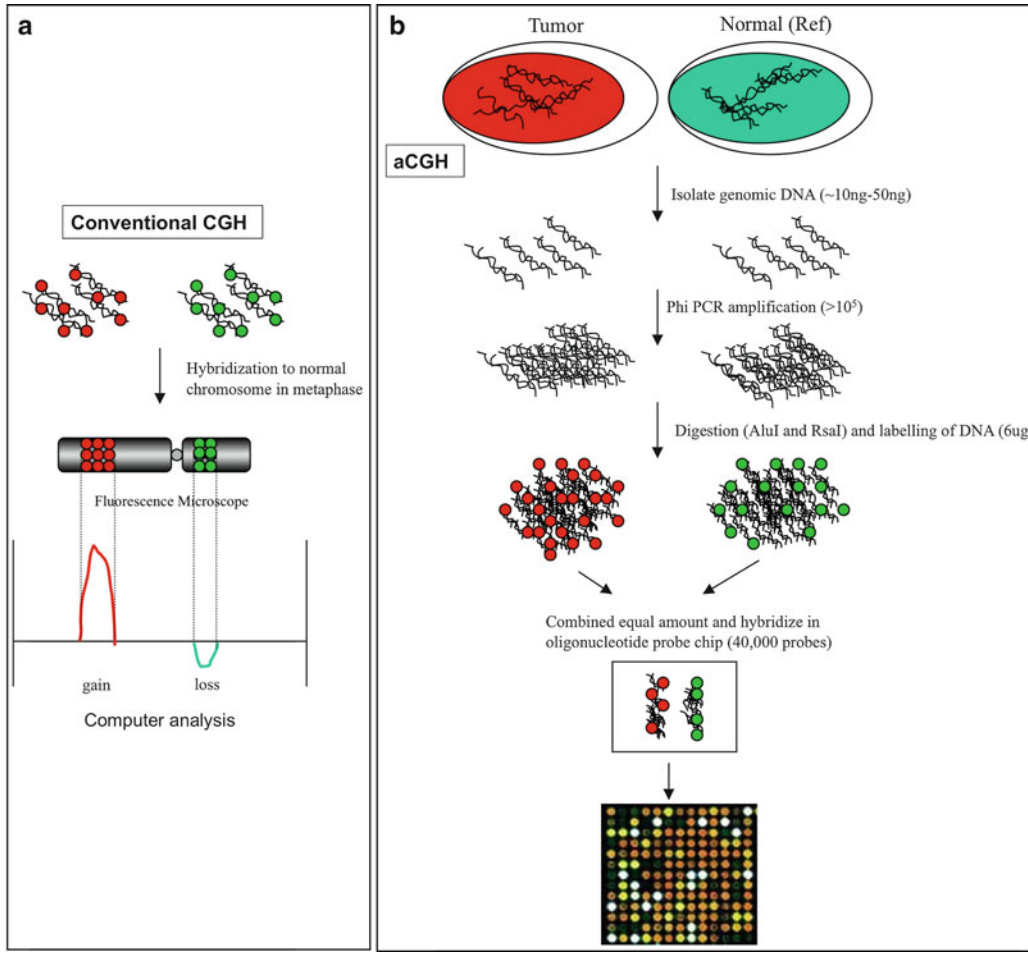


Fig. 33.5 Schematic representation of the conventional CGH (**a**) and array-CGH (**b**) methods

different fluorochromes (i.e., green for normal DNA and red for tumor DNA), and then hybridized with metaphase chromosomes (Fig. 33.5a)

– If the tumor sample has a gain in a particular chromosomal region, a predominance of tumor DNA (red signal) will be detected, and if there is a lost, a predominance of reference DNA (green signal) will be detected
– An oligonucleotide array-based CGH approach with higher resolution and accuracy has been developed (Fig. 33.5b)
– Oligonucleotide array probes can be designed in silico for any sequenced region of the genome (custom arrays), thus allowing region-specific coverage with a high number of probes and making the array disease specific
– Another advantage of the array-based CGH is that it allows the detection of genomic imbalances at the level of single genetic loci (Fig. 33.6). This methodology is frequently used for research, but it is being implemented in some clinical laboratories

33.2.3.1 B Cell and T Cell Clonality by Southern Blot

• Principle: in lymphocytes, the process of antigen receptor gene rearrangement would delete

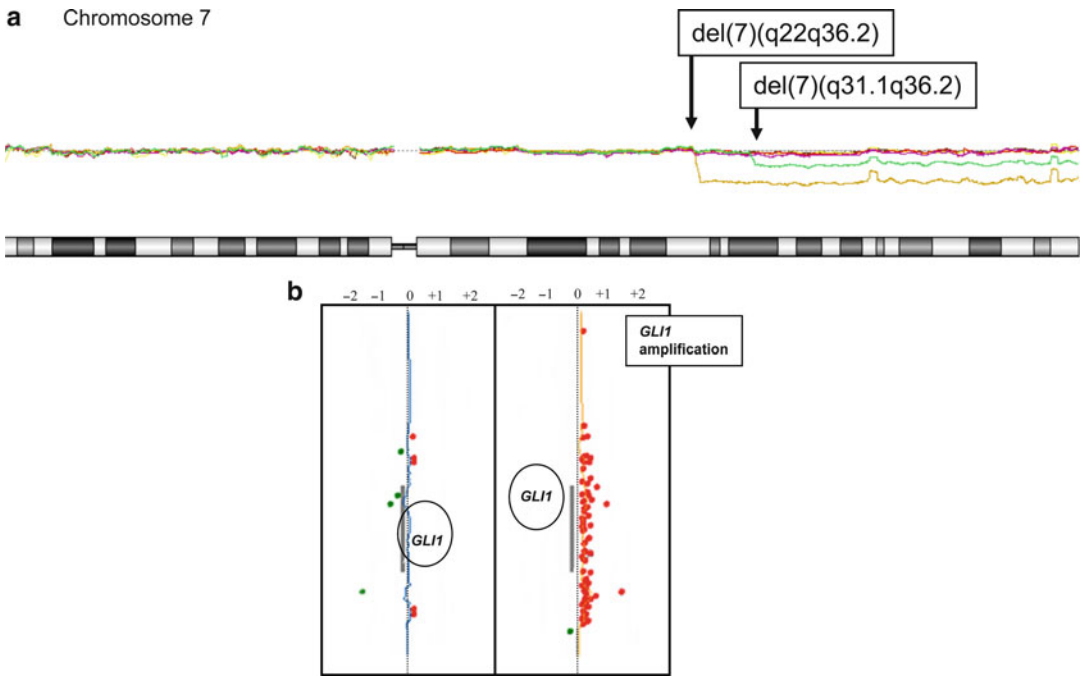


Fig. 33.6 (**a**) Shows genomic imbalances in chromosome 7 detected by aCGH, in a series of patients with splenic marginal zone lymphoma (SMZL). Two cases showed long deletions involving the long arm of chromosome 7: in one case, the deletion starts at 7q22 (*yellow line*), and in the other, it starts at 7q31 (*green line*). No major imbalances were seen in the rest of cases. aCGH can also detect genomic imbalances at the level of single genetic loci. (**b**) Shows that one of the SMZL (*right*) has amplification of the 12q13.2 region. The array used contained a high density of probes covering the region where the proto-oncogen *GLI1* is located. *Red dots* indicate a significant gain for each single probe, and *green dots* indicate a significant loss. Note that the other case (*left*) shows no amplification of *GLI1* gene

one or more restriction sites and thus alter the size of detected fragments (Fig. 33.7)

- The clonality of nearly all T cell and B cell tumors can be determined using probes specific for the *IGH*, *IGK*, and *TCRB* genes. In general, J region probes are superior to C region probes
- If a polyclonal lymphoid population is present, altered fragments would be of different sizes and thus detected as a smear on gel
- A monoclonal population would show discrete bands different from the germline (unrearranged) bands
- For the *IGH* and *TCR* loci, Southern blot should show altered bands in at least two of three restriction digests to be considered positive for a monoclonal rearrangement

- Sensitivity: 1–5% of the total cells (i.e., 1–5 clonal cells in a population of 100 cells)
- Advantages
  - This technique examines the entire gene and should thus find virtually any clonal rearrangement if it is present
  - Partial D-J rearrangements can be detected (common in lymphoblastic leukemia [LL])
  - Clones can be detected in tumors showing extensive V_H somatic mutation
- Disadvantages
  - Laborious, time-consuming, and requires significant amounts of nondegraded DNA
  - Insufficient sensitivity for assessing clonality in tumors with few neoplastic cells or for assessing MRD after therapy

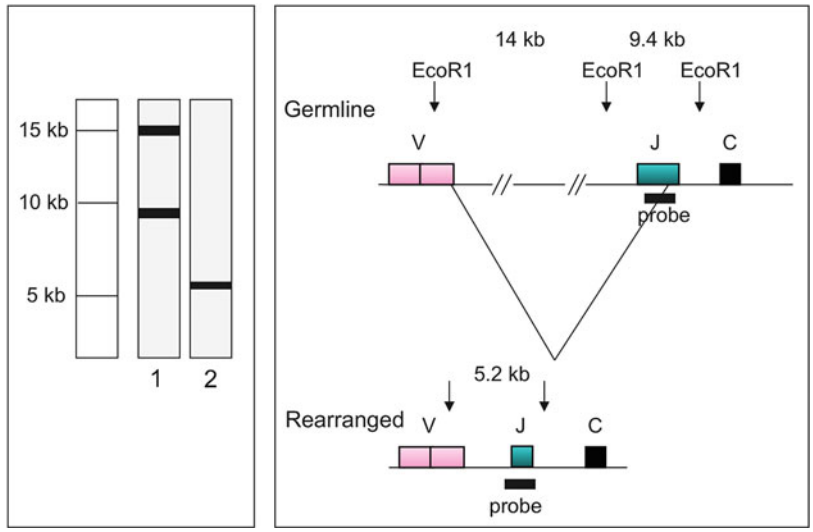


Fig. 33.7 Southern blot analysis of the *IGH* gene. (*Left*) *VDJ* gene rearrangement deletes an *Eco*RI restriction enzyme site changing the size of the DNA fragment from 9.4 to 5.2 kb. (*Right*) Two specimens are shown, with genomic DNA digested with the *Eco*RI restriction enzyme, run on an agarose gel, transferred to a nylon membrane, and hybridized with a radioactively labeled JH probe. *Specimen 1* reveals only the germline bands. *Specimen 2* has evidence of clonal *IGH* gene rearrangement, as evidenced by loss of one band and a shift in the remaining band size

  - Partial restriction digest due to poor quality sample or DNA methylation may lead to spurious results

33.2.3.2 B Cell and T Cell Clonality by PCR

- Principle: in the germline configuration, the discontinuous V and J regions are widely separated, precluding PCR amplification. In contrast, a rearranged gene has V and J regions close together allowing amplification
  - Monoclonal lymphocytes contain 1 or 2 rearranged antigen receptor alleles; therefore, PCR will amplify 1 or 2 predominant bands
  - Polyclonal cells each carry a distinctive gene rearrangement of slightly different sizes due to CDR3 diversity, which will show a Gaussian/normal distribution PCR size range by capillary electrophoresis or a smear pattern by gel electrophoresis
- Sensitivity: using conventional PCR, *IGH*, and *TCRG* PCR can identify 1 monoclonal B or T cell in up to 10^2–10^4 B or T cells depending on the number of polyclonal lymphocytes present in the sample
  - Detection of products by agarose or polyacrylamide gel electrophoresis with ethidium bromide staining has a sensitivity of approximately 1–10%. Capillary electrophoretic analysis of fluorescent PCR products allows for approximately one-log increase in sensitivity
  - PCR assays designed to detect chromosomal translocations are much more sensitive and can identify a single tumor cell among 10^5 cells
- Advantages: rapid and can be performed on poor-quality DNA

33.2.3.3 IGH PCR

- Forward PCR primers are derived from conserved sequences in the framework regions of the V segments (frameworks 1, 2, and 3)
- Reverse PCR primers are derived from a conserved sequence within the J_H region

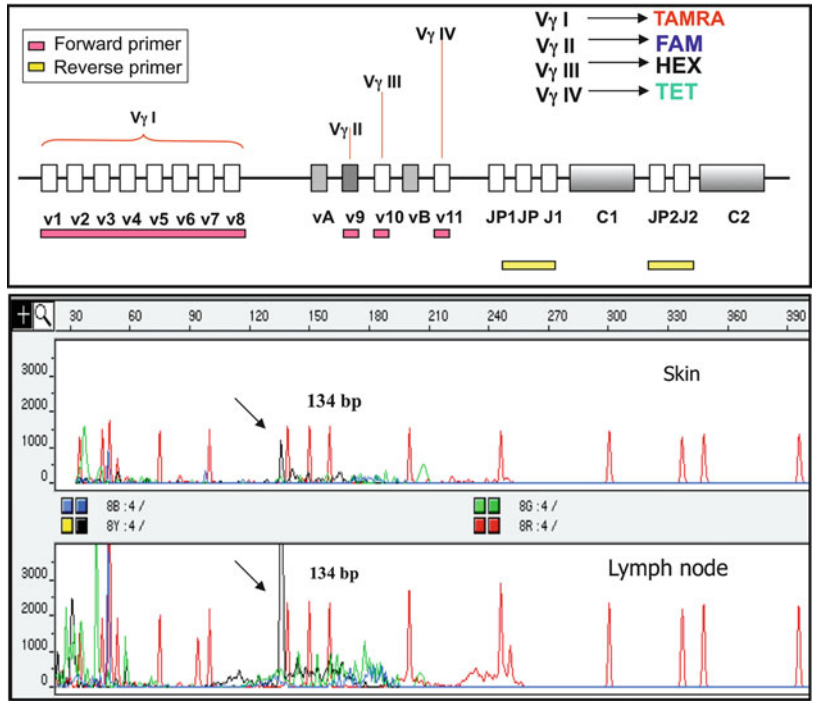


Fig. 33.8 *TCRG* PCR using multicolor capillary electrophoretic detection. (*Top*) In this four-color *TCRG* PCR assay, each Vγ family primer is labeled with a different fluorescent dye. If a VγI segment is used in the tumor *TCRG* rearrangement, a *red peak* will appear in the electropherogram. If VγII, III, or IV is used, *blue*, *black*, and *green peaks* will appear, respectively. The monoclonal population using the VγIII primers (*black peak*, 134 bp) is detected by capillary electrophoresis in two samples from lymph node and skin in a patient with a cutaneous T cell lymphoma. Regularly spaced *red peaks* are size standards

- Disadvantages: B cell tumors with a high intrinsic rate of somatic mutation, such as follicular lymphoma (FL) and plasma cell myeloma, may be undetected by *IGH* PCR in up to 40% of cases due to failure of primers to bind to mutated sequences

33.2.3.4 TCRG and TCRB PCR

- *TCRG* represents the most useful marker for T cell clonality, and *TCRG* PCR detects >90% of clonal T cell neoplasms
- The *TCRG* gene is rearranged at an early stage and has 11 V segments that can be grouped into four homologous families and 5 J segments that can be grouped into two highly homologous groups (Fig. 33.8)
  - Forward PCR primers usually include four consensus primers from conserved sequences within the several VG families. False-negative results are rare
  - Reverse PCR primers usually include two to four consensus primers from conserved sequences within the JG regions
- Disadvantages: limited V-D-J diversity of the *TCRG* gene may give rise to false-positive results for T cell tumors
  - Reactive oligoclonal T cell proliferations are common in blood and bone marrow and are another source of false-positive results
- *TCRB* PCR requires a large number of primers owing to the complexity of *TCRB* locus, which has 75–100 V segments, 2 D segments, and 13 J segments. Multiplex PCR assays have now been developed to assess the *TCR* β chain gene successfully with a very low false-negative rate
- *TCRD* PCR is useful for LL MRD analysis and can be used to support γ/δ-lineage for T cell tumors but is rarely used in routine practice

33.2.3.5 FISH Studies

- The ability to localize fluorescent signals to specific interphase nuclei in nondividing cells is a principal advantage of FISH
- Two strategies can be used to detect chromosomal translocations, break-apart and fusion strategies (Figs. 33.9a and b)

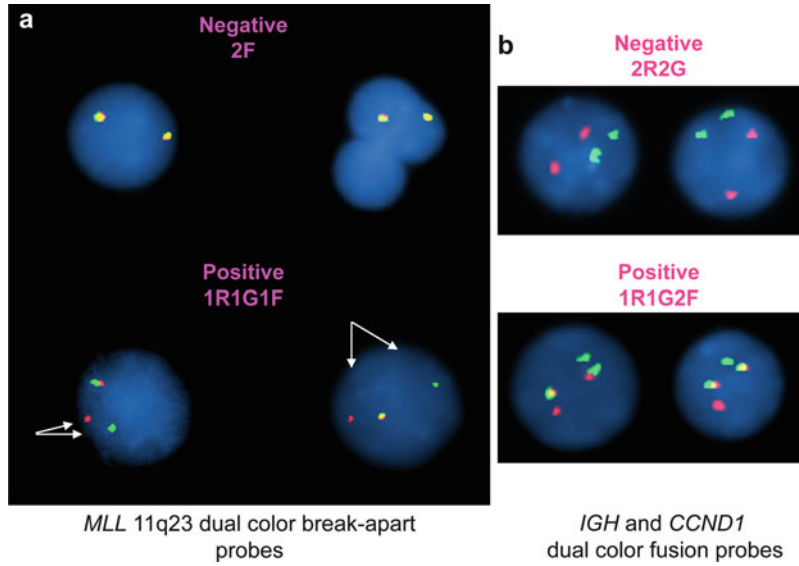


Fig. 33.9 (**a**) Example of dual color (*red*/*green*) break-apart probes. The *MLL* locus at both chromosomes is intact in the two depicted upper nuclei, and two overlapping pairs of *red* and *green* signals are seen (yielding two *yellow signals*). If the *MLL* locus is involved in a chromosomal translocation, the cells carrying the translocation (lower nuclei) will show the expected one *red*, one *green*, and one *red*/*green* fusion signal pattern (1R1G1F). (**b**) Example of a dual color (*red*/*green*) fusion probe to detect t(11;14). The two upper nuclei show no evidence of rearrangements between *IGH* (at 14q32) and *CCND1* (at 11q13) genes; thus, two *red* and two *green signals* are seen per nuclei (normal pattern). The two lower interphase nuclei show the abnormal 1R1G2F signal pattern indicative of the presence of the t(11;14)

- With the break-apart strategy, probes labeled with two different fluorescent chromophores, often orange/red and green hybridize to sequences that flank a known chromosomal breakpoint region
  - When a translocation occurs, the signals are split, and separated orange/red and green signals are detected along with the remaining intact yellow signal representing the nonrearranged locus
- With the fusion strategy, two probes are used (labeled with two distinct chromophores)
  - The probes are targeting two loci that are known to recombine
  - When the translocation occurs, the signals are fused generating nuclei with three or four signals, one or two yellow (indicating the recombination product) and one red and one green signal (corresponding to the other two loci not involved in the translocation)
- FISH methods can be applied easily to touch preparations, fresh tissue, karyotype preparations, frozen specimens, and formalin-fixed paraffin-embedded samples

33.3 Clinical and Molecular Genetic Features of Specific Lymphoid Malignancies

33.3.1 Immature B Cell and T Cell Leukemia/Lymphomas

33.3.1.1 Lymphoblastic Leukemia/Lymphoma (LL/LBL)

- Clinical features: wide age range but more common in children
  - B cell LL/LBL primarily involves the peripheral blood/bone marrow, with skin and lymph node being the most common sites of extramedullary lymphomatous involvement

  - T cell LL/LBL frequently presents as a mediastinal mass, with variable marrow and blood involvement
  - Therapy is multiagent "induction" chemotherapy followed by consolidation/maintenance chemotherapy
    - Imatinib is used for maintenance in Ph+ LL/LBL
    - Rituximab (anti-CD20 therapeutic antibody) used for CD20+ LL/LBL
  - Common locations of relapse include sanctuary sites like central nervous system (CNS) and testes
- Pathological features: blasts have fine nuclear chromatin and a high mitotic rate, with tumor cell infiltration of tissues in a single-file fashion
- Immunologic features: LL/LBL express markers of immaturity including TdT (90%) and CD34
  - B cell LL/LBL can be divided into the pre-B type that expresses CD10 (CALLA), a more immature pro-B cell type, and a more mature CD20+ type that overlaps immunophenotypically with Burkitt lymphoma (BL)
  - T cell LL/LBL can be subclassified into immature tumors that lack CD3, CD4, CD5, and CD8, as well as the more common "double-positive" (CD4+ CD8+) CD1a+ cases and more mature CD4+ or CD8+ "single-positive" tumors that lack CD1a
- Molecular features: the presence of defining reciprocal chromosomal translocations is currently used to define B-LL/LBL subtypes in the World Health Organization (WHO) classification (Table 33.1). Those translocations define distinct entities with distinctive clinical or phenotypic characteristics and prognostic implications. Hyperdiploid and hypodiploid subtypes are also recognized
  - Detection of most LL/LBL translocations can be accomplished by dual probe FISH
  - t(12;21)(p13;q22) *ETV6–RUNX1* fusion is frequently a cryptic translocation, and its detection requires FISH or RT-PCR methods
    - Break-apart FISH probes have also been used for loci with multiple translocation partners such as the *MLL* gene at chromosome 11q23
  - Several multiplex PCR assays have also been developed that detect up to 10 common LL/LBL translocations in a single amplification reaction
  - Because LL/LBL tumors continuously express the machinery for V(D)J recombination (i.e., Rag 1/2 and TdT), many can exhibit *lineage infidelity* in which *IGH* and *IGK* are clonally rearranged in precursor T-LL/LBL and *TCRG* and *TCRB* are rearranged in precursor B-LL/LBL
  - In B-LL/LBL, additional genetic alterations that are not currently incorporated in the category of B-LL/LBL with recurrent genetic abnormalities have been described
    - The *E2A-HLF* oncoprotein arisen from the t(17;19) in B-LL/LBL
    - Intrachromosomal amplification of chromosome 21 (iAML21) in ~2% of the childhood B-LL cases without recurrent genetic abnormalities is defined by multiple copies of *RUNX1* and *miR-802* and is associated with poor prognosis
    - Alterations in genes that regulate B cell development are seen in more than 60% of B-LL/LBL. These include alterations involving *PAX5*, *EBF1*, and *IKZF1*. Deletions or mutations of *IKZF1* that encodes the transcription factor Ikaros are associated with very poor outcome. Deletions of *IKZF1* are frequently present in *BCR–ABL1*-positive (PH+) B-LL/LBL
    - High frequency of somatic alterations in *TP53*/*RB* tumor suppressor pathway (54%), RAS signaling (50%), and Janus kinases (11%), in newly diagnosed high-risk pediatric B-LL/LBL
  - Abnormal karyotype is found in 50–70% of the T-LL/LBL (Table 33.2)

Table 33.1 Recurrent genetic abnormalities in B-LL (2008 WHO classification)

Translocation	Fusion product	Frequency (%)	Prognosis	Immunophenotype
t(9;22)(q34;q11.2)	*BCR–ABL1*	25% in adults 2–4% in children	Poor in both adults and children	CD25+, CD117–, CD13+, CD33+, CD10+, CD19+, TdT+, CD34+, CD9+
t(v;11q23) t(4;11)(q21;q23) t(19;11)(p13;q23) t(9;11)(p22;q23)	*MLL* *AF4–MLL* *ENL–MLL* *AF9–MLL*	Most common Also seen in T-ALL Common in AML	Poor	CD19+, CD10–, CD24–, CD15/65+, CD9+
t(12;21)(p13;q22)	*TEL–AML1* (*ETV6–RUNX1*)	25% of cases, common in children, only 3% adults	Favorable	CD34+, CD13+, CD9–
Hyperdiploidy >50 chromosomes (most common +21, +X, +14, +4)	–	25% of cases, common in children	Favorable	CD34+, CD45-
Hypodiploidy <46 chromosomes	–	1%	Poor	
t(5;14)(q31;q32)	*IL3–IGH*	<1%	Unknown	Eosinophilia
t(1;19)(q23;p13.3)	*E2A–PBX1* (*TCF3–PBX1*)	6%	Improved with current intensive therapies	CD9+ strong, CD34–, cμ heavy chain

Abbreviations: *BCR* breakpoint cluster region, *ABL1* Abelson murine leukemia viral oncogene holog 1, *MLL* mixed lineage leukemia, *AF4* AF4/FMR2 family, member 1, *ENL*, *MLLT1* myeloid/lymphoid/leukemia or mixed lineage leukemia translocated to 1, *AF9*, *ALL-1* fused gene from chromosome 9, *TEL* (*ETV6*) translocation, *ETS* leukemia, *AML1* (*RUNX1*) acute myeloid leukemia 1 gene (RUNT-related transcription factor), *IL-3* interleukin 3, *IGH@* immunoglobulin heavy chain, *E2A* (*TCF3*) immunoglobulin enhancer-binding factor E12/E47 (transcription factor 3), *PBX1* pre-B leukemia transcription factor 1

- The most common recurrent cytogenetic abnormalities include translocations that result in upregulation of transcription of the translocated gene by juxtaposition with the regulatory region of one of the *TCR* loci
- The genes most commonly involved include the transcription factors *HOX11 (TLX1)* at 10q24, *HOX11L2 (TLX3)* at 5q35, *MYC* at 8q24.1, *TAL1* at 1p32, *LMO1 (RBTN1)* at 11p15, *LMO2 (RBTN2)* at 11p13, *LYL1* at 19p13, and *LCK* at 1p34.3–35
- Other translocations include the t(10;11), *PICALM–MLLT10*, and translocations involving *MLL*, the most frequent partner being *ENL* at 19p13
- Additional genetic aberrations that result in activation or deregulation of signaling pathways include:
  - 50% of T-LL/LBL is characterized by genetic alterations of *NOTCH1*. The t(7;9) juxtaposes the *NOTCH1* gene next to the *TCRB*. The Notch1 pathway can also be deregulated by missense mutations of *hCDC4* gene, negative regulator of *NOTCH1*
  - Deletion of *CDKN2A*/*p16INK4A* at 9q in 30% of the cases
  - WT1 mutations have also been described in 11% of the pediatric and adult T-LL/LBL cases associated with aberrant HOX expression
- Gene expression profiling studies have identified groups of LL/LBL that correspond to different stages of T cell development: overexpression of *LYL1* corresponds to pro-T cell LL/LBL, overexpression of *HOX11* to cortical stage, and overexpression of *TAL1* to late cortical thymocyte stage

Table 33.2 Common genetic aberrations in T-LL/LBL

Gene(2) involved	Rearrangement/ other genetic alteration	Frequency(%)	Prognosis
TLX1/HOX11[a]	t(10;14)(q24;q11)	8	Good
	t(7;14)(q34;q24)		
TLX3/HOX11L[a]	t(5;14)(q35;q32)	24	Undefined[d]
HOXA[a]	t(7;7)(p15;q34)	3	Undefined
	inv(7)(p15q34)		
TAL1[a]	t(1;14)(p32;q11)	4	Good
	t(1;7)(p32;q34)	15	
	1p32 deletion		
TAL2[a]	t(7;9)(q34;q32)	1–2	Unknown
LMO1[a]	t(11;14)(p15;q11)	1–2	Unknown
	t(7;11)(q34;p15)		
LMO2[a]	t(11;14)(p13;q11)	6	Unknown
	t(7;11)(q34;p13)	3	
	11p13 deletion		
LMO3[a]	t(7;12)(q34;p12)	<1	Unknown
LYL1[a]	t(7;19)(q34;p13)	2	Poor
CALM-AF10[c]	t(10;11)(p13;q14)		Poor
MLL-ENL[c]	t(11;19)(q23;p13)	2–5	Unknown
EML1-ABL1	t(9;14)(q34;q32)	<1	Unknown
ETV6-ABL1	t(9;12)(q34;p13)	<1	Unknown
BCR-ABL1	t(9:22)(q34;q11)	<1	Poor
ETV6-JAK2	t(p24;p13)	<1	Unknown
SET-NUP214	9q34 deletion	<1	Unknown
NUP214-ABL1	9p34 amplification	4	Poor
MYB	t(6;7)(q23;q34)	3	Unknown
	Duplication	8–15	Unknown
CDKN2A/2B	9p21 deletion, mutation or hypermethylation	70	Unknown
NOTCH1	t(7;9)(q34;q34)[a]	<1	Good initial response
	Activating mutation[b]	50–60	Poor
hCDC4	Missense mutation	30	Unknown
FLT3	Activating mutation	2–4	No impact
NRAS/KRAS	Activating mutation	4–10	Unknown
PI3K	Activating mutation	7	No impact
PTEN	10q23.31 deletion or	8	Undefined[e]
	inactivating mutation	17–27	
NF1	17q11.2 deletion	3	Unknown
JAK1	Activating mutation	20	Unknown
WT1[f]	Frameshift mutations	11–13	No impact

[a]Translocations involving TCR loci (TCRA and TCRD, 14q11; TCRB, 7q35; and TCRG, 7p15). In many cases, the translocations are not detected by conventional cytogenetics but my molecular studies

[b]Approximately 50% of the cases show activating mutations in the extracellular heterodimerization domain (HD) and/or C terminal proline, glutamic acid, serine, threonine (PEST) domain

[c]Neither of these translocations is specific; the CALM-AF10 occurs in AML and the MLL-ENL occurs in B-LL

[d]Undefined prognostic significance indicates conflicting results that have shown either poor or good outcome or no impact in prognosis

[e]Contradictory results that indicate poor outcome or no impact in prognosis

[f]WT1 mutations are most prominently found in T-LL cases with aberrant rearrangements of the oncogenic TLX1, TLX3, and HOXA transcription

- Ongoing recombination can also occur in LL/LBL with generation of new clonal *IGH* or *TCR* products over the course of disease, or at relapse

33.3.1.2 Burkitt Lymphoma

- Clinical features: BL accounts for 30% of childhood lymphomas but is rare in adults, with three different clinical variants including

  - Endemic form: commonly observed in equatorial Africa, in children with frequent involvement of the jaw and kidneys
  - Sporadic form: 1–2% of all adult lymphomas in Western Europe and the United States. Most patients present with abdominal involvement
  - Immunodeficiency associated: observed in the setting of HIV infection and is usually extranodal disease
  - Unlike other HIV-related lymphomas, BL is frequently noted in patients with CD4 counts exceeding 200 cells/μL
  - BL is an extremely chemosensitive malignancy, with current overall response rates of 50–70% in adults
  - As CNS involvement is common, CNS prophylaxis (intrathecal chemotherapy) is standard
- Pathological features: medium-sized lymphocytes with moderately abundant basophilic cytoplasm, with cytoplasmic lipid vacuoles noted on smears
  - Characteristics include a very high mitotic index and a *starry sky* growth pattern, produced by macrophages ingesting apoptotic tumor cells
- Immunologic features: CD19+ CD20+ B cells that express CD10, BCL6, and sIg but lack BCL2 and TdT
  - The cellular proliferation rate as measured by Ki67 approaches 100%
  - The subcellular localization of MYC protein expression as detected by immunohistochemistry predicts the presence of *MYC* translocations
- Molecular features: BL is characterized by chromosomal translocations that overexpress the *MYC* gene at 8q24 through insertional activation by the Ig enhancer sequences
  - The *IGH* gene at 14q32 is the most frequent breakpoint partner (75–85% of cases), with t(2;8) and t(8;22) variant translocations utilizing the *IGK* and *IGL* enhancers seen in the other cases
  - On the cytogenetic level, around 70% of BL has only one secondary alteration. Using 2.8 K array CGH, the median number of imbalances per patient was only two
  - Gains or amplifications of 1q (∼20%) and gains in 7q (∼10%) and an increasing genetic complexity have been reported associated with a poor clinical outcome
  - Given the wide range of breakpoints, *MYC* alterations are usually detected by conventional cytogenetic methods or by FISH rather than by PCR (Fig. 33.10). Southern blot can also be used. The location of different *MYC* breakpoints correlates with clinical patterns of disease
  - Using a gene expression profiling approach, independent groups have been able to recognize specific gene expression signatures for BL that identify almost the very same lymphomas. The cases that carry this signature are described as molecular BL (mBL) and can be clearly differentiated from non-mBL, a group mainly composed of diffuse large B cell lymphoma (DLBCL)
  - All three variants of BL demonstrate similar microRNA profiles that are also distinct from DLBCL. This microRNA signature includes microRNAs that are regulated by *MYC* gene or are associated with NFκB pathway
  - *TP53* mutations have been reported in ∼30% of sporadic BL

33.3.2 Mature B Cell Lymphomas

- B cell lymphoproliferative disorders are largely classified according to a combination of clinical, morphologic, immunophenotypic, and molecular findings
  - This approach is reflected in the current WHO classification
  - Earlier lymphoma classifications were primarily based on the morphological features of a tumor (e.g., "well-differentiated" small cell)
- Lymphoma clinical staging does not utilize the TNM system used for other malignant

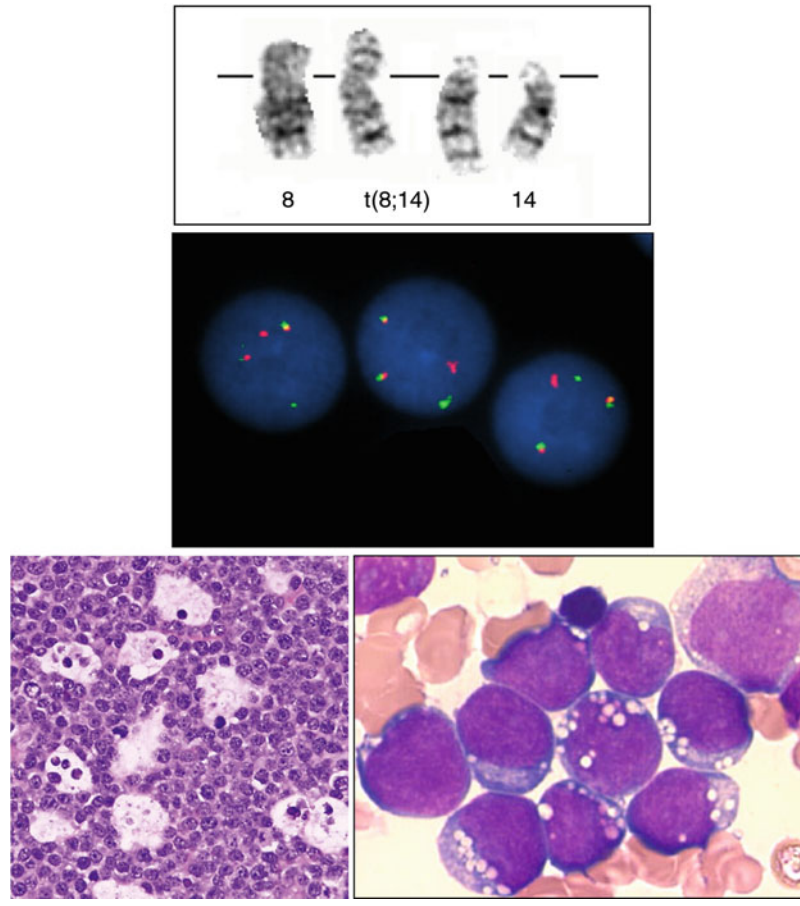


Fig. 33.10 *MYC* translocation. Burkitt lymphoma. t(8;14)(q24; q32) G-banding (*top*). FISH methods using *MYC*/*IGH* dual color fusion probes (*middle*). The presence of dual fusion signals indicates the presence of the t(8;14). Classical morphological features of BL with the classical starry sky pattern (*bottom left*). Typical blasts of BL with deeply basophilic and vacuolated cytoplasm (*bottom right*)

neoplasms but is based on the pattern of nodal and/or extranodal involvement

- The Ann Arbor/American Joint Committee on Cancer (AJCC) clinical staging criteria include
  - Stage I: localized nodal or extranodal (E)
  - Stage II: groups of lymph nodes on same side of diaphragm or an extranodal site and adjacent lymph node
  - Stage III: different groups of lymph node on both sides of the diaphragm
  - Stage IV: bone marrow involvement or extensive extranodal dissemination
- Most regimens for B cell lymphomas currently include rituximab, which is a recombinant humanized monoclonal antibody against the pan-B cell marker CD20 or other therapeutic monoclonal antibodies (anti-CD19, -CD22, and -CD30)
- Many B cell lymphoma/leukemia types have defining chromosomal translocations, which arise early in the disease course and can be used for diagnosis and molecular monitoring for MRD
- Molecular monitoring is complementary and in some cases superior to multiparameter flow cytometric analysis
- For B cell tumors that have characteristic chromosomal translocations, highly sensitive molecular MRD tests have been developed
- For other B cell tumors, *IGH* PCR remains the mainstay of molecular testing for confirmation at diagnosis and for MRD

33.3.2.1 Chronic Lymphocytic Leukemia (CLL)/Small Lymphocytic Lymphoma (SLL)

- CLL is a common leukemia in the Western world and is characterized by a progressive accumulation of small, mature-appearing B cells in the blood, bone marrow, and secondary lymphoid tissues
- Clinical features: CLL/SLL occurs predominantly in middle-aged and elderly persons, and its incidence increases with age
  - Rai and Binet staging systems are used to estimate prognosis based on sites of disease and degree of suppression of platelet and red blood cell counts
  - Median survival in CLL/SLL varies between 2 and >10 years, depending on the stage
  - Decision to treat a patient with CLL/SLL is based on a combination of clinical staging, the presence of symptoms, and molecular features
  - Parameters associated with aggressive disease independent of the disease stage include
    - Elevated serum beta2-microglobulin
    - Expression of the tyrosine kinase ZAP70
    - Expression of CD38 in >30% of tumor cells
    - Short lymphocyte doubling time (<6 months)
    - Elevated serum levels of soluble CD23
    - Elevated serum thymidine kinase activity
    - Absence of somatic hypermutations
    - Deletions in chromosome regions 11q23 and 17p
  - About 10% of patients with CLL have an associated Coombs-positive autoimmune hemolytic anemia
- Pathological features: small lymphocytes with minimal cytoplasm and clumped nuclear chromatin that infiltrate the bone marrow in an interstitial, nodular, or diffuse pattern
  - Cases with lymph node involvement and $<5 \times 10^9$/L lymphocytes in marrow or blood are classified as SLL rather than CLL
  - Proliferation centers are areas with prolymphocytes and paraimmunoblasts that represent areas of tumor proliferation. Proliferation centers are characteristic of CLL/SLL. Increasing number of *prolymphocytes* in blood and/or bone marrow represent the most common pattern of disease progression
- Immunologic features: CD19+ CD20+ (dim) B cells that coexpress CD5, CD43, and CD23 and are negative for CD10
  - Atypical CLL cases may show dim or absent CD5 and/or loss of CD23
- Molecular features: CLL lacks reciprocal chromosome alterations in nearly all cases but can be grouped into molecular subtypes based on their pattern of chromosomal deletions or additions (unbalanced cytogenetic abnormalities, 80% of the cases)
  - The most common genetic abnormalities are deletion of 13q14 (50%), trisomy 12 (20%), del11q (10–20%), isochromosome 17q/–17 or del17p (10%), and del6q (5%)
    - Multiprobe FISH panels to detect these loci on interphase or metaphase chromosomal spreads have become the most common method of screening
    - Conventional cytogenetic analysis typically misses some of these abnormalities (particularly del13q14) and is therefore most useful to assess for clonal evolution
  - Assessment of the degree of somatic hypermutation of the V_H gene in the CLL clone has been shown to have prognostic significance
    - CLL with unmutated V_H genes, defined as <2% bp changes compared with germline V_H, requires treatment earlier in the disease course than CLL with mutated V_H
    - ZAP-70 expression is a useful surrogate for unmutated IGH status, but discordances occur in 15% of cases
- A subset of patients has chromosomal translocations (~30% of cases after stimulation of CLL cells with CD40 ligand or with combination of CpG-oligodeoxynucleotides and IL2).

The presence of these translocations independently correlated with aggressive disease
- CLL with some specific translocations has distinctive morphological, immunophenotypic, and genetic features, i.e., t(14;19) involving *IGH* and *BCL3* loci
- Specific microRNA genes (miR15 and miR16), mapped to 13q14, are downregulated or deleted in the majority of CLL cases
- Four genes have been found recurrently mutated in CLL: *NOTCH1*, exportin 1 (*XPO1*), myeloid differentiation primary response gene 88 (*MYD88*), and kelch-like 6 (*KLHL6*)
  - Mutations in *MYD88* and *KLHL6* are detected in cases of CLL with mutated immunoglobulin genes, whereas *NOTCH1* and *XPO1* mutations are mainly detected in patients with unmutated immunoglobulins
  - *NOTCH1*, *MYD88*, and *XPO1* mutations are oncogenic changes that contribute to the clinical evolution of the disease

33.3.2.2 Mantle Cell Lymphoma (MCL)

- An aggressive lymphoma bearing the t(11;14) chromosomal translocation that activates the cyclin D1 (*CCND1*) gene through the *IGH* enhancer
- Clinical features
  - MCL frequently involves the gastrointestinal (GI) tract and bone marrow
    - MCL accounts for approximately 5–10% of non-Hodgkin lymphomas with a predominance in elderly males
    - Current therapies currently include intensive chemotherapeutic regimens, such as HyperCVAD, and have dramatically improved survivals in recent years
      - High relapse rates in MCL support the use of intensive molecular monitoring
    - Novel effect therapeutic strategies include radiolabeled antibodies and proteasome inhibitors (bortezomib; PS-341, Velcade®)
    - A splenic/leukemic variant of MCL is defined, which may morphologically overlap with B cell prolymphocytic leukemia
- Pathological features: classically, medium-sized lymphocytes, with mantle zone, nodular, or diffuse patterns of nodal infiltration
  - Blastoid variant with fine nuclear chromatin and increased mitotic rate has adverse outcome and complex molecular aberrations
  - Pleomorphic variant with large cells also has poor prognosis
    - The proliferation rate as determined by Ki67 stain may predict outcome
- Immunologic features
  - CD19+ CD20+ B cells that express CD5 but lack CD23 in contrast to CLL
  - Strong nuclear expression of cyclin D1 is characteristic
  - Rare phenotypic variants may have absent/dim CD5
  - SOX11, a neuronal transcription factor, is positive in MCL as well as B and T lymphoblastic leukemias including a subset of Burkitt lymphoma and hairy cell leukemia
- Molecular findings
  - The genetic hallmark of MCL is the (11;14)(q14;q32) translocation leading to overexpression of cyclin D1 protein through the *IGH* enhancer, which can be detected by FISH or PCR
  - For diagnosis, dual probe FISH to detect fusion signals is preferred and can detect t(11;14) in up to 95% of the cases
  - Approximately 35% of t(11;14) translocations cluster within an 80–100 bp region of the *CCND1* gene on chromosome 11q13 known as the major translocation cluster (MTC) (Fig. 33.11). The remainder of chromosome 11 breakpoints is widely scattered over approximately 120 kb
    - In the MTC+ group, qPCR for t(11;14) can be used for MRD monitoring
    - *IGH* PCR can be used for MRD detection
  - In rare cases, typical MCL without t(11;14) may occur by activation of other cyclin genes.

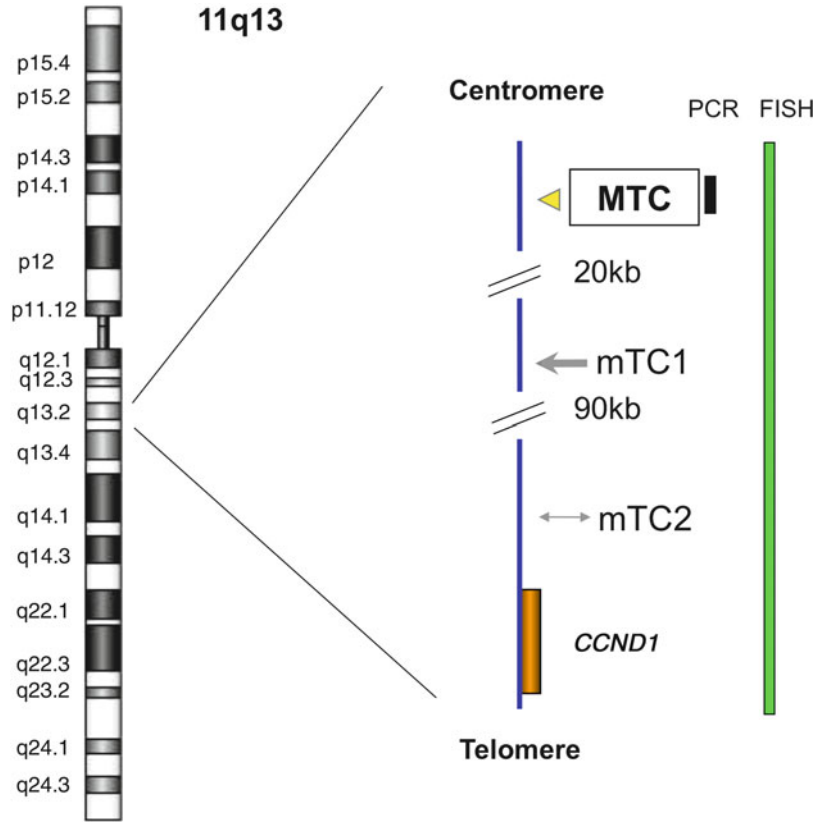


Fig. 33.11 Schematic representation of the chromosome 11q13 region involved in t(11;14) in MCL. The cyclin D1/*CCND1* gene is indicated in an *orange box*. The MTC region accounts for 30–40% of all translocations. Minor breakpoint clusters are indicated as mTC1 and mTC2. The range of detection of this translocation using PCR and FISH is indicated as *black* and *green rectangles*

These cases are cyclin D1-negative and positive for SOX11 by immunohistochemistry
  - *CCND1* gene amplification without translocation has been documented in cases of plasma cell myeloma but not in MCL
  - 15–40% of MCL carries VH somatic hypermutations and has strong bias in the IGH gene repertoire with VH3–21, VH4–34, V1–8, and V3–23 used by 45% of cases
    - The subtype utilizing VH3–21 may have an adverse prognosis
- Genetic events in disease progression
  - Aberrations in genes of the p53/MDM2 pathway are frequently seen, in particular, in tumors with high proliferative activity and clinical aggressive behavior
  - The *ATM* gene is deleted or mutated in 40–75% of cases
  - Mutations of *CHK2* are found in a subset of MCL with a high number of chromosomal aberrations. CHK2 prevents cell cycle progression in response to DNA damage signals
  - Other genomic secondary alterations detected include +3/3q, +12/12q, −13/13q, −6q, −9/9p, −1p, −11q, −Y, and −17p
  - Blastoid and pleomorphic variants have more complex karyotypes and more high-level DNA amplifications than classic variants
    - Gains of 3q, 7p, and 12q and losses of 17p are more frequent in blastoid and pleomorphic variants

33.3.3 Follicular Lymphoma

- The most common low-grade lymphoma in the United States, which bears the t(14;18) chromosomal translocation that activates the antiapoptotic *BCL2* gene through the *IGH* enhancer in most cases

- Clinical features
  - Wide age range with median onset in fifth decade with a median survival of 8–10 years, although both more indolent and more aggressive courses are common
  - 80% of FL is stage III or IV at presentation
  - The FL international prognostic index identifies five poor prognostic factors including age >60, stage III or IV disease, hemoglobin <12 g/dL, the presence of more than four nodal sites of involvement, and an elevated serum lactate dehydrogenase (LDH) level
  - Treatment options include rituximab immunotherapy, localized radiotherapy, combination chemotherapy, or allogeneic or autologous stem cell transplantation
- Pathological features
  - B cell lymphoma that homes to lymphoid follicles and is comprised of a mixture of centrocytes (small and large cleaved follicle center cells) and centroblasts (large noncleaved follicle center cells)
  - Growth around the bone trabeculae (paratrabecular) is common if the bone marrow is involved
- Histological grading is based on the number of centroblasts (grades 1–3A and B)
  - Histological progression includes increasing numbers of large cells and a shift from follicular to diffuse growth pattern
  - The median survival from the time of histological transformation to DLBCL is only 18 months
  - Variants of FL with characteristic clinical or pathological features include pediatric FL, primary intestinal FL, and intrafollicular neoplasia/in situ FL
    - Pediatric FL is usually grade 3A
      - They express BCL6 and CD10, often coexpress CD43
      - Typically do not express BCL2 protein and lack t(14;18) translocation
      - Good prognosis
    - Primary intestinal FL is frequent in the duodenum, in the second portion, and may present with biliary obstruction
      - They have similar morphologic, immunophenotypic, and genetic features than nodal FL
- Immunologic features
  - CD19+ CD20+ B cells that express CD10, BCL6, LMO2, and BCL2, in contrast to nonneoplastic GC B cells that are BCL6-positive but negative for BCL2
- Molecular features
  - The t(14;18)(q32;q21), which activates *BCL2* through the *IGH* enhancer, is present in 85–90% of FL and can be detected by FISH, PCR, or by uniform overexpression of BCL2 in tumor cells (Fig. 33.12)
  - For diagnosis, FISH using *BCL2* and *IGH* probes to detect fusion signals is the preferred method
  - The majority of breakpoints on chromosome 18 are tightly clustered. Two well-known clusters, the major breakpoint cluster region (MBR) and the minor breakpoint cluster region (MCR), are involved in 60–70% and 5% of the cases of FL, respectively. A third breakpoint cluster region is located between MBR and MCR has been designated the intermediate cluster region (icr)
  - The breakpoints on chromosome 14 are tightly clustered, occurring immediately 5′ to the *IGH* joining regions
  - MBR and MCR as well as PCR can be used to detect the t(14;18) in 60–75% of the cases. Higher detection rates can be achieved by using long-range PCR
  - The t(14;18) is also found in 20–30% of DLBCL cases
    - These tumors are presumably of follicle center cell origin
  - Rare variant translocations, t(2;18)(p12;q21) and t(18;22)(q21;q11), involve the *IGK* or *IGL* genes, respectively
  - Approximately 40% of FL also have alterations involving the *BCL6* gene (3q27) that include rearrangements or mutations and have been more frequently described in t(14;18)-negative FL
  - Genes upregulated in FL include cell cycle regulators, such as *CDK10*, *p120*,

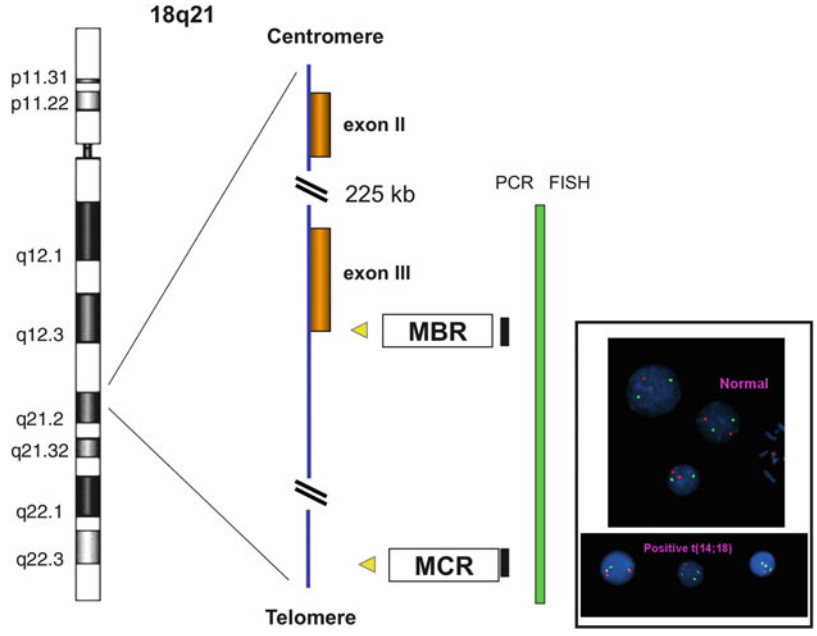


Fig. 33.12 Schematic representation of the chromosomal 18q21 region involved in t(14;18) in FL. (**a**) Most of the *BCL2* breakpoints occur within the MBR with a smaller subset in the MCR (*yellow arrowheads*). The *black* and *green rectangles* represent the range of detection using PCR and FISH, respectively. (**b**) FISH studies using a dual fusion translocation probe to detect t(14;18). The two upper interphase nuclei show no evidence of rearrangements between *IGH* (at 14q32) and *BCL2* (at 18q21) genes; thus, two *red* and two *green* signals are seen per nuclei (normal pattern). The three lower interphase nuclei show the abnormal 1R1G2F signal pattern indicating the presence of the (14;18)

p21CIP1, and *p16INK4A*; transcription factors such as *PAX5* and *Id-2*; and genes involved in cell–cell interactions such as *tumor necrosis factor*, *IL-2R*, and *IL-4R*

– Downregulated genes in FL include cell adhesion genes such as *MRP8* and *MRP14*
– FL is characterized by a different expression profile of microRNA when compared to DLBCL and reactive lymph nodes
– Genetic events involved in disease progression
  - Deletions of the *CDKN2A* and *CDKN2B* genes at chromosomes 7p and 9p that encode the p16 and p15 proteins, respectively
  - *MYC* deregulation and mutated *TP53*
  - Chromosomal gains at 8q, 12q, 18q, and trisomy 7 (candidate genes include *MDM2, CDK4*, and *GLI*)
  - Chromosomal losses at 1p, 6q, 10q, and 17p

33.3.3.1 Marginal Zone Lymphoma (MZL)

- A heterogeneous category that includes at least three different types of indolent B cell lymphomas (splenic, nodal MZL, and MALT). Overall, MZL accounts for between 5% and 17% of all lymphomas, with median age at presentation of 60 years. Etiologies include uncontrolled autoimmune and foreign antigen-driven B cell proliferation
- Splenic MZL
  – Clinical features: presents with splenomegaly but minimal peripheral lymphadenopathy
    - Bone marrow and blood involvement in 95% of the patients (stage IV)
    - Low to moderate levels of monoclonal IgM paraprotein in blood
    - Normal serum LDH level but increased serum beta2-microglobulin is typical

  - Pathological features: small- to intermediate-sized tumor cells with moderate amount of cytoplasm that usually partially replace the white pulp of spleen
    - Intrasinusoidal infiltration of bone marrow is characteristic
    - In 15% of the cases, villous lymphocytes displaying cytoplasmic protrusions will be seen in blood
  - Immunologic features: sIgM+ B cells that are negative for CD5 (unlike CLL), CD10 (unlike FL), and CD103 (unlike hairy cell leukemia)
    - Rare CD5+ and CD10+ variants have been described
  - Molecular features recurring reciprocal chromosomal translocations are rare in splenic MLZ (SMZL) so conventional *IGH* PCR is most useful for diagnosis and followup
    - Deletions involving the distal portion of the long arm of chromosome 7 (q22–36) are the most characteristic genomic alteration (45% of cases)
    - Trisomy of chromosome 3 (~10–20%)
    - Occasional cytogenetic abnormalities involving 14q32, such as t(6;14)(p12;q32) and t(10;14)(q24;q32) or 7q21 (with deregulation of cyclin-dependent kinase 6)
    - Deletion of 17p13 has been observed in ~3–15% of cases
    - Molecular predictors of poor outcome include presence of *TP53* mutations and chromosome 7q deletion
- Nodal MZL (NMZL)
  - Clinical features: initially localized to peripheral lymph nodes, most frequently cervical lymph nodes, with spread to bone marrow (50%), cytopenias are rare
    - 5-year overall survival: 50–70%
  - Pathological features: small lymphocytes with monocytoid or plasmacytoid appearance infiltrating lymph node in a parafollicular, perisinusoidal, or interfollicular distribution
  - Pediatric nodal marginal zone lymphoma
    - Distinctive morphological and clinical features
    - Interfollicular polymorphic infiltrate, plasma cells, monocytoid cells, and centrocytes-like cells associated with follicular expansion
    - Striking male predominance
  - Immunologic features: CD19+ CD20+ B cells that are negative for CD5, CD10, and CD103, also negative for IgD expression (in contrast to SMZL)
  - Molecular features
    - Numerical abnormalities are most common and include trisomies 3 (most common; 50–70%), 7, 12, and 18
    - *IGH* genes are clonally rearranged with selected usage of VH3 and VH4 gene segments
    - Most cases show evidence of somatic hypermutation
    - Recurrent cytogenetic aberrations have not been described
- MALT lymphoma
  - The prototypic antigen-driven B cell lymphoma, with different etiologies and molecular pathogenesis at different tissue sites (Table 33.3), including postulated autoantigen and microbial causes
  - Clinical features
    - Indolent lymphoma with localized extranodal presentation (stage IE)
    - Multifocal lesions present in 30–40%
    - Common sites include
      - Stomach: the most common subtype
        - It is driven by gastric *Helicobacter pylori* infection
        - Triple antibiotic treatment given as initial therapy
        - Another largely distinct subset is driven by t(11;18) and has a more aggressive course
      - Intestine: immunoproliferative small intestinal disease typically found in the Mediterranean areas that is associated with *Campylobacter jejuni* infection
      - Salivary glands and conjunctivae: progressing from myoepithelial

Table 33.3 Well-characterized chromosomal translocations in MALT lymphomas

Translocation	Activated gene	Mechanism of action	Tissue sites (frequency)
t(11;18) (q21;q21)	*API2/ MALT1*	NFκB activation	Lung (45%)
			Stomach (20%)
			Intestine (15%)
			Skin (7%)
t(1;14) (p22;q32)	*BCL10*	NFκB activation	Intestine (10%)
			Ocular (10%)
			Lung (7%)
			Stomach (3%)
t(3;14) (p14;q32)	*FOXP1*	Unclear	Thyroid (50%)
			Ocular (20%)
			Skin (10%)
t(14;18) (q32;q21)	*MALT1*	NFκB activation	Skin (14%)
			Salivary gland (5%)

Abbreviations: *AP12* apoptosis inhibitor 2, *NFκB* nuclear factor kappa B, *BCL10* B cell CLL/lymphoma, *FOXP1* forkhead box P1, *MALT1* mucosa-associated lymphoid tissue lymphoma translocation gene 1

sialoadenitis, with or without associated Sjögren syndrome
      - Lung: consequence of long-term exposure to a variety of antigenic stimuli – including smoking, inflammatory disorders, or autoimmune diseases
      - Orbit/conjunctivae: some associated with *Chlamydia psittaci* infection or connective tissue disease
      - Thyroid: progressing from Hashimoto thyroiditis
      - Skin: a minority of cases associated with *Borrelia burgdoferi* infection (Lyme disease)
    - Regardless of site, MALT lymphoma has a good prognosis with 5-year overall survival ranging from 86% to 95%
  - Pathological features
    - Small lymphocytes including monocytoid forms surrounding reactive lymphoid follicles, also forming lymphoepithelial lesions
      - Dutcher bodies (eosinophilic intranuclear pseudoinclusions of cytoplasm) are common with tumors with plasmacytoid differentiation, particularly in thyroid and bowel sites
      - Low-level bone marrow involvement in approximately 10–20% at presentation
      - Histological transformation to large cell lymphoma in <10%
  - Immunologic features: the neoplastic monocytoid/plasmacytoid component is CD19+ CD20+/− B cells that are usually negative for CD5 and CD10
  - Molecular features
    - Both the *IGH* and light chain genes are clonally rearranged with somatic hypermutation and ongoing mutations indicating direct antigenic stimulation and clonal evolution
    - A total of 10 chromosomal translocations have been reported, of which 4 are well characterized [t(11;18)(q21;q21), t(14;18)(q32;q21), t(3;14)(p14.1;q32), and t(1;14)(p22;q32)] and implicated in pathogenesis, with 6 more recently described and variable defined. FISH methods are useful in detecting all the translocations
    - These chromosomal translocations are mutually exclusive and correlate with the site of disease
      - t(11;18)(q21;q21)*api2-malt1* in ∼40% of the MALT lymphomas of the lung and ∼30% of the gastric MALT lymphomas. Gastric MALT lymphomas with t(11;18) are resistant to *H. pylori* eradication therapy
      - t(14;18)(q32;q21)/*IGH-malt1*, seen more frequent in MALT lymphomas involving skin, ocular adnexa, salivary glands, and liver
      - t(3;14)(p14.1;q32)/*IGH-foxp1* in MALT lymphomas involving thyroid gland, skin, and ocular adnexa
      - t(1;14)(p22;q32)/*IGH-bcl10* in MALT lymphomas with predilection for the intestines. BCL10 is typically expressed in the cytoplasm of

nonneoplastic B cells but in the nuclei of lymphoma cells with the t(1;14) and t(11;18)
- Other common genetic changes in MALT are trisomy 3, mutated p53, or loss of heterozygosity at the *TP53* locus
- A20, an essential global NF κB inhibitor, is found to be inactivated by somatic deletion and/or mutation in a subset of translocation-negative MALT lymphomas

33.3.3.2 Diffuse Large B Cell Lymphoma (DLBCL)

- The most common lymphoma type, which may arise de novo or from preexisting low-grade lymphoma (most commonly FL). DLBCLs are a group of morphologically, immunohistochemically, and clinically heterogeneous tumors rather than one single entity
- Clinical features
  - Wide age range and heterogeneous pattern of disease presentation, frequent presence of B symptoms, and a generally aggressive clinical course
  - Extranodal DLBCL accounts for 40% of cases, including GI tract, genital tract, skin, bone, lung, and CNS
  - Mediastinal large B cell lymphoma is a distinct variant with unusual pattern of metastasis (e.g., to kidney and CNS) and likely distinct genetic origin
  - Risk stratification using the international prognostic index (including age, performance status, LDH levels, Ann Arbor stage, and extranodal involvement) effectively predicts outcome with conventional chemotherapy
  - Addition of rituximab to combination chemotherapy (e.g., cyclophosphamide, doxorubicin, vincristine, and prednisone [CHOP]) has improved response rate above 50%
- Pathological features
  - Heterogeneous category unified only by the large size of the tumor cell and diffuse pattern of growth
  - Morphological variants: centroblastic, immunoblastic, and anaplastic
  - Subtypes/entities: T cell/histiocyte-rich, primary DLBCL of the central nervous system, primary cutaneous DLBCL, leg-type, EBV-positive DLBCL of the elderly, primary mediastinal (thymic) large B cell lymphoma, DLBCL associated with chronic inflammation, lymphomatoid granulomatosis, plasmablastic, and anaplastic lymphoma kinase (ALK)-positive DLBCL between others
- Immunologic features
  - The neoplastic cells usually express CD19, CD20, CD22, and CD79a
  - Phenotypic subtypes of DLBCL include CD10+ (related to FL), CD5+, and CD138+ plasmablastic/immunoblastic cases that lack most/all surface B cell markers
  - All subgroups can express BCL6 as a result of translocation or mutation
  - The proliferation index is variable but usually higher than 40%
- Molecular features
  - DLBCL has complex genetic alterations reflective of both its aggressive biology and its heterogeneous derivation from other lymphoma types
  - Chromosomal translocations affecting *BCL6* are one of the most common genetic abnormalities in DLBCL
  - Somatic mutations of *BCL6* gene are also frequent
  - Translocations of the *BCL2* gene next to the *IGH* gene through t(14;18)(q32;a21) are detected in 30–40% of DLBCL (those of germinal center [GC] type, whereas it is absent in DLBCL of activated B cell [ABC] type)
  - *MYC* translocation occurs in ~15% of classical DLBCL cases
  - Cases with both *MYC* and *BCL2* gene rearrangements with or without *BCL6* translocations are classified as B cell lymphomas, unclassified, with features intermediate between DLBCL and Burkitt lymphoma (see below)

- Expression microarray studies have generally established two major molecular groups of DLBCL: those with a profile similar to GC B cells and those mimicking ABC
  - CD10 and LMO2 are markers of the GC subset, whereas MUM1 and FOXP1 expressions reflect an activated post-GC B cell phenotype
  - DLBCL with an ABC gene expression profile does significantly worse than GC-type tumors with conventional R–CHOP-type therapy
  - Molecular subtypes seem to arise by distinct oncogenic pathways
    - Abnormalities seen in ABC DLBCL include
      - Amplifications/gains of 19q (*SPIB*)
      - Deletions of INK4a/ARF
      - Upregulation of *FOXP1* (by trisomy 3 or focal high-level amplifications)
    - Abnormalities seen in GC DLBCL include
      - Amplification of mir17–92 cluster
      - *PTEN* deletion
- Activation of proto-oncogenes through aberrant hypermutation is detected in more than 50% of DLBCLs and includes the following genes: *PIM1*, *MYC*, *RhoH/TTF* (*ARHH*), and *PAX5*

33.3.3.3 ALK-Positive Diffuse Large B Cell Lymphoma (DLBCL)

- DLBCL expressing ALK protein and associated with *ALK* gene abnormalities
  - Identified by its characteristic lack of CD30 expression in an otherwise large series of classical ALK-positive anaplastic large cell lymphomas
  - ALK-positive DLBCL displays clinicopathologic features that distinguish them from classical DLBCL
- Clinical features
  - ALK-positive DLBCL is a rare tumor and represents less than 1% of all cases of DLBCL
  - There are approximately 80 cases reported to date
  - Patient median age is 43 years with 30% of cases occur in the pediatric population
  - It is more frequent in male than female with a ratio of 5 to 1
  - There is no apparent ethnic predisposition
  - There is no association with immunosuppression
  - Usually presents with high-stage nodal disease (60–70% of the cases) with widespread lymphadenopathies, systemic (B-type) symptoms, and aggressive clinical course
  - The patients can also present with enlarged mediastinal lymph nodes and also with leukemic involvement
- Pathological features
  - ALK-positive DLBCL shows a partial or diffuse effacement of the lymph node architecture with lymphoma cells seen infiltrating sinusoids along with focal necrosis, binucleated HRS-like cells, and multinucleated giant lymphoma cells
  - The tumor cells are relatively monomorphic with immunoblastic or plasmablastic appearance
  - They can appear deceptively cohesive and thus may be misinterpreted as carcinoma cells
- Immunologic features
  - By definition, express ALK
  - The pattern of ALK expression predicts *ALK* partner in fusion gene
    - Granular and cytoplasmic (most frequent pattern): *CLTC* and *SEC31A*
    - Nuclear, nucleolar, and cytoplasmic: *NPM*
  - The tumor cells are usually positive for CD138, VS38, EMA, IgA, and CD45/LCA (weak) and can be positive for CD4 (40%) and CD43 (rare)

  - Usually negative for CD30, CD20, CD79a, and IRF4/MUM1
  - HHV8 and EBV–LMP are also negative
- Molecular features
  - Four chromosomal translocations described up to date, t(2;17)(p23;q23), t(2;5)(p23;q35), t(2;3)(p23;q27), and t(2;5)(p23.1;q35)
    - Most frequent: t(2;17)(p23;q23) (70% of cases) resulting in clathrin heavy chain gene (CLTC)–ALK fusion protein
    - The t(2;5)(p23;q35) leading to NPM–ALK fusion is present in 10% cases
    - The t(2;3)(p23;q27) results in SEC31A–ALK fusion protein
    - The t(2;5)(p23.1;q35) results in sequestosome1 (SQSTM1)–ALK fusion protein
  - The fusion transcript consequence of these translocations can be detected by RT-PCR or FISH

33.3.4 B Cell Lymphomas, Unclassifiable ("Gray-Zone" Lymphomas)

- The term "gray-zone" lymphomas has been used to indicate a group of high-grade lymphomas that cannot be reliably classified into a single distinct disease entity and show overlapping morphological, immunophenotypic, biologic, and clinical features between various types of lymphomas
- The 2008 WHO classification of tumors of hematopoietic and lymphoid tissues recognized this issue and introduced two new provisional entities
  - B cell lymphoma, unclassifiable, with features intermediate between DLBCL and BL
  - B cell lymphoma, unclassifiable, with features intermediate between DLBCL and classical Hodgkin lymphoma

33.3.4.1 B Cell Lymphoma, Unclassifiable, with Features Intermediate Between Diffuse Large B Cell Lymphoma (DLBCL) and Burkitt Lymphoma (BL)

- Provisional category that includes aggressive B cell lymphomas with morphological and genetic features of both DLBCL and BL
  - This is a heterogenous category and not a distinct disease entity
  - Previously, these cases were classified as Burkitt-like lymphomas (Revised European American Lymphoma [REAL] classification)
- Clinical features
  - These lymphomas are relatively uncommon, but true frequency is not yet known
  - They occur predominantly in adults, median age sixth decade, and more frequently in males than females
  - They are aggressive tumors
  - 50% of the patients present with lymphadenopathies and/or extranodal masses
  - Bone marrow and CNS are frequently involved, ~50% and ~33%, respectively
  - B-type symptoms are common, while leukemia presentation with elevated white cell count is rare but can occur
  - 10–20% of patients have a history of FL
  - Heterogeneous prognosis: a large subset of patients have poor clinical outcome despite intensive chemotherapy regimen
- Pathological features
  - Diffuse proliferation of intermediate-sized cells or a spectrum of intermediate-sized and large cells with high mitotic and apoptotic rates and relatively few admixed small reactive lymphocytes
  - A "starry sky" is common and sclerosis is uncommon
  - Cases that should be considered in this category include
    - Neoplasms resembling BL with "starry sky" appearance but with marked variation in cytomorphology

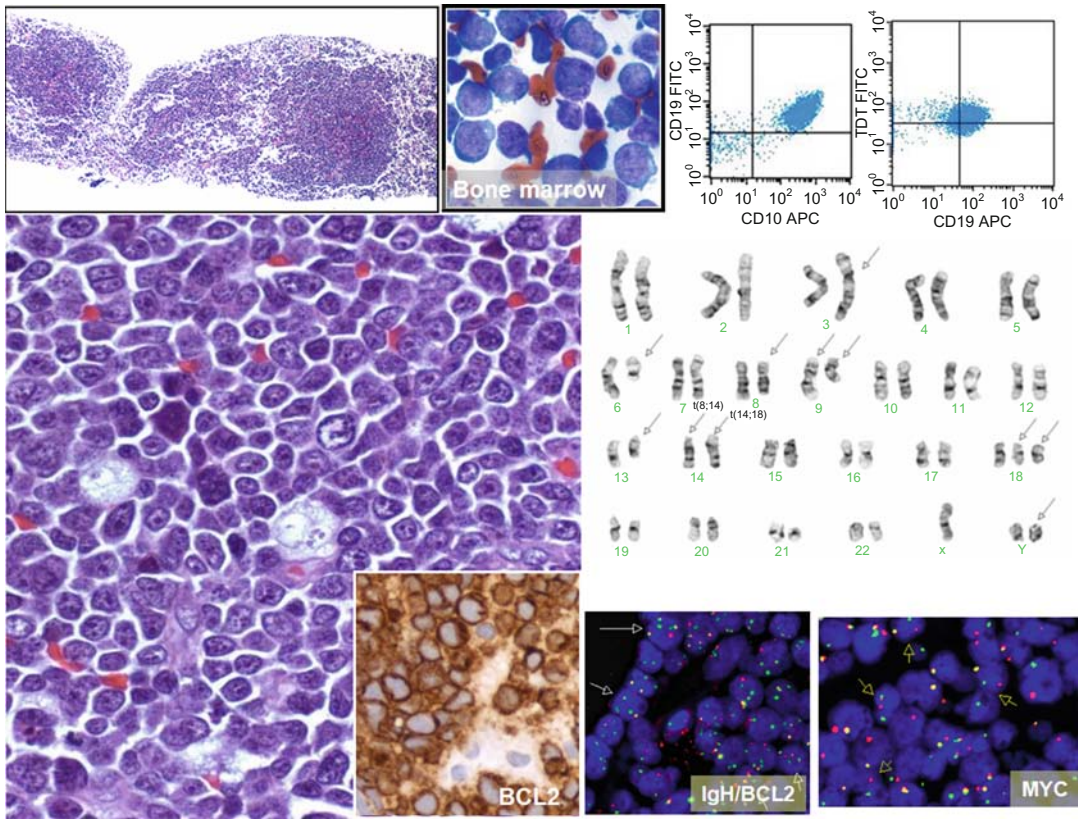


Fig. 33.13 B cell lymphoma, unclassifiable with features intermediate between diffuse large B cell lymphoma and Burkitt lymphoma (case of double-hit B cell lymphoma). The patient was initially diagnosed with low-grade follicular lymphoma (*top left*). However, almost 2 years later, in a lymph node biopsy specimen, a diagnosis of aggressive diffuse large B cell lymphoma was rendered (*bottom left*). The tumor cells were positive for B cell markers and strongly positive for BCL2. The bone marrow was also involved by malignant lymphoma (*top middle*). The tumor cells were intermediate in size with dispersed nuclear chromatin (*blastic-like appearance*). They were positive for CD19, CD10, and TdT (*top right*). G-banding karyotyping revealed a complex karyotype with the t(8;14) as well as t(14;18). FISH studies confirmed the presence of rearrangements of *BCL2* and *MYC* confirming the diagnosis of double-hit B cell lymphoma

    - Neoplasms resembling BL (uniform or mildly pleomorphic medium-sized cells) but have atypical immunophenotype and/or genetic findings
- Immunologic features
  - They express pan-B cell markers and include cases with immunophenotype consistent with BL (CD10+, BCL6+, IRF4/MUM1–, BCL2–, and TCL1+)
  - Strong expression of BCL2 and a wide range of Ki67 proliferation index (70–100%) are present in so-called "double-hit" or "triple-hit" lymphomas with *MYC* as well as *BCL2* and/or *BCL6* gene rearrangements
  - They usually are negative for *EBER*
- Molecular features
  - A complex karyotype (>3 chromosomal abnormalities) is frequent when compared with typical BL (Fig. 33.13)
  - 35–50% of the cases have 8q24/*MYC* translocations with many of them (approximately 60%) involving a non-Ig partner (non-*IG–MYC*)

  - Non-*IG–MYC* translocations include the t(8;9)(q24;p13) and t(3;8)(q27;q24); the breakpoints in the *MYC* locus have been described to be mostly telomeric
  - ~15% of the cases have *BCL2* translocation; when present along with *MYC* rearrangements, defines "double-hit" lymphomas. Presence of *BCL6* translocations is less frequent (along with *BCL2* translocations, defines "triple-hit" lymphomas)
  - FISH assays are very useful for detecting these translocations
  - Gene expression profile is either intermediate between BL and DLBCL or more similar to BL

33.3.4.2 B Cell Lymphoma, Unclassifiable, with Features Intermediate Between Diffuse Large B Cell Lymphoma (DLBCL) and Hodgkin Lymphoma (HL)

- This provisional category in the updated WHO classification (2008) includes B lineage lymphomas with clinical, morphological, and/or immunophenotypic features, between classical HL and DLBCL, especially primary mediastinal large B cell lymphoma (PMLBCL)
  - Composite lymphomas of classical HL and PMLBCL are excluded from this category
- Clinical features
  - More common in young patients, between the ages of 20–40
  - There is a male predominance, and they are most common in Western countries
  - They frequently present as an anterior mediastinal mass with or without involvement of supraclavicular lymph nodes
  - Cases without mediastinal involvement are currently included also in this category
- Pathological features
  - Broad spectrum of cytologic appearances with areas of confluent sheets of pleomorphic cells that resemble DLBCL and others more closely resembling Hodgkin cells, Reed–Sternberg cells, or lacunar cells
  - Focal fibrotic bands and sparse inflammatory infiltrate may be present
  - When necrosis is seen, it is not usually associated with neutrophils
- Immunologic features
  - Mixed immunophenotype with expression of markers of classical HL (CD30, CD15 in most cases, PAX5, IRF4/MUM1) and expression of markers usually absent in classic HL that include CD45, CD20 uniformly strong, CD79a, OCT2, and BOB.1
  - The lymphoid infiltrate in the background is predominantly composed of CD3- and CD4-positive T cells
  - They can present with a PMLBCL-like morphology and a classical HL phenotype with expression of CD30 and CD15 and loss of CD20 and CD79a
  - Alternatively, they can have a classical HL morphology and a phenotype of large B cell lymphoma (positive for CD20 (strong), CD79a and/or CD45/LCA and CD15-negative)
- Molecular features
  - Studies have shown a genetic similarity between classical HL (nodular sclerosis subtype) and PMLBCL
  - A number of common genetic and epigenetic aberrations in PMLBCL and classical HL further underscore their close relationship
    - Most cases have monoclonal *IGH* gene rearrangement
    - Few cases have rearrangements involving *BCL6* and most cases lack t(14;18)(q32;q21)
    - In almost all cases assessed, *TP53* was in germline configuration
    - Class prediction models utilizing methylation status of genes such as *HOXA5*, *MMP9*, *EPHA7*, and *DAPK1* allow the distinction between mediastinal gray zone lymphoma, classical HL, and PMLBCL
    - Gains/amplifications of 2p16.1 in ~33% of cases

- Gains/amplifications of 9p24.1 (*JAK2*, *PDL1*, and *PDL2*) in ~55% of cases
- Rearrangements involving 16p (*CIITA*) in ~27% of cases

33.3.4.3 Plasma Cell Myeloma (PCM)

- A tumor of plasma cells associated with antibody production (M protein) in serum or urine as well as tumor-associated skeletal destruction
- Clinical features
  - PCM accounts for 10% of hematologic malignancies
  - Median survival is 3 years and depends on disease stage
  - 3% of the patients have no detectable M protein and are considered to have nonsecretory PCM
  - Novel effective therapeutic strategies include thalidomide or lenalidomide (immunomodulatory agents), bortezomib (proteosome inhibitor), and FGFR-specific tyrosine kinase inhibitors
  - A preceding premalignant plasma cell proliferation termed monoclonal gammopathy of undetermined significance occurs in about 3% of individuals over the age of 50
- Pathological features
  - PCM requires 10% or more plasma cells on bone marrow examination or biopsy-proven plasmacytoma, M protein in the serum and/or urine (except in patients with true nonsecretory myeloma), and evidence of end-organ damage (hypercalcemia, renal insufficiency, anemia, or bone lesions)
  - Anaplastic or blastic morphology of plasma cells is an adverse prognostic factor
- Immunologic features
  - Neoplastic plasma cells are usually positive for CD38, CD138, CD56 (50–60%), CD79a, and cytoplasmic κ- or λ-Ig light chains and negative for CD19 and surface Igs
  - CD20 is positive in 10%
- Molecular features
  - Broad molecular heterogeneity and structural and numerical chromosomal abnormalities are common

Table 33.4 Common recurrent translocations involving IGH in PCM

Chromosome region	Activated gene(s)	Frequency(%)
11q13	*CCND1* and *MYEOV*	15–20
4p16.3	*FGFR3* and *MMSET*	15
16q23	*MAF*	5
6p21	*CYCLIN D3*	3
20q11	*MAFB*	2

Abbreviations: *MYEOV* myeloma overexpressed gene, *FGFR3* fibroblastic growth factor receptor 3, *MMSET* multiple myeloma set domain, *MAF* musculoaponeurotic fibrosarcoma gene

  - There is no clear landscape of the molecular changes underlying this disease and their impact on classification, prognosis, and in the therapeutic approach
  - FISH methods are the most useful approach for detecting translocations
  - Chromosome 14q32 translocations involving *IGH* are the most frequent event in PCM involving a variety of partner chromosomal loci (Table 33.4)
  - Although cytogenetics is difficult to assess, patients can be divided into two categories: hyperdiploidy and nonhyperdiploidy (about one-half in each group)
  - Monosomy 13 in ~50% of patients
  - t(4;14) confers a poor prognosis
    - These patients may benefit from use of bortezomib
  - Recurrent secondary events: del(17p) in ~10% of patients (poor prognosis) and gains of the long arm of chromosome 1 in 30% (poor prognosis)
  - Activating *RAS* mutations have been noted in 35–50% of patients

33.3.5 Mature T Cell Lymphoma

- General features in contrast to B cell tumors; the current WHO classification for postthymic

T cell malignancies is largely based on clinical pattern of disease and not on histogenesis
  - T cell clonality studies by PCR are the mainstays of molecular diagnosis
  - T cell lymphomas are rare in the United States, but relatively common in South America and common in Asia, due to a role for several viruses in pathogenesis
- Important T cell tumor subtypes recognized in the WHO classification include
  - Anaplastic large cell lymphoma (ALCL): includes a major subtype expressing the ALK receptor tyrosine kinase, usually through a fusion transcript with nucleophosmin (NPM) as a result of the t(2;5) translocation
    - The ALK+ group has better prognosis than the ALK–ALCL subset, which appears to be a heterogeneous entity
  - Mycosis fungoides: the most common cutaneous T cell lymphoma that progresses through patch, plaque, and tumor stages and disseminates outside skin to lymph node and blood as *Sézary syndrome*
  - Enteropathy-associated T cell lymphoma: associated with or without celiac disease, and have the immunophenotype of gut-homing T cells
- Immunologic features: CD4+ T cell tumors predominate
- Molecular features: except for the t(2;5) in ALCL, reciprocal chromosomal translocations are rare in T cell lymphomas
  - *TCRG* PCR is the principle modality in diagnosis and MRD detection
    - The limited diversity of the *TCRG* loci can lead to false-positive pseudoclonality and benign reactive and oligoclonal T cell expansions
  - *TCRB* PCR is less prone to false-positive results but is technically more challenging due to large numbers of V and J segments
  - The t(2;5) in ALCL can be detected by immunostaining for ALK, karyotyping, FISH with fusion signal probes, or long-range DNA PCR across the breakpoint

33.3.5.1 Hepatosplenic T Cell Lymphoma (HSTL)

- Clinical features
  - Rare subtype of non-Hodgkin lymphomas (<1%)
  - Occurs in young adults with a male predominance
  - Up to 20% of cases have been reported in patients with chronic immunosuppression especially following solid organ transplant
  - An association between the use of tumor necrosis factor (TNF) blockers and HSTCL has been postulated (mostly young patients with Crohn disease)
  - Splenomegaly, hepatomegaly, and bone marrow involvement but no lymphadenopathy
  - Systemic/B symptoms and cytopenias with marked thrombocytopenia are characteristic
  - The outcome is poor with median survival less than 2 years
- Pathological features
  - Diffuse involvement of the splenic red pulp characterized by marked sinusoidal infiltration and atrophy of the white pulp
  - Intrasinusoidal pattern of involvement is also seen in the liver and bone marrow
  - The infiltrate is relatively monotonous, and the cells are medium in size with inconspicuous nucleoli
  - With disease progression, the pattern of bone marrow involvement becomes increasingly interstitial and the neoplastic cells become larger and blastic
- Immunologic features
  - Positive for CD3, CD2, and sometimes CD8, TIA1, and granzyme M
  - Negative for CD5, CD56, CD4, granzyme B, and perforin
  - Most cases express γδ TCR, and some express the αβ TCR
  - Loss of γδ TCR expression has occasionally been observed during the course of the

disease, resulting in a TCR silent immunophenotype
  - EBV infection has been reported rarely and has been associated with cytologic features of transformation, suggesting that EBV might be involved secondarily
- Molecular features
  - Isochromosome 7q [i(7q)] is a consistent cytogenetic abnormality and might be the primary cytogenetic event
  - Other cytogenetic changes include trisomy 8 and loss of chromosome Y
  - Gene expression profile studies have shown a γδ TCR signature with high expression of NK cell-associated molecules, such as killer cell Ig-like receptor (KIR) genes (*KIR3DL1*, *KIR2DL4*, and *KIR2DL2*) and lectin-like receptors (*KLRC4*, *KLRD1*, and *KLRC2*)
  - Upregulation of CD16 genes (*FCGR3B* and *FCGR3A*)
  - γδ T cell lymphomas of other sites show similar gene expression profile to HSTCL
  - PCR studies have demonstrated monoclonal *TCR*δ, *TCR*β, and *TCR*γ chain gene rearrangements

33.3.5.2 Angioimmunoblastic T Cell Lymphoma (AILT)

- AITL is the most common specific type of peripheral T cell lymphoma (PTCL) (accounting for approximately 15–20% of cases), represents 1–2% of all non-Hodgkin lymphomas, and accounts for the 6.6% of PTCL cases. AILT is a tumor of follicular helper T cells
- Clinical features
  - More common in middle-age or old adults (median age, 59–65 years) and with a slight male predominance
  - The clinical presentation includes generalized lymphadenopathy, hepatosplenomegaly, and frequently skin rash (>50% of patients)
  - Polyclonal hypergammaglobulinemia with Coomb-positive hemolytic anemia is common
  - Aggressive clinical course with a median survival of less than 3 years
- Pathological features
  - Lymph node: partial or complete effacement of architecture with paracortical distribution of the neoplasm
    - Marked proliferation of high endothelial venules (HEV)
    - Usually small- to medium-size neoplastic cells with clear cytoplasm
    - Polymorphic background, small reactive lymphocytes, eosinophils, plasma cells, and histiocytes
- Immunologic features
  - Positive for pan-T cell antigens (CD2, CD3, and CD5) and CD4
  - Follicular helper T cell immunophenotype in most cases (CD10, BCL6, CXCL13, and/or PD1)
  - CD21, CD23, and CD35 highlight a characteristic expanded follicular dendritic cell meshworks around HEV
- Molecular features
  - *TCR* genes are clonally rearranged (75–90% of the cases), as well as the Ig genes in 25–30% of the cases. The latter finding correlates with the presence of EBV-positive cells
  - Chromosomal breakpoints affecting *TCR* gene loci appear to be rare
  - Most common cytogenetic aberrations include trisomies of chromosome 3, 5, and 21, gains of 5q, 21, and 3q, gain of X, and loss of 6q
  - CGH studies have also shown gains at 22q, 19, and 11q13 and losses at 13q in a subset of cases
  - Gene expression profiling studies
    - Overexpression of CXCL13, BCL6, PDCD1, CD40L, NFATC1, and CD200
    - Strong microenvironment imprint (overexpression of B cell and follicular dendritic-related genes)

- High expression of *VSIG4*, PDGFRα, and β correlates with poor prognosis
- Mutations of *TP53* are infrequent
- AILT cases tend to have an activated NFκB pathway

33.3.5.3 Anaplastic Large Cell Lymphoma (ALCL), ALK-Positive

- CD30-positive systemic lymphoma of T or null cell lineage with chromosomal abnormalities involving 2p23 and ALK
  - Current WHO classification distinguishes two types of systemic ALCL
    - ALK-positive
    - ALK-negative
- Clinical features
  - It is more frequent in children and young adults; male predominance
  - Most patients present with clinical stages III/IV disease
  - Extranodal involvement is common (60%), skin soft tissue, lungs, and bone marrow (∼30%)
  - Relative good prognosis. ALCL–ALK-positive patients have a favorable prognosis comparing to ALCL–ALK-negative
- Pathological features
  - Large neoplastic cells with irregular bizarre, irregular, horseshoe-shaped or kidney-shaped nuclei, and abundant eosinophilic cytoplasm (hallmark cells)
  - Characteristic intrasinusoidal involvement with partial to complete effacement of the nodal architecture
- Immunologic features
  - Strongly and uniformly positive for CD30 and ALK
  - The majority of cases show a T cell phenotype (CD3 is frequently negative)
  - Null cell phenotype in 10–20% of the cases
  - Cytotoxic molecules (granzyme B, TIA1, and perforin)-positive
  - Usually BCL2-negative
  - Type of ALK staining correlates with type of underlying genetic abnormality

Table 33.5 Translocations involving *ALK* in ALCL

Translocation	Pattern gene	Frequency (%)	Pattern of staining on tissue
t(2;5)(p23;q35)	*NPM1*	75–80	Cytoplasmic and nuclear
t(1;2)(q25;p23)	*TMP3*	12–18	Cytoplasmic
Inv(2)	*ATIC*	2	Cytoplasmic
t(2;3)(p23;q21)	*TFG*	2	Cytoplasmic
t(2;17)(p23;q23)	*CLTL*[a]	2	Cytoplasmic granular
t(2;17)(p23;q25)	*ALO17*	<1	Cytoplasmic
t(2;19)(p23;p13)	*TPM4*	<1	Cytoplasmic
t(2;22)(p23;q11.2)	*MYH9*	<1	Cytoplasmic
t(2;X)(p23;q11-12)	*MSN*	<1	Membranous

Abbreviations: *ALK* anaplastic lymphoma kinase, *NPM1* nucleophosmin, *TMP3* tropomyosin 3, *ATIC* 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase, *TFG* TRK-fused gene, *CLTL* clathrin heavy chain 1, *AL017* ALK lymphoma oligomerization pattern on chromosome 17, *TMP4* tropomyosine 4, *MYH9* nonmuscle myosin heavy chain; MSN, moesin

[a]CLTL is the most common pattern gene in ALK-positive diffuse large B cell lymphoma

  - Cytoplasmic and nuclear: t(2;5)
  - Cytoplasmic, not coarsely granular: variant translocations, except t(2;X) and t(2;17)
  - Cytoplasmic, coarsely granular: t(2;17)
  - Membranous: t(2;X)
- Molecular features
  - Clonal *TCR* gene rearrangement in 90% of cases
  - Chromosomal translocations involving *ALK* gene at 2p23 (Table 33.5)
  - The most frequent (∼80% of cases) is the translocation between *ALK* gene and nucleophosmin (*NPM*) gene on chromosome 5
    - t(2;5) drives expression of novel fusion protein NPM–ALK

  - Variant chromosomal abnormalities (25%), *ALK* gene rearranged with other genes
    - Tropomyosin 3 (*TPM3*), t(1;2)(p25;p23)
    - TRK-fused gene (*TFG*), t(2;3)(p23;q21)
    - *ATIC*, inv(2)(p23;q11–12)
    - Moesin (*MSN*), t(2;X)(p23;q11–12)
    - Clathrin heavy chain (*CTCL*), t(2;19)(p23;q13.1)
    - *ALO17*, t(2;17)(p23;q25)
    - *MYH9*, t(2;22)(p23;q11.2)
  - CGH has shown secondary chromosomal alterations that include loses of chromosomes 4q13–q21, 11q14, and 13q and gains of 7, 17p, and 17q
  - Genes overexpressed in ALK-positive ALCL include *BCL6*, *PTPN12*, *serpinA1*, and *C/EBP*

33.3.5.4 Anaplastic Large Cell Lymphoma, ALK-Negative

- Lymphoma resembling morphologically and immunophenotypically ALK-positive ALCL with strong and uniform expression of CD30 but lacking ALK expression. ALK-negative ALCL accounts for the 5.5% of PTCL cases
- Clinical features
  - Older patients
  - No sex predilection
  - B symptoms frequent
  - Poorer prognosis than ALK-positive ALCL but better than peripheral T cell lymphoma not otherwise specified (NOS)
- Pathological features
  - Large neoplastic cells with irregular bizarre, irregular, horseshoe-shaped or kidney-shaped nuclei, and abundant eosinophilic cytoplasm (hallmark cells)
  - Intrasinusoidal involvement with partial to complete effacement of the nodal architecture
  - Anaplasia usually greater in ALK-negative ALCL than in ALK-positive ALCL
- Immunologic features
  - CD30 is uniformly and strongly positive
  - Aberrant T cell phenotype (loss of CD3, CD5, and T cell receptors)
  - Cytotoxic molecules (granzyme B, TIA1, and perforin)-positive
- Molecular features
  - Heterogenous category but molecularly distinguishable from PTCL, NOS
  - Gains in 1q, 6p21, and 3p and losses of 16pter, 6q13–21, 15, 16qter, and 17p13
  - Overexpressed genes include *CCR7*, *CNTFR*, *IL22*, and *IL21*

33.3.5.5 Enteropathy-Associated T Cell Lymphoma

- Enteropathy-associated T cell lymphoma (EATL) is an intestinal tumor with an immunophenotype of intraepithelial T cells. There is a classical form associated with celiac disease and a monomorphic variant (type II) that may occur in patients without celiac disease
- Clinical features
  - EATL occurs in older individuals (sixth and seventh decades of life)
  - Most patients have adult-onset celiac disease; in some patients, celiac disease and EATL are diagnosed at the same time
  - The most common usual presentation is abdominal pain associated with intestinal perforation
  - The jejunum or ileum is most commonly involved. The tumor usually presents as multiple ulcerated masses in the intestinal mucosa
  - The type II monomorphic variant has similar clinical presentation, but most patients do not have evidence of celiac disease
  - Poor prognosis
- Pathological features
  - Multiple ulcerating raised mucosal masses or large ulcer or exophitic mass
  - Tumor cells are usually medium to large in size
  - Sometimes tumor cells with marked pleomorphism, anaplastic morphology, and multinucleated tumor cells
  - In the type II EATL (monomorphic form), the tumor cells are small- to medium-sized, round, and monomorphic, with dark-stained nuclei and pale cytoplasm

  - Noninvolved mucosa often shows enteropathy-associated changes
- Immunologic features
  - Classic form
    - Positive for CD3, CD7, CD103, and cytotoxic proteins
    - Negative for CD4, CD5, CD8, and CD56
  - Type II EATL
    - Positive for CD3, CD8, and CD56
    - Negative for CD4
    - Rare cases are reported to be TCR γ/δ-positive
- Molecular features
  - Amplifications at 9q31.3–q34 (60–70%) (in both subtypes)
  - Recurrent deletions in 16q12.1, 8p22–23.2, 11q14.1–q14.2, and 9p21.2–p21.3
  - Gains at 5p15.33 and 6p25.2
  - Gains at 9q13.3–q34 and losses at 16q12.1 are mutually exclusive and are seen in more than 80% of cases
  - Gains at 1q32.2–q41, 5q34–q35.2, 7q11.23–q21.3, and 8q13.3–q21 tend to occur within the same tumor
  - Potential genes involved in gene amplifications include *ABL1*, *NOTCH1*, and *h-TERT*
  - Potential genes involved in recurrent losses include *CDKN2A/B*
  - In type II EATL, gains in *MYC* gene are common and gains in 1q and 5q are rare
  - Type II EATL shows a HLA–DQB1 genotype pattern

33.3.5.6 Peripheral T Cell Lymphoma, Not Otherwise Specified (PTCL, NOS)

- Mature T cell lymphomas that cannot be classified into specific T cell categories
- Clinical features
  - PTCL, NOS is a heterogeneous category and represents ∼50% of all NK/T cell lymphomas
  - More common in middle-aged adults and are very rare in children
  - Patients present with generalized lymphadenopathy and B symptoms
  - Cytokine-related paraneoplastic syndromes can occur including pruritus, eosinophilia, and hemophagocytic syndrome
  - Extranodal involvement occurs with skin and gastrointestinal tract representing the most common affected sites
  - They are aggressive lymphomas with low survival and usually poor response to therapy
- Pathological features
  - Paracortical infiltrate or diffuse effacement of the lymph node architecture
  - Proliferation of postcapillary venules can be seen
  - Polymorphic reactive background with eosinophils and plasma cells
  - Spectrum of tumor cells, small to intermediate or large size
  - Three morphological variants
    - Lymphoepithelioid (Lennert lymphoma) with clusters of epithelioid histiocytes and CD8+ T cells
    - Follicular pattern composed of atypical clear cells forming intrafollicular aggregates
    - T zone with an perifollicular or interfollicular pattern
- Immunologic features
  - Frequent aberrant T cell phenotype (∼80% of cases) with loss of CD2, CD3, CD5, and/or CD7 expression, dim expression of T cell markers or coexpression or absence of both CD4 and CD8
  - PTCL is most commonly positive for CD4 and TCRαβ
  - Variable expression of CD56 and CD30
  - CXCR3 and CCR4 are expressed in 63% and 34% of PTCL, NOS cases
  - Expression of CD30 confers bad prognosis, in particular if they also coexpress CD15
  - Cytotoxic markers can be expressed in a subset of cases
- Molecular features
  - Clonal rearrangement of *TCR* genes in most cases
  - Cytogenetic studies usually reveal a very complex karyotype

- Cases with complex karyotype have shorter overall survival
- Recurrent chromosomal abnormalities have been described
  - Gains in chromosome 7q22–31, 1q, 3p, 5p, 8q (*MYC*), 17q, and 22q
  - Deletions in chromosomes 4q, 5q, 6q22–24, 9p, 10p13pter, 12q, and 13q
  - del 5q, 10q, and 12q are associated with better prognosis
- PTCL, NOS is heterogeneous at the molecular level
- Overexpression of PDFRα in a subset of cases
- A subgroup has a molecular signature of cytotoxic T cells and has an inferior survival than other PTCL, NOS cases
- CXCR3+/CCR4– phenotype is an independent prognostic factor

33.3.5.7 Extranodal NK/T Cell Lymphoma, Nasal Type

- Lymphoma of either NK cell or T cell lineage characterized by necrosis, cytotoxic immunophenotype, and EBV infection
- Clinical features
  - Rare in the United States but relatively common in South America and common in Asia
  - Commonly involve the naso-oropharynx (nasal type) causing progressive destructive and ulcerative lesions (so-called midfacial destructive disease)
  - Extranasal type (skin, testis, gastrointestinal tract, kidney, eye, and salivary glands) is less common and defined as neoplasms involving any site but without nasal involvement
  - Lung and skin are common sites of metastasis
  - Occasional cases have mixed lymphoma-leukemic pattern
  - Frequently radiosensitive
- Pathological features
  - Often deceptively low-grade histological appearance
  - Occasional NK-like T cell lymphomas with clonal *TCRG* rearrangement may have identical clinical features
  - Angiocentric and angiodestructive growth patterns
  - Zonal geographic necrosis is common
  - Background rich in inflammatory cells
- Immunologic features
  - Typically express CD2 and cytoplasmic but not surface CD3 and lack CD5 expression
  - CD56 usually positive and CD16 and CD57 usually negative
  - Negative for both TCR γδ and αβ
  - Epstein–Barr virus (EBV) infection of tumor cells seen in majority of cases, usually type II latency with *EBNA*1 and *LMP*1 viral gene expression
  - High expression of survivin (∼90%); survivin is a caspase-9 inhibitor
- Molecular features
  - Lack clonal *IGH* and *TCRG*/*B* rearrangements in the majority of cases
  - EBV in situ hybridization using EBER-1/2 probes is the most common diagnostic test used. EBV clonality studies can be done by Southern blot to detect a fixed number of terminal repeats in the viral episome
  - EBV titers in blood may be useful as a tool for monitoring for disease relapse
  - Deregulation of *TP53* and *MYC* genes with high expression of MYC and NFκB signatures in comparison with nonneoplastic NK cells
  - Mutations involving *TP53* (20–60%). Its presence correlates with large cell morphology and advanced stage (secondary genetic event)
  - MYC activation is not due to *MYC* gene rearrangements
  - Promoter methylation of *TP73* (94% cases), retinoid acid receptor B, and death-associated protein kinase genes
  - Complex chromosomal abnormalities
    - Losses of 6q21–25 (most common: *PRDM1*, *ATG5AIM1*, and *HACE1*), 8p, 1q, 11q23.1, 11q24–q25, 12q, 13q14.11, and 17p13.3
    - Gains of 1q21–q44, 2q13–q14, 6p25–p11.1, 7q11.2–q34, 7q35–q36, 11q, 15q, 17q21.1, and 22q

  - Translocations involving 8p23 have been reported in a subset of cases
  - Upregulation of *BIRC5*, *EZH2*, and *STMN1* genes

33.3.6 Classical Hodgkin Lymphoma

- Classical HL is a lymphoid neoplasm composed of Hodgkin and Reed–Sternberg (HRS) cells in a variable inflammatory background
- Clinical features
  - Bimodal age-specific incidence with a low prevalence of EBV in younger classical HL cases
- Pathological features
  - Diagnosed based on presence of HRS cells or variant forms in the appropriate histological background of nonneoplastic reactive inflammatory cells
  - Histological types include nodular sclerosis, mixed cellularity, lymphocyte-depleted, and lymphocyte-rich types
  - All subtypes are currently believed to arise from B cells of GC origin
  - Subsets of classical HL have latent EBV infection of tumor cells
- Immunologic features
  - Positive for CD30, CD15, PAX5 (characteristically weak), and MUM1
  - CD20 can be variable expressed
  - Rarely T cell markers can be also expressed
  - Negative for CD45/LCA and CD79a
- Molecular features
  - Not useful from the diagnostic perspective except to exclude other lymphoma types
  - By single cell analysis, clonal Ig heavy chain gene rearrangements can be demonstrated
  - Despite the presence of rearranged Ig genes, lacks Ig expression on the messenger RNA and protein level
  - Susceptibility loci for classical HL have been identified at 2p16.1 (*REL*), 8q24.21 (*PVT1*), and 10p14 (*GATA 3*)
  - Molecular analysis of microdissected HRS has revealed genetic lesions that are associated with deregulated signaling pathways
    - NFκB pathway: gains of *REL* (50% of the cases), gains of *BCL3* (~10%), point mutations or deletions of *NFKBIA* (20%), *NFKBIE* (15%), and *TNFAIP3* (40%)
    - JAK/STAT pathway: amplification of 9p24/*JAK2* region (40% of the cases) and point mutations and deletions of the negative regulator *SOCS1* (45% of the cases)
    - Point mutations or deletions of *TP53*
    - Point mutations of CD95/FAS
  - STAT6 appears to be activated in an autocrine fashion through expression of IL-13 and IL-13R by HRS cells
  - Multiple pathways are activated in classical HL including NFκB, PI3K/AKT, MAPK/ERK, AP1, NOTCH1, and JAK/STAT pathways. Numerous external signals contribute to the activation of these pathways in an autocrine and paracrine fashion
  - Constitutive activation of NFκB is essential for HRS cell survival and proliferation
  - Frequent gains: 2p13 (*REL*), 2q37.3, 7p21.1, 8q24.3, 9p24 (*JAK2*) 16q23.3–q24.3, 9q34.13–q34.3 (*ABL1*, *CDK9*, *ENG*, *LCN2*, *PTGES*, and *TSC1*), 14q32.33, 19q13.33, and 20q13.33
  - Losses of *AMACR*, *GDNF*, and *SKP2* at 5p13.2, *ID4* at 6p22.3, *TXNIP*, and *PPARGC1A* (also known as *COX2*)
  - Losses of *IL21*/*IL2* at 4q27, 11p14.3/*SLC17A6*, and 17p12 with gains of *BCL3* at 19q13.31 were described associated with primary refractory classical HL
  - Tumors showing 16q13 gains were associated with a chemosensitivity phenotype
  - Recently, rare cases showing rearrangements of *JAK2* including the recurrent t(4;9)(q21;p24)/*SEC31A–JAK2* have been described

33.3.7 Primary Cutaneous B Cell Lymphomas

- This category comprises a relatively heterogeneous group of disorders that include B cell lymphomas that occur in the skin without evidence of extracutaneous disease at presentation

33.3.7.1 Primary Cutaneous Marginal Zone Lymphoma (PCMZL)

- Clinical features
  - One on the most common B cell lymphomas of skin
  - More common in middle-aged adults
  - There is a predilection for the upper extremities and trunk
  - Tumors may be multiple
  - No ulceration
  - Some cases have been associated with *Borrelia burgdorferi* in reports mainly from Europe
- Pathological features
  - Reactive GC surrounded by a proliferation of marginal zone cells, plasma cells, and reactive T cells (sometimes very numerous)
  - The epidermis is not involved (grenz zone of uninvolved papillary dermis)
  - Follicular colonization can be prominent; these lesions closely mimic FL
  - Plasmacytic differentiation, plasma cells with intranuclear pseudoinclusions (Dutcher bodies)
- Immunologic features
  - Positive for pan-B cell markers: CD20, CD19, and CD22
  - Negative for CD5, CD10, CD23, LMO2, and BCL6
  - In colonized follicles, the neoplastic cells are characteristically BCL2-positive and BCL6- and CD10-negative admixed with nonneoplastic follicular center cells (BCL2-negative and positive for LMO2, BCL6, and CD10)
  - Monotypic plasma cells (usually more prominent beneath the epidermis)
- Molecular features
  - Clonal rearrangement of Ig heavy chain genes is detected in 70% of the cases using PCR-based methods
  - BCL2 protein expression is common but is not associated with the t(14;18)(q32;q21)
  - Trisomy 13 (∼20%), 18 (∼4%), and *CDKN2A (p16)* deletions
  - Inactivation of *CDKN2B (p15)* and *CDKN2A* genes as a result of promoter hypermethylation
  - t(14;18)(q32;q21), involving *MALT1* gene and *IGH*, has been described in a small subset of cases
  - t(3;14)(p14.1;q32) involving *FOXP1* and *IGH* has also been described in a small subset of cases
  - FISH can be used to detect MALT lymphoma-associated translocations

33.3.7.2 Primary Cutaneous Follicle Center Lymphoma (PCFCL)

- Lymphoma arising in the skin composed of follicular center cells
- Clinical features
  - Most common primary cutaneous B cell lymphoma
  - Affects adults, median age of 65 years, and affects slightly more men than women
  - Most patients present with multiple or solitary erythematous papules, plaques, or nodules in the head and region often involving the scalp
  - No ulceration
  - Favorable prognosis, even in patients with multiple skin lesions
- Pathological features
  - Dermal infiltrate composed of intermediate to large centrocytes and variable number of centroblasts in follicular, follicular and diffuse, or diffuse (most common) patterns
  - The neoplastic follicles are usually ill-defined and lack tangible body macrophages and mantle zones
  - Reactive T cells can be numerous
  - The epidermis is not involved (grenz zone of uninvolved papillary dermis)
  - Sclerosis is often present
  - Those arising in the legs or those expressing FOXP1 appear to have worse prognosis
- Immunologic features
  - Positive for CD20, CD79a, LMO2, and BCL6
  - BCL2 is negative or dimly positive
  - CD10 is positive in cases with follicular pattern and negative in cases with diffuse pattern
  - Negative for CD5, CD43, and MUM1/IRF4

  - CD21, CD23, or CD35 highlights follicular dendritic cell meshworks
- Molecular features
  - t(14;18)(q32;q21) frequently occurs in PCFCL
    - By PCR is detected in 60% of the cases
    - By FISH is detected in 41%
  - The cases positive for t(14;18) are more frequently positive for BCL2 (54.5% vs. 25%)
  - The t(3;14)(q27;q32) involving *BCL6* has been detected in a small subset of cases (~7%). These cases were negative for the t(14;18), suggesting that these aberrations are mutually exclusive
  - CGH studies in PCFCL have shown amplification in 2p13–15
  - Upregulation of *SPINK2* (GC gene, most differentially expressed gene between PCFCL and primary cutaneous DLBCL, leg type)

33.3.7.3 Primary Cutaneous Diffuse Large B Cell Lymphoma-Leg Type (PCDLBCL-LT)

- PCDLBCL-LT is a primary cutaneous DLBCL composed exclusively of large transformed B cells without significant admixture of centrocytes, most commonly arising on the lower leg. These lymphomas can arise at other skin sites
- Clinical features
  - PCDLBCL-LT is a rare tumor and represents 4% of all cutaneous lymphomas and 20% of primary cutaneous B cell lymphomas
  - It occurs late in life, with more than 80% of cases occurring in patients older than 70 years and is more frequent in women (male to female ratio: 1:1.6, as high as 1:4 in some studies)
  - Usually present with red or blue-red lesions on skin, often with ulceration
  - Most cases, ~85% of all cases, arise in skin of lower leg, one or both legs
  - A subset of cases, ~15% of cases, arises in skin of other sites (trunk, arms, head, and neck)
  - The neoplasm can present with single or multiple (~20%) lesions
  - B symptoms are present in 10–20% of patients
  - This is the cutaneous B cell lymphoma associated with the worst prognosis, with a 5-year disease survival rate of 50%
  - European studies have shown that in cutaneous lymphoma, leg location is an adverse prognostic factor
- Pathological features
  - Monomorphic population of large lymphoid cells with vesicular chromatin and round nuclei resembling centroblasts or immunoblasts
  - Diffuse pattern of growth
  - Usually deep lymphoid infiltrate involving subcutaneous adipose tissue
  - Mitotic figures are usually numerous
  - There are few small reactive T cells in the background and no epidermotropism
  - Important features are the monomorphism of the infiltrate and the lack of background inflammatory cells
- Immunologic features
  - Positive for pan-B cell markers, IgM, BCL2 (strong), IRF4/MUM1, FOXP1, and BCL6
  - The intensity of the BCL2 staining may exceed that of the T cells
  - Negative for CD10
  - No FDC meshworks are usually seen
  - They have a high proliferation index
- Molecular features
  - Monoclonal *IGH* gene rearrangements are frequently detected
  - The t(14;18) is absent
  - Translocations or gains of *MYC* in a subset of cases
  - Somatic mutations of *BCL6* in a subset of cases that also may contribute to deregulation of BCL6
  - Highly complex genetic changes that correlate with the high histological grade and unfavorable outcome of this neoplasm
  - CGH studies have showed frequent imbalances that include gains in 18q, 1q, 7, 12q, and Xp and losses in 6q
  - Deletions of *CDKN2A* (67%), *TP53*, and *RB* genes are frequent

- Gene expression studies identified a profile consistent with activated B cell phenotype
- Upregulation of genes associated with cellular proliferation such as *cyclin E*, *CDC6*, *PCNA*, *DP1*, *MYC*, and *MYBL2*
- Upregulation of genes involved in B cell signaling, such as *IgM heavy chain*, *LYN*, *BLK*, and the B cell transcription factors *MUM1/IRF4*, *OCT2, PIM1*, and *PIM2*

33.3.8 Cutaneous T Cell Lymphomas

- This category comprises a heterogeneous group of disorders that include T cell lymphomas occurring in the skin without evidence of extracutaneous disease at presentation
  - Represents 70% of all primary cutaneous lymphomas
  - Most frequent cutaneous T cell lymphoma is mycosis fungoides (MF)

33.3.8.1 Mucosis Fungoides (MF)

- Primary cutaneous T cell lymphoma characterized by epidermotropism and with a clinical course showing stepwise evolution of patches, plaques, and tumors
- Clinical features
  - Adults usually in the fifth to sixth decades, although any age group may be involved
  - Incidence is higher among African-Americans than among whites
  - More frequent in men than women
  - Patches, plaques, and eventually tumors that typically occur on sun-protected skin (breast, axilla, buttocks being the most common sites)
  - The most common presentation is one of patches and plaques
- Pathological features
  - Lichenoid lymphoid infiltrate with "toy soldier" lining up of atypical lymphoid cells at the dermoepidermal junction
  - Epidermotropism with little spongiosis
  - Usually the epidermotropic lymphocytes have more atypical nuclear cytologic features than those in the dermis
  - Pautrier microabscesses (clustering of T cells around Langerhans cells)
  - Wiry collagen in the papillary dermis
- Immunologic features
  - MF is usually a neoplasm of CD4-positive helper T cells of α/β phenotype
  - CD8-positive MF is rare and usually presents with hypopigmented lesions
  - Pagetoid reticulosis is another CD8-positive MF
  - Double-negative CD4/CD8 MF is rare but exists. Loss of CD4 may occur during the clinical course
- Molecular features
  - *TCR* genes are frequently clonally rearranged
  - Notch1 is expressed in advanced MF and positively correlates with disease stage
  - Gains of 17q and 18, with minimal common regions of deletion at 1p31–p36 and 10q26, have been described as common alterations in MF
  - Losses involving 1p, 17p, 10q/10, and 19
  - Gains of 7q36 and 7q21–7q33 and losses of 5q13 and 9p21 have also been described in transformed MF
  - Large cell transformation of MF: two groups, one with low number of chromosomal abnormalities and another with high number and chromosomal instability that is associated with short survival and disease progression
    - Deletion of 9p21.3 (*CDKN2A*, *CDKN2B*, and *MTAP*)
    - Gain of 8q24.21 (*MYC*)
    - Deletion of 10q26qter (*MGMT* and *EBF3*)
  - *NAV3* (neuron navigator 3) translocation/deletion at 12q21–q22 is rare
  - Hypermethylation of *CDKN2A* and *2B* and *BCL7a* genes has been described
  - *IRF4* gene rearrangements are detected in a subset of transformed MF (~20%)
  - Sézary syndrome has a different genomic profile than transformed MF

33.3.8.2 Primary Cutaneous Anaplastic Large Cell Lymphoma (C-ALCL)

- Primary cutaneous lymphoma composed of large T cells that express CD30
  - Part of the spectrum of CD30-positive lymphoproliferative disorders involving the skin (that includes lymphomatoid papulosis)
- Clinical features
  - Adults usually in the sixth decade, although any age group may be involved
  - More frequent in males than females
  - Solitary nodule or tumors frequently ulcerated (face, trunk, and extremities)
  - They may show partial or complete spontaneous regression, as in lymphomatoid papulosis
  - Extracutaneous dissemination may occur
- Pathological features
  - Diffuse infiltrate of large neoplastic cells mainly located in dermis usually extending into the subcutis
  - Neutrophilic-rich variant in which a confluence of neutrophils may obscure the large atypical tumor cells
- Immunologic features
  - Positive for CD30, CD4, and cytotoxic phenotype
  - Rare cases positive for CD8
  - Variable loss of pan-T cell markers: CD2, CD3, and CD5
  - Negative for ALK
- Molecular features
  - Most cases have monoclonal TCR gene rearrangements
  - Complex and heterogeneous karyotypic abnormalities
  - The t(2;5) has been described only rarely
  - Deletions at 9p21–22 (*P16*), 16q22.11 (*CTCF*), 16q24.3 (*ANKRD11*), 13q34 (*ING1*), 17p13 (*P53*), and 20q13.13 (*CEBPB*)
  - Gains at 16p13.3pter and 1p38.32pter (*MIB2*, *SKI*, or *PRDM16*)
  - Translocations involving *IRF4* (~30% of the cases)

33.3.8.3 Subcutaneous Panniculitis-Like T Cell Lymphoma (SPTCL)

- SPTCL is a T cell lymphoma that preferentially involves subcutaneous tissue and expresses TCRαβ and cytotoxic proteins
- Clinical features
  - SPTCL affects adults and children; it has a broad age range (median age of 30 years [range, 9–79]) and slight predominance in women
  - Suggestion of an association with lupus erythematosus
  - Multiple subcutaneous nodules in the extremities and the trunk
  - Ulceration is rare
  - Systemic symptoms (fever, fatigue, weight loss) can be seen 60% of the cases
  - Hemophagocytic syndrome can also be present (~15%) and is associated with aggressive clinical course
  - Usually clinically indolent and prolonged remission with therapy
- Pathological features
  - Primarily involves subcutaneous adipose tissue
  - The dermis is usually not involved
  - Dense lymphoid infiltrate with involvement of the lobules and sparing the septa
  - Neoplastic cells are usually small to intermediate in size and often rim individual dipocytes
  - Some cases have large cells
  - Karyorrhexis and histiocytes with apoptotic debris are common
- Immunologic features
  - The tumor cells have an αβ cytotoxic T cell phenotype
  - Positive for CD3, CD8, granzyme B, TIA1, and perforin
  - Usually negative for CD4 and CD56
- Molecular features
  - Clonal rearrangement of TCR genes is detected in most cases
  - One CGH study showed a relatively uniform DNA copy number changes
    - Losses in chromosomes 1p, 2p, 2q, 5p, 7p, 9q, 10q, 11q, 12q, 16, 17q, 19, 20, and 22

- Gains in chromosomes 2q, 4q, 5q, 6q, and 13q
- 10q, 17p, and 19 are likely to harbor genes common for the molecular pathogenesis of several subtypes of primary cutaneous T cell lymphomas, whereas 5q and 13q gains characterize SPTL
- Deletion of *NAV3* gene at 12q21

Further Reading

Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B cell lymphoma identified by gene expression profiling. Nature. 2000;403:503–11.

Blum KA, Lozanski G, Byrd JC. Adult Burkitt leukemia and lymphoma. Blood. 2004;104:3009–20.

Garcia JF, Mollejo M, Fraga M, Forteza J, et al. Large B cell lymphoma with Hodgkin's features. Histopathology. 2005;47:101–10.

Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999;94:1848–54.

Herling M, Khoury J, Washington L, Duvic M, Keating M, Jones D. A systematic approach to diagnosis of mature T cell leukemias reveals heterogeneity among WHO diagnostic categories. Blood. 2004;104:328–35.

Hervé AL, Florence M, Philippe M et al. Molecular heterogeneity of multiple myeloma: pathogenesis, prognosis and therapeutic implications. J Clin Oncol. 2011;29:1893–7.

Isaacson PG, Du MQ. MALT lymphoma: from morphology to molecules. Nat Rev Cancer. 2004;8:644–53.

Jaffe ES, Harris NL, Stein H, Vardiman JW, editors. World health classification of tumours. Pathology and genetics of tumours of haematopoietic and lymphoid tissues. Lyon: IARC Press; 2001.

Klein U, Dalla-Favera R. Germinal centres: role in B cell physiology and malignancy. Nat Rev Immunol. 2008;8:22–33.

Korsmeyer SJ, Hieter PA, Ravetch JV, Poplack DG, Waldmann TA, Leder P. Developmental hierarchy of immunoglobulin gene rearrangements in human leukemic pre-B cells. Proc Natl Acad Sci USA. 1981;78:7096–100.

Kuppers R. The biology of Hodgkin's lymphomas. Nature Rev Cancer. 2009;9:15–27.

Kuppers R, Klein U, Hansmann ML, Rajewsky K. Cellular origin of human B cell lymphomas. N Engl J Med. 1999;341:1520–9.

Medeiros LJ, Carr J. Overview of the role of molecular methods in the diagnosis of malignant lymphomas. Arch Pathol Lab Med. 1999;123:1189–207.

Mollejo M, Camacho FI, Algara P, Ruiz-Ballesteros E, Garcia JF, Piris MA. Nodal and splenic marginal zone B cell lymphomas. Hematol Oncol. 2005;23:108–18.

Rajewsky K. Clonal selection and learning in the antibody system. Nature. 1996;381:751–8.

Rosenwald A, Wright G, Leroy K, et al. Molecular diagnosis of primary mediastinal B cell lymphoma identifies a clinically favorable subgroup of diffuse large B cell lymphoma related to Hodgkin lymphoma. J Exp Med. 2003;198:851–62.

Ruiz-Ballesteros E, Mollejo M, Rodriguez A, et al. Splenic marginal zone lymphoma: proposal of new diagnostic and prognostic markers identified after tissue and cDNA microarray analysis. Blood. 2005;106:1831–8.

Salaverria I, Perez-Galan P, Colomer D, Campo E. Mantle cell lymphoma: from pathology and molecular pathogenesis to new therapeutic perspectives. Haematologica. 2006;91:11–6.

Salavierra I, Siebert R. The gray zone between Burkitt's lymphoma and diffuse large B cell lymphoma from a genetics perspective. J Clin Oncol. 2011;29:1835–43.

Sigal S, Ninette A, Rechavi G. Microarray studies of prognostic stratification and transformation of follicular lymphomas. Best Pract Res Clin Haematol. 2005;18:143–56.

Stewart AK, Fonseca R. Prognostic and therapeutic significance of myeloma genetics and gene expression profiling. J Clin Oncol. 2005;23:6339–44.

Tonegawa S. Somatic generation of antibody diversity. Nature. 1983;302:575–81.

Vega F, Medeiros LJ. Chromosomal translocations involved in non-Hodgkin lymphomas. Arch Pathol Lab Med. 2003;127:1148–60.

Wiestner A, Rosenwald A, Barry TS, et al. ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. Blood. 2003;101:4944–51.

Witzig TE. Current treatment approaches for mantle-cell lymphoma. J Clin Oncol. 2005;23:6409–14.

Wu G, Keating A. Biomarkers of potential prognostic significance in diffuse large B cell lymphoma. Cancer. 2006;106:247–57.

Myeloproliferative Neoplasms and Myelodysplastic Syndromes: Molecular Diagnostics

34

Gerlinde Wernig and Daniel A. Arber

Contents

34.1 Myeloproliferative Neoplasms: Introduction 930

34.2 Clinical and Molecular Genetic Features of Myeloproliferative Neoplasms 930
34.2.1 Chronic Myelogenous Leukemia 931
34.2.2 Polycythemia Vera, Essential Thrombocythemia, and Primary Myelofibrosis 932
34.2.3 Chronic Eosinophilic Leukemia, NOS 933
34.2.4 Chronic Neutrophilic Leukemia 934
34.2.5 Systemic Mastocytosis 934

34.3 Molecular Diagnostics of Myeloproliferative Neoplasms and Implications for Clinical Management 934
34.3.1 Chronic Myelogenous Leukemia 934
34.3.2 Polycythemia Vera, Essential Thrombocythemia, and Primary Myelofibrosis 935
34.3.3 Chronic Eosinophilic Leukemia, NOS 936
34.3.4 Chronic Neutrophilic Leukemia 936
34.3.5 Systemic Mastocytosis 936

34.4 Molecular Genetic Features and Diagnosis of Myeloid and Lymphoid Neoplasms with Eosinophilia and Abnormalities of *PDGFRA*, *PDGFRB*, and *FGFR1* 936

34.5 Molecular Genetic Features and Diagnosis of Myelodysplastic/Myeloproliferative Neoplasms 937
34.5.1 Chronic Myelomonocytic Leukemia 937
34.5.2 Juvenile Myelomonocytic Leukemia 937
34.5.3 Atypical Chronic Myeloid Leukemia, *BCR–ABL1* Negative 937
34.5.4 Myelodysplastic/Myeloproliferative Neoplasm, Unclassifiable 938
34.5.5 Refractory Anemia with Ring Sideroblasts Associated with Marked Thrombocytosis 938

34.6 Myelodysplastic Syndromes: Introduction 938

34.7 Molecular Genetic Features of Myelodysplastic Syndromes 938
34.7.1 Chromosome 5q Deletions 938
34.7.2 Trisomy 8 939
34.7.3 Chromosome 7 and 7q Deletions 940
34.7.4 Chromosome Y and 20q Deletions 940
34.7.5 Abnormalities Involving Chromosome 3q26 941
34.7.6 Other Cytogenetic Abnormalities 941
34.7.7 Gene Mutations 941
34.7.8 Epigenetic Changes in Myelodysplastic Syndromes 943
34.7.9 The Bone Marrow Microenvironment 943

34.8 Molecular Diagnostics and Implications for Clinical Management of Myelodysplastic Syndromes 943

Further Reading 944

G. Wernig, MD (✉)
Department of Pathology, Stanford University School of Medicine, Stanford Hospital and Clinics, Stanford, CA, USA

D.A. Arber, MD
Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_34,

34.1 Myeloproliferative Neoplasms: Introduction

- Myeloproliferative neoplasms (MPN) are clonal proliferations of one or more maturing bone marrow cell lines that may progress to marrow fibrosis or acute leukemia
- Sixty years ago, clinical and morphologic similarities between chronic myelogenous leukemia (CML), polycythemia vera (PV), essential thrombocythemia (ET), and primary myelofibrosis (PMF) were first noticed and categorized as myeloproliferative disorders (MPD) by William Dameshek
- In 2005, William Vainchenker and several others described the *JAK2*V617F mutation, which is present in high frequency in PV, ET, and PMF and subsequently was shown to cause a MPN-like phenotype in mice
- In 2008 the WHO classification system subcommittee changed the terminology from MPD to MPN
- MPN category includes now CML; PV; ET; PMF; chronic neutrophilic leukemia; chronic eosinophilic leukemia, not otherwise specified; mastocytosis; and MPN unclassifiable (Table 34.1)

Table 34.1 Classification of myeloid neoplasms according to the 2008 World Health Organization Classification scheme

Myeloproliferative neoplasms (MPN)
Chronic myelogenous leukemia, *BCR–ABL1* positive (CML)
Chronic neutrophilic leukemia (CNL)
Polycythemia vera (PV)
Primary myelofibrosis (PMF)
Essential thrombocythemia (ET)
Chronic eosinophilic leukemia, not otherwise specified (CEL-NOS)
Mastocytosis
Myeloproliferative neoplasm, unclassifiable (MPN-U)
Myeloid and lymphoid neoplasms with eosinophilia and abnormalities of PDGFRA, PDGFRB, and FGFR1
Myeloid and lymphoid neoplasms with *PDGFRA* rearrangement
Myeloid neoplasms with *PDGFRB* rearrangement
Myeloid and lymphoid neoplasms with *FGFR1* abnormalities
Myelodysplastic/myeloproliferative neoplasms
Chronic myelomonocytic leukemia
Atypical chronic myeloid leukemia, *BCR–ABL1* negative
Juvenile myelomonocytic leukemia
Myelodysplastic/myeloproliferative neoplasm, unclassifiable
Myelodysplastic syndromes
Refractory cytopenia with unilineage dysplasia
Refractory anemia with ring sideroblasts
Refractory cytopenia with multilineage dysplasia
Refractory cytopenia with excess blasts
Myelodysplastic syndrome with isolated del(5q)
Myelodysplastic syndrome, unclassifiable
Childhood myelodysplastic syndrome

34.2 Clinical and Molecular Genetic Features of Myeloproliferative Neoplasms

- CML originates from an abnormal fusion gene t(9;22), *BCR–ABL1* (the Philadelphia chromosome) in a pluripotent bone marrow stem cell
- Initially CML presents as a neutrophilic leukocytosis, but the *BCR–ABL1* is found in all myeloid lineages, some lymphoid cells, and endothelial cells
- PV, ET, and PMF are stem cell-derived monoclonal or oligoclonal diseases harboring mutations at the progenitor cell level
- Currently known MPN-associated mutations involve:
  - Janus kinase 2 (*JAK2*; exon 14)
  - Myeloproliferative leukemia (*MPL*; exon 10)
  - TET oncogene family member 2 (*TET2*)
  - Additional sex combs-like 1 (*ASXL1*)
  - Isocitrate dehydrogenases 1 and 2 (*IDH1* and *IDH2*)
  - Casitas B-lineage lymphoma (*CBL*)
  - IKAROS family zinc finger 1 (*IKZF1*)
  - SH2B adaptor protein 3 (*SH2B3*)
  - Enhancer of zeste homolog 2 (*EZH2*)
- Chronic eosinophilic leukemia, not otherwise specified (NOS), is a clonal proliferation of eosinophil precursors with eosinophilia in the peripheral blood, bone marrow, and peripheral tissues, negative for the Philadelphia

chromosome, *BCR–ABL1* fusion gene, or rearrangement of *PDGFRA, PDGFRB*, or *FGFR1*
- Similarly, chronic neutrophilic leukemia is a clonal proliferation of mature neutrophils in the absence of these same abnormalities
- Mastocytosis is a clonal, neoplastic proliferation of abnormal mast cells accumulating in bone marrow and/or other extracutaneous organs that frequently demonstrate mutations of *KIT*

34.2.1 Chronic Myelogenous Leukemia

- CML is a clonal hematopoietic stem cell disorder with an annual incidence of one to two cases per 100,000 per year
- The target cell that initiates CML is a granulocyte–macrophage-like progenitor cell
- Three clinical phases
  - Indolent chronic phase confined to hematopoietic organs with an average of 4–6 years
  - Accelerated phase
  - Blast crisis with involvement of extramedullary tissues including lymph nodes, skin, and soft tissues
- Chronic phase is characterized by peripheral blood leukocytosis $\sim 100 \times 10^9$/L with peaks in the differential percentage of myelocytes and segmented neutrophils, $<2\%$ blasts in the peripheral blood, and markedly decreased neutrophil alkaline phosphatase
- Major pathogenic event is translocation (9;22)(q34;q11), known as the Philadelphia chromosome, resulting in the *BCR–ABL1* fusions p210 and p190
- Frequency of t(9;22)(q34;q11) by karyotype in CML patients: 90–95%
- Rare cases harbor variant translocations involving a third/fourth chromosome or a cryptic translocation of 9q34 and 22q11.2 only detectable by FISH, RT-PCR, or Southern blot
- *BCR–ABL1* fusion is essential for initiation, maintenance and progression of CML, yet the transformation of CML from chronic phase to blast phase requires additional genetic and/or epigenetic abnormalities; for example, deletion of IKAROS family zinc finger 1 (*IKZF1*), shown to be an acquired lesion at the time of transformation of CML to lymphoid blast crisis (ALL), has been identified in 83.7% of *BCR–ABL1* ALL, but not in chronic-phase CML
- Various *BCR–ABL1* breakpoints are associated with different disease phenotypes
  - Breakpoint in the major breakpoint cluster region (M-BCR) creates an abnormal fusion protein $p210^{BCR–ABL1}$
    - $p210^{BCR–ABL1}$ is detected in the majority of CML and associated with excessive proliferation of the granulocytic lineage
    - This fusion protein is also detected in one-third of ALL
  - Breakpoint in the minor breakpoint cluster region (m-BCR) results in the $p190^{BCR–ABL1}$ transcript
    - $p190^{BCR–ABL1}$ is detected in two-thirds of ALL and only rare cases of CML and acute myeloid leukemia (AML)
    - In CML it is associated with prominent monocytosis and basophilia
  - A 3′ breakpoint in BCR creates $p230^{BCR–ABL1}$
    - $p230^{BCR–ABL1}$ is rarely identified and results in CML with predominantly mature neutrophils
    - This form of CML may be mistaken for chronic neutrophilic leukemia
- The BCR–ABL1 protein has increased tyrosine kinase (TK) activity with constitutive activation of many signal transduction pathways, e.g., RAS–RAF–ERK, JAK–STAT, and PI3K kinase
  - Hundreds of tyrosine-phosphorylated proteins in CML cells are dependent on BCR–ABL1 activity as has been revealed by recent phosphoproteomic analyses
  - Unlike c-Abl, which shuttles between the nucleus and cytoplasm, BCR–ABL1 is largely cytoplasmic; its unregulated TK activity allows it to use signaling pathways normally activated by growth factor receptor TKs, such as EGF and PDGF receptors, or cytokine receptors that utilize

nonreceptor TKs, such as Src and JAK family kinases, for signaling
- Pathogenetic defects include
  - Increased proliferation due to activation of the RAS pathway
  - Decreased apoptosis of hematopoietic stem cell or progenitor cells mediated in part through STAT5 upregulation of the antiapoptotic molecule BCLXL
  - Phosphorylation of and inactivation of the proapoptotic molecule BAD by AKT which leads to massive increase in myeloid cell numbers
  - Premature release of immature myeloid cells into the circulation due to a defect in adherence of myeloid progenitors to marrow stroma due to reduced adhesion to fibronectin
  - Genetic instability resulting in disease progression

34.2.2 Polycythemia Vera, Essential Thrombocythemia, and Primary Myelofibrosis

34.2.2.1 JAK2V617F Mutations

- *JAK2*V617F (Janus kinase 2; 9p24) is the most prevalent mutation in these disorders (Fig. 34.1)
- Frequency of mutation in patients: 96% in PV, 55% in ET, and 65% in PMF
- Mutation affects the noncatalytic (pseudokinase) domain of *JAK2*, disrupting kinase regulatory activity
- Forced expression of *JAK2*V617F induces a PV-like phenotype in mice
- Specific characteristics of *JAK2*V617F in MPN are older age, higher hemoglobin level, leukocytosis, and lower platelet count
- Higher mutant allele burden is associated with pruritus and fibrotic transformation

34.2.2.2 JAK2 Exon 12 Mutations

- *JAK2* exon 12 mutations are relatively specific to *JAK2*V617F-negative PV
- Frequency in patients: 3% in PV; *JAK2* N542-E543del is the most frequent among the *JAK2* exon 12 mutations
- Specific characteristics of *JAK2* exon 12 mutation-positive patients: mutations are heterozygous and disease presents at a younger age

34.2.2.3 MPL Mutations

- MPL virus (*MPL* oncogene; 1p34) W515L, *MPL*W515K and other exon 10 *MPL* mutations
- Frequency in patients: 3% in ET and 10% in PMF
- Specific characteristics of *MPL* mutations are older age, female sex, lower hemoglobin level, and higher platelet count

34.2.2.4 TET2 Mutations

- Frequency of *TET2* (TET oncogene family member 2; 4q24) mutations in patients: 16% in PV, 5% in ET, 17% in PMF, 14% in post-PV myelofibrosis (MF), 14% in post-ETMF, and 17% in blast-phase MPN
- Occur in *JAK2*V617F-positive and negative MPN
- Occur before, after, or independently of a *JAK2* mutation
- *TET2* and *ASXL1* probably contribute to epigenetic regulation of hematopoiesis

34.2.2.5 ASXL1 Mutations

- Frequency of additional sex combs-like 1 (*ASXL1*; 20q11.1) mutations: 8% of *JAK2*V617F-negative MPN, 11% of myelodysplastic syndromes (MDS), 43% of chronic myelomonocytic leukemia (CMML), and 7% of primary and 47% of secondary AML

34.2.2.6 IDH1 and IDH2 Mutations

- Frequency of isocitrate dehydrogenase (*IDH1* R132 and *IDH2* R140, R172; 2q33.3 and 15q26.1, respectively) mutations in MPN: 0.8% in ET, 1.9% in PV, 4.2% in PMF, 1% in post-PV/ET MF, and 21.6% in blast-phase MPN
- Mutant *IDH1/2* is detected in the presence or absence of *JAK2*, *MPL*, and *TET2* mutations
- Functional characterization of *IDH1/2* mutations suggests neoenzymatic activity in converting alpha ketoglutarate to the putatively oncogenic 2-hydroxyglutarate

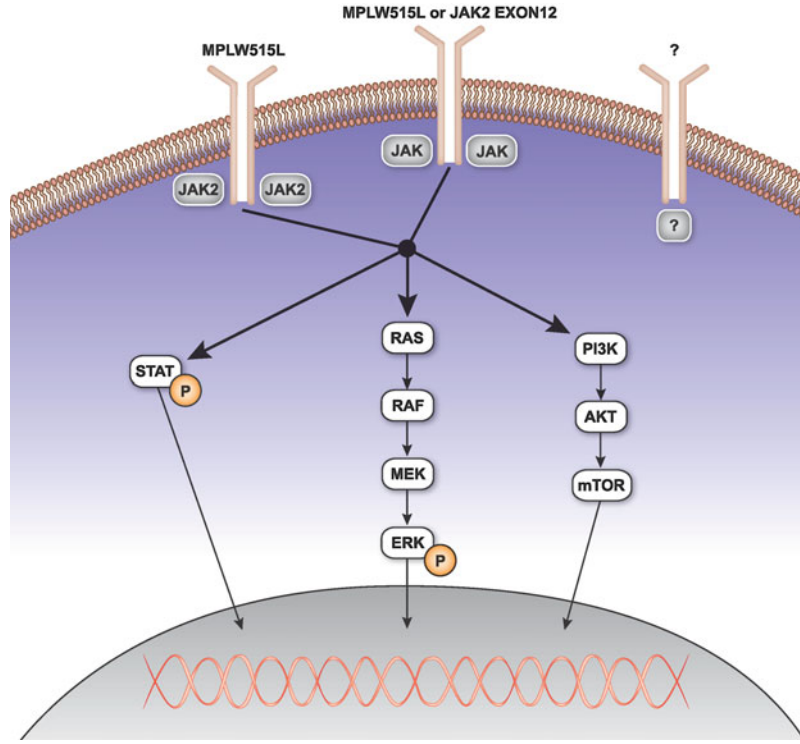


Fig. 34.1 Point mutations in myeloproliferative neoplasms. Mutations in *JAK2* and *MPL* resulting in constitutive activation of the JAK/STAT, RAS, and PI3K pathways have been implicated in the pathogenesis of myeloproliferative neoplasms

34.2.2.7 CBL Mutations

- Frequency of casitas B-lineage lymphoma (CBL proto-oncogene; 11q23.3) mutations: 17% of juvenile myelomonocytic leukemia, 11% of CMML, and 6% (exon 8 or 9) in PMF
- Homozygous mutations in juvenile myelomonocytic leukemia suggest a tumor suppressor function for the normal protein

34.2.2.8 IKZF1 Mutations

- Frequency of *IKZF1* (7p12) mutation: occurs in blast-phase CML or *BCR–ABL1*-positive ALL, 19% of *BCR–ABL1*-negative MPN, and <0.5% in blast- and chronic-phase disease

34.2.2.9 SH2B3 Mutations

- SH2B adaptor protein 3 (*SH2B3*) (also known as *LNK*) (12q24.12) encodes a plasma membrane-bound lymphocyte-specific adaptor protein that inhibits wild-type and mutant JAK2 signaling
- *SH2B3* exon 2 loss of function mutations described in *JAK2*V617F-negative ET or PMF

34.2.2.10 EZH2 Mutations

- *EZH2* (7q36.1) encodes a catalytic subunit of the polycomb repressive complex 2, a histone H3 lysine 27 methyltransferase with putative epigenetic effect
- Frequency of *EZH2* mutations: 12% of MDS/MPN and 13% of MF

34.2.3 Chronic Eosinophilic Leukemia, NOS

- Eosinophilia $>1.5 \times 10^9$/L
- Negative for Ph chromosome, *BCR–ABL1* fusion gene, or myeloproliferative neoplasm (PV, ET, PMV) or MDS/MPN (CMML or aCML)
- Negative for t(5;12) or other rearrangement of *PDGFRB*

- Negative for *FIP1L1–PDGFRA* fusion gene or other rearrangement of *PDGFRA*
- Negative for rearrangement of *FGFR1*
- Blast cell count in peripheral blood and bone marrow <20%
- Negative for inv(16) or t(16:16) or other diagnostic features of AML

34.2.4 Chronic Neutrophilic Leukemia

- Peripheral blood neutrophilia with leukocytosis >25 × 10^9/L, bone marrow hypercellularity due to neutrophilic granulocyte proliferation, and hepatosplenomegaly
- Negative for reactive neutrophilia
- Negative for Ph chromosome and *BCR–ABL1* fusion gene
- Normal karyotype in 90% of patients, remainder patients with trisomies 8 and 9 and deletions 20q and 11q
- Prognosis variable from 6 months to >20 years
- Rare disease

34.2.5 Systemic Mastocytosis

- Major criterion
  - Multifocal, dense infiltrates of mast cells (>15 mast cells in aggregates) detected in bone marrow and/or other extracutaneous organ
- Minor criteria
  - >25% of mast cells are immature or atypical
  - Activating point mutation at codon 816 of *KIT*, D816V, D816Y, D816H, and D816F
  - Abnormal expression of CD2 and/or CD25 in addition to normal mast cell markers tryptase and CD117
  - Serum total tryptase >20 ng/mL
- Required for the diagnosis are one major and one minor, or at least three minor criteria
- Frequency of *KIT* D816V: >90% of patients

34.3 Molecular Diagnostics of Myeloproliferative Neoplasms and Implications for Clinical Management

34.3.1 Chronic Myelogenous Leukemia

- Clinical and laboratory features required for a diagnosis of CML: presence of *BCR–ABL1* fusion
- Imatinib, an inhibitor of the TK activity of *BCR–ABL1*, has been successfully used to treat patients with chronic-phase CML; imatinib resistance occurs in particular in blast crisis usually as a result of mutations in *BCR–ABL1* in Thr315Ile (T315I) mutation
- Overall, survival for newly diagnosed chronic-phase patients treated with imatinib at 5 years is 89%, an estimated 93% of imatinib-treated patients remain free from disease progression to the accelerated phase or blast crisis, and additional 6% of patients have shown loss of response to imatinib; however their disease has not progressed to the accelerated phase or blast crisis
- 70–90% of patients achieve a complete cytogenetic response, but the majority of these patients have detectable *BCR–ABL1* by RT-PCR
- 50% of patients who relapse on imatinib therapy have *BCR–ABL1* point mutations in at least 40 different amino acids in the *ABL1* kinase domain
- Two main mechanisms of relapse
  - Mutations at contact points between the ABL1 kinase domain and imatinib, e.g., T315I and F359V
  - Mutations involving conformational changes in the P-loop which bridges the ATP-binding pocket of the kinase domain (M244V, G250E, Q252H, Y253F/H, and E255K/V) and the activation loop (H396R/P)
- Second generation *BCR–ABL1* inhibitors can block activity of imatinib-resistant mutant

forms of *BCR–ABL1* except Thr315Ile; for example, nilotinib has 10–30-fold increased potency against the major resistant mutants; dasatanib, a c-Src and c-Abl inhibitor, is effective against all resistant mutants except Thr315Ile
- Curative treatment with allogeneic stem cell transplantation (allo-SCT)

34.3.2 Polycythemia Vera, Essential Thrombocythemia, and Primary Myelofibrosis

- Clinical and laboratory features required for a diagnosis of PV, ET, or PMF
  - Exclusion of the presence of *BCR–ABL1* fusion
  - Bone marrow examination, peripheral blood mutation screening for *JAK2*V617F, subnormal serum erythropoietin levels
  - Peripheral blood mutation screening for *JAK2*V617F is highly sensitive (97% sensitivity) and specific (~100%) for distinguishing PV from nonneoplastic causes of increased hematocrit
  - Subnormal serum erythropoietin level in the absence of *JAK2*V617F warrants additional mutational analysis for *JAK2* exon 12 mutations
  - Bone marrow histopathology necessary for accurate morphologic diagnosis of ET and distinction from other myeloid neoplasms including prefibrotic PMF
  - Bone marrow fibrosis associated with *JAK2*V617F, trisomy 9, or 13q– consistent with the diagnosis of PMF
  - Presence of *JAK2*V617F confirms the presence of an underlying MPN
  - Absence of J*AK2*V617F does not rule out an underlying MPN because 45% of patients with ET are *JAK2*V617F negative
  - Dwarf megakaryocytes are a pathological hallmark in CML, should warrant *BCR–ABL1* fluorescent in situ hybridization or polymerase chain reaction analysis
  - Differential diagnosis of PMF: acute panmyelosis with MF, fibrotic MDS, CMML, and fibrotic/spent phase of PV
- Risk factors for shortened survival and leukemic or fibrotic transformation in PV and ET: history of thrombosis, leukocytosis, advanced age, and anemia
- International Prognostic Scoring System (IPSS) for PMF (Table 34.2) uses five independent predictors of inferior survival: age >65 years, hemoglobin <10 g/dL, leukocyte count >25 × 10^9/L, circulating blasts >1%, and presence of constitutional symptoms
  - Patients with post-PV/ET MF are managed similarly to PMF patients
  - Current drug therapy for PV, ET, or PMF is not curative and has no benefit on survival
  - Allogeneic stem cell transplantation (allo-SCT) is potentially curative in PMF (or post-ET/PV MF), but has high incidence of treatment-related mortality and morbidity
  - Therapy goals in PV and ET: prevent thrombohemorrhagic complications and alleviate anemia, symptomatic splenomegaly, or constitutional symptoms in PMF

34.3.2.1 Conventional Therapies in PV and ET

- PV and ET: low-dose aspirin and phlebotomy are indicated in all patients with PV (target hematocrit of 45%)
- High-risk patients with PV or ET receive hydroxyurea as first-line treatment, IFN-alpha if nonresponders and younger than 65 years and busulfan if older than 65 years

34.3.2.2 Conventional Therapies in PMF

- Low-risk patients: no therapeutic intervention
- High- or intermediate-2-risk patients: consideration for investigational drug therapy or allo-SCT
- Intermediate-1-risk patients: observation, conventional drug therapy, or participation in investigational drug trials
- Indications for treatment are anemia and splenomegaly

Table 34.2 International Prognostic Scoring System for PV, ET, and PMF (IPSS and DIPSS)

IPSS risk category for PV and ET[a]		
Low risk		Age <60 year; no thrombosis history
Low risk with extreme thrombocytosis		Platelet count >1,000 × 10^9/L
High		Age >60 year; positive thrombosis history
DIPSS risk category for PMF[b]		***Median survival (year)***
Low risk	No risk factors[c]	15.4
Intermediate-1	1 risk factor	6.5
Intermediate-2	2 or 3 risk factors	2.9
High	>4 risk factors	1.3

[a]Finazzi et al. (2008)
[b]Adapted from Gangat et al. (2011)
[c]8 risk factors for inferior survival in PMF: age >65 year, hemoglobin <10 g/dL, leukocyte count >25 × 10^9/L, circulating blasts >1%, presence of constitutional symptoms, presence of unfavorable karyotype, platelet count <100 × 10^9/L, red cell transfusion dependence

- Anemia treated with intramuscular androgens or thalidomide and lenalidomide, response rate of approximately 20%
- Symptomatic splenomegaly in PMF is treated with hydroxyurea, splenectomy, or splenic irradiation
- Three investigational drugs (multicentric phase I/II randomized studies) under evaluation in MF and post-PV/ET MF: pomalidomide, a second-generation immunomodulatory drug, and two JAK inhibitor ATP mimetics (TG101348 and INCB018424); other investigational drugs currently in clinical trials for MF, PV, or ET include other kinase inhibitors (e.g., CYT387, CEP-701, AZD1480, SB1518) and histone deacetylase inhibitors (e.g., ITF2357, MK-0683, panobinostat; http://ClinicalTrials.gov)

34.3.2.3 Allogeneic Hematopoietic Stem Cell Transplantation (Allo-SCT)

- Several studies evaluated allo-SCT in PMF and post-PV/ET MF: disease-free survival ranging from 30% to 50% and treatment-related mortality at approximately 50%

34.3.3 Chronic Eosinophilic Leukemia, NOS

- Novel therapies, such as anti-IL-5 antibody therapy may improve current treatment of the idiopathic hypereosinophilic syndrome and chronic eosinophilic leukemia, not otherwise specified

34.3.4 Chronic Neutrophilic Leukemia

- Curative treatment only with allo-SCT

34.3.5 Systemic Mastocytosis

- Cytoreductive therapies for aggressive systemic mastocytosis
- Potential treatment with investigational TK inhibitors other than imatinib
- Treatment of associated hematologic non-mast cell neoplasm when present

34.4 Molecular Genetic Features and Diagnosis of Myeloid and Lymphoid Neoplasms with Eosinophilia and Abnormalities of *PDGFRA*, *PDGFRB*, and *FGFR1*

- "New disease category" in 2008 WHO classification of myeloid neoplasms associated with eosinophilia (and sometimes with neutrophilia or monocytosis) and rearrangements of *PDGFRA*, *PDGFRB*, or *FGFR1*

- *FIP1L1–PDGFRA* fusion is the most common TK fusion identified in myeloid malignancies associated with eosinophilia
- Specific t(5;12) resulting in an abnormal fusion gene, *ETV6* (*TEL*)/*PDGFRB*, occurs in 1–2% of cases previously diagnosed as CMML
- Patients with rearrangements of *PDGFRA* or *PDGFRB* have an excellent and durable response to low-dose imatinib treatment
- Patients with *FGFR1* abnormalities, most often t(8;13)(p11;q12), may have myeloid or lymphoid neoplasms that do not respond to TK inhibitor therapies

34.5 Molecular Genetic Features and Diagnosis of Myelodysplastic/ Myeloproliferative Neoplasms

34.5.1 Chronic Myelomonocytic Leukemia

- Persistent peripheral blood monocytosis $>1 \times 10^9$/L
- Negative for Philadelphia chromosome or *BCR–ABL1* fusion gene
- Negative for *PDGFRA* or *PDGFRB*
- Less than 20% blasts in blood and bone marrow
- Dysplasia in one or more myeloid lineages
- Persistence of monocytosis for at least 3 months and exclusion of other causes of monocytosis
- 20–40% of patients carry clonal nonspecific cytogenetic abnormalities: trisomy 8, monosomy 7, and structural abnormalities of 12p
- 20% of patients with point mutations of *NRAS or KRAS* genes
- 3% of patients with *FLT3–ITD* mutation
- 3% and 13% of patients carry *JAK2* mutations
- Median survival 20–40 months
- Progression to AML in 15–30%
- No specific therapies: cytoreductive treatment with 5-azacitidine and decitabine
- Allo-SCT is the only curative regimen

34.5.2 Juvenile Myelomonocytic Leukemia

- Clonal hematopoietic disorder of childhood (0–14 years)
- Spontaneous proliferation of granulocyte–-macrophage colonies due to acquired hypersensitivity by leukemic progenitor cells to granulocyte–macrophage colony-stimulating factor (GM-CSF)
- Blasts and monocytes account for less than 20%
- Negative for *BCR–ABL1*
- 25% monosomy 7 and 60% normal karyotype
- Associated with two hereditary syndromes
  - Neurofibromatosis 1 (NF1) 500x increased risk for JMML
  - Noonan syndrome, mutations in *PTPN11* encoding the protein tyrosine phosphatase SHP2
- Aberrant signal transduction due to mutations in RAS/MAPK signaling pathway present in 70% of patients: 35% of patients with *PTPN11* mutation, 20% of patients with mutations in *NRAS, KRAS2*, and *NF1*
- *PTPN11* mutations are found to be mutually exclusive with *RAS* or *NF1* mutations
- Rapidly fatal if untreated
- Curative treatment with allo-SCT

34.5.3 Atypical Chronic Myeloid Leukemia, *BCR–ABL1* Negative

- Overlapping clinical features of MDS and MPN
- Leukemic disorder with dysplastic peripheral blood leukocytosis $>13 \times 10^9$/L
- Negative for Ph chromosome or *BCR–ABL1* fusion gene
- Negative for *PDGRFA* or *PDGFRB*
- Neutrophil precursors >10% of leukocytes
- Minimal absolute basophilia <2% of leukocytes
- Minimal absolute monocytosis <10% of leukocyte
- Hypercellular bone marrow with granulocytic proliferation and granulocytic dysplasia;

variable dysplasia in the erythroid and megakaryocytic lineages
- Less than 20% blasts in the blood and bone marrow
- Poor prognosis with median survival 14–29 months
- Treatment with allo-SCT

34.5.4 Myelodysplastic/Myeloproliferative Neoplasm, Unclassifiable

- Overlapping clinical features of MDS and MPN
- Diagnosis of exclusion: negative for *BCR–ABL1*, *PDGFRA*, *PDGFRB*, or *FGFR1*

34.5.5 Refractory Anemia with Ring Sideroblasts Associated with Marked Thrombocytosis

- 60% of patients positive for *JAK2V617F* or less frequently *MPLW515L* mutation
- Subcategory of myelodysplastic/myeloproliferative neoplasm, unclassifiable

34.6 Myelodysplastic Syndromes: Introduction

- MDS are clonal disorders of hematopoiesis characterized by inefficient hematopoiesis, peripheral blood cytopenias, and risk of progression to AML
- Clinical heterogeneity of MDS is due to the genetic complexity with a variety of genetic lesions contributing to disease pathogenesis
- Genetic lesions known to date that cause MDS: copy number changes (genetic amplifications or deletions), mutations that alter the sequence or expression of individual genes, and epigenetic abnormalities
- Although balanced translocations are rare, chromosomal abnormalities evident by standard karyotypic analysis are present in approximately half of patients with MDS
  - Loss of 5q (5q–), trisomy 8, loss of 7 or 7q (−7/7q–), loss of 20q (20q–), and loss of Y (−Y)
  - Copy number changes or acquired uniparental disomy in 75% of patients
  - Mutations alter the sequence and function of oncogenes or tumor suppressor genes (Fig. 34.2)
  - Abnormal epigenetic profiles result in aberrant gene expression
- Molecular lesions are used to guide diagnosis, prognosis, and treatment of MDS
- The IPSS (Table 34.3) incorporates the common karyotypic abnormalities for therapeutic implications: for example, the use of lenalidomide is guided by the presence of a chromosome 5q deletion as it increases the likelihood of cytogenetic and hematologic responses
- Pathogenesis of MDS includes enhanced self-renewal of a hematopoietic stem cell or acquired self-renewal in a progenitor cell, increased proliferative capacity, impaired or block in differentiation, genetic and epigenetic instability, antiapoptotic mechanisms in the disease-sustaining cell, evasion of the immune system, suppression of normal hematopoiesis, and an abnormal bone marrow microenvironment
- MDS occurs when one of the molecular lesions causes dysplastic differentiation of at least one myeloid lineage resulting in ineffective hematopoiesis

34.7 Molecular Genetic Features of Myelodysplastic Syndromes

34.7.1 Chromosome 5q Deletions

- Deletions of chromosome 5q (5q–): most common cytogenetic abnormality, occurring in approximately 15% of patients with MDS
- Two commonly deleted regions (CDRs): more distal CDR in 5q33.1 associated with 5q– syndrome, characterized by absence of other karyotype abnormalities, severe macrocytic anemia, relative thrombocythemia, female predominance, and a lower risk of progression to AML

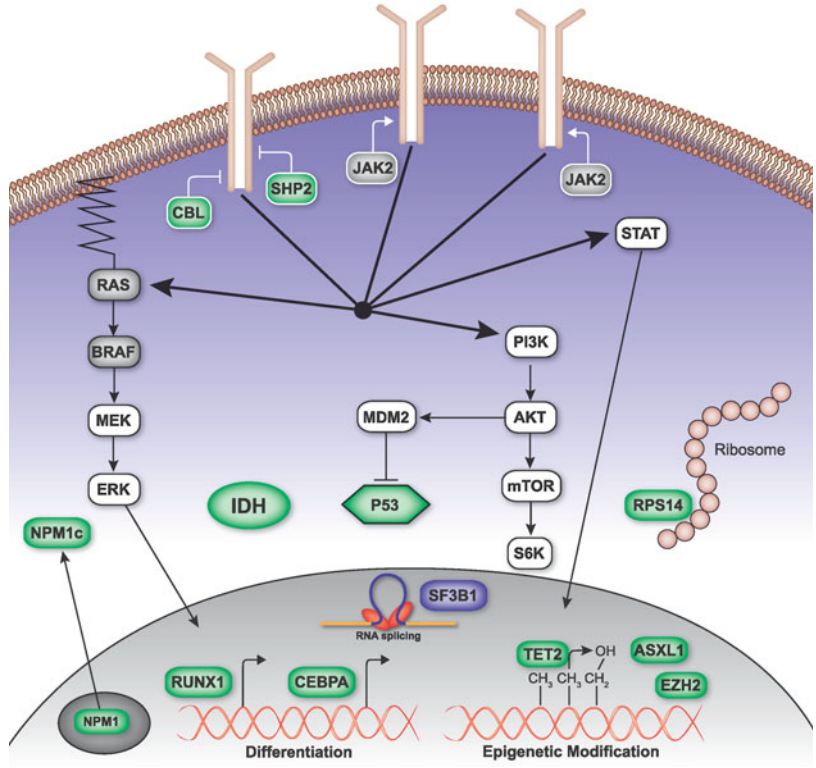


Fig. 34.2 Point mutations and deletions in myelodysplastic syndromes. Mutations in multiple pathways contribute to the pathogenesis of myelodysplastic syndromes. *Gray* icons are labeling activating mutations, whereas *green* icons are highlighting mutations or deletions resulting in a loss of function. *SF3B1* is a newly discovered point mutation involved in RNA splicing

- The proximal CDR at 5q31 associated with therapy-related MDS and a more aggressive MDS and AML phenotype
- Multiple genes on 5q have been implicated in the pathogenesis of MDS *RPS14*, a critical gene for the erythroid phenotype of the 5q– syndrome, identified in a systematic functional screen of the 5q33 CDR
- Haploinsufficiency for two microRNAs located on chromosome 5q33, miR-145 and miR-146, causes elevated platelet counts and selective advantage to the 5q– clone
- Several candidate MDS genes are contained in the proximal CDR at 5q31: early growth response gene (*EGR1*) increases stem cell self-renewal when one copy is deleted
- Alpha catenin, *CTNNA1*, is under expressed in patients with 5q– syndrome, and hypermethylation of the remaining allele is associated with transformation to AML
- *APC* and *NPM1* genes are often lost with deletions of 5q
- Isolated 5q– is associated with a relatively favorable prognosis due to a high level of complete cytogenetic remissions after lenalidomide therapy
- Two cell cycle–regulating phosphatases encoded by the genes on 5q, *CDC25C* and *PP2A*, have been implicated in the favorable response to lenalidomide

34.7.2 Trisomy 8

- Frequency of trisomy 8: 8%
- Intermediate-risk cytogenetic abnormality: 22 months survival compared to 53.4 month in patients with normal karyotype
- Subset of patients with MDS and isolated trisomy 8 respond well to immunosuppressive therapy

Table 34.3 *International prognostic scoring system for myelodysplastic syndrome (MDS)*. A score is assigned to three prognostic parameters as outlined in the table. The sum of the scores determines the international prognostic scoring system (IPSS) risk group in MDS

IPSS for MDS					
	Score value				
Prognostic parameter	0	0.5	1.0	1.5	2.0
Bone marrow blast percentage (%)	<5	5–10		11–20	21–30
Karyotype category[a]	Good	Intermediate	Poor		
Number of cytopenias[b]	0 or 1	2 or 3			
[a]Karyotype categories					
Good: normal, –Y, del(5q), del(20q) Intermediate: other (not good or poor) Poor: chromosome 7 anomalies or complex (≥3 abnormalities) [b]Cytopenia definitions Hemoglobin: <10 g/dL Absolute neutrophil Count: <1,800 per μl					
Platelet count: <100,000 per μl					

IPSS risk category	***Total score***
Low	0
Intermediate-1	0.5–1.0
Intermediate-2	1.5–2.0
High	≥2.5

IPSS risk category	***Total Score***	***Leukemic death***	***Median time to AML (year)***	***Median survival (year)***
Low	0	19%	9.4	5.7
Intermediate-1	0.5–1.0	30%	3.3	3.5
Intermediate-2	1.5–2.0	33%	1.1	1.2
High	≥2.5	45%	0.2	0.4

Adapted from: Greenberg et al. (1997)

34.7.3 Chromosome 7 and 7q Deletions

- Frequency of monosomy 7: 10% either isolated or as part of a complex karyotype and 50% in patients with therapy-related MDS
- Three common deleted regions on 7q have been identified so far, but the underlying molecular lesions are not well defined: *EZH2* is frequently mutated and is located at 7q36, chromosome 7 abnormalities co-occur with 5q– or mutations in *RUNX1* in therapy-related MDS

34.7.4 Chromosome Y and 20q Deletions

- Isolated 20q– or –Y abnormalities are considered a favorable cytogenetic risk group similar to patients with a normal karyotype
- These abnormalities alone are insufficient to make the diagnosis of MDS
- Loss of chromosome Y is often unrelated to hematological disease; interstitial deletion of 20q is a common cytogenetic abnormality in myeloproliferative disorders and AML

34.7.5 Abnormalities Involving Chromosome 3q26

- Recurrent translocations and inversions of 3q26 have been identified in AML and rare cases of MDS and are associated with a poorer prognosis
- Breakpoints typically include the *MDS1-EVI1* gene locus (MECOM)
- *EVI1* impairs the activity of several transcription factors including MDS1-EVI1, GATA1, PU.1, and RUNX1 leading to impaired hematopoiesis
- In mouse models, forced EVI1 expression results in an MDS-like phenotype and increased self-renewal of hematopoietic progenitors

34.7.6 Other Cytogenetic Abnormalities

- Several rare but recurrent cytogenetic abnormalities are diagnostic of MDS in persistently cytopenic patients with minimal dysplasia and are lumped into the intermediate-risk group: abnormalities of chromosome 17 (presumably disrupting *TP53*), isodicentric chromosome Xq13 (associated with the presence of ring sideroblasts), and t(6;9)(p23;q34) that generates the *DEK-NUP214* fusion gene

34.7.7 Gene Mutations

34.7.7.1 *TET2* Mutations

- Frequency of mutated *TET2*: approximately 20% of patients with MDS, 10% of MPN, 30–50% of CMML, and 25% of secondary AML
- *TET2* encodes a dioxygenase that alters the epigenetic mark created by DNA methyltransferases (DNMTs)
- Its homolog, the 10–11 translocation 1 gene (*TET1*) is reported as a fusion partner with the mixed-lineage leukemia (*MLL*) gene in rare cases of AML
- *TET2* mutations are not generally associated with other recurrent mutations or cytogenetic abnormalities
- The prognostic impact of *TET2* mutations in MDS and MPN is unclear which is in contrast to *TET2* mutations in AML and CMML which are associated with a relatively poor prognosis

34.7.7.2 *ASXL1* Mutations

- Frequency of mutations in the additional sex combs-like 1 (*ASXL1*) gene: 10% of MDS and MPN, 17% of AML, and 40% of patients with CMML
- *ASXL1* is part of the polycomb family of chromatin-binding proteins involved in epigenetic regulation of gene expression
- Functions as a ligand-dependent coactivator of the retinoic acid receptor through a direct interaction with the histone acetyltransferase encoded by *NCOA1* or the histone demethylase encoded by *LSD1*
- Prognostic significance of *ASXL1* mutations in MDS is unclear

34.7.7.3 *RUNX1* Mutations

- *RUNX1*, member of the transcriptional core-binding factor gene family (also known as *CBFA2* or *AML1*)
- Frequency of *RUNX1* mutations: 7–15% of de novo patients and higher frequency in therapy-related disease
- *RUNX1* point mutations are found in MDS, AML, CMML, and, more rarely, in MPN
- *RUNX1* mutations are markers of poor prognosis in MDS and AML
- The RUNX1 protein contains a proximal Runt homology domain, important for DNA binding, and a distal transactivation domain responsible for protein–protein interactions and recruitment of cofactors
- Missense mutations of *RUNX1* are clustered in the Runt domain, whereas stop codon mutations and frameshifts are found throughout the length of the gene and almost always disrupt the transactivation domain
- In MDS, *RUNX1* mutations often accompanied by activation of the RAS pathway or mutations in these genes
- A distal frameshift mutation of *RUNX1* sparing the Runt domain results in a gain

of function and MDS-like phenotype with marked erythroid dysplasia, pancytopenia, and only rare development of AML in mice

34.7.7.4 *IDH1* and *IDH2* Mutations

- Mutations of isocitrate dehydrogenase genes 1 and 2 (*IDH1* and *IDH2*) have been confirmed in AML, leukemic transformation of MPN, and in rare cases of MDS
- *IDH1* and *IDH2* encode homodimeric, nicotinamide adenine dinucleotide phosphate (NADP)-dependent enzymes that convert isocitrate to ketoglutarate and are specific missense mutations of conserved codons

34.7.7.5 Tyrosine Kinases, *RAS*, and *CBL* Mutations

- TK signaling pathways are infrequently mutated in MDS and involve activating mutations of the downstream *RAS* genes
- Frequency of mutations: *NRAS* mutations in 10–15% of patients, *KRAS* mutations in 1–2%, *BRAF* in <5%, and *PTPN11* and *CBL* less than 3%, all associated with a poor prognosis and progression to AML
- *CBL* mutations result in increased receptor TK levels and phosphorylation of *STAT5* mediating hypersensitivity of these mutant cells to a wide variety of growth factors and cytokines
- *CBL* mutations cluster in exons 8 and 9 and encode a dominant-negative protein that inhibits the ubiquitin ligase activity of the wild-type gene product and of its homolog, CBLB
- *CBL* mutations appear to be mutually exclusive of several other commonly mutated genes including *FLT3*, *KIT*, *NPM1*, *CEBPA*, *PTPN11*, and *NRAS*
- *JAK2*V617F has been described in 5% of MDS and has no prognostic significance; this mutation is most common in the MDS subtype refractory anemia with ring sideroblasts and thrombocythemia (RARS-T) where *JAK2*V617F mutations are present in 50% of patients
- *FLT3* mutations are very rare in MDS and play a critical role in the progression to AML in a subset of MDS patients, regarded as a "second hit" in the development of MDS

34.7.7.6 *TP53* Activation and Mutation

- In 5q– MDS, activation of p53, the gene product of *TP53*, appears to be essential for the erythropoietic defect associated with haploinsufficiency for *RPS14*
- Frequency of *TP53* mutation: 5–15% of de novo patients and more frequently in patients with prior exposure to alkylating agents or radiation (therapy-related MDS)
- Loss of wild-type p53 associated with advanced disease, complex karyotype, and resistance to treatment
- Mutations of *TP53* associated with a poor prognosis in MDS

34.7.7.7 *EZH2* Mutations

- Frequency of mutations in the enhancer of zeste homolog 2 (*EZH2*) gene: 6% of MDS and over 10% of MDS/MPN cases
- Encodes a histone methyltransferase that serves as the catalytic subunit of the polycomb repressive complex 2 (PRC2)
- Loss of PRC2 function shown to increase hematopoietic stem cell activity and expansion
- Mutations of *EZH2* in MDS are associated with a poor prognosis

34.7.7.8 *SF3B1* Mutations

- Frequency of mutations in the RNA splicing factor 3B, subunit 1 (*SF3B1*): present in 20% of patients with myelodysplastic syndrome
- Particularly common in subtypes with ring sideroblasts detectable in 65% of patients
  - Frequency in refractory anemia with ring sideroblasts 68%
  - Frequency in refractory cytopenia with multilineage dysplasia and ring sideroblasts 57%
- Present in lower frequency in other subtypes: 10% in refractory anemia, 6% in refractory cytopenia with multilineage dysplasia, and 5% in refractory anemia with excess blasts

- *SF3B1* mutations are not specific for myelodysplastic syndrome and occur in low frequency in other myeloid neoplasms: 5% in AML, 4% in PMF, 3% in ET, 5% in CMML, 1–5% in other cancers including breast cancer (1%), renal cancer (3%), chronic lymphocytic leukemia (5%), multiple myeloma (3%), adenoid cystic carcinoma (4%), and are present in a few cancer cell lines (8 out of 746 cancer cell lines screened)
- *SF3B1* mutations are associated with longer overall survival, leukemia-free survival, and event-free survival and mark relatively low-grade myelodysplastic syndromes characterized phenotypically by the presence of ring sideroblasts

34.7.8 Epigenetic Changes in Myelodysplastic Syndromes

- The term "epigenetic" refers to methylation of cytosine residues in DNA and covalent modification of histones
- Methylation of CpG islands in promoters leads to silencing of neighboring genes and represents a mechanism for loss of tumor suppressor gene expression
- Target genes of DNA hypermethylation in MDS and AML: cell cycle regulators *CDKN2A* (p14 and p16) and *CDKN2B* (p15), *CTNNA1*, E-cadherin (*CDH1*), and many others
- Suggests a rational for use of demethylating agents, azacitidine and decitabine, for treatment of MDS with response rates from 30% to 73% and prolonged survival in high-risk patients treated with azacitidine
- Inhibition of histone-modifying enzymes represents another potential epigenetic target for MDS therapy
- Histones are protein multimers associated with DNA that help form chromatin structure and influence the level of transcription of nearby genes
- The acetylation of lysine residues on histones is associated with an open chromatin conformation and increased transcription
- Histone deacetylases (HDACs) can remove these acetyl groups and silence nearby genes

34.7.9 The Bone Marrow Microenvironment

- Important role in the development of MDS and other myeloid malignancies; for example, levels of vascular endothelial growth factor and several inflammatory cytokines are elevated in the bone marrow of patients with MDS
- Autocrine or cell contact-mediated interactions with the stroma can negatively influence normal hematopoiesis and result in cytopenias even when MDS cells occupy only a fraction of the bone marrow
- Primary stromal cell abnormalities cause dysplasia in normal hematopoietic cells and drive clonal evolution
- Selective deletion of the microRNA processing gene, *DICER1*, in osteoprogenitor cells of mice causes a phenotype that closely resembles MDS: mice have cytopenias, multilineage dysplasia, bone marrow hypercellularity, and increased apoptosis
- Acquired chromosomal abnormalities in nonhematopoietic bone marrow cells from patients with MDS lend support for a role of primary stromal changes in MDS

34.8 Molecular Diagnostics and Implications for Clinical Management of Myelodysplastic Syndromes

- Diagnostic requirements for diagnosis of MDS: peripheral blood and bone marrow morphologic examination and metaphase karyotyping
  - MDS-specific fluorescent in situ hybridization panels are commercially available
  - More sensitive techniques are available in the research setting: spectral karyotyping and single nucleotide polymorphism array genotyping

- MDS is clinically heterogeneous; a combination of two or more mutations in cooperation with more global changes in epigenetic states and cellular environment is usually present in the disease
- Evaluation of individual genes for mutations may become useful for diagnosis, prediction of prognosis, and therapeutic management in patients with MDS

Further Reading

Bejar R, Levine R, Ebert BL. Unraveling the molecular pathophysiology of myelodysplastic syndromes. J Clin Oncol. 2011a;29:504–15.

Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med. 2011b;364:2496–506.

Finazzi G, et al. Evidence and expertise in the management of polycythemia vera and essential thrombocythemia. Leukemia. 2008;22:1494–502.

Gangat N, et al. Dynamic international prognostic scoring system (DIPSS Plus). J Clin Oncol. 2011;29:392–7.

Greenberg P, Cox C, LeBeau MM, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. Blood. 1997;89:2079–88.

James C, Ugo V, Le Couedic JP, et al. A unique clonal JAK2 mutation leading to constitutive signaling causes polycythaemia vera. Nature. 2005; 434:1144–8.

Levine RL, Wernig G. Role of JAK-STAT signaling in the pathogenesis of myeloproliferative disorders. Hematology Am Soc Hematol Educ Program. 2006; 2006:233–9. 510.

Papaemmanuil E, Cazzola M, Boultwood J, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med. 2011;365:1384–95.

Tefferi A, Vainchenker W. Myeloproliferative neoplasms: molecular pathophysiology, essential clinical understanding, and treatment strategies. J Clin Oncol. 2011;29:573–82.

Tefferi A, Vardiman JW. Myelodysplastic syndromes. N Engl J Med. 2009;361:1872–85.

Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009;114:937–51.

Lymphoproliferative Disorders: Molecular Diagnostics

35

Timothy C. Greiner

Contents

35.1 Overview 946

35.2 Posttransplant Lymphoproliferative Disorders 946
35.2.1 Classification of Posttransplant Lymphoproliferative Disorders 946
35.2.2 Development of Posttransplant Lymphoproliferative Disorder 946
35.2.3 Risk Factors for Poor Survival 946
35.2.4 Late Posttransplant Lymphoproliferative Disorder 947
35.2.5 Epstein-Barr Virus Substrains 947
35.2.6 Epstein-Barr Virus Detection in Tissue Slides by Epstein-Barr Virus-Encoded RNA In Situ Hybridization 947
35.2.7 Epstein-Barr Virus Detection in Blood 947
35.2.8 Kappa and Lambda In Situ Probes 948
35.2.9 *IGH* Gene Rearrangement 948
35.2.10 T Cell Receptor Gamma and Beta Rearrangement 948
35.2.11 Cytogenetic/Molecular Findings in Posttransplant Lymphoproliferative Disorder 948

35.3 HIV Lymphoproliferations 948
35.3.1 Common Lymphomas 948
35.3.2 Uncommon Tumors/Proliferations 948
35.3.3 In Situ Hybridization 949
35.3.4 Cytogenetics 949

35.4 Other Immunodeficiencies 949
35.4.1 Autoimmune Lymphoproliferative Syndrome 949
35.4.2 X-Linked Lymphoproliferative Disease 949
35.4.3 Molecular Diagnosis of X-Linked Lymphoproliferative Disease 949
35.4.4 Common Variable Immunodeficiency 950

Further Reading 950

T.C. Greiner, MD
Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_35,

35.1 Overview

- There are several challenges of diagnosing lymphoproliferative disorders in immunodeficiency syndromes
- Epstein–Barr virus (EBV) causes proliferation and transforms cells
  - It is thought to allow greater chances for random molecular events to occur above the baseline expected for the patient's age
  - T-cell immunosuppression contributes to the allowance of proliferation
- Patients are at variable risk throughout their lives for lymphoproliferations due to the immune deficiencies, whether hereditary or subsequent to an acquired immune deficiency such as transplantation, human immunodeficiency virus (HIV) infection, or in other diseases where induced immune suppression occurs secondary to steroid or methotrexate treatment
- Determining whether the proliferation is a hyperplasia or a lymphoma is difficult because both events, or rather a pectrum of events, can occur in such patients
  - While an EBV-positive atypical lymphoid hyperplasia can occur at one site, a clonal neoplasm may occur at another site in the same patient
- The lymphoproliferations are often polymorphous, at least initially and then develop a monomorphous histology similar to a typical lymphoma
  - Histology is not easily predictive of when a committed clonal translocation develops
    - Multiple modalities are needed to aid in the establishment of a lymphoma
  - It is important to separate it from the differential diagnosis of infectious mononucleosis or an EBV-positive lymphoproliferation that mimics infectious mononucleosis

35.2 Posttransplant Lymphoproliferative Disorders

- Posttransplant lymphoproliferative disorders (PTLD) occur in 1–10% of transplant patients depending upon the solid organ which was transplanted
- A high degree of immune suppression in heart and lung transplant patients and EBV naivety in pediatric patients contribute to the higher incidence in these settings
- PTLD can be lethal regardless of whether they are polymorphous or monomorphous in histology

35.2.1 Classification of Posttransplant Lymphoproliferative Disorders

- Early lesions (Swerdlow et al. 2008)
  - Plasmacytic hyperplasia
  - Infectious mononucleosis-like lesions
- Polymorphic PTLD
- Monomorphic PTLD (describe histology similar to usual morphology)
  - Diffuse large B cell
  - Burkitt
  - Classical Hodgkin
  - Peripheral T cell

35.2.2 Development of Posttransplant Lymphoproliferative Disorder

- PTLD develops as early as 6 weeks
  - EBV positive in >90% of cases in the first 2 years
  - B cell lineage predominates in 90% of cases

35.2.3 Risk Factors for Poor Survival

- Monomorphic histology
- High stage or septic clinical picture

- Clonality established by one or more methods
  - Light chain restriction
  - Immunoglobulin or T-cell receptor gene rearrangement
  - Abnormal cytogenetic karyotype
  - EBV integration by Southern blot
  - Fluorescent in situ hybridization (FISH) showing an *IGH* translocation

35.2.4 Late Posttransplant Lymphoproliferative Disorder

- EBV positivity seen in <50% of cases
- Frequency of a T cell lineage increases after 10 years, especially in renal transplantation

35.2.5 Epstein-Barr Virus Substrains

- After the description of a 30-base pair deletion in the latent membrane protein 1 gene (*LMP1*), which was shown to transform cells, it was hypothesized that this substrain of EBV was more tumorigenic or might be important in the prediction of prognosis
  - It was subsequently determined that the *LMP1* deletion tends to reflect the geographical incidence of EBV substrains and was found not to be important in prognosis
- Most EBV substrains in PTLD are type A
  - Type B strains have also been observed and have no prognostic significance
- Variations on single nucleotide polymorphisms have been described in *BZLF1* in EBV
  - Similar to *LMP1*, no prognostic significance has been described to detecting the *BZLF1* substrains of EBV in PTLD or in other immunodeficiencies

35.2.6 Epstein-Barr Virus Detection in Tissue Slides by Epstein-Barr Virus-Encoded RNA In Situ Hybridization

- EBV early RNA (EBER)-specific probes
- Decalcification in hydrochloric acid may affect RNA integrity
- An RNA Integrity Slide is required when using probes
- When tissue is poorly preserved by fixation or necrosis, nucleoli may still be positive
- Negative control is needed to rule out nonspecific staining
- If RNA is degraded, a backup method of LMP1 immunohistochemical stain may be used
- However, LMP1 only has 75% sensitivity compared to EBERs
- PCR is not recommended, as latent infected cells in blood can cause a positive PCR result in the tissue
- Background level in *immunocompetent* patients' lymphoid tissue is 0–1 positive EBERs cells/hpf
- Background level in *immunosuppressed* patients' tissue is 1–5 positive EBERs cells/hpf to focally 10/hpf
- Usually, sheets of cells (>15 positive EBERs cells/hpf) are seen in PTLD or in lymphoproliferations of immunodeficiency patients

35.2.7 Epstein-Barr Virus Detection in Blood

- Quantitative PCR (qPCR)
  - Monitoring quantitative levels in blood is generally useful to predict the development of PTLD
- Plasma is better than whole blood as the fluctuating number of white blood cells affect levels
- Log changes are more important than small incremental changes
- Caveat: not all patients have increased levels when PTLD occurs

- Myth: negative EBV blood level means the patient cannot have PTLD; this scenario occurs in patients with
  - Single lesions
  - CNS disease
  - Hodgkin type
- Quantitative standards are available for PCR

35.2.8 Kappa and Lambda In Situ Probes

- Cytoplasmic location
- Cleaner for interpretation than immunohistochemical stains
- Normal ratio: kappa to lambda is 2–3:1 but may be as high as 5:1
- Clonal kappa suggested by ratio of 10:1
- Clonal lambda suggested by a reversed 1:3 ratio
- Different lesions may have a different light chain restriction in the same patient

35.2.9 *IGH* Gene Rearrangement

- Clonality tends to correlate with monomorphous histology
  - Polyclonal pattern may be seen in monomorphous lesions
  - Monoclonal pattern may be seen in polymorphous lesions
- Different *IGH* gene rearrangements may occur at different sites in one patient
- Clonal cases less likely to respond to reduction in immunosuppression and have an increased need for chemotherapy
- See Chap. 30

35.2.10 T Cell Receptor Gamma and Beta Rearrangement

- Important in the diagnosis of T-cell PTLD
- T-cell lineage increases in frequency 10 years after transplant
- Various subtypes of T-cell lymphomas have been observed

35.2.11 Cytogenetic/Molecular Findings in Posttransplant Lymphoproliferative Disorder

- There is no one singular recurrent translocation in all lesions
- *BCL6* and *CMYC* translocations have been described in diffuse large B-cell and Burkitt lymphoma, respectively
- Karyotype analysis shows scattered deletions, gains, and complex abnormalities in individual cases
- RNA expression pattern by array
  - Demonstrates ABC subtype in monomorphic lesions
- *TP53* mutations are rare
- FISH with an *IGH* break apart probe is useful to establish clonality in B cell lesions
- FISH for a specific gene translocation is not recommended due to the low yield
- *BCL6* mutations have been described as predictive of prognosis, but this finding has never been confirmed
- Comparative genomic hybridization (aCGH) and single nucleotide polymorphous arrays demonstrate (none are seen in more than 15% of cases)
  - Gains: 2p24p25, 3q27, 5p, 8q24, 9q22q34, 11,11p, 12q22q24, 14q32, 17q, 18q21
  - Losses: Xp, 1p36, 2p16, 4q, 6q, 17p, 17p13, 17q23q25, 1p, 9p

35.3 HIV Lymphoproliferations

35.3.1 Common Lymphomas

- Diffuse large B-cell lymphoma (DLBCL)
- Burkitt lymphoma (BL)
- Hodgkin lymphoma (HL) (not an AIDS-defining lesion)

35.3.2 Uncommon Tumors/ Proliferations

- Plasmablastic lymphoma (PBL)
- Primary effusion lymphoma (PEL)

- Kaposi sarcoma (KS)
  - Often coexistent with lymphoma
- Multicentric Castleman disease (MCD)

35.3.3 In Situ Hybridization

- EBV – 50% of DLBCL, BL, HL, PBL; 100% of CNS lymphomas
- HHV8 – 100% of KS, PEL
  - Positive in microlymphomas associated with MCD

35.3.4 Cytogenetics

- Translocations with *BCL6*, *CMYC*, and *BCL2* have been described

35.4 Other Immunodeficiencies

- The intent of this section is not to provide a description of the genetic basis of each of the immunodeficiency disorders – the goal is to identify the workup that is necessary to evaluate a lymphoproliferation that may occur in a known immunodeficiency disorder or when evaluating a patient when a disorder is not suspected (e.g., autoimmune lymphoproliferative syndrome) and where the proliferating cell type may suggest a malignancy
- Lymphoproliferative disorders occur at an increased risk in immunodeficiency especially in patients with
  - X-linked lymphoproliferative syndrome
  - Ataxia-telangiectasia syndrome
  - Wiskott–Aldrich syndrome
  - Common variable immunodeficiency
- The most common subtype is diffuse large B cell lymphoma; uncommon subtypes include
  - Marginal zone lymphoma
  - Peripheral T cell lymphoma
  - T-prolymphocytic leukemia
  - Hodgkin lymphoma

35.4.1 Autoimmune Lymphoproliferative Syndrome

- Young patients have an expansion of double-negative T cells (CD8– and CD4–) in the peripheral blood and lymph nodes
- Reduced apoptosis is identified in these proliferating T cells
- Germline mutations have been described in several genes, and identifying them helps avoid an incorrect diagnosis
  - ALPS-FAS (type 0), germline homozygous mutations in CD95 (*FAS* gene)
  - ALPS-FAS (type 1a), heterozygous mutations in the same gene
  - ALPS-FASLG (type 1b), germline mutations in FAS ligand gene
  - ALPS-sFAS (type 1m), heterozygous *FAS* gene mutations seen only in the somatic T cell
  - ALPS-CASP10 (type 2a), mutations in caspase 10 gene
  - ALPS-U (type 3), no germline mutation
- T cell receptor gene rearrangements are not identified in the CD4, CD8 double-negative cells, thereby avoiding an unnecessary diagnosis of malignancy
- Mutations in *CASP8*, *NRAS*, and *SH2DIA* are now placed in ALPs-related disorders

35.4.2 X-Linked Lymphoproliferative Disease

- Infection in young males with XLP may lead to chronic or fatal infectious mononucleosis (50%) or to non-Hodgkin lymphoma (25%)
- Other manifestations of the disease may include
  - Lymphomatoid granulomatosis in the lung
  - Hemophagocytic syndrome
  - Cytopenias
  - Vasculitis

35.4.3 Molecular Diagnosis of X-Linked Lymphoproliferative Disease

- The *SH2D1A* gene was identified in 1998

- The SH2D1A protein is important in the regulation of the immune responses by interacting with the SLAM molecule (CD150)
- Mutations have also been described in X-linked inhibitor of apoptosis (*XIAP*), aka *BIRC4*
- Mutations include missense mutation, nonsense mutations, and truncating deletions
- The severity of the phenotype does not correlate directly with the type of mutation
- EBV positivity is seen is approximately 25% of tumor specimens

35.4.4 Common Variable Immunodeficiency

- Patients often develop reactive lymphoid hyperplasia, and this alternates with a lymphoma
- While individuals usually present in their 20s, other individuals may present later in life without a typical history of recurrent infections
- Extranodal sites are more often involved by lymphoma than in immunocompetent patients
- In situ hybridization for EBER is positive about 25% of cases of lymphoid hyperplasia and about 50% of cases of lymphoma
- The diagnosis should be suspected in a patient with lymphadenopathy when hypogammaglobulinemia is identified for all antibodies
- Germline mutations have been identified in *CD19*, *ICOS*, *TNFRSF13B* (*TACI*), and *TNFRSF13C* (*BAFFR*)

Further Reading

Cesarman E, Chang Y, Moore PS, et al. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in AIDS-related body-cavity-based lymphomas. N Engl J Med. 1995;332:1186–91.

Cesarman E, Chadburn A, Liu YF, et al. BCL-6 gene mutations in posttransplantation lymphoproliferative disorders predict response to therapy and clinical outcome. Blood. 1998;92:2294–302.

Chadburn A, Chen JM, Hsu DT, et al. The morphologic and molecular genetic categories of posttransplantation lymphoproliferative disorders are clinically relevant. Cancer. 1998;82:1978–87.

Djokic M, Le Beau MM, Swinnen LJ, et al. Post-transplant lymphoproliferative disorder subtypes correlate with different recurring chromosomal abnormalities. Genes Chromosomes Cancer. 2006;45:313–8.

Finn L, Reyes J, Bueno J, et al. Epstein-Barr virus infections in children after transplantation of the small intestine. Am J Surg Pathol. 1998;22:299–309.

Gaidano G, Carbone A, Dalla-Favera R. Genetic basis of acquired immunodeficiency syndrome-related lymphomagenesis. J Natl Cancer Inst Monogr. 1998;23:95–100.

Gartner BC, Schafer H, Marggraff K, et al. Evaluation of use of Epstein-Barr viral load in patients after allogeneic stem cell transplantation to diagnose and monitor posttransplant lymphoproliferative disease. J Clin Microbiol. 2002;40:351–8.

Greiner TC, Abou-Elella AA, Smir BN, et al. Molecular epidemiology of EBNA-1 substrains of Epstein-Barr virus in posttransplant lymphoproliferative disorders which have infrequent p53 mutations. Leuk Lymphoma. 2000;38:563–76.

Gutierrez MI, Ibrahim MM, Dale JK, et al. Discrete alterations in the BZLF1 promoter in tumor and non-tumor-associated Epstein-Barr virus. J Natl Cancer Inst. 2002;94:1757–63.

Ishihara M, Tanaka E, Sato T, et al. Epstein-Barr virus load for early detection of lymphoproliferative disorder in pediatric renal transplant recipients. Clin Nephrol. 2011;76:40–8.

Jaffe ES, Harris NL, Stein H, et al., editors. Pathology and genetics of tumours of haematopoietic and lymphoid tissues. Lyon: IARC Press; 2001.

Knowles DM, Cesarman E, Chadburn A, et al. Correlative morphologic and molecular genetic analysis demonstrates three distinct categories of posttransplantation lymphoproliferative disorders. Blood. 1995;85:552–65.

Niesters HG, van Esser J, Fries E, et al. Development of a real-time quantitative assay for detection of Epstein-Barr virus. J Clin Microbiol. 2000;38:712–5.

Oliveira JB, Bleesing JJ, Dianzani U, et al. Revised diagnostic criteria and classification for the autoimmune lymphoproliferative syndrome (ALPS): report from the 2009 NIH International Workshop. Blood. 2010;116:e35–40.

Poirel HA, Bernheim A, Schneider A, et al. Characteristic pattern of chromosomal imbalances in posttransplantation lymphoproliferative disorders: correlation with histopathological subcategories and EBV status. Transplantation. 2005;80:176–84.

Raab-Traub N, Flynn K. The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation. Cell. 1986;47:883–9.

Rinaldi A, Kwee I, Poretti G, et al. Comparative genome-wide profiling of post-transplant lymphoproliferative disorders and diffuse large B-cell lymphomas. Br J Haematol. 2006;134:27–36.

Rinaldi A, Capello D, Scandurra M, et al. Single nucleotide polymorphism-arrays provide new insights in the pathogenesis of post-transplant diffuse large B-cell lymphoma. Br J Haematol. 2010;149:569–77.

Sayos J, Wu C, Morra M, et al. The X-linked lymphoproliferative-disease gene product SAP regulates signals induced through the co-receptor SLAM. Nature. 1998;395:462–9.

Smir BN, Hauke RJ, Bierman PJ, et al. Molecular epidemiology of deletions and mutations of the latent membrane protein 1 oncogene of the Epstein-Barr virus in posttransplant lymphoproliferative disorders. Lab Invest. 1996;75:575–88.

Swerdlow SH, Campo E, Harris NL, et al., editors. WHO classification of tumours of haematopoietic and lymphoid yissues. Lyon, France: IARC Press; 2008.

Vakiani E, Nandula SV, Subramaniyam S, et al. Cytogenetic analysis of B-cell posttransplant lymphoproliferations validates the World Health Organization classification and suggests inclusion of florid follicular hyperplasia as a precursor lesion. Hum Pathol. 2007;38:315–25.

Section 6

Hematological Disorders and Specialized Applications

Molecular Testing for Coagulation Abnormalities

36

Veshana Ramiah and Thomas L. Ortel

Contents

36.1 Normal Hemostasis 956
36.1.1 Primary Hemostasis 956
36.1.2 Secondary Hemostasis 957
36.1.3 Natural Anticoagulant Proteins 958
36.1.4 Fibrinolysis 958

36.2 Factor V Leiden 958
36.2.1 General 958
36.2.2 Clinical Manifestations 958
36.2.3 Acquired Risk Factors for Venous Thrombosis 958
36.2.4 Prevalence 959
36.2.5 Differential Diagnosis 959
36.2.6 Genetics and Biochemistry 959
36.2.7 Relative Risk 959
36.2.8 Functional Testing 959
36.2.9 Molecular Testing 960
36.2.10 Indications for Factor V Leiden DNA Test 961
36.2.11 Benefits and Limitations of APC-Resistance Testing and DNA Testing 961
36.2.12 Testing Not Indicated in the Following Situations 962
36.2.13 Management of Homozygotes with Thrombosis 962
36.2.14 Management of Heterozygotes with Thrombosis 962
36.2.15 Management of Asymptomatic Carriers 962

36.3 Prothrombin G20210A Mutation 962
36.3.1 General 962
36.3.2 Clinical Manifestations 962
36.3.3 Acquired Risk Factors for Venous Thrombosis 962
36.3.4 Prevalence 963
36.3.5 Genetics and Biochemistry 963
36.3.6 Relative Risk 963
36.3.7 Functional Testing 963
36.3.8 Molecular Testing 963

36.4 Testing Not Indicated in the Following Situations 963
36.4.1 Management of Homozygotes with Thrombosis 963
36.4.2 Management of Heterozygotes with Thrombosis 964
36.4.3 Management of Asymptomatic Carriers 964

36.5 Methylenetetrahydrofolate Reductase (MTHFR) C677T Thermolabile Polymorphism 964
36.5.1 General 964
36.5.2 Clinical Manifestations 964
36.5.3 Prevalence 964
36.5.4 Genetics 964
36.5.5 Relative Risk 964
36.5.6 Diagnostic Assays for Homocysteine 965
36.5.7 Who Should Be Tested for Hyperhomocysteinemia? 965
36.5.8 Molecular Testing 965
36.5.9 Management 965

36.6 Plasminogen Activator Inhibitor-1 (PAI-1) 4 G/5 G Polymorphism 965
36.6.1 Clinical Manifestations 965
36.6.2 Genetics 965
36.6.3 Functional Testing 966
36.6.4 Molecular Testing 966
36.6.5 Management 966

V. Ramiah, MD (✉)
Department of Hematology and Oncology, Danville Regional Medical Center, Danville, VA, USA

T.L. Ortel, MD, PhD
Departments of Medicine/Hematology and Pathology, Duke University Medical Center, Durham, NC, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_36,

36.7 Platelet Surface Glycoprotein IIIA (Human Platelet Antigen 1A and 2A) 966
36.7.1 General 966
36.7.2 Clinical Manifestations 966
36.7.3 Prevalence 967
36.7.4 Genetics 967
36.7.5 Antigenic Testing 967
36.7.6 Molecular Testing 967
36.7.7 Management 967

36.8 Hemophilia Mutations 967
36.8.1 General 967
36.8.2 Clinical Manifestations 967
36.8.3 Prevalence 968
36.8.4 Differential Diagnosis 968
36.8.5 Genetics 968
36.8.6 Functional Testing 968
36.8.7 Molecular Testing 968
36.8.8 Indications for Testing 968
36.8.9 Management 969
36.8.10 Other Coagulation Factor Mutations 969

36.9 Anticoagulant Therapy 969
36.9.1 General 969
36.9.2 Genetics 969
36.9.3 Clinical Significance 970
36.9.4 Molecular Testing 970
36.9.5 Indications for Testing 970

36.10 Antiplatelet Therapy 971
36.10.1 General 971
36.10.2 Genetics 971
36.10.3 Clinical Significance 972
36.10.4 Functional Testing 973
36.10.5 Molecular Testing 973
36.10.6 Indications for Testing 973

Further Reading 973

36.1 Normal Hemostasis

- Sudden and severe loss of blood can lead to shock and death. When blood vessels are damaged, hemostasis (clot formation) will arrest bleeding. This process is divided into primary and secondary hemostasis

36.1.1 Primary Hemostasis

- Vascular phase
  - Cutting or damaging blood vessels leads to vascular spasm of the smooth muscle in the vessel wall
  - This produces a vasoconstriction which will slow or even stop blood flow. This response will last up to 30 min and is localized to the damaged area
- Platelet phase
  - Damaged endothelial cells lining the blood vessel release von Willebrand factor (VWF). This substance makes the surfaces of the endothelial cells "sticky"
  - This condition may, by itself, be enough to close small blood vessels
  - In larger blood vessels, platelets begin to stick to the surfaces of endothelial cells. This effect is called platelet adhesion
  - Platelet adhesion is mediated by subendothelial VWF binding to platelet surface receptor glycoprotein Ib and subendothelial collagen binding to platelet collagen receptors
  - The platelets that adhere to the vessel walls now begin to secrete adenosine diphosphate, which is released from "stuck" platelets. This material causes the aggregation of nearby free platelets, which attach to the fixed platelets and each other
  - Platelet aggregation is mediated by fibrinogen and VWF binding to a second platelet receptor, glycoprotein IIb/IIIa
  - This aggregation of platelets leads to the formation of a platelet plug

36.1.2 Secondary Hemostasis

- Begins 30 s to several minutes after primary hemostasis starts
- The overall process involves the formation of the insoluble protein fibrin from the plasma protein fibrinogen through the action of the enzyme thrombin
- Fibrin forms a network of fibers, which traps blood cells and platelets forming a thrombus or clot
- This process depends on the presence in the blood of 11 different clotting factors and calcium
- Ultimately, these factors will generate the production of factor Xa
- Depending on the initial trigger for the clotting reactions, there are two pathways leading to the formation of the thrombus: the extrinsic pathway and the intrinsic pathway
  - Extrinsic pathway
    - Is initiated with material outside of or "extrinsic" to the blood
    - Damaged tissue releases tissue factor
    - Tissue factor activates factor VII (calcium-dependent step)
    - Factor VIIa activates factor X (calcium-dependent step)
    - Factor VIIa can also activate factor IX, which is also a calcium-dependent step (see Fig. 36.1)
  - Intrinsic pathway
    - Is initiated by the blood coming in contact with exposed collagen in the blood vessel wall
    - Factor XII is activated by making contact with exposed collagen underlying the endothelium in the blood vessel wall
    - Factor XIIa activates factor XI
    - Factors XIIa and XIa jointly activate factor IX
    - Factor IXa converts factor X to factor Xa
    - Factor Xa generates factor IIa (thrombin) from factor II (prothrombin)
    - It should be noted that both pathways lead to the same reaction, namely, the activation of factor X

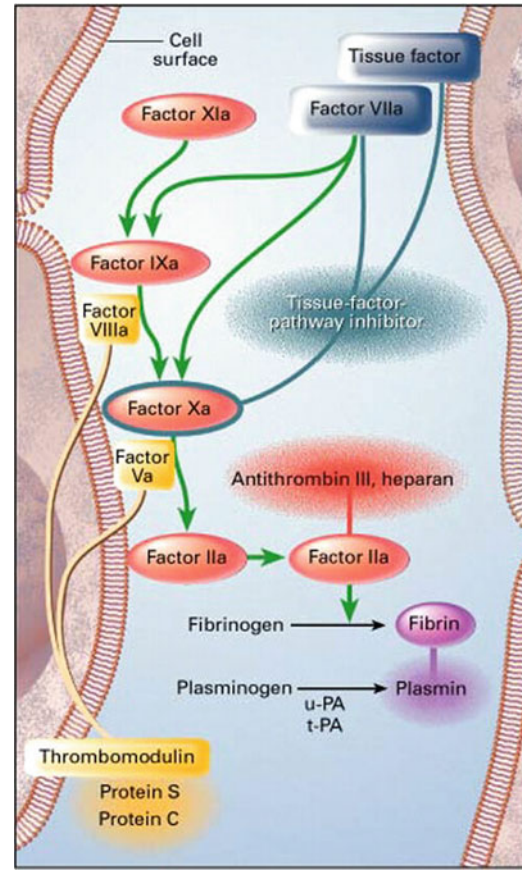


Fig. 36.1 Coagulation pathway. The extrinsic pathway begins with the binding of factor VIIa to tissue factor, which can then activate factor IX to factor IXa and factor X to factor Xa. The intrinsic pathway begins with factor XIIa (not shown) activating factor XI to factor XIa, which activates factor IX to factor IXa. The common pathway begins at the level of factor Xa, which, in the presence of factor Va, membrane surfaces, and calcium ions, converts factor II (prothrombin) to factor IIa (thrombin). The natural anticoagulants include antithrombin, which inactivates factor IIa, as well as factor Xa and factor Xia (not shown), and protein C, which inactivates factor Va and factor VIIIa. The fibrinolytic pathway begins with the conversion of plasminogen to plasmin by either the urokinase-plasminogen activator (u-PA) or the tissue-type plasminogen activator (t-PA) (Rosenberg RD, et al. N Engl J Med 1999;340:1555–1564)

- From this point on, both pathways follow the same course to fibrin formation
- For this reason the steps from factor X activation to fibrin formation are referred to as the common pathway

- Common pathway
  - Factor Xa engages in a series of reactions with factor Va, calcium ions, and phospholipids derived from platelets. This composite of clotting factors and their reactions is referred to as the prothrombinase complex
  - The prothrombinase complex initiates the conversion of prothrombin to active form of the enzyme thrombin
  - Thrombin accelerates the formation of fibrin strands from fibrinogen (Fig. 36.1)

36.1.3 Natural Anticoagulant Proteins

- These proteins counterbalance the procoagulant protein cascade and prevent excessive, unregulated fibrin production
- All these proteins are synthesized in the liver
- Antithrombin inhibits factors XIIa, XIa, IXa, Xa, and IIa (thrombin)
- Protein C-activated form inactivates cofactors VIIIa and Va
- Protein S required as a protein C cofactor
- Tissue factor pathway inhibitor inhibits TF/VIIa/Xa

36.1.4 Fibrinolysis

- Tissue-type plasminogen activator and urokinase-type plasminogen activator convert plasminogen to plasmin
- Once generated, plasmin proteolytically degrades fibrin. Patients with hemorrhagic and/or thromboembolic disorders may have either inherited and/or acquired defects in normal hemostasis, natural anticoagulant pathways, or fibrinolysis. An increasing number of inherited risk factors, particularly for thrombosis, can be tested with molecular diagnostic strategies

36.2 Factor V Leiden

36.2.1 General

- Factor V is a cofactor in the activation of prothrombin to thrombin by factor Xa
- Factor V is activated to factor Va by thrombin and is inactivated by activated protein C
- Factor V Leiden is a common hereditary hypercoagulable syndrome resulting from a single point mutation (R506Q) in the factor V gene, which results in arginine (R) being replaced by glutamine (Q) at residue 506. This mutant glutamine renders the factor V protein resistant to cleavage by activated protein C (referred to as "APC [activated protein C] resistance"), resulting in a longer half-life of this cleavage-resistant factor V, leading to a hypercoagulable state

36.2.2 Clinical Manifestations

- Deep vein thrombosis (DVT)
- Pulmonary embolism
- Clinical characteristics associated with an inherited hypercoagulable syndrome
  - Recurrent thrombotic episodes
  - Thrombosis at a young age (<50 years)
  - Thrombosis at unusual anatomic sites (cerebral, mesenteric, portal, or hepatic veins)
  - Pregnancy-related venous thrombosis (or in association with oral contraceptives or hormone replacement therapy)
  - Family history

36.2.3 Acquired Risk Factors for Venous Thrombosis

- The following acquired conditions can work synergistically to increase the risk of

thrombosis in a patient with the factor V Leiden mutation
  - Pregnancy
  - Long periods of immobility
  - Postsurgical state
  - Use of oral contraceptives
  - Use of hormone replacement therapy
  - Trauma
  - Cancer
  - Smoking
  - Obesity

36.2.4 Prevalence

- Factor V Leiden is the most common cause of inherited thrombophilia in Caucasians
- Heterozygosity for factor V Leiden is found in 3–8% of the US population, and the incidence varies by ethnicity. Homozygosity for factor V Leiden is found in 1 in 5,000 Caucasians (Table 36.1)
- Factor V Leiden is found in 15–20% of patients experiencing their first episode of DVT
- Factor V Leiden is found in 50–60% of patients with recurrent or estrogen-related thrombosis

36.2.5 Differential Diagnosis

- Of patients with APC resistance, 5% do not have the factor V Leiden mutation, a condition referred to as "acquired" APC resistance
- Acquired APC resistance can be seen in certain situations, including pregnancy or in the presence of a lupus anticoagulant
- Acquired APC resistance also appears to be associated with an increased thrombotic risk, even in the absence of factor V Leiden
- Venous thromboembolism can also be caused by other inherited thrombophilic disorders, such as the prothrombin G20210A mutation, or inherited deficiencies in protein C, protein S, and antithrombin

36.2.6 Genetics and Biochemistry

- The gene for factor V is located on chromosome 1
- Factor V Leiden is caused by a well-conserved point mutation in the gene for coagulation factor V
- The G to A transition at nucleotide 1691 replaces an arginine (R) with glutamine (Q) at residue 506 (R506Q) (Fig. 36.2)
- Normal factor V protein is cleaved by APC at arginine 506
- Mutant glutamine (Q) 506 is cleaved much less efficiently by APC
- Cleavage-resistant factor V has longer half-life, resulting in a hypercoagulable state

36.2.7 Relative Risk

- Factor V Leiden heterozygotes have a four-to tenfold increased risk of venous thrombosis
- Although the factor V Leiden mutation predisposes the carrier to increased thrombosis, most factor V Leiden heterozygotes remain asymptomatic
- Factor V Leiden homozygotes have an 80-fold increased risk of spontaneous venous thrombosis

36.2.8 Functional Testing

- Assays testing for APC resistance can be used as screening tests prior to DNA testing
  - Ratio of activated partial thromboplastin times (aPTT) measured in the presence and absence of exogenous APC (aPTT + APC/aPTT-APC)
  - Normal APC-resistance ratio may be approximately 2 or greater, depending on assay configuration
  - Failure of added APC to prolong the aPTT indicates APC resistance. Factor V Leiden

Table 36.1 Ethnic and racial distribution of common inherited thrombophilic disorders analyzed by molecular diagnostic testing

	Ethnic or racial group					
Thrombophilic disorder	Caucasian (%)	Hispanic American (%)	African American (%)	Native American (%)	Asian American (%)	African or Asian (%)
Factor V Leiden	4.8	2.21	1.23	1.25	0.45	0.05
Prothrombin G 20210A	2.7	–	–	–	–	0.06
MTHFR C677T	56	52	10	32	40	–
MTHFR A1298C	42	38	–	–	–	–
PAI-1 4 G/5G[a]	49	–	24	–	–	–
Platelet GP-IIIb PI^{A1}/PI^{A2b}	84–89	–	92	–	99	–

Frequency of the heterozygous state for each specific disorder in normal individuals from each of the individual ethnic or racial groups is provided
[a]Frequency provided is for the 4 G haplotype
[b]Frequency provided for the PlA1 haplotype

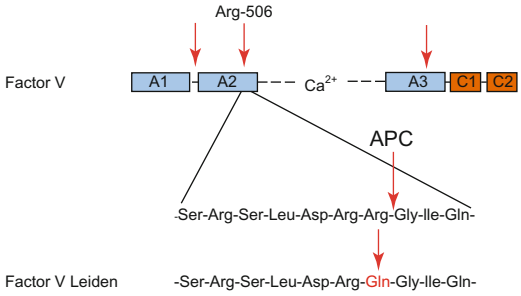


Fig. 36.2 Factor V Leiden. Conversion of the arginine at position 506 in the mature protein to a glutamine renders the activated cofactor resistant to cleavage by activated protein C at this position. Other activated protein C cleavage sites are unaffected

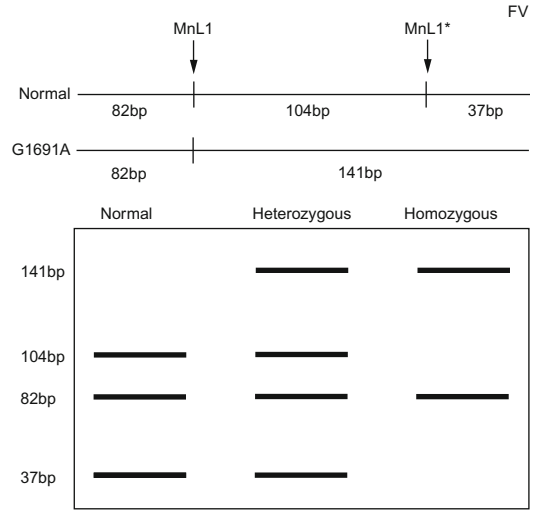


Fig. 36.3 RFLP analysis for G1691A mutation in factor V

homozygotes may have an APC ratio as low as 1.0–1.5

- Of patients with APC resistance, 90–95% will have the factor V Leiden mutation (heterozygous or homozygous)
- Factors which may cause APC resistance in the absence of the factor V Leiden mutation include heparin, lupus anticoagulants, or elevated factor VIII levels

36.2.9 Molecular Testing

- Direct DNA testing for factor V Leiden mutation is now the gold standard
- May be performed by the polymerase chain reaction with restriction fragment length polymorphism (PCR-RFLP) (Fig. 36.3) or

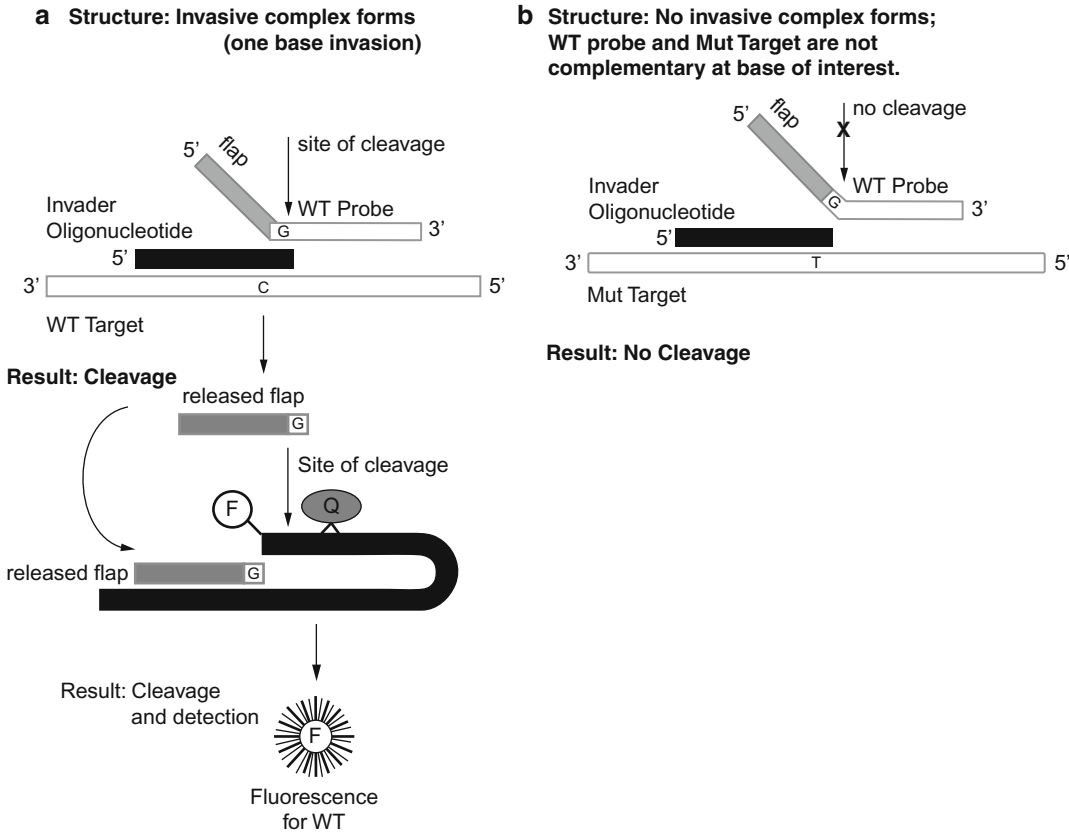


Fig. 36.4 Schematic outline of the invader assay for FVL

with fluorescence resonance energy transfer (PCR-FRET) (Fig. 36.4) or by other methods. Assays used include the Invader assay shown in Fig. 36.4 as well as other including the LightCycler® assays (Roche, Indianapolis, IN)

- Pseudohomozygosity has been described in patients with coinheritance of factor V Leiden for one allele and a mutation resulting in the loss of expression of factor V on the second allele

36.2.10 Indications for Factor V Leiden DNA Test

- Confirm molecular diagnosis in patients with APC resistance
- Evaluate patients with personal history of thrombophilia
- Evaluate asymptomatic family members of patients with the factor V Leiden mutation, if clinically indicated

36.2.11 Benefits and Limitations of APC-Resistance Testing and DNA Testing

- APC-resistance test is less costly and is relatively widely available
- Functional test can be inaccurate in cases of prolonged baseline aPTT (assay needs to be modified by using factor V-deficient plasma)
- DNA test is more definitive and results are unambiguous. Specificity and sensitivity approximate 100% for the presence of the genetic mutation

- DNA testing alone may miss 5–10% of APC resistance due to other, acquired causes

36.2.12 Testing Not Indicated in the Following Situations

- As general population screen
- As a routine initial test during pregnancy
- As a routine initial test before or during oral contraceptive use, hormone replacement therapy, or serum estrogen receptor modifier

36.2.13 Management of Homozygotes with Thrombosis

- Treat acute thrombotic episode with heparin (or low molecular weight heparin), followed by warfarin
- Recommendation for long-term treatment and prophylaxis with warfarin to prevent recurrent thrombosis

36.2.14 Management of Heterozygotes with Thrombosis

- Those with first DVT event and presence of another reversible risk factor (pregnancy, oral contraceptives, and immobility) generally do not need long-term warfarin
- For those with a first DVT and no obvious acquired risk factor, consider undetermined additional genetic risk factor(s) and consider long-term warfarin therapy
- Patients with recurrent thrombotic events usually require long-term anticoagulation therapy and prophylaxis
- Patients should be counseled to avoid high-risk thrombotic situations and to get prophylactic anticoagulants before exposure to oral contraceptives, immobility, surgery, or pregnancy

36.2.15 Management of Asymptomatic Carriers

- Although some authorities recommend thromboprophylaxis in high-risk situations (e.g., postoperative state, extended plane flights), this is controversial and generally considered unnecessary

36.3 Prothrombin G20210A Mutation

36.3.1 General

- Prothrombin (factor II) is the precursor of thrombin, the final enzyme of the coagulation cascade, which converts fibrinogen to fibrin
- Prothrombin is a vitamin K-dependent protein, which is synthesized in the liver and circulates with a half-life of approximately 3–5 days
- The prothrombin G20210A mutation is in the 3′-untranslated region of prothrombin and is associated with increased levels of prothrombin in the circulation
- Increased prothrombin levels are associated with an increased risk of thrombosis

36.3.2 Clinical Manifestations

- DVT
- Pulmonary embolism
- Arterial thromboembolic complications are rare
- Indications of an inherited hypercoagulable syndrome are the same as for factor V Leiden

36.3.3 Acquired Risk Factors for Venous Thrombosis

- These are the same conditions that are described for factor V Leiden

36.3.4 Prevalence

- 1–3% in Caucasians and uncommon in individuals of Asian or African descent (Table 36.1)
- Threefold increased risk of venous thrombosis in heterozygotes
- Polymorphism present in 5–18% of patients with spontaneous venous thromboembolism

36.3.5 Genetics and Biochemistry

- The gene for factor II (prothrombin) is located on chromosome 11
- Prothrombin G20210A gene polymorphism is located in the 3′-untranslated region of the prothrombin gene
- The polymorphism is a single base-pair substitution at position 20210 of a guanine (G) for an adenine (A) nucleotide
- This polymorphism results in increased levels of prothrombin, which is associated with increased risk of venous thrombosis

36.3.6 Relative Risk

- Prothrombin gene heterozygotes have a threefold increased risk of venous thrombosis
- Risk of thrombosis substantially increased in patients with additional genetic risk factors including factor V Leiden, hyperhomocysteinemia, antithrombin III deficiency, protein C deficiency, and/or protein S deficiency

36.3.7 Functional Testing

- Functional or antigenic assays are not useful to detect the prothrombin G20210A polymorphism

36.3.8 Molecular Testing

- Direct DNA testing via PCR using either PCR-RFLP (Fig. 36.5) or PCR-FRET

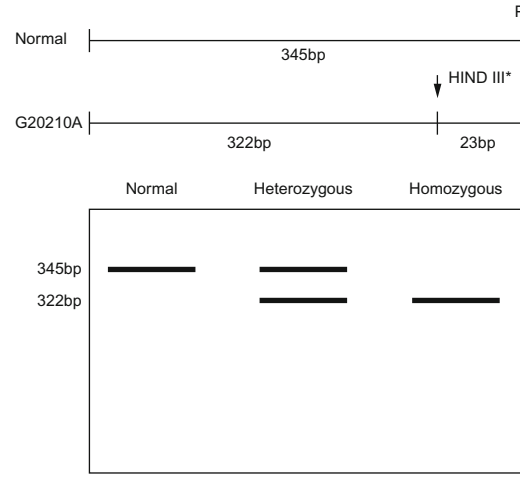


Fig. 36.5 RFLP analysis for G2021A mutation in prothrombin gene

- Indications for prothrombin G20210A testing include
  - Patients with a personal history of thrombophilia
  - Asymptomatic family members of patients with the prothrombin G20201A, if clinically indicated

36.4 Testing Not Indicated in the Following Situations

- As general population screen
- As a routine initial test during pregnancy
- As a routine initial test before or during oral contraceptive use, hormone replacement therapy, or serum estrogen receptor modifier therapy

36.4.1 Management of Homozygotes with Thrombosis

- Treat acute thrombotic episodes with heparin or low molecular weight heparin, followed by warfarin
- Consider long-term treatment with warfarin to prevent recurrent thrombosis

36.4.2 Management of Heterozygotes with Thrombosis

- Those with first DVT event and presence of another reversible risk factor (pregnancy, oral contraceptives, and immobility) may not need long-term warfarin
- Those with first DVT and no obvious acquired risk factors should be considered for an undetermined genetic risk factor and treatment with long-term warfarin
- Patients with recurrent thrombotic events usually require long-term anticoagulation therapy
- Patients should be counseled to avoid high-risk thrombotic situations and to use prophylactic anticoagulants with exposure to oral contraceptives, prolonged immobility, surgery, or during pregnancy

36.4.3 Management of Asymptomatic Carriers

- Similar to recommendations for asymptomatic carriers of factor V Leiden

36.5 Methylenetetrahydrofolate Reductase (MTHFR) C677T Thermolabile Polymorphism

36.5.1 General

- Homocysteine is an amino acid, derived from methionine, and may be converted to cysteine
- Homocysteine metabolic pathways require vitamins B_{12}, B_6, and folate; elevated homocysteine levels may be hereditary (due to mutations in these pathways) or acquired (due to deficiencies of vitamins B_{12}, B_6, or folate; renal failure; carcinoma; hypothyroidism; or medications)
- Elevations in homocysteine are associated with increased risk of arterial and venous thrombosis and atherosclerosis, based on retrospective case control studies; prospective studies show a weak positive association with arterial thrombosis, and no definite association for venous thrombosis
- Homozygosity or heterozygosity for the C677T mutation in the *MTHFR* gene, which is involved in homocysteine metabolic pathway, does not appear to be an independent risk factor for thrombosis. However, homozygosity for the C677T mutation may be significant in folate-deficient patients

36.5.2 Clinical Manifestations

- Severe MTHFR deficiency is a rare cause of homocystinuria
- The thermolabile polymorphism for MTHFR can result in mild-to-moderate elevations in the homocysteine level
- Moderate hyperhomocysteinemia typically manifests when folate levels are in lower end of normal range. Usually result of low intake of folate, B_6, or B_{12}

36.5.3 Prevalence

- 12% of US population is homozygous for the MTHFR C677T mutation (Table 36.1)

36.5.4 Genetics

- Gene for MTHFR is located on chromosomal 1, at region 1p36
- Thermolabile MTHFR variant has reduced activity at 37 °C and increased lability at 46 °C
- The C677T mutation is due to a C to T substitution at nucleotide 677, which encodes a change in alanine to valine
- Another common polymorphism in MTHFR is A1298C, which encodes for a change in glutamic acid to alanine

36.5.5 Relative Risk

- Controversial whether elevated homocysteine is a risk factor for venous thromboembolism

- No evidence that C677T or A1298C heterozygosity is a risk factor for venous or arterial thrombotic disease
- Homozygosity for C677T mutation in MTHFR is associated with higher plasma homocysteine levels, but is not, in itself, an independent risk factor for thrombosis

36.5.6 Diagnostic Assays for Homocysteine

- Both high-performance liquid chromatography and immunoassay are acceptable methods for measurement of plasma homocysteine levels
- Gender- and local population-specific reference ranges are strongly recommended
- Samples drawn in EDTA should be kept on ice if not analyzed within 30 min
- Homocysteine levels may remain elevated for several months following myocardial infarction or stroke
- Secondary causes of hyperhomocysteinemia such as B_{12} deficiency should be considered

36.5.7 Who Should Be Tested for Hyperhomocysteinemia?

- Patients with documented atherosclerotic disease (coronary artery, cerebrovascular, or peripheral vascular disease), particularly in younger individuals
- Controversial whether testing is indicated in patients with venous thromboembolism
- Routine screening for hyperhomocysteinemia in asymptomatic individuals is not recommended

36.5.8 Molecular Testing

- Direct DNA testing via PCR with RFLP or FRET analysis
- Genotyping for either 677 or 1,298 mutations in MTHFR is generally not recommended in subjects without first testing for elevated homocysteine

36.5.9 Management

- Still not clear regarding benefit of homocysteine-lowering therapy (i.e., with vitamin B_{12} or B_6 therapy)
- Selected patients (i.e., those with history of or at high risk for premature cardiovascular disease, stroke, or venous thromboembolism) may benefit from detection and treatment of hyperhomocysteinemia
- Treatment with either folic acid (0.4–1.0 mg/day) or vitamin B_{12} (0.5–1.0 mg/day) or both is relatively inexpensive and safe
- Goal to maintain plasma total homocysteine level <10 μmol/L
- Lowering homocysteine levels does not replace anticoagulant therapy in patients with a thromboembolic event

36.6 Plasminogen Activator Inhibitor-1 (PAI-1) 4 G/5 G Polymorphism

36.6.1 Clinical Manifestations

- PAI-1 inhibits tissue plasminogen activator
- High levels of PAI-1 may be associated with increased risk of arterial thrombosis due to inhibition of fibrinolysis
- Low levels have been reported to cause a rare familial bleeding disorder

36.6.2 Genetics

- Gene for PAI-1 is located on chromosome 7
- Gene for PAI-1 codes for protein of 50-kDal. It has several polymorphic loci including
  - A 3′ *Hind*III site
  - A CA(n) dinucleotide repeat in intron 3
  - A 4 G/5 G insertion/deletion—675 bp from the start site of the promoter
- The 4 G/5 G promoter site has been reported to exhibit genotype-specific responses to triglyceride with highest levels of PAI-1 in 4 G/4 G persons with elevated triglyceride levels

- PAI-1 4 G/5 G polymorphism modulates the basal PAI-1 levels
- PAI-1 4 G allele was associated with significantly increased PAI-1 levels and with myocardial infarction. This polymorphism is also a risk for severity of disease
- PAI-1 4 G allele also appears to increase the risk of venous thromboembolism, particularly in subjects with other thrombophilic defects

36.6.3 Functional Testing

- Diagnosed by functional (activity) assays or antigen (enzyme-linked immunosorbent) assays
- Not commonly performed clinical assay
- May be considered in patients with strong evidence for familial bleeding disorder
- Should not be measured in acute phase following thrombosis as it is an acute phase reactant
- Also elevated in pregnancy

36.6.4 Molecular Testing

- Direct DNA diagnosis by PCR and FRET exists for the PAI-1 4 G/5 G polymorphism
- Molecular testing is not available for the other polymorphisms or mutations

36.6.5 Management

- Venous thromboembolism is treated with heparin or low molecular weight heparin, followed by warfarin, as described above

36.7 Platelet Surface Glycoprotein IIIA (Human Platelet Antigen 1A and 2A)

36.7.1 General

- Expression of different platelet surface antigens is genetically determined
- Platelet surface glycoprotein GP IIIa is the most abundantly expressed platelet membrane glycoprotein
- Platelet glycoprotein GP IIIa exists on the platelet surface in association with glycoprotein IIb
- Upon activation, GPIIb-IIIa binds fibrinogen or VWF and mediates platelet aggregation
- A common polymorphism occurs at position 33 of GP IIIa, with either a leucine (referred to as HPA 1A or $P1^{A1}$) or proline (referred to as HPA 2A or $P1^{A2}$) at this position

36.7.2 Clinical Manifestations

- $P1^{A1}$ is implicated in neonatal alloimmune thrombocytopenia
  - Occurs when fetal platelets have an antigen from the father ($P1^{A1}$) that is absent in the mother ($P1^{A2}$)
  - Mother forms antibodies to $P1^{A1}$ that cross the placenta and destroy fetal platelets
  - Newborn platelet counts are often $<100{,}000/\mu L$ at birth, returning to normal within 2 weeks
- $P1^{A1}$ is also implicated in posttransfusion purpura (PTP)
  - A rare condition that occurs when a patient is transfused with platelets that express an antigen ($P1^{A1}$) that is absent in the patient ($P1^{A2}$)
  - The patient forms antibodies against the donor platelets
  - PTP is characterized by the sudden onset of thrombocytopenia 5–12 days after transfusion of a platelet-containing fraction
  - The thrombocytopenia is typically severe ($<10{,}000/\mu L$), and it usually begins to resolve within 14 days after the transfusion
- $P1^{A2}$ and cardiovascular disease
  - Weak association with coronary artery disease overall (reported in some studies, not in others)
  - Weak association with restenosis after revascularization procedures

36.7.3 Prevalence

- The wild-type Leu ($P1^{A1}$) is found in approximately 85% of the white population, whereas the Pro33 substitution ($P1^{A2}$) is present in 15% (Table 36.1)

36.7.4 Genetics

- The gene encoding GPIIIa is located on chromosome 17
- Polymorphism(s) of platelet antigens usually involve single amino acid substitution(s) caused by single nucleotide substitutions in the coding gene
- Expression of $P1^{A1}$ or $P1^{A2}$ is determined by whether leucine or proline is in position 33 of platelet surface glycoprotein IIIa, respectively

36.7.5 Antigenic Testing

- Platelet antigen typing by antigen-capture immunoassays
  - Monoclonal antibodies are used to immobilize the patient's platelet antigens onto a solid phase
  - Various antibodies of known antigen specificity are added
  - If an antibody binds, the patient's platelets have that particular antigen. For example, if an anti-$P1^{A1}$ antibody binds to the patient's platelet antigens, then the patient is found to carry the $P1^{A1}$ antigen
- Flow cytometry

36.7.6 Molecular Testing

- Alternatively, PCR can be used to identify the patient's platelet antigens. Many of the alterations in DNA sequence that account for these polymorphisms are known and can be identified by PCR

36.7.7 Management

- The treatment of choice for neonatal alloimmune thrombocytopenia or PTP is intravenous immunoglobulin at a dose of 400 mg/kg/day for 5 days
- Future transfusions should be washed or from an HPA-1a-negative donor

36.8 Hemophilia Mutations

36.8.1 General

- Hemophilia A (factor VIII deficiency) is the most common X-linked hereditary bleeding disorder involving secondary hemostasis
- Hemophilia B (factor IX deficiency) is also an X-linked hereditary bleeding disorder
- von Willebrand factor (VWF) is a protein that complexes with and stabilizes factor VIII in plasma. Mutations in the factor VIII binding site of VWF cause a phenotype similar to mild hemophilia A called von Willebrand disease, type 2 N (Normandy)
- In contrast to the previously described polymorphisms, molecular diagnostic testing for hemophilia consists of identifying mutations that may be scattered throughout the gene

36.8.2 Clinical Manifestations

- The clinical phenotype of hemophilia A or B depends on the factor level in the blood
- In patients with hemophilia, $<1\%$ factor VIII or IX activity results in a severe clinical phenotype, characterized by spontaneous bleeding in the head, gastrointestinal tract, joints, and so on
- Activity level of 1–5% is associated with moderate bleeding symptoms
- Activity level of $>5\%$ and $<35\%$ is considered mild hemophilia, in which bleeding occurs primarily with trauma or surgery rather than spontaneously

36.8.3 Prevalence

- Hemophilia A affects 1 in 5,000–10,000 males
- Hemophilia B affects 1 in 25,000–30,000 males

36.8.4 Differential Diagnosis

- Hemophilia A and hemophilia B are identical clinically and must be distinguished from each other by specific factor assays
- Factor XI deficiency can also result in bleeding with an isolated prolonged aPTT

36.8.5 Genetics

- The genes for factor VIII and factor IX are located on the X chromosome
- The gene for factor VIII is quite large, including 26 exons and spanning 186 kb
- Typically only males are clinically symptomatic
- Females with a hemophilia mutation on one of their two X chromosomes are carriers
- Female carriers with factor VIII levels <50% and bleeding symptoms have been reported
- If a family history is present, the inheritance pattern is X-linked recessive
- Up to 30% of hemophilia A or B cases arise from new mutations

36.8.6 Functional Testing

- The initial diagnostic tests for hemophilia are factor VIII and factor IX functional assays
- Mixing studies are used to determine if a specific factor inhibitor is present

36.8.7 Molecular Testing

- Factor VIII
  - Numerous missense, nonsense, deletion, and insertion mutations causing hemophilia A have been identified, making genetic testing difficult
  - An inversion mutation involving intron 22 has been shown to cause up to 40% of severe hemophilia A in Caucasians, which simplifies genetic testing in these families
  - Factor VIII intron 22 inversions can be detected by Southern blotting or by a DNA amplification assay that combines overlapping PCR with long-distance PCR
  - The specific intron 22 inversion assay has a sensitivity of >99% and a specificity of >97%. The sensitivity and specificity for carrier detection and for prenatal diagnosis for families with identified factor VIII gene inversions are both estimated to be >99%
  - For patients with mild or moderate disease or for those with severe disease and not having an intron 22 inversion, the coding regions and splice junctions of the factor VIII gene need to be analyzed by full gene sequencing
  - If full gene sequencing is negative, further analyses for the factor VIII gene intron 1 inversion, for example, or VWF Normandy should be considered
- Factor IX
  - Numerous mutations causing hemophilia B have also been identified
  - Like hemophilia A, genetic testing for female carrier status or prenatal detection can often be achieved with RFLP analysis or methods that directly demonstrate the mutation

36.8.8 Indications for Testing

- Individuals with a diagnosis of hemophilia A or B
- Appropriate at-risk female relatives of probands with identified mutations
- Hemophilia A or B carriers with previously identified factor VIII or IX gene mutations desiring prenatal diagnosis

36.8.9 Management

- The traditional goal of hemophilia management has been to recognize the earliest signs of bleeding and to treat promptly with the appropriate product to stop bleeding and avoid resulting chronic complications
- This usually requires intravenous infusion of the missing clotting factor. Available therapies for these patients include recombinant as well as highly purified plasma-derived factor VII and IX preparations
- There is an increasing emphasis on primary prophylaxis, or preventing bleeding episodes, especially in young children
- Primary prophylaxis involves the regular infusion of factor replacement products from an early age to prevent bleeding
- The goal of prophylaxis is to maintain factor VIII or factor IX levels above 1–2% to prevent spontaneous bleeding
- Secondary prophylaxis is a similar strategy that is implemented after repeated bleeding in a particular joint or other location, in an effort to prevent further bleeding in this area

36.8.10 Other Coagulation Factor Mutations

- Multiple other coagulation factors (procoagulant and anticoagulant) have been analyzed at the molecular level, but most of these analyses are not available as clinical laboratory tests. Those that have been analyzed most extensively are summarized in Table 36.2

36.9 Anticoagulant Therapy

36.9.1 General

- Anticoagulant therapy is efficacious in the prevention and treatment of venous thromboembolism, in the prevention of systemic embolism in patients with atrial fibrillation or prosthetic heart valves, as an adjunct in the prophylaxis of systemic embolism following myocardial infarction, and for the reduction of the risk for recurrent myocardial infarction
- The vitamin K antagonists (e.g., warfarin) produce an anticoagulant effect by interfering with the cyclic interconversion of vitamin K with its 2,3 epoxide, which modulates the γ-carboxylation of glutamate residues on the N-terminal regions of prothrombin and other vitamin K-dependent proteins (Fig. 36.6)
- The decreased functional levels of the vitamin K-dependent factors II (prothrombin), VII, IX, and X result in an anticoagulant effect that is monitored by the prothrombin time, which is converted to a specific ratio (international normalized ratio, or INR). Excessive anticoagulation with the vitamin K antagonists leads to an increased risk for hemorrhagic complications
- Warfarin, the most common vitamin K antagonist in clinical use, is a racemic mixture of two optically active isomers, the R and S enantiomers
- The S enantiomer of warfarin is more potent than the R enantiomer and is metabolized primarily by the CYP2C9 enzyme of the cytochrome P450 system. The less potent R enantiomer is metabolized primarily by CYP1A2 and CYP3A4 (Fig. 36.5)

36.9.2 Genetics

- Multiple variants in the CYP2C9 gene have been identified that alter the ability of this hepatic microsomal enzyme to metabolize S-warfarin. The alleles CYP2C9*2 and CYP2C9*3 exhibit a decreased ability to metabolize S-warfarin, resulting in a decreased dose of warfarin needed to have the same effect on the INR
- The target for warfarin's inhibitory effect on the vitamin K cycle is the vitamin K oxide reductase (VKOR) enzyme. Several mutations

Table 36.2 Outcome of investigations of hemostatic polymorphisms not typically available in clinical laboratories

Polymorphism	Phenotype	Relationship of phenotype to disease	Association of genotype with disease
Fibrinogen			
β-chain-455 G/A	Associated with altered level	Established but causal?	Inconsistent data
β-chain-Bc1 I	Associated with altered level	Established but causal?	Inconsistent data
Factor VII			
HVR4	Altered level	Inconsistent data	Unknown
402 G/A	Altered level	Inconsistent data	Unknown
401 G/T	Altered level	Inconsistent data	Unknown
323 0/10	Altered level	Inconsistent data	Unknown
Factor XIII			
Val 34Leu	Increased activation rate	Unknown	Preliminary results suggest 34Leu is protective against myocardial infarction
Thrombomodulin			
Ala 455Thr	Unknown	Suggestive	Inconsistent data
Ala25Thr	Unknown	Suggestive	Preliminary results suggest 25Thr associated with myocardial infarction

have been identified in this gene that result in enzymes that vary in their sensitivity to inhibition by warfarin
- The frequencies of these variant CYP2C9 alleles and VKOR genotypes vary in different ethnic and racial groups (Table 36.3)

36.9.3 Clinical Significance

- Several studies have shown that CYP2C9 and VKOR genetic variations contribute to maintenance dose requirements and are able to predict therapeutic doses during the initial weeks of therapy. In addition, certain variations of CYP2C9 and VKOR have been associated with an increased risk for excessive anticoagulation
- Pharmacogenetics-guided dosing has been shown in smaller studies to improve warfarin response in terms of more time spent in the therapeutic range, decreased bleeding events, and faster attainment of therapeutic INR in two small clinical trials

36.9.4 Molecular Testing

- Several approaches have been used for genotype determination of CYP2C9 and VKOR, including RT-PCR, high-resolution melting profile analysis, allelic discrimination, direct sequencing, as well as with various platforms

36.9.5 Indications for Testing

- Patients initiating warfarin therapy would potentially benefit the most from this information, which could be used to guide initial dosing recommendations
- A multicenter, prospective, double-blind, randomized clinical trial to determine the clinical benefit of pharmacogenetic information on warfarin dosing is currently open and enrolling patients (Clarification of Optimal Anticoagulation through Genetics [COAG]; Clinicaltrials.gov #NCT00839657)

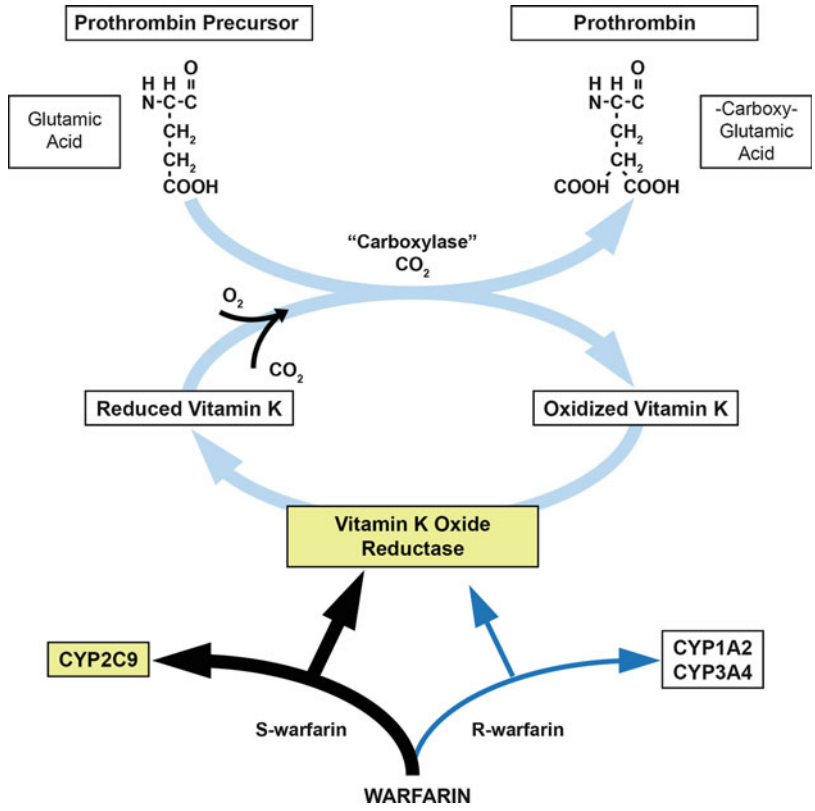


Fig. 36.6 The effect of warfarin on the synthesis of vitamin K-dependent coagulation factors. Reduced vitamin K is essential for the posttranslational conversion of selected glutamic acid residues in prothrombin (and other vitamin K-dependent coagulation factors) to γ-carboxylglutamic acid residues, which are necessary for normal reactivity of the zymogen. Vitamin K oxide reductase, the major warfarin-sensitive enzyme in this reaction, regenerates the reduced vitamin K, and this enzyme is primarily inhibited by the S enantiomer of warfarin. S-warfarin is metabolized by the P450 cytochrome enzyme CYP2C9 (Adapted from Ansell J, et al. Chest 2008;133:160S–198S)

36.10 Antiplatelet Therapy

36.10.1 General

- The main antiplatelet agents that are currently available include acetylsalicylic acid (aspirin), which irreversibly inhibits cyclooxygenase-1-dependent production of thromboxane A2; the thienopyridines (e.g., ticlopidine, clopidogrel, prasugrel), which irreversibly inhibit the platelet P2Y12 receptor for ADP; and inhibitors of the glycoprotein IIb/IIIa complex (e.g., abciximab)
- Following myocardial infarction or intracoronary stent placement, dual antiplatelet therapy with aspirin and clopidogrel is recommended to prevent future adverse cardiac events
- Clopidogrel is a prodrug, bioactivated predominantly by CYP2C19 (Fig. 36.7), and the presence of polymorphisms in this P450 isoenzyme has been associated with variable pharmacological response to this agent

36.10.2 Genetics

- A loss-of-function polymorphism, CYP2C19*2, results in the poor-metabolizer trait that is associated with a decreased inhibitory effect on platelet function

Table 36.3 Distribution of CYP2C9 and VKOR polymorphisms in different racial groups

	CYP2C9 polymorphisms					VKOR polymorphisms	
	*1	*2	*3	*4	*5	−1,639 G/A	1,173 C/T
Caucasians	0.743	0.008–0.143	0.109	0	0	–	0.422
African Americans	0.953	0	0.005–0.023	0	0.008	–	0.086
Chinese	0.967	0	0.021–0.045	–	–	0.157	0.14
Japanese	0.984	0	0.011–0.068	0	0	–	0.891

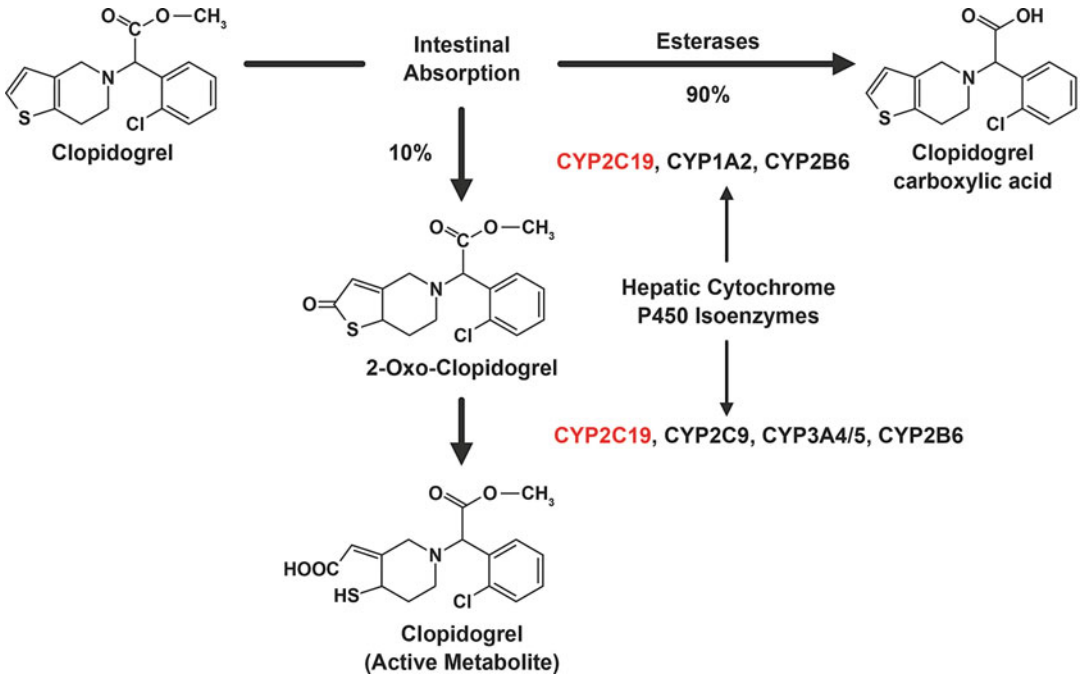


Fig. 36.7 Metabolism of clopidogrel. Clopidogrel is a prodrug that is converted to the active metabolite by CYP2C19 and other cytochromes (Adapted from Cattaneo M. Hematology Am Soc Hematol Educ Program. 2011;2011:70–75)

- The frequency of CYP2C19*2 varies in different racial groups and has been reported to be ~17% in African Americans, 30% in Chinese, and ~15% in Caucasians
- Other genetic polymorphisms (e.g., CYP2C19*3) have also been implicated in decreased efficacy of clopidogrel therapy. Conversely, the polymorphism CYP2C19*17 encodes a protein with increased function as compared to the wild-type allele

36.10.3 Clinical Significance

- High platelet reactivity while receiving clopidogrel therapy during coronary stent implantation, detected by various platelet function tests (see below), has been associated with an increased risk of death, nonfatal acute myocardial infarction, stent thrombosis, and ischemic stroke
- Reduced function CYP2C19 genotypes (e.g., CYP2C19*2) is associated with an increased

risk of major adverse cardiovascular events, particularly stent thrombosis
- Limited data exist concerning appropriate management for patients with laboratory evidence for either high on-treatment platelet reactivity and/or reduced function genotypes. Alternatives that have been proposed include an increased dose of clopidogrel or using prasugrel instead of clopidogrel, but prospective clinical trials documenting the efficacy of these approaches are not available

36.10.4 Functional Testing

- Several laboratory assays are available for evaluating the effect of clopidogrel on platelet function, including light-transmittance aggregometry with adenosine diphosphate (ADP), the VerifyNow P2Y12 assay, the Plateletworks assay using ADP tubes, and the Innovance PFA P2Y assay

36.10.5 Molecular Testing

- Multiple approaches have been used for determining the CYP2C19 genotype
- Loss-of-function mutations of CYP2C19 account for a minority of the variability of the functional response to clopidogrel, however

36.10.6 Indications for Testing

- Routine testing for platelet function or for CYP2C19 genotype is currently not recommended. Testing may be useful in certain high-risk clinical settings

Further Reading

Bagnall RD, Waseem N, Green PM, et al. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A. Blood. 2002;99:168–74.

Bertina RM, Koeleman BP, Koster T, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. Nature. 1994;369:64.

Bowen DJ, Keeney S. Unleashing the long-distance PCR for detection of the intron 22 inversion of the factor VIII gene in severe haemophilia A. Thromb Haemost. 2003;89:201–2.

Brinke A, Tagliavacca L, Naylor J, et al. Two chimaeric transcription units result from an inversion breaking intron 1 of the factor VIII gene and a region reportedly affected by reciprocal translocations in T-cell leukaemia. Hum Mol Genet. 1996;5:1945–51.

Cattaneo M. The clinical relevance of response variability to antiplatelet therapy. Hematology Am Soc Hematol Educ Program. 2011;2011:70–5.

Dahlback B. Inherited resistance to activated protein C, a major cause of venous thrombosis, is due to a mutation in the factor V gene. Haemostasis. 1994; 24:139.

Doggen CJM, Kunz G, Rosendaal FR, et al. A mutation in the thrombomodulin gene, 127 G to A coding for A1a25Thr, and the risk of myocardial infarction. Thromb Haemost. 1998;80:743.

Gardemann A, Schwartz O, Haberbosch W, et al. Positive association of the b fibrinogen H1/H2 gene variation to basal fibrinogen levels and to increase in fibrinogen concentration during acute phase reaction but not to coronary artery disease and myocardial infarction. Thromb Haemost. 1997;77:1120.

Gong IY, Schwarz UI, Crown N, et al. Clinical and genetic determinants of warfarin pharmacokinetics and pharmacodynamics during treatment initiation. PLoS One. 2011;6:e27808.

Hooper WC, Lally C, Austin H, et al. The role of the tissue-type plasminogen activator I/D and PAI-1 4 G/5G polymorphisms in African-American adults with a diagnosis of myocardial infarction or venous thromboembolism. Thromb Res. 2000;1:223–30.

Kangelaris KN, Bent S, Nussbaum RL, et al. Genetic testing before anticoagulation? A systematic review of pharmacogenetic dosing of warfarin. J Gen Intern Med. 2009;24:656–64.

Key N, McGlennen RC. Hyperhomocysteinemia. Arch Pathol Lab Med. 2002;126:1367–75.

Kohler HP, Futers T, Grant P. Prevalence of three common polymorphisms in the a-subunit gene of factor XIII in patients with coronary artery disease: association with FXIII activity and antigen levels. Thromb Haemost. 1999;81:511.

Koster T, Blann AD, Briet E, et al. Role of clotting factor VIII in effect of von Willebrand factor on occurrence of deep vein thrombosis. Lancet. 1995;345:152.

Lane DA, Grant PJ. Role of hemostatic gene polymorphisms in venous and arterial thrombotic disease. Blood. 1999;95:1517–32.

Lindh JD, Holm L, Andersson ML, et al. Influence of CYP2C9 genotype on warfarin dose requirements – a systematic review and meta-analysis. Eur J Clin Pharmacol. 2009;65:365–75.

Liu Q, Nozari G, Steve S. Sommer single tube polymerase chain reaction for rapid diagnosis of the inversion hotspot of mutation in hemophilia A. Blood. 1998;92:1458–9.

McGlennen RC, Key N. Clinical and laboratory management of the prothrombin G20210A mutation, archives of pathology and laboratory. Medicine. 2002;126:1319–25.

Mega JL, et al. Reduced-function *CYP2C19* genotype and risk of adverse clinical outcomes among patients treated with clopidogrel predominantly for PCI. A meta-analysis. JAMA. 2010;304:1821–30.

Nguyen V, et al. Robust dosage (RD)-PCR protocol for the detection of heterozygous deletions. Biotechniques. 2004;37:360–4.

Tybjaerg-Hansen A, Agerholm-Larsen B, Humphries SE, et al. A common mutation (G_{-455}-A) in the b-fibrinogen promoter is an independent predictor of plasma fibrinogen, but not of ischemic heart disease: a study of 9,127 individuals based on the Copenhagen City Heart Study. J Clin Invest. 1997;99:3034.

Voorberg J, Roelse J, Koopman R, et al. Association of idiopathic venous thromboembolism with single point-mutation at Arg506 of factor V. Lancet. 1994;343:1535.

Wang XL, Wang J, McCredie RM, et al. Polymorphisms of factor V, factor VII, and fibrinogen genes: relevance to severity of coronary artery disease. Arterioscler Thromb Vasc Biol. 1997;17:246.

Molecular Hemoglobinopathies

37

Tina Y. Fodrie, Shaobo Zhang, Mina Patel, Jodi A. Parks, and Liang Cheng

Contents

37.1 Normal Hemoglobin (Hb) 976

37.2 Hemoglobinopathies 977

37.3 Hemoglobin S 978

37.4 Hemoglobin C 980

37.5 Hemoglobin SC 981

37.6 Hemoglobin D 981

37.7 Hemoglobin E 981

37.8 Hemoglobin Constant Spring 982

37.9 Thalassemia 982

37.10 α-Thalassemia 982

37.11 β-Thalassemia 983

37.12 Complex β-Thalassemia 985

37.13 Hereditary Persistence of Fetal Hb Hemoglobin (HbFH) 985

37.14 Traditional Laboratory Techniques for Diagnosis 986

37.14.1 Complete Blood Count 986

37.14.2 Hemoglobin Electrophoresis 986

37.14.3 Isoelectric Focusing 987

37.14.4 High-Performance Liquid Chromatography 987

37.15 Molecular Techniques 987

37.15.1 Direct DNA Tests 987

37.15.2 Polymerase Chain Reaction-Based Test 988

Further Reading 991

T.Y. Fodrie, MT, MP(ASCP) (✉)
Laboratory for Personalized Molecular Medicine, San Diego, California, USA

S. Zhang, MD • J.A. Parks, MD
Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

M. Patel, MT, MP(ASCP)
Molecular Pathology Laboratory, Roudebush VA Medical Center, Indianapolis, IN, USA

L. Cheng, MD
Indiana University School of Medicine, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_37,

37.1 Normal Hemoglobin (Hb)

- Functional characteristics
  - Capable of carrying large quantities of oxygen
    - Hb is an iron-containing oxygen transporting metalloprotein
    - Makes up 97% of red blood cell (RBC) dry content
  - Able to act as a buffer
  - Remains soluble
  - At appropriate pressures, is able to take up and release oxygen
- Structural characteristics
  - Hb molecule is the combination of *heme* and *globin*
    - Globulins are globular proteins that are synthesized by ribosomes in the cytosol
    - Types of globin chains (Table 37.1)
      - Alpha (α)
      - Beta (β)
      - Delta (δ)
      - Epsilon (ε)
      - Gamma (γ)
      - Zeta (ζ)
    - Each globulin chain is covalently linked to a heme group
    - Heme is an iron-containing pigment that is synthesized by ribosomes in the cytosol
    - There are four hemes in each Hb molecule allowing it to bind to four molecules of oxygen
    - Numerous amino acids in or near the heme pocket facilitate the Hb molecule's ability to take up and release oxygen
  - Hb is a tetramer of 2α- and 2β-globin chains in normal adults
  - Types of normal Hb structures
    - HbA ($\alpha_2\beta_2$)
      - Normal adult Hb (97–98%)
      - Consists of two α- and two δ-globin chains (most common)
    - HbA_2($\alpha_2\delta_2$)
      - Adult (dependent upon method but usually 2%)
      - Consists of two α- and two δ-globin chains
    - HbF ($\alpha_2\gamma_2$)
      - Present in fetal period, approximately 1% in adults (in reference to Hb chains, adulthood is reached at 12 months of age)
      - Restricted to a few erythrocytes called F cells
      - Consists of two α- and two γ-globin chains
- Genetic characteristics (Fig. 37.1)
  - The α-cluster located on chromosome 16p with sequential genes of ζ_1, ζ_2, α_2, and α_1
  - The globin cluster located on chromosome 11p with sequential genes of ε, Gγ, Aγ, δ, and β
  - The globin genes are arranged from 5′–3′ according to the order of expression
  - Each globin chain is under separate genetic controls and is regulated completely independent of one another
  - Two globin gene switches occur during development
    - Embryonic to fetal switch-yolk sac erythropoiesis is completed by 10 weeks of fetal development and is superseded by α- and γ-chain production
    - Fetal to adult switch—γ-chain production decreases at approximately 6 weeks fetal development, and β-production increases and continues into adult life (Fig. 37.2)
  - Mismatched globin genes and errors in the switching mechanisms are the basis of hemoglobinopathies

Table 37.1 Types of Hb

Hbs	Globin chains	Mutation
HbA	2α and 2β	–
HbA_2	2α and 2δ	–
HbF	2α and 2γ	–
HbH	4β	–
Hb Bart	4γ	–
HbS	2α and 2β	Glutamic acid to valine at β_6
HbC	2α and 2β	Glutamic acid to lysine at β_6
HbE	2α and 2β	Glutamic acid to lysine at β_{26}
HbD	2α and 2β	Glutamic acid to glutamine at β_{121}

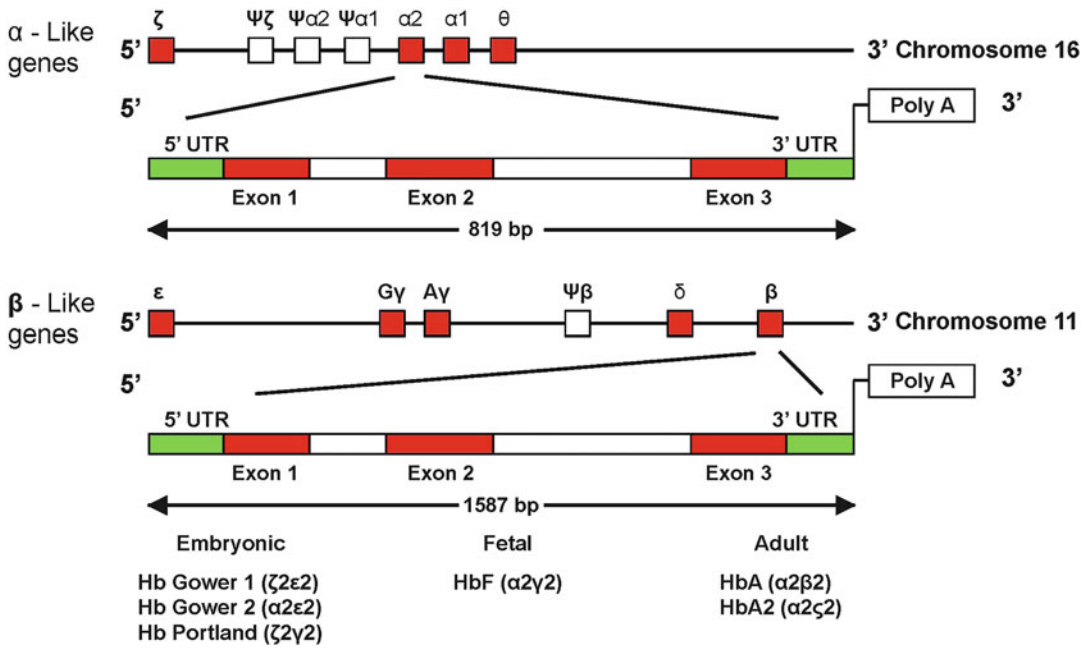


Fig. 37.1 Sequences of human globin genes within the α and β loci are located on chromosomes 16 and 11, respectively. Coding exons are *red*, introns are not shaded, and untranslated regions (UTRs, common to all globin genes) are *green*. The β genes are in the order of their developmental expression. The globins' expression shifts along with the development of individual. There are two major switches, one from embryonic to fetal form of Hb and another from fetal to adult form

37.2 Hemoglobinopathies

- General
  - Disorders of Hb caused by mutations
  - Approximately 5% of the world's population carries genes responsible for hemoglobinopathies
  - Each year about 30,000 infants are born with major hemoglobin disorders in the world population. Africa accounts for over 200,000 of these cases, specifically HbS anemia
  - Most are autosomal recessive (Table 37.2)
  - Subdivided into three categories based on the effects of mutation
    - Qualitative or structural (alteration of amino acid sequence in one or more of the globin chains)
    - Quantitative (reduced synthesis of one or more of the globin chains)
    - Hereditary persistence of fetal hemoglobin (HPFH)
- General characteristics
  - Changes in globin structure do not affect the rate of its synthesis
  - Usually caused by a point mutation (single-nucleotide substitution)
  - RBCs are more prone to lysis due to the instability of the Hb tetramers
  - Hundreds of clinically significant variants exist
  - Hb has increased or decreased oxygen affinity
- Types of hemoglobinopathies (Table 37.3)
  - HbS
  - HbC
  - HbD
  - HbE
  - Hb Constant Spring
  - α-Thalassemia
  - β-Thalassemia

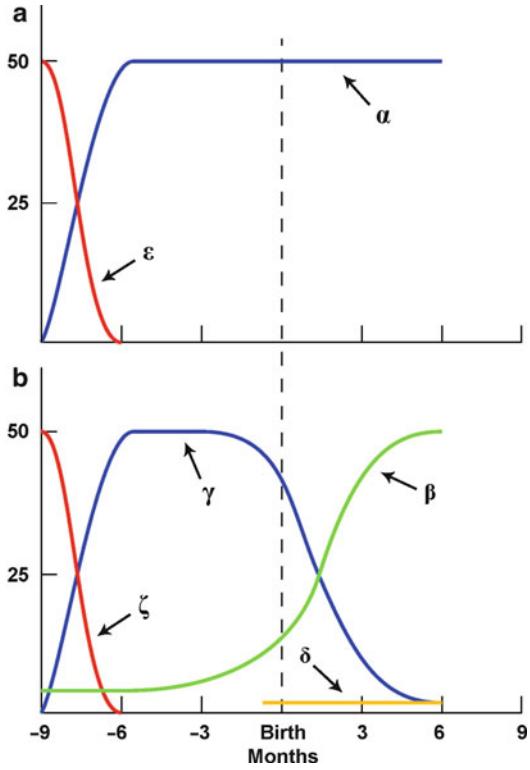


Fig. 37.2 Graphic illustration of the relative concentrations of each globin chain produced from conception to age 6 months (when Hb composition becomes like that of an adult). *Vertical axis* represents percentage of total globin chains

Table 37.2 Hb structural variants

Hb	Effect of mutation	Inheritance
S	HbS polymerizes under low oxygen tension → cells sickle → vascular occlusion	AR
C	HbC crystallizes within RBCs; cells less deformable, tend to fragment → hemolysis	AR
SC	Compound heterozygotes have mild sickle-cell disease symptoms	AR
E	Abnormal RNA splicing → decreased synthesis and mild thalassemia	AR

AR autosomal recessive

Table 37.3 Hb compositions and biologic conditions

Hb	Composition	Biologic condition
HbA	$(\alpha_2\beta_2)$	Normal adult
HbA_2	$(\alpha_2\delta_2)$	Normal level 2.5%
HbF	$(\alpha_2\gamma_2)$	Fetal period
HbH	(β_4)	α-Chain limitation
Hb Gower 1	$(\zeta_2\varepsilon_2)$	Embryonic stage
Hb Gower 2	$(\alpha_2\varepsilon_2)$	Embryonic stage
Hb Portland	$(\zeta_2\gamma_2)$	Embryonic stage
Hb Bart	(γ_4)	α-Thalassemia fetus
HbS	$(\alpha_2\beta_2{}^S)$	Sickle-cell disease
HbC	$(\alpha_2\beta_2{}^C)$	Hb C disease
HbCS	$(\alpha^{CS}\beta_2)$	Hb Constant Spring

  - Complex β-thalassemia
  - HPFH

37.3 Hemoglobin S

- General
  - Hemoglobin S (HbS) is the most common Hb variant
  - HbS differs from normal adult HbA only by a single amino acid substitution at the sixth position of the β-globin chain (A to T mutation)
    - Glutamic acid replaced with a valine
    - Autosomal recessive trait
    - HbS carriers are asymptomatic
  - First molecular disease to be recognized
  - Molecular pathogenesis is well characterized but still not completely understood
  - Hb composition in the HbS-related diseases
    - HbS >80%, HbF <20%, and HbA_2 3–8%
    - HbF, a tetramer of 2α- and 2γ-globulin, does not incorporate into the HbS polymer
    - Higher concentrations of HbF indicates lower concentration of HbS, resulting in an overall favorable outcome, decreased RBC destruction, and systemic damage
    - Studies have indicated that the production of HbF is largely regulated by several quantitative trait loci and genetic elements linked to the *HBB* gene-like cluster
    - Regulation occurs on chromosome 2p and 6q and upstream of *HBG2*

  - HbS molecules are only 20% as soluble as HbA in deoxygenated blood; therefore, under conditions of low oxygen tension, HbS molecules polymerize out of solution into insoluble strands
  - This polymerization changes the shape of the RBC containing the HbS from a biconcave disk to that of a sickle
    - HbS has normal oxygen affinity/carrying capacity until sickling occurs
  - Repeated sickling weakens the RBC membrane, altering its lipid content
  - These cells occlude the microvasculature, leading to ischemia
  - Lifespan of RBCs are also greatly reduced from 120 days to 10–20 days due to sequestration and extravascular hemolysis in the spleen
- Prevalence
  - Most common in equatorial Africa
    - The incidence of the mutation associated with HbS gene of HbS-related disease is increased in areas of the world where malaria (*Plasmodium falciparum*) is endemic
    - The mutation confers some resistance to falciparum malaria during a critical period of childhood development. The HbS trait favors the survival of the host, thus increasing survival and transmission of the HbS mutation. However, the inheritance of two HbS genes results in sickle-cell anemia, which will cause ill-health and possible death
    - Approximately 24% of the population in Nigeria carry one HbS gene and 20 per 1,000 births suffer from sickle-cell anemia, resulting in more 100,000 children per year inflicted with the disease
    - The prevalence of the HbS trait ranges between 10% and 40% across equatorial Africa to <2% in South Africa and the northern African coast
  - Median survival in the United States is estimated to be 42 years old for a man and 48 years for a woman. Survival in Africa is not known, as data collected is isolated only to those whom regularly visit the clinics. Data collected indicates that half of those with HbS anemia have died by the age of 5 years. Death is attributed to malaria, pneumococcal sepsis, or from the disease itself
  - One in 600 African-Americans are homozygous, and 8% are heterozygous
- Clinical symptoms
  - Sickle-cell trait: one allele contains (A to T mutation) on the β-globin gene, which results in a glutamic acid being substituted for a valine at position 6
    - Asymptomatic
    - Reduced susceptibility to malaria due to natural selection for the heterozygote advantage
  - Sickle-cell anemia: both alleles contain the mutation
    - Moderate to severe anemia due to markedly reduced lifespan of RBCs
    - Autosplenectomy occurs after repeated spleen infarction
    - Vaso-occlusive crisis is caused by sickle-shaped RBCs that obstruct capillaries and restrict blood flow to an organ, resulting in ischemia, pain, and organ damage
    - Bones are also a common target of vaso-occlusive damage, which may result from ischemia
    - Aplastic crisis is an acute worsening of the patient's baseline anemia producing pallor, tachycardia, and fatigue
    - Splenic sequestration crisis is an acute, painful enlargement of the spleen
    - Multiorgan system failure
    - Sickle cells may form in vivo resulting in sudden death during times of strenuous exercise at high altitudes, although this is very rare
    - An acute crisis is often precipitated by infection (usually urinary tract infection or occult pneumonia)
- Molecular pathogenesis
  - Autosomal recessive trait
  - Point mutation (adenine to thymine) at position 6 (of 146) in the β-globin gene

  - Codon GAG to GTG, glutamic acid to valine at position 6 (of 146) of the β-globin chain
- Common laboratory findings
  - Refer to Sect. 37.14
  - Abnormal complete blood cell count (CBC) results
  - Abnormal RBC morphology (poikilocytosis – variation in shape with the presence of target cells and sickle cells)
  - Positive newborn screen
  - Positive solubility test
- Molecular testing
  - Polymerase chain reaction (PCR)-restriction fragment length polymorphism (RFLP)
  - Advantages: solubility test for HbS is subjective and difficult to interpret and cannot be performed on newborns due to presence of high fetal proteins in blood sample
  - A confirmation test for positive solubility test and to determine carrier status
  - Performed on amniotic fluid from a fetus that is at high risk for sickle-cell anemia based on the carrier status of the parents
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Exposure of RBCs to sodium metabisulfate: produces sickling
  - Solubility test: HbS cells polymerize in the presence of a reducing agent resulting in increased solution turbidity
  - Refer to Sect. 37.14

37.4 Hemoglobin C

- General
  - Hemoglobin C (HbC) is the second Hb variant discovered; the second most common in the United States
  - The "C" designation for HbC is from the name of the city where it was discovered – Christchurch, New Zealand
  - Autosomal recessive trait
  - This mutation reduces the normal plasticity of host erythrocytes
    - HbC crystallizes in the RBC, due to decreased solubility
  - RBCs become more rigid, often fragment (microspherocytes form) as they attempt to transverse microvasculature
  - RBC lifespan is 30–35 days
  - Hb composition
    - HbC >90%, HbF <10%
- Prevalence
  - The HbC mutant allele is common in West Africa and found in 1% of African-Americans
  - Genetic compounds (heterozygotes for both HbS and HbC or thalassemia) are not infrequent, due to significant geographic overlap
- Clinical symptoms
  - Splenomegaly (from sequestration of rigid cells), mild to moderate normocytic, normochromic anemia (Hb level 8–12 g/dL)
  - In homozygotes, nearly all Hb is in the HbC form, resulting in moderate normocytic normochromic anemia
  - In those who are heterozygous for the mutation, about 28–44% of total Hb is HbC and anemia does not develop (it is considered a benign condition)
- Molecular pathogenesis
  - HbC is an abnormal Hb with a missense (mutation substitution of a lysine residue for glutamic acid residue) at the sixth position of the β-globin chain
- Common laboratory findings
  - Refer to Sect. 37.14
  - The presence of HbC crystals in RBCs
  - Abnormal CBC results
  - RBC morphology may reveal target cells, spherocytes, fragments, and obelisk-shaped cells (if patient has spleen)
  - Positive newborn screen
- Molecular testing
  - PCR-RFLP
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Osmotic fragility normal to decreased
  - Refer to Sect. 37.14

37.5 Hemoglobin SC

- General
  - Hemoglobin SC are red cells containing both HbS and HbC
  - HbSC exhibits a moderately severe phenotype in spite of being a mixture of HbS and HbC trait
  - Hb composition
    - HbS >40%, HbC >40%, and HbF <10%
- Prevalence
  - HbSC disease has an incidence of about 1:833 live births in African-Americans
- Clinical symptoms
  - Symptoms common to HbS (such as sickling and in some cases vaso-occlusive episodes) are found in individuals with HbSC disease but seen less often (rare or absent) and are not as severe
  - Moderate normocytic, normochromic anemia, and splenomegaly
- Molecular pathogenesis
  - Mutations consistent with HbS and HbC are found
- Common laboratory findings
  - Consistent with those for HbS and HbC
- Molecular testing
  - Consistent with those for HbS and HbC
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Consistent with those for HbS and HbC
  - Refer to Sect. 37.14

37.6 Hemoglobin D

- General
  - Hemoglobin D (HbD) occurs in four forms
    - Homozygous HbD disease
    - Heterozygous HbD trait
    - HbD-thalassemia
    - HbS-D disease
  - Several other autosomal recessive mutations result in HbD variants
  - Autosomal recessive trait
- Prevalence
  - HbD occurs mainly in northwest India, Pakistan, and Iran and in African-Americans
- Clinical symptoms
  - Homozygotes usually present with mild hemolytic anemia and mild to moderate splenomegaly
- Molecular pathogenesis
  - Missense mutation changes the glutamic acid at position 121 of the β-globin chain to glutamine
- Common laboratory findings
  - Possible abnormal CBC
- Molecular testing
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Hb electrophoresis: the electrophoretic mobility of HbD is identical to that of HbS at alkaline pH in cellulose acetate but is different on the citrate agar at pH 6.2
  - Solubility: HbD can be distinguished from HbS by its normal solubility
  - Refer to Sect. 37.14

37.7 Hemoglobin E

- General
  - Hemoglobin E (HbE) is the second most common structurally abnormal Hb in the world
  - Autosomal recessive trait
- Prevalence
  - Occurs mainly in Southeast Asia and is common in Thailand
- Clinical symptoms
  - Mild asymptomatic anemia
  - The effects of this mutation will be discussed in Sect. 37.9
- Molecular pathogenesis
  - Missense mutation replaces the glutamic acid residue at position 26 of the β-globin chain with a lysine
  - The β-chain of HbE is synthesized at a reduced rate compared with that of normal HbA as the mutation creates an alternative splicing site within an exon

- Common laboratory findings
  - Possible abnormal CBC
- Molecular testing
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Refer to Sect. 37.14

37.8 Hemoglobin Constant Spring

- General
  - Hemoglobin Constant Spring (HbCS) is a variant in which a mutation in the α-globin gene produces a α-globin chain that is abnormally long
  - HbCS α-chain protein is unstable because it contains 31 additional amino acids
  - Autosomal recessive trait
- Clinical symptoms
  - Anemia
- Molecular pathogenesis
  - Point mutation converts stop codon (UAA or UAG) to "coding" codon (CAA or CAG)
  - Messenger RNA (mRNA) for HbCS is unstable and is degraded prior to protein synthesis
- Common laboratory findings
  - Possible abnormal CBC (low Hb in RBCs)
- Molecular testing
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Refer to Sect. 37.14

37.9 Thalassemia

- General
  - Reduced rate of synthesis of normal globin chains resulting in frail RBCs which are vulnerable to mechanical injury (easily destroyed)
  - Autosomal recessive trait
  - The thalassemias are classified according to which chain of the globin molecule is affected
    - The production of α-globin in α-thalassemia is deficient
    - The production of β-globin in β-thalassemia is deficient
  - Thalassemias are a diverse group of diseases of Hb synthesis (chains are structurally normal)
- Prevalence
  - The estimated prevalence is 16% in Cyprus, 3–14% in Thailand, and 3–8% in India, Pakistan, Bangladesh, and China
  - Also common in Africa and Southeast Asia (distribution corresponds to prevalence of malaria)
  - Thalassemia was first identified in persons of Mediterranean descent
- Molecular pathogenesis
  - Various mutations decrease the synthesis of or destabilize either the α- or β-globin chain
  - In the absence of one chain, the complementary chain, which is present in excess forms tetramers (α_4, β_4, or γ_4)
  - These tetramers precipitate, accumulate in the nucleus, and then bind to the cytoskeleton
  - Cell division is blocked and the red cell membrane is compromised
  - Many cells are destroyed in the marrow and those that enter the circulation are sequestered by the spleen and remodeled
- General clinical symptoms
  - Asymptomatic to severe anemia
  - Severity of anemia is in direct proportion to the degree of chain imbalance

37.10 α-Thalassemia

- General
  - Excess β-chain production in adults and excess γ-chains in newborns
  - Excess β-chains form unstable tetramers that have abnormal oxygen dissociation curves
  - Thalassemia can coexist with other hemoglobinopathies
- Clinical symptoms
  - Clinical severity depends on how many of the four α-globin alleles are altered and if the change leads to partial (α+) or total (α°) deletion of α-globin (total is most common)

  - Disease can occur before or after birth
  - If two genes are deleted, anemia is prominent
  - With little to no α-globin present (3–4 deleted genes), β- or γ-globins form tetramers
    - β-Globin tetramer forms HbH
    - γ-Globin tetramer forms Hb Bart
    - The tetramers cannot release oxygen to tissues
    - Hb Bart leads to severe intrauterine hypoxia and massive generalized fluid accumulation (hydrops fetalis) that is incompatible with life
- Molecular pathogenesis
  - Genes *HBA1* and *HBA2* are involved
  - Four genetic loci for α-globin; as these loci that are deleted or mutated the manifestations of the disease become more severe
  - Gene deletion is the most common cause of α-thalassemia
  - There are two identical α genes on each chromosome 16p, both contain and are flanked by regions of very similar DNA sequence
  - Misalignment of these similar sequences leads to recombination of the α-1 gene on one chromosome with the α-2 on the other
- Forms of α-thalassemia (Table 37.4)
  - α-Thalassemia (silent)
    - Genotype − α/αα
  - α-Thalassemia minor
    - Genotype − −/αα or − α/−α
    - Anemia is minimal or absent
  - α-Thalassemia inter-media (HbH disease)
    - Genotype α−/− −
    - H bodies found in RBCs
    - Microcytic hypochromic anemia with hemolysis
    - Screening of prospective parents and genetic counseling of great importance
  - Hb Bart (γ-globin tetramer)
    - Genotype − −/− −
    - This Hb has a very high affinity for oxygen (see Table 37.4)
    - Incompatible with life – fetuses are hydropic and die in utero or soon after premature birth
- Common laboratory findings

Table 37.4 Forms of α-thalassemia and their genotype

Thalassemia	Genotype	Hb H (%)
α-Trait	−α/αα	0
α-Minor	−α/−α or − −/αα	Absent – trace
α-Intermedia (Hb H disease)	α−/− −	10–25
Hb Bart	− −/− −	γ-Tetramer

  - Highly abnormal RBC morphology found in the HbH and Hb Bart diseases
- Molecular testing
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Refer to Sect. 37.14

37.11 β-Thalassemia

- General
  - One or both β-globin genes have mutations that cause partial (β+) or total (β°) loss of β-chain production
  - The shift from γ-to β-chain production does not occur until after birth, so β-thalassemia does not cause hydrops fetalis
  - Autosomal recessive
  - Mutation present in the *HBB* gene on chromosome 11
  - Severity of disease depends on the type of mutation present
    - Reduction in function and alleles do not contain mutation are characterized as (β)
    - Complete absence of β chain due to mutation is characterized as (β°)
    - Mutations are characterized as (β+) if some production still occurs. (Note: the "+" in β+ is relative to β°, not β)
  - Relative excess amount of α chains present that do not form tetramers. They bind to cell membranes, producing membrane damage
- Forms of β-thalassemia (Table 37.5)
  - β-Thalassemia minor or commonly referred to as β-thalassemia trait

- Genotype (b+/b) or (b/b)
- Mild or no anemia
- One allele has the mutation
- Causes mild microcytic anemia (<75%fL)
  - Normal to increased RBC counts
    - Microcytosis
    - Electrophoresis
    - Mild increase in HbF
- Increase in HbA_2

– β-Thalassemia intermedia (β+ β+)
- Wide spectrum of disease, some patients may have very few clinical abnormalities, and others may experience severe consequences as their systems attempt to compensate for ineffective erythropoiesis and anemia
- Homozygous for β-thalassemia but able to maintain hemoglobin of 6–10 g/dL without regular transfusions
- Growth retardation with bony abnormalities (secondary to bone marrow hyperplasia)
- HbF 20–40%, increased HbA_2

– β-Thalassemia major (β°/β°)
- Also called Cooley anemia
- Microcytic, hypochromic anemia (Hb 2–3 g/dL)
- Increased number of RBCs
- Anisocytosis and poikilocytosis
- Presence of nucleated RBCs
- Hepatosplenomegaly (secondary to lysed RBCs)
- Bony abnormalities, failure to thrive
- HbF >90%, HbA_2 3–8%
- Patient is transfusion dependent, susceptible to iron overload
- Iron overload may cause growth retardation and failure to sexually mature
- May cause venous thrombosis, osteoporosis, and cardiac dysfunction

– Most patients have simple β-thalassemia, in which only the production of β-globin chains is affected
- Decreased synthesis of β-globin chains disturbs the balance between the two chains and α-chains precipitate
- The α-chains bind to the RBC membranes producing membrane damage

– Complex thalassemia involves deletion of both the β-globin gene and one or more of the other genes at the β-globin locus

Table 37.5 Forms of β-thalassemia and their genotype

Thalassemia	Genotype	Hb F (%)
Minor/trait	β°/β or β+/β	<10% (rarely exceeds 7%)
Intermedia	β+/β+	20–40
Major	β°/β°	>90

- Prevalence
  – The frequency of different types of mutations is specific to geographical location and ethnic group
- Clinical symptoms
  – Patients do not become symptomatic until HbF synthesis wanes at about 2 years of age
  – Depending on the form of β-thalassemia acquired (refer to the above section that provides description of the different types of β-thalassemia for details)
- Molecular pathogenesis
  – Due to mutations in the β-globin gene on chromosome 11
  – Overwhelming majority of β-thalassemias arise from point mutations in the β-globin gene (over 100 are known)
  – Sequence deletions are also recognized, for instance, a 619-bp deletion within the β-globin gene is common in patients of Indian heritage
  – mRNA splicing error is the most common mutation
    - >24 have been identified
    - The examination of the mutant mRNAs has yielded a great deal of data indicating which sequences are crucial to proper RNA processing
    - If the mutation involves either the 5′ GT or the 3′ AG (e.g., at a splice junction of an intron), splicing cannot occur (this particular type of mutation is seen in African-Americans)

    - Mutations within an exon or intron may also create a cryptic splice site (very similar in sequence to the true splice site)
    - Cryptic splice sites occur frequently and can be anywhere within the gene; the amount of normal mRNA present depends upon how often splicing takes place at the true site vs the cryptic site
    - This mechanism often underlies β+ thalassemia, because the true splice site is still utilized to some extent
  - Defects in posttranscriptional modification
    - A single-nucleotide substitution at the 5′ cap site of mRNA (A–C) or in the 3′ polyadenylation sequence (T–C) renders mRNA susceptible to degradation
    - Nonsense and frameshift mutations can both lead to the production of truncated mRNA, while a frameshift can also elongate mRNA, resulting in instability in both cases
    - Single-nucleotide point mutation halts translation at codon 39 (CAG or glutamine changed to UAG or stop) of the β-globin chain → β° thalassemia (Mediterranean)
    - Alternatively, a single base pair deletion at position 16 alters reading frame; translational apparatus encounters a stop codon too soon → β° thalassemia (Indian)
  - Promoter mutations
    - In the Japanese population, a single-nucleotide change within the ATA box (promoter sequence) leads to β+ thalassemia
- Common laboratory findings
  - Morphology: microcytic RBCs, although the number of RBCs present would be within normal range
  - Laboratory findings are dependent on the severity of the disease
- Molecular testing
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Refer to Sect. 37.14

37.12 Complex β-Thalassemia

- These forms of thalassemia are, fortunately, much less common
- Involve large deletions from the β-globin gene cluster
- If at least one of the γ genes is still intact, HbF will persist after birth
- For molecular and other laboratory tests available refer to Sects. 37.15 and 37.14

37.13 Hereditary Persistence of Fetal Hb Hemoglobin (HbFH)

- General
  - Fetal Hb (HbF) is the main oxygen transport protein in the fetus during the last 7 months of development
  - HbF binds oxygen with greater affinity than the adult form
  - HbF is nearly completely replaced by HbA by approximately the 12th week of postnatal life
  - Decreased β-globin chain synthesis is compensated for by the production of γ-globin
  - Homozygotes have 100% HbF
  - Heterozygotes have 70% HbA and 30% HbF
- Clinical symptoms
  - HPFH clinically similar to β-thalassemia but milder
- Molecular pathogenesis
  - Point mutations in the promoter region of one or another γ-globin gene alters interactions between various transcription factors and the promoter
  - A 27-kb deletion in the β-globin gene brings normally distant *cis*-acting factors into the vicinity of the genes, deregulating normal development
- Molecular testing
  - Refer to Sect. 37.15
- Miscellaneous laboratory testing
  - Kleihauer–Betke or acid elution test: red cells are fixed in alcohol and treated with

buffered citric acid; the cells are then stained with eosin
    - Cells that contain HbF will stain bright red (this Hb remains within the cells under these conditions)
    - Cells that contain HbA will not stain (this Hb elutes out of the cells)
    - This test can be performed to assess the extent of fetomaternal hemorrhage in Rh-negative women to determine RhoGam dosage or to detect HPFH
  - Flow cytometry: HbF identified by fluorescently labeled murine monoclonal antibodies in an instrument, which utilizes RBC gating parameters

37.14 Traditional Laboratory Techniques for Diagnosis

- Clinical indications for testing
  - Patient's clinical history may be suspicious
  - Abnormalities noted in complete blood count
  - Screening programs: all newborns born in the United States are screened for variant Hbs

37.14.1 Complete Blood Count

- Hb concentration: determined by cyanoHb method
- Hematocrit (Hct): MCV × RBC (or roughly 3 × Hb)
- Mean corpuscular volume (MCV): Hct/RBC
- Mean corpuscular Hb (MCH): Hb/RBC
- Mean corpuscular Hb concentration (MCHC): Hb/Hct × 100
- Absolute number of RBC
  - Absolute number of red cells: if this number is normal but cells are microcytic → possible β-thalassemia minor
- RBC distribution width: indicates range of RBC sites microcytosis
- RBC morphology
  - Anisocytosis: variation in size (microcytosis and macrocytosis)
  - Poikilocytosis: variation in shape (target cells, sickle cells, etc.)
  - Inclusions
    - HbH
      - In α-thalassemia (deletion of three α genes), excess β-globin chains precipitate, forming HbH inclusions
      - These inclusions, also called H bodies, are visible only with supravital staining (cannot be seen on standard Wright–Giemsa staining)
      - Appear within red cells in a regular distribution, like dimples on a golf ball (golf ball cells)
      - May be present in fewer than 50% of erythrocytes
      - H bodies (and α-inclusion bodies, in β-thalassemia) are not Heinz bodies; must be distinguished on supravital staining
    - Heinz bodies
      - Composed of precipitated Hb molecules, rather than globin chains
      - Can result from exposure of normal Hb to oxidant drugs or from the precipitation of unstable Hb
      - Heinz bodies are larger, less numerous, irregularly distributed, and exhibit more variation in size than do H bodies
      - More common to see both H bodies and Heinz bodies in patients who have had splenectomies
    - HbC
      - Tends to forms crystals

37.14.2 Hemoglobin Electrophoresis

- Method of choice for traditional laboratories for qualitative and quantitative analysis
- Used to examine globin chain composition and globin synthesis ratio
- RBC lysate analyzed
- Cellulose acetate electrophoresis performed at pH 8.6
- Citrate agar or agarose gel electrophoresis performed at pH 6.0–6.2

- Urea triton gel electrophoresis (allows the rapid analysis of small quantities of Hb)
- Hb separates into bands that migrate based on charge
- Acidic and alkaline gels yield different sets of bands. If two bands migrate together, both gels are needed
- The quantity of each Hb is determined with a densitometer (a spectrophotometer that measures the intensity of the stain taken up by each Hb fraction; the uptake is proportional to the Hb present)
- Advantages: very simple and fully automated
- Disadvantages: poor precision and accuracy of Hb quantitation by densitometer

37.14.3 Isoelectric Focusing

- Sensitive enough to separate Hb variants with isoelectric points that differ by as little as 0.02 pH
- Performed on agarose gel and employed to reveal Hb fractions, variants, and globin chains (polyacrylamide gel can be used for greater resolution)
- Combination with capillary electrophoresis has shown improved resolution and more accurate quantification

37.14.4 High-Performance Liquid Chromatography

- High-performance liquid chromatography (HPLC) is widely used for Hb quantification and to screen for Hb variants
- Separates Hb based on charge (porous cation-exchange column)
- Has replaced Hb electrophoresis in primary screening for clinically significant Hbs and acts as an adjunct for the detection of Hb variants
- Types
  - Microcolumn chromatography: sensitive method for HbA_2 quantitation
  - Cation-exchange HPLC: method of choice to quantify the Hb fractions
  - Reverse-phase HPLC: used to quantitate γ-chain levels, more sensitive and higher resolution than electrophoresis
  - Disadvantages: the presence of a Hb variant could alter the quantification result

37.15 Molecular Techniques

- Targets different levels of Hb expression at genomic/DNA, RNA, and protein levels
- Advantages: assays are very specific and are widely used in clinical settings
- Clinical indications for testing
  - Patient's clinical history may be suspicious
  - Abnormalities noted in complete blood count
  - Screening programs: all newborns born in the United States are screened for variant Hbs
- Sample requirements
  - Whole blood in ethylenediaminetetraacetic acid
  - Fetal DNA collected from the chorionic villi, amniotic fluid, or maternal circulation

37.15.1 Direct DNA Tests

37.15.1.1 Southern Blot

- Best method to determine if there is a deletion
- Applies RFLP to demonstrate a mutant allele (Table 37.6)
- Genomic DNA digested and then separated by electrophoresis
- Separated DNA is then blotted on a membrane
- DNA probes hybridize to the target DNA if they are complementary
- Advantage: ideal for screening for large deletions or rearrangements
- Disadvantage and limitations
  - The disease-causing mutations could only be identified through the genomic DNA library of affected individuals

Table 37.6 Hb genotype profiles identified by restriction endonuclease digestion

Genotype	Lost one *Mnl 1* site	Cuts with *Ddel*	*Mnl 1* fragments produced	*Dde 1* fragments
AA	No	Yes	107, 61, 21, 16	149, 56
AS	A no; S yes	A yes; S no	107, 77, 61, 21, 16	205, 149, 56
AC	A no; C yes	A yes; C yes	107, 77, 61, 21, 16	149, 56
SS	Yes	No	107, 77, 21	205
SC	S yes; C yes	S no; C yes	107, 77, 21	205, 149, 56
CC	Yes	Yes	107, 77, 21	149, 56

37.15.1.2 Direct Sequence Analysis of Amplified DNA

- Best method for definitive identification of mutations
- Sequencing is automated
- There are well-established protocols for the globin chain variants
- Disadvantage and limitations
  - DNA sequencing should always be coupled with a mutation screen method
  - Great caution should be exercised when working on the sequences with homologs, such as HBA1/HBA2 and HBG1/HBG2

37.15.1.3 Denaturing Gradient Gel Electrophoresis

- Identification (screening) for unknown globin gene sequence mutations
- Denaturing gradient gels are used to detect non-RFLP polymorphisms
- The small genomic restriction fragments are run on a low to high denaturing gradient acrylamide gel
- The fragments initially move according to molecular weight, but as they progress into higher denaturing conditions, each reaches a point where the DNA begins to "melt"
- The "melting" is due to the breaking of the weakest intrastrand bonding
- The structural changes in the DNA severely retards the progress of the molecule in the gel, and a change in mobility is observed
- Minor differences in genetic sequence can cause significant mobility shifts
- By comparing the melting behavior of the polymorphic DNA fragments side by side on denaturing gradient gels, it is possible to detect fragments that have mutations

37.15.1.4 Single-Stranded Conformation Polymorphism

- Single-stranded conformation polymorphism is defined as conformational difference of single-stranded nucleotide sequences of identical length as induced by differences in the sequences under certain experimental conditions
- This property allows the ability to distinguish the sequences by means of gel electrophoresis, which separates the different conformations
- Previously used as a tool to discover new DNA polymorphisms apart from DNA sequencing
- Can detect homozygous individuals of different allelic states, as well as heterozygous individuals, which demonstrate distinct patterns in electrophoresis
- Disadvantages and limitations
  - Must be coupled with DNA sequencing to demonstrate the mutation

37.15.2 Polymerase Chain Reaction-Based Test

- PCR is used to amplify specific regions of DNA, a gene or gene fragment
- DNA primers flank the region of interest at the 5′ and 3′ ends and are used to amplify the fragment
- The amplified DNA fragment is then further processed for electrophoresis, restriction enzyme digestion, sequencing, or hybridization

- Advantages
  - Can be adapted to any type of mutation
  - Needs very little DNA
  - Relatively less time consuming

37.15.2.1 Allele-Specific Priming or Amplification Refractory Mutation System

- Amplification is achieved by employing a primer which perfectly matches the 3′ terminal nucleotide
- Target DNA amplified in two reactions
  - Common forward primer
  - Second primer is either complementary to the wild type or the mutant
  - Can now be done in a single tube: tetra primer allele-specific priming or amplification refractory mutation system-PCR
  - Coamplify an unrelated sequence to serve as an internal control
- False negatives avoided by using internal control
- Advantage: can potentially detect any *known* mutation

37.15.2.2 Restriction Fragment Length Polymorphism Analysis

- Used to detect mutations of the β-globin gene, which is associated with HbS and HbC (Fig. 37.3a and b)
- The mutation region is amplified using PCR
- The amplified product contains restriction endonuclease sites; the number of cutting sites is dependent on the presence or absence of the point mutation
- After enzyme digestion (both *Mnl1* and *Dde1* are used), the PCR products are analyzed by gel electrophoresis (Fig. 37.4)
  - Various Hb genotypes can be identified by characteristic restriction fragment lengths of DNA (Table 37.6)

37.15.2.3 Allele-Specific Oligonucleotide Hybridization or Dot Blot Analysis

- The dot-blotting method requires binding the PCR-amplified target DNA sequence to a nylon membrane
- The DNA fixed to the membrane is then hybridized to the allele-specific oligonucleotide (ASO) probes that are labeled either with ^{32}P-labeled deoxynucleoside triphosphate, biotin, horseradish peroxidase, or a fluorescent marker at the 5′ end
- For mutation screening, a panel of ASO probes must be adapted to the mutations found in the ethnic group of the individual, which is tested
- For genotyping homozygous patients and for prenatal diagnosis, two oligonucleotide probes are required for each mutation
  - One complementary to the mutant DNA sequence
  - The other complementary to the normal gene sequence at the same position
- The patient's genotype is determined by the presence or absence of the hybridization signal of the mutation-specific and/or normal probe
- Disadvantages and limitations
  - The method is not suited for screening populations carrying a large number of different mutations, since each mutation requires a separate hybridization and washing step

37.15.2.4 Reverse Dot Blot Analysis

- The reverse dot-blotting technique allows detection of mutations with a single hybridization reaction
- DNA is bound to a nylon membrane strip with dots or slots
- Labeled amplified genomic DNA is then hybridized to the filter
- This procedure may require the use of several filters; the first will detect more frequent mutations observed in the patient's ethnic group and the others to less frequent abnormalities
- There are thalassemia detection strips commercially available, which correspond to the most frequent mutations observed in the various regions of the world

37.15.2.5 Gap-Polymerase Chain Reaction

- Applications: HPFH, δβ-thalassemia, and common α-thalassemia deletions/rearrangements

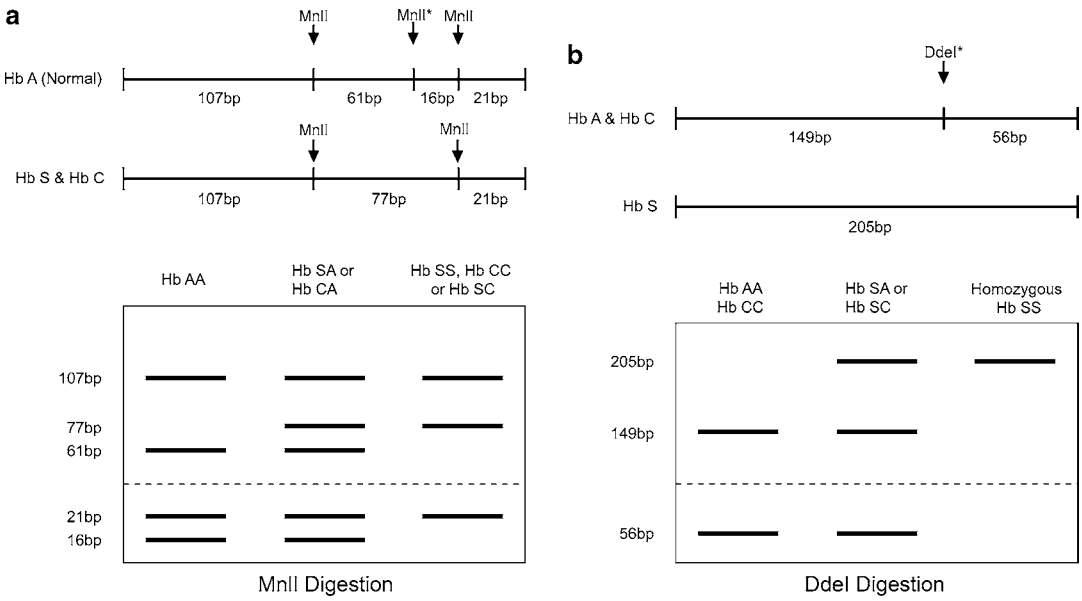


Fig. 37.3 The DNA fragments are the result of PCR amplification and digestion with the enzymes *MnlI* (*panel A*) and *DdeI* (*panel B*). In the lanes labeled AA (homozygous normal), *MnlI* produces four fragments or bands (107, 61, 21, and 16 bp), but the 16- and 21-bp fragments are too small to be visualized. *DdeI* (*panel B*) produces two bands (149 and 56 bp). In the lanes marked SA (sickle-cell trait or heterozygous mutant), *MnlI* produces five bands (107, 77, 61, 21, and 16 bp). *DdeI* digestion leads to three bands (205, 149, and 56 bp). In a person homozygous for S (lane SS, homozygous mutant), *MnlI* produces three fragments (107, 77, and 21 bp). *DdeI* does not cleave this DNA, since its site is lost (1 band, 205 bp). In SC disease (lane SC, double heterozygous), *MnlI* produces three bands (107, 77, and 21 bp), while *DdeI* also yields three (205, 149, and 56 bp, identical to AS). In Hb C disease (homozygous), *MnlI* yields three bands (107, 77, and 21) and *DdeI* produces two bands (149 and 56 bp)

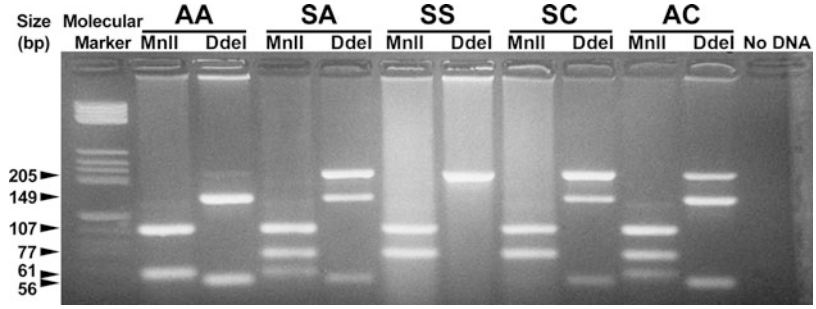


Fig. 37.4 Electrophoresis gel of PCR product after restriction enzyme digestion. The DNA fragments are the result of PCR amplification and digestion with the restriction endonucleases *MnlI* and *DdeI*. *Lane 1* on the far left contains a DNA molecular marker, whereas *lane 12* to the far right is a no DNA control. In *lanes 2* and *3*, labeled AA, (normal genotype), *MnlI* produces four DNA fragments or bands, (107, 61, 21, and 16 bp). The 21- and 16-bp fragments are too small to be seen. Digestion with *DdeI* produces two bands (149 and 56 bp). In *lanes 4* and *5* labeled SA, (sickle-cell trait genotype), *MnlI* produces five bands (107, 77, 61, 21, and 16 bp). Digestion with *DdeI* produces three bands (205, 149, and 56 bp). In *lanes 6* and *7* labeled SS (sickle-cell anemia genotype), *MnlI* digestion produces three fragments (107, 77, and 21 bp). The cleavage site for *DdeI* is lost in the S allele, leaving one band 205-bp long. In *lanes 8* and *9* labeled SC (SC disease genotype), *MnlI* produces three bands (107, 77, and 21 bp) whereas digestion with *DdeI* yields 3 (205, 149, and 56 bp). *Lanes 10* and *11* are labeled AC (Hb C trait genotype), digestion with *MnlI* yields five bands (107, 77, 61, 21, and 16 bp). Digestion with *DdeI* produces two bands (149 and 56)

- Pair of primers is employed; they are complementary to the flanking regions of the wild-type DNA sequence
- A third primer is complementary to the DNA sequence produced by a deletion (control)
- Wild-type PCR product larger than mutant, differentiate by electrophoresis

37.15.2.6 Polymerase Chain Reaction-Fluorescence Resonance Energy Transfer Probes

- Method is based on the fluorescence-labeled probes that are specifically designed for each mutation
- The probe yields a melting curve
- The genotype is determined by comparing the melting curves against normal reference
- Real-time PCR-fluorescence resonance energy transfer allows quick assigning of heterozygosity or homozygosity for the gene alleles
- Amplified DNA does not need more manipulations
- Disadvantage and limitations
  - Expensive
  - Specific probes are required; therefore, sequence must be known

Further Reading

Alli NA. Acquired hemoglobin H disease. Hematology. 2005;10:413–8.

Breda L, Gambari R, Rivella S, et al. Gene therapy in thalassemia and hemoglobinopathies. Mediterr J Hematol Infect Dis. 2009;1:e2009008.

Chui DH, Hardison RC, Riemer C, et al. An electronic database of human hemoglobin variants on the World Wide Web. Blood. 1998;91:2643–4.

Clark BE, Thein SL. Molecular diagnosis of haemoglobin disorders. Clin Lab Haematol. 2004;26:159–76.

Gu X, Zeng Y. A review of the molecular diagnosis of thalassemia. Hematology. 2002;7:203–9.

Higgs DR, Garrick D, Anguita E, et al. Understanding alpha-globin gene regulation: aiming to improve the management of thalassemia. Ann N Y Acad Sci. 2005;1054:92–102.

Oneal P, Gantt N, Schwartz G, et al. Fetal hemoglobin silencing in humans. Blood. 2006;108:2081–6.

Papasavva T, Kalakoutis G, Kalikas I, et al. Noninvasive prenatal diagnostic assay for the detection of beta-thalassemia. Ann N Y Acad Sci. 2006;1075:148–53.

Patrinos GP, Panagoula K, Papadakis MN. Molecular diagnosis of inherited disorders: lessons from hemoglobinopathies. Hum Mutat. 2005;26(5):399–412.

Paulin L, Itano HA, Singer SJ, et al. Sickle cell anemia, a molecular disease. Science. 1949;110:543–8.

Pauling L, Itano HA, Singer SJ, Wells IC. Sickle cell anemia: a molecular disease. Science. 1949;110:543–8.

Rees D, Porter J, Clegg J. Why are hemoglobin F levels increate HbE/β-thalassemia? Blood. 1999;94:3199–204.

Sadelain M, Lisowski L, Samakoglu S, Rivella S, May C, Riviere I. Progress toward the genetic treatment of the beta-thalassemias. Ann N Y Acad Sci. 2005;1054:78–91.

Solovieff N, Milton J, Hartley S, et al. Fetal Hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. Blood. 2010;115:1815–22.

Stamatoyannopoulos G. Control of globin gene expression during development and erythroid differentiation. Exp Hematol. 2005;33:259.

Walter P, Fung E, Killilea D, et al. Oxidative stress and inflammation in iron-overloaded patients with β-thalassemia or sickle cell disease. Br J Haematol. 2006;135:254–63.

HLA System and Transfusion Medicine: Molecular Approach

38

S. Yoon Choo

Contents

38.1 The Human Leukocyte Antigen System 994
38.1.1 Genomic Organization of the Human Major Histocompatibility Complex 994
38.1.2 Human Leukocyte Antigen Haplotypes 994
38.1.3 Expression of Human Leukocyte Antigen 995
38.1.4 Structure and Polymorphism of Human Leukocyte Antigen Molecules 995
38.1.5 Functional Implications of the Human Leukocyte Antigen Polymorphism 996
38.1.6 Clinical Human Leukocyte Antigen Testing 997
38.1.7 The Human Leukocyte Antigen System and Transplantation 998
38.1.8 The Human Minor Histocompatibility Antigens 1000
38.1.9 The Human Leukocyte Antigen System in Transfusion Medicine 1000
38.1.10 Human Leukocyte Antigen and Disease Association 1001
38.1.11 Parentage Human Leukocyte Antigen Testing 1002
38.1.12 Human Leukocyte Antigen in Anthropological Studies 1002

38.2 Transfusion Medicine 1002
38.2.1 Human Blood Group Systems 1002
38.2.2 Human Platelet Antigen System 1004
38.2.3 Human Neutrophil Antigen System 1005
38.2.4 Blood Donor Screening for Infectious Diseases 1005

38.3 Posttransplant Chimerism Study 1005
38.3.1 Quantification of Chimerism 1006

Further Reading 1007

S.Y. Choo, MD
Mount Sinai School of Medicine, New York, NY, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_38,

38.1 The Human Leukocyte Antigen System

- The genetic system involved in the rejection of foreign ("nonself") organs is called the major histocompatibility complex (MHC), and highly polymorphic cell surface molecules are encoded by the MHC
- The human MHC is called the human leukocyte antigen (HLA) system because these antigens were first identified and characterized using alloantibodies developed against allogeneic leukocytes
- The HLA system has been well known as transplantation antigens, but the primary biologic role of HLA molecules is to facilitate immune response to foreign antigens

38.1.1 Genomic Organization of the Human Major Histocompatibility Complex

- The human MHC maps to the short arm of chromosome 6 (6p21) and spans approximately 3,600 kb
- The human MHC is divided into three regions (Fig. 38.1)
  - The class I region is located at the telomeric end of the complex
  - The class II region at the centromeric end
  - The class III region in the center

38.1.1.1 The Class I Region

- Consists of the classical genes (*HLA-A, HLA-B, HLA-C*), the nonclassical genes (*HLA-E, HLA-F, HLA-G*), pseudogenes (*HLA-H, HLA-J, HLA-K, HLA-L*), and gene fragments (*HLA-N, HLA-S, HLA-X*)
- The *HLA-A, HLA-B*, and *HLA-C* loci encode the heavy α chains of class I antigens and they define HLA-A, B, and C antigens, respectively. The class I gene has an exon–intron structure and separate exons encode different domains of the class I heavy chain (Fig. 38.2)
- Some of the nonclassical class I genes are expressed with limited polymorphism, and their functions are not well known

38.1.1.2 The Class II Region

- Consists of a series of subregions, each containing *A* and *B* genes encoding α and β chains, respectively. The *DR, DP*, and *DQ* subregions encode the major class II molecules
- The *DR* gene family consists of a single *DRA* gene and nine *DRB* genes (*DRBI–DRB9*). Different HLA haplotypes contain particular numbers of *DRB* loci. The *DRB1, DRB3, DRB4*, and *DRB5* genes are usually expressed, and the other *DRB* loci contain pseudogenes. The *DRA* locus encodes an invariable α chain and it binds various β chains. HLA-DR antigen specificities (i.e., DR1–DR18) are determined by the polymorphic β chains encoded by *DRB1* alleles
- The *DQ* and *DP* families each have one expressed gene for α and β chains and additional pseudogenes. The *DQA1* and *DQB1* gene products associate to form the DQ molecules, and the *DPA1* and *DPB1* products form DP molecules
- The nonclassical class II gene, *HLA-DO* and *HLA-DM*, may play a role in antigen processing and presentation

38.1.1.3 The Class III Region

- Does not encode HLA molecules but contains genes for complement components (C2, C4, and factor B), 21-hydroxylase, and tumor necrosis factors

38.1.2 Human Leukocyte Antigen Haplotypes

- HLA loci are closely linked, and the entire MHC is inherited as an HLA haplotype in a Mendelian fashion from each parent. Recombination within the HLA system occurs with a frequency $<1\%$, and it appears to occur most frequently between the *DP* and *DQ* loci

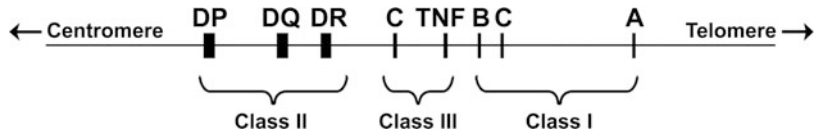


Fig. 38.1 The human MHC on the short arm of chromosome 6. The HLA-DR, DP, and DQ regions consist of one or more A and B genes, respectively. *TNF* tumor necrosis factors, *C* complement genes

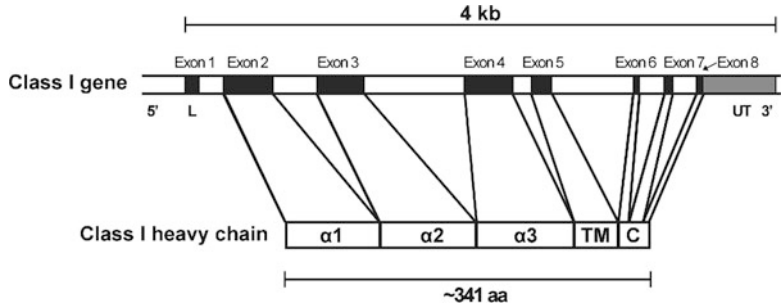


Fig. 38.2 Exon–intron structure of the HLA class I gene and its encoded heavy chain. *L* leader peptide, *UT* untranslated region, α1, α2, α3 extracellular domains, *TM* transmembrane, *C* cytoplasmic, *aa* amino acids

- Possible combinations of antigens from different HLA loci on an HLA haplotype are enormous, but some HLA haplotypes are found more frequently than expected by chance in certain populations. This phenomenon is called the linkage disequilibrium. For example, HLA-A1, Cw7, B8, DR17 is the most common HLA haplotype among Caucasians, with a frequency of 5%
- The segregation of HLA haplotypes within a family can be assigned by family studies (Fig. 38.3). Two siblings have a 25% chance of being genotypically HLA identical (sharing both haplotypes), a 50% chance of being HLA haploidentical (sharing one haplotype), and a 25% chance that they share no HLA haplotypes

38.1.3 Expression of Human Leukocyte Antigen

- HLA class I molecules are expressed on the surface of almost all nucleated cells. They can also be found on red blood cells and platelets
- Class II molecules are expressed on B lymphocytes, antigen-presenting cells (monocytes, macrophages, and dendritic cells), and activated T lymphocytes

38.1.4 Structure and Polymorphism of Human Leukocyte Antigen Molecules

38.1.4.1 Human Leukocyte Antigen Class I

- Class I molecules consist of glycosylated heavy chains of approximately 44,000 Da (44 kDa) encoded by the HLA class I genes and a noncovalently bound extracellular 12 kDa β_2-microglobulin. Human β_2-microglobulin is invariant and is encoded by the *B2M* gene located on the long arm of chromosome 15
- The class I heavy chain has three extracellular domains (α_1, α_2, and α_3), a transmembrane region, and an intracytoplasmic domain (Fig. 38.4). Each extracellular domain comprises about 90 amino acids. The α_1 and α_2 domains contain variable amino acid

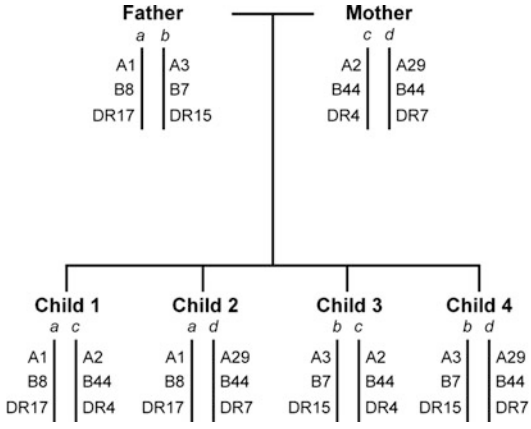


Fig. 38.3 Mendelian inheritance of HLA haplotypes demonstrated in a family study. HLA haplotypes and genotypes can be inferred from phenotype data in an informative family study as illustrated. For example, the father's HLA phenotype is HLA-A1, 3; B7, 8; DR15, 17. From the family study, his genotype is A1, B8, DR17/A3, B7, DR15. The paternal HLA haplotypes are A1, B8, DR17 ("*a*") and A3, B7, DR15 ("*b*"); and the maternal HLA haplotypes are A2, B44, DR4 ("*c*") and A29, B44, DR7 ("*d*"). The four possible genotypes among the offsibling are *a*/*c*, *a*/*d*, *b*/*c*, and *b*/*d*

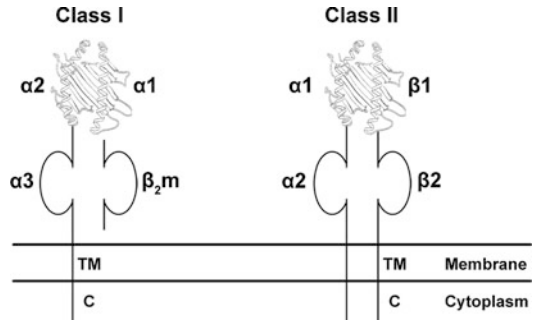


Fig. 38.4 Schematic diagram of the HLA class I and class II molecules. The α_1 and α_2 domains of class I molecule and the α_1 and β_1 domains of class II molecule form the peptide-binding sites

sequences, and these domains determine the serologic specificities of the HLA class I antigens
- The heavy chain α_1 and α_2 domains together form a unique structure consisting of a platform of eight antiparallel β strands and two antiparallel α-helices on top of the molecule. A groove formed by the two α-helices and the β-pleated floor is the binding site for processed peptide antigens

38.1.4.2 Human Leukocyte Antigen Class II

- The products of the class II loci DR, DP, and DQ are heterodimers of two noncovalently associated glycosylated polypeptide chains: α (30–34 kDa) and β (26–29 kDa) (Fig. 38.4). The α- and β-chains are transmembrane, and they have the same overall structures. An extracellular portion composed of two domains (α_1 and α_2, β_1 and β_2), respectively, is anchored on the membrane by a short transmembrane region and a cytoplasmic domain
- The extent of class II molecule variation depends on the subregion and the polypeptide chain. Most polymorphisms occur in the first amino terminal α_1 domain of *DRB1, DQB1*, and *DPB1* gene products
- The three-dimensional structure of the HLA-DR molecule is similar to that of the class I molecule. The α_1 and β_1 domains together form a peptide-binding groove

38.1.5 Functional Implications of the Human Leukocyte Antigen Polymorphism

- The HLA system is known to be the most polymorphic in humans. The polymorphic residues are not evenly spread throughout the molecule but are clustered in the antigen-binding groove. T cells recognize intracellularly processed peptides that are bound to the self-MHC molecules on the cell surface. This phenomenon is called the MHC restriction. Amino acid variations in several regions change the fine shape (pockets) of the groove and, thus, provide binding specificity to different peptides. An individual's HLA repertoire will thus determine what peptide antigens can be presented to elicit the immune system in a given individual
- The distribution and frequency of HLA antigens vary greatly among different ethnic

Table 38.1 Pretransplant HLA workups and donor selection

	Kidney transplantation	Hematopoietic stem cell transplantation
Pretransplant HLA testing	HLA antigen typing	HLA allele typing
	HLA antibody test	HLA antibody test and lymphocyte crossmatch (optional)
	Lymphocyte crossmatch	
Final selection criteria	ABO major compatibility	HLA allele matching
	Negative lymphocyte crossmatch	Negative lymphocyte crossmatch (optional)

Table 38.2 Numbers of recognized HLA private antigen specificities and alleles

Locus	Antigen specificities	Alleles
HLA-A	24	1,698
HLA-B	55	2,271
HLA-C	9	1,213
HLA-DRB1	17	975
Total	105	6,157

groups. It has been postulated that the diversity of HLA polymorphism was derived and evolved by unique selective pressure in different geographic areas. This could be related to the role of the HLA molecule in the presentation of significant infectious agents in the different areas of the world

38.1.6 Clinical Human Leukocyte Antigen Testing

- HLA testing in the transplant workup includes HLA typing of the recipient and the potential donor, screening and specificity identification of preformed HLA antibodies in the recipient, and detection of HLA antibodies in the recipient that are specifically reactive with lymphocytes of a prospective donor (crossmatch) (Table 38.1)

38.1.6.1 Serologic Typing of Human Leukocyte Antigen

- The complement-mediated microlymphocytotoxicity technique had been used historically as the standard serologic typing of HLA antigens. Lymphocytes to be typed are incubated with various antibodies with known specificities against different HLA antigens, and complement is added to mediate the lysis of antibody-bound cells
- Peripheral blood lymphocytes express HLA class I antigens and are used for the serologic typing of HLA-A, HLA-B, and HLA-C. HLA class II typing is done with isolated B lymphocytes because these cells express class II molecules. Cells must have good viability, and thus requiring the blood specimens to be drawn fresh
- Formal nomenclature of serologically defined antigens is given by the World Health Organization HLA Nomenclature Committee

38.1.6.2 Molecular Typing of Human Leukocyte Antigen Alleles

- The extent of HLA polymorphism at the gene level is far higher than known antigen specificities. Clinical molecular typing has been developed to identify serologically undistinguishable but functionally discrete HLA alleles (Table 38.2)
- The first molecular HLA typing technique introduced in the mid-1980s was a restriction fragment length polymorphism Southern blotting analysis
- Polymerase chain reaction (PCR)-based clinical HLA typing was developed using sequence-specific oligonucleotide probe (SSOP) methods. The hypervariable exon 2 sequences encoding the first amino terminal domains of the *DRB1, DQB1*, and *DPB1* genes are amplified from genomic DNA by PCR reaction. The HLA class I polymorphism is located in the two domains, α_1 and α_2, requiring amplification of two exons. Based on the HLA sequence database, a panel of synthetic oligonucleotide sequences corresponding to variable regions of the genes are designed and used as SSOP in hybridization with the amplified PCR products

- Alternatively, the sequence-specific primer (SSP) method can be used. Polymorphic DNA sequences are used as amplification primers. Only alleles containing sequences complementary to these primers will anneal to the primers and amplification will proceed. It detects sequence polymorphism at given areas by the presence of a particular amplified DNA fragment (Fig. 38.5)
- Actual DNA sequencing of amplified products of multiple HLA loci is increasingly used as clinical HLA typing in support of the unrelated donor hematopoietic stem cell transplantation
- HLA alleles are designated by the locus followed by an asterisk (*), a two-digit number corresponding to the antigen specificity, and ":" followed by the assigned allele number. For example, *HLA-A*02:10* represents the tenth of *HLA-A*02* alleles with the serologically defined HLA-A2 antigen specificity

38.1.6.3 Human Leukocyte Antigen Antibody Screening and Lymphocyte Crossmatch

- Preformed HLA antibodies can be detected by testing the patient's serum against a panel of lymphocytes of known HLA phenotypes or purified HLA molecules. This test is called HLA antibody screening. With a panel of well-selected cells expressing various HLA antigens or multiple purified HLA molecules, antibody specificities can be assigned
- When a potential donor is identified, a final crossmatch is performed between the recipient's serum and donor's lymphocytes to determine the compatibility. The positive crossmatch results are predictive of the risks of allograft rejection and shorter graft survival

38.1.7 The Human Leukocyte Antigen System and Transplantation

- HLA-A, HLA-B, HLA-C, and HLA-DR have long been known as major transplantation antigens. Both T cell and B cell (antibody) immune responses against HLA alloantigens are important

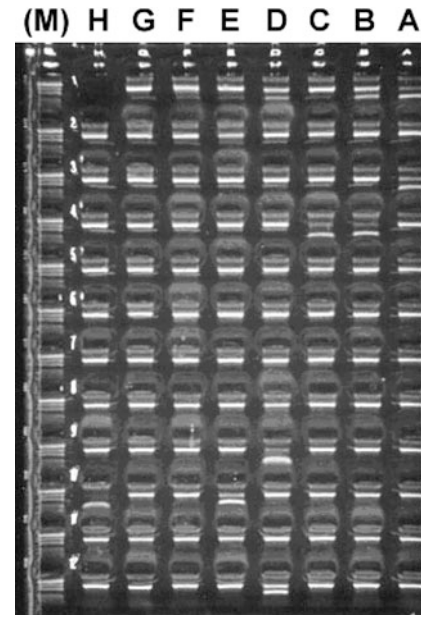


Fig. 38.5 Gel picture of PCR-SSP typing results. After PCR reactions with a set of sequence-specific amplification primers, amplified products are visualized on agarose gel electrophoresis by staining with ethidium bromide and exposure to ultraviolet light. Interpretation of PCR-SSP results is based on the presence (positive reaction) or absence (negative reaction) of a specific amplified DNA fragment. An internal control band is used to verify the integrity of the PCR reactions, and it should be amplified in all reactions except in the negative ("no DNA") control (1H). The internal control primer pair amplifies a conserved region of the human β-globin gene in this test system. A positive band will be observed if a specific HLA gene sequence was amplified (1A, 1B, 1D, 3A, 4B, 4C, 9D, 10E, 10H, and 12D). HLA typing results are interpreted from the patterns of positive wells. In the presence of a positive band, the amplified internal control product may be weaker or absent due to the differences in concentration and melting temperatures between the specific primer pairs and the internal control primer pair

38.1.7.1 Solid Organ Transplantation

- In solid organ transplantation, blood group ABO system is the most important transplantation antigen. Preexisting natural anti-A and/or anti-B antibodies in the recipient cause hyperacute rejection of the organ from

an ABO major incompatible donor because ABO antigens are expressed on endothelial cells. Preformed HLA antibodies as a result of pregnancy, transfusion, or previous organ transplantation also cause hyperacute rejection. These antibodies bind to the corresponding alloantigens on vascular endothelium of the graft, fix complement, and cause damage. The problem of hyperacute rejection can be prevented when transplantation is performed from a selected donor whose lymphocytes are not reactive with recipient's serum (i.e., negative or compatible lymphocyte crossmatch)

- The benefits of HLA antigen matching between the recipient and the donor are well established in kidney transplantation. There is a clear relationship between the degree of HLA matching and kidney graft survival in transplants from living-related donors. The influence of HLA matching on the survival of liver and thoracic organs is yet uncertain

38.1.7.2 Allogeneic Hematopoietic Stem Cell Transplantation

- Allogeneic hematopoietic stem cell transplantation is used to treat hematologic malignancy, severe aplastic anemia, severe congenital immunodeficiencies, and selected inherited metabolic diseases. The sources of hematopoietic stem cells are bone marrow, mobilized peripheral blood stem cells, and umbilical cord blood
- The HLA system is the major transplantation antigen in stem cell transplantation, and the degree of HLA matching is predictive of the clinical outcome. HLA antigen or allele mismatch between a recipient and a stem cell donor represents a risk factor not only for graft rejection/failure but also for acute graft versus host disease (GVHD) because immunocompetent donor T cells contained in the allograft attack the targets of mismatched allogeneic HLA molecules on the recipient's organs (skin, liver, and gastrointestinal tract)
- The best compatible stem cells are from an identical twin or a genotypically HLA-identical sibling. For those who do not have an HLA-identical sibling, an alternative related family member who is HLA-haploidentical and partially mismatched for the nonshared HLA haplotypes may serve as a suitable donor, but these transplants have a higher risk of developing acute GVHD and graft rejection or failure

38.1.7.3 Unrelated Donor and Cord Blood Transplantation

- When an HLA-matched sibling or a partially mismatched acceptable related donor is not available, phenotypically matched unrelated donors can be considered. The National Marrow Donor Program (NMDP) was founded in the United States in 1986 to establish a volunteer marrow donor registry and to serve as a source of HLA-matched unrelated stem cell donors. The chance of finding an HLA-matched unrelated donor depends on the patient's HLA phenotype. The NMDP registry now contains more than ten million HLA-typed adult donors and more than 165,000 cord blood units. There are also international donor registries, and most of these registries share their donors each other (20 million adult donors and 550,000 cord blood units globally)
- Unrelated donor transplants are associated with an increased incidence of more severe acute GVHD compared with HLA-matched sibling transplants. Such an increase may result partly from mismatch in HLA alleles and from unrecognized minor histocompatibility antigen disparity. For this reason, *HLA-A, B, C*, and *DRB1* allele matching is strongly recommended for unrelated donor transplants. Some patients do not find a perfectly allele-matched unrelated donor for multiple HLA loci. A partially mismatched cord blood transplants can be considered for some of these patients because partial HLA mismatch is better tolerated with this source of stem cells. The major limitations of cord blood transplantation, however, are increased incidence of graft failure and delayed hematopoietic recovery because of the lower number of stem cells in cord blood units

38.1.8 The Human Minor Histocompatibility Antigens

- Minor histocompatibility antigens are intracellularly processed peptides naturally derived from normal cellular proteins that associate with HLA molecules. Minor histocompatibility antigens are inherited and have allelic forms. The number of minor histocompatibility loci is probably high, and the extent of polymorphism for each locus is not known
- Minor histocompatibility antigens have been defined by both class I and class II MHC-restricted T cells. Examples include the male-specific H-Y antigens and a series of HA antigens (HA-1 through HA-8). The H-Y antigens are encoded by multiple Y-specific genes, with differences of 1–4 amino acids from their X homologs
- Minor histocompatibility antigen disparity can be associated with GVHD in HLA-identical transplants (e.g., H-Y antigen in a male recipient and a female donor who has been immunized to H-Y antigen by having a male baby)
- Whether minor histocompatibility antigen disparity can have a significant impact as a risk factor for graft rejection or GVHD might depend on the tissue-specific expression of proteins, the frequency of different allelic forms, and the immunogenicity of peptides. The expression of some antigens (e.g., HA-1 and HA-2) is limited to the hematopoietic system, while others (e.g., H-Y and HA-3) are ubiquitously expressed on tissues

38.1.9 The Human Leukocyte Antigen System in Transfusion Medicine

- The HLA system can cause adverse immunologic effects in transfusion therapy. These effects are primarily mediated by "passenger" donor leukocytes contained in the cellular blood components. HLA antibodies can develop from previous alloimmunization episodes and can cause platelet immune refractoriness, febrile transfusion reaction, and transfusion-related acute lung injury

38.1.9.1 Transfusion-Associated Graft Versus Host Disease

- When functionally competent, allogeneic T lymphocytes are transfused into an individual who is severely immunocompromised; these T lymphocytes are not removed and can mount an immune attack against the recipient's cells, causing transfusion-associated graft-versus-host disease (TA-GVHD). TA-GVHD is not common and typically occurs in patients with congenital or acquired immunodeficiencies or immunosuppression that affects T lymphocytes
- TA-GVHD has also occurred in patients without apparent evidence of immunodeficiency. The majority of these studied cases involved a blood donor who was homozygous for one or more HLA loci for which the recipient was heterozygous for the same antigen and a different one (e.g., HLA-A2 vs. HLA-A2, 3). This relationship is called a one-way HLA mismatch in the GVH direction and a one-way HLA match in the rejection direction. As a result, the donor cells will not be recognized as foreign by the recipient's lymphocytes, while the donor's lymphocytes will recognize HLA alloantigens present in the recipient. The one-way match more likely occurs when an HLA haplotype is shared by a donor and a recipient (HLA-haploidentical), such as in directed donation from blood relatives and among populations with relatively homogeneous HLA phenotypes. The latter possibility may account for the observation that more cases of TA-GVHD have been reported among Japanese patients
- The clinical features of TA-GVHD are similar to those of GVHD following an allogeneic hematopoietic stem cell transplant, i.e., fever, rash, diarrhea, and liver dysfunction. TA-GVHD is further characterized by prominent pancytopenia due to marrow aplasia

- Demonstration of donor-derived lymphocytes in the circulation of a patient with characteristic clinical findings is diagnostic for TA-GVHD. The persistence of donor lymphocytes (mixed lymphoid chimerism) can be tested by molecular HLA typing, by cytogenetic analysis if donor and patient are of different gender, and by other molecular marker polymorphisms. The demonstration of donor-derived lymphohematopoietic cells in a transfusion recipient is not diagnostic of TA-GVHD per se, because donor lymphocytes can be normally detected in the recipient's circulation a few days after transfusion
- Similarly, GVHD can occur following a solid organ, especially liver, transplant. The clinical pictures and diagnosis are the same as in TA-GVHD

38.1.10 Human Leukocyte Antigen and Disease Association

- Certain diseases, especially of autoimmune nature, are associated more frequently with particular HLA types. However, the association level varies among diseases, and there is generally a lack of a strong concordance between the HLA phenotype and the disease. Thus, definite diagnosis or assessing risk for most diseases cannot be made by HLA tying alone. The exact mechanisms underlying the HLA-disease association are not well understood, and other genetic and environmental factors may play roles as well
- The degree of association between a given HLA type and a disease is often described in terms of relative risk, which is a measure of how much more frequently a disease occurs in individuals with a specific HLA type when compared with individuals not having that HLA type
- Among the most prominent associations are ankylosing spondylitis with HLA-B27, narcolepsy with *HLA-DQB1*06:02*/*HLA-DRB1*15:01*, and celiac disease with *HLA-DQB1*02*. The HLA-A1, B8, DR17 haplotype is frequently involved in autoimmune disorders. Rheumatoid arthritis is associated with a particular sequence of the amino acid positions 66–75 in the DR β-chain that is common to the major subtypes of DR4 and DR1. Type 1 diabetes mellitus is associated with DR3, 4 heterozygotes, and the absence of asparagine at position 57 on the DQ β-chain appears to render susceptibility to this disease

38.1.10.1 Human Leukocyte Antigen and Pharmacogenetics

- Certain HLA alleles can predict the side effects of drugs. Severe hypersensitivity reactions can occur in patients receiving abacavir for the treatment of HIV, and the *HLA-B*57:01* allele is strongly associated with this immunologically mediated side effect. Screening for this allele is recommended prior to starting abacavir therapy, and this approach of avoiding abacavir treatment in the *HLA-B*57:01*-positive patients has reduced the incidence of hypersensitivity and offered alternative personalized treatment
- The presence of the *HLA-B*15:02* allele was reported as a risk for Steven–Johnson syndrome and toxic epidermal necrolysis in Han Chinese patients receiving carbamazepine treatment for epilepsy. Since this allele is present almost exclusively in Asian populations, the US FDA recommends *HLA-B*15:02* allele screening in this group

38.1.10.2 Hereditary Hemochromatosis

- One of the most common inherited diseases manifested by an increased absorption of dietary iron, resulting in excess iron deposition in the liver, heart, and endocrine organs and finally organ failure
- Determined by an autosomal recessive gene, *HFE*, up to 10% of the population is heterozygous (carriers) and 0.5% homozygous. The *HFE* gene is located approximately 5 Mb telomeric to the *HLA-A* locus. The HFE protein is similar to the HLA class I heavy chain, associates with β_2-microglobulin, and regulates iron absorption. Hereditary

hemochromatosis had been known to be associated with HLA-A3 antigen before the *HFE* gene was identified. More than 20 mutations have been identified, but a particular single amino acid substitution, C282Y, is involved most frequently

38.1.11 Parentage Human Leukocyte Antigen Testing

- In parentage testing, genetic markers of a child, biological mother, and alleged father are compared to determine exclusion or nonexclusion of the alleged father. An alleged father would be excluded if he does not share an HLA haplotype with the child; i.e., 0% probability of paternity (POP). Conversely, a man who has one haplotype identical to the child's would not be excluded, and the probability of being a biological father varies with the frequency of that particular haplotype in the population. A POP of 99.9% indicates a 99.9% probability that the alleged father is indeed the biological father of the child
- There are some advantages of using HLA types in parentage testing. The HLA system is inherited in a Mendelian manner and is extensively polymorphic; its recombination rate is low; mutation has not been observed in family studies; and antigen frequencies are known for many different ethnic groups
- However, the HLA system does not provide a high exclusion probability when the case involves a paternal HLA haplotype that is common in the particular ethnic group. Molecular techniques using non-HLA genetic systems are now widely used, and there is decreasing use of HLA typing for paternity testing

38.1.12 Human Leukocyte Antigen in Anthropological Studies

- Analysis of differential distribution of HLA alleles and haplotypes is an invaluable tool in the study of the evolutionary origins and migration of human populations

38.2 Transfusion Medicine

38.2.1 Human Blood Group Systems

- A blood group system includes those antigens that are encoded by alleles at a single genetic locus or those produced by a complex of two or more very closely linked homologous genes with virtually no or extremely rare recombination (crossing over) occurring between them (e.g., three of the systems, MNS, Rh, and Chido/Rogers, comprise at least two loci each, so closely linked that recombination between them is extremely rare)
- In some systems, the gene encodes directly the blood group determinants (i.e., protein determinants), whereas in others, where the antigen is carbohydrate in nature, the gene encodes a glycosyltransferase enzyme, which catalyzes biosynthesis of the carbohydrate determinants
- Some antigens are expressed only on red cells (e.g., Rh), whereas others throughout the body (e.g., ABO)
- The biologic function of most blood group antigens is mostly unknown
- The polymorphism in blood group antigens can be detected by
  - Serologic typing: blood group antigens are defined by reagent antibodies with known specificities
  - Molecular determination: DNA analysis to detect allelic polymorphism, based on single nucleotide polymorphism (SNP) and allele-specific PCR amplification
- The clinical significance of the blood group system relates to the potential of alloantibodies to cause destruction of transfused incompatible red blood cells (hemolytic transfusion reaction) or to cross the placenta (IgG antibodies are capable of crossing the placenta) and destroy incompatible fetal red cells (hemolytic disease of the newborn [HDN])

Table 38.3 Blood group systems

Number	System name	System symbol	Gene name(s)	Chromosomal location
001	ABO	ABO	*ABO*	9q34.2
002	MNS	MNS	*GYPA, GYPB, GYPE*	4q31.21
003	P	P1	*P1*	22q11.2–qter
004	Rh	RH	*RHD, RHCE*	1p36.11
005	Lutheran	LU	*LU*	19q13.32
006	Kell	KEL	*KEL*	7q34
007	Lewis	LE	*FUT3*	19p13.3
008	Duffy	FY	*DARC*	1q23.2
009	Kidd	JK	*SLC14A1*	18q12.3
010	Diego	DI	*SLC4A1*	17q21.31
011	Yt	YT	*ACHE*	7q22.1
012	Xg	XG	*XG, MIC2*	Xp22.33
013	Scianna	SC	*ERMAP*	1p34.2
014	Dombrock	DO	*ART4*	12p12.3
015	Colton	CO	*AQP1*	7p14.3
016	Landsteiner–Wiener	LW	*ICAM4*	19p13.2
017	Chido/Rodgers	CH/RG	*C4A, C4B*	6p21.3
018	H	H	*FUT1*	19q13.33
019	Kx	XK	*XK*	Xp21.1
020	Gerbich	GE	*GYPC*	2q14.3
021	Cromer	CROM	*CD55*	1q32.2
022	Knops	KN	*CR1*	1q32.2
023	Indian	IN	*CD44*	11p13
024	Ok	OK	*BSG*	19p13.3
025	Raph	RAPH	*CD151*	11p15.5
026	John Milton Hagen	JMH	*SEMA7A*	15q24.1
027	I	I	*GCNT2*	6p24.2
028	Globoside	GLOB	*B3GALT3*	3q26.1
029	Gill	GIL	*AQP3*	9p13.3
030	Rh-associated glycoprotein	RHAG	*RHAG*	6p21–qter

38.2.1.1 Terminology for the Blood Group Systems

- Defined by the International Society of Blood Transfusion (ISBT) Working Party on terminology for red cell surface antigens
- A numerical terminology for red cell surface antigens. By definition, these antigens must be defined serologically by the use of a specific antibody (Table 38.3)

38.2.1.2 Hemolytic Disease of the Newborn

- Maternal IgG antibodies against red cell antigens following alloimmunization from transplacental fetomaternal hemorrhage or transfusions can cross placenta and coat the fetal red cells, causing accelerated destruction (immune hemolysis) and resulting fetal anemia
- During any pregnancy, a small amount of the fetal blood can enter the mother's circulation. Fetal red cells possessing paternal antigen foreign to mother can cause alloimmunization. Obstetrical events that increase the risk of transplacental hemorrhage include spontaneous abortion, therapeutic abortion, ectopic pregnancy, amniocentesis, intrauterine surgery, abdominal trauma, and hemorrhage in the peripartum
- Antigens involved in HDN

– D antigen of the Rh blood group system is best known, but many others are implicated: other Rh blood group system (c, C, e, and E), Kell, Duffy, Kidd, and Ss systems

38.2.1.3 Prenatal Determination of RhD Type of Fetus

- Early and safe prenatal diagnosis of RhD status of fetus is advantageous for the management of pregnancies at risk of HDN due to RhD alloimmunization
- Fetal *RHD* genotyping by PCR amplification: *RhD* gene is absent in most RhD-negative chromosome 1p34–p36
  - Amniocyte DNA typing from amniocentesis as early as 10 weeks of gestation
  - Chorionic villus biopsy samples in first trimester
  - Fetal *RHD* genotype detection from circulating cell-free fetal DNA in maternal plasma (noninvasive method)
- PCR-based amplification assays have been also developed to determine other *RH*, *KEL*, *FY*, and *JK* genotypes

38.2.2 Human Platelet Antigen System

- The human platelet antigen (HPA) system is expressed specifically on platelets. These platelet-specific antigen specificities are determined by platelet glycoproteins (Table 38.4)
- HPA alloantibodies are responsible for the following clinical conditions: neonatal alloimmune thrombocytopenia, post-transfusion purpura, and immune refractoriness to platelet transfusions

38.2.2.1 Neonatal Alloimmune Thrombocytopenia

- Neonatal alloimmune thrombocytopenia (NAIT) develops as a result of maternal alloimmunization during pregnancy against fetal platelet antigens inherited from the father and absent in the mother. Antiplatelet IgG antibodies cross the placenta and cause fetal and neonatal immune thrombocytopenia. The major complication of severe thrombocytopenia is intracranial hemorrhage, which leads to death or neurologic sequelae. About half of cases involve the first child. NAIT is considered to be the platelet counterpart of the HDN
- Different HPA are implicated in NAIT of different races. The antigens most frequently implicated in NAIT are HPA-1a and HPA-5a in Caucasians and HPA-4a and HPA-3a in Asians
- Platelet-specific antigens are generally weak immunogens, and genetic factors may influence whether HPA-1a-negative women will develop anti-HPA-1a antibody. Individuals with certain HLA haplotypes with *HLA-DRB3*01:01* allele are more likely to develop antibodies against HPA-1a antigen. The incidence is estimated approximately 1/1,000 live births
- Diagnosis is confirmed when a maternal antiplatelet alloantibody is demonstrated to be directed against a paternal antigen present in the affected fetus or newborn. Platelet antigen typing of the newborn and parents are performed either by phenotyping or genotyping
- HPA genotyping
  - PCR using sequence-specific primers (PCR-SSP) is currently the most widely used technique for HPA genotyping
  - The risk of a subsequent pregnancy being affected is 100% if the father is homozygous for the implicated antigen and 50% if heterozygous

Table 38.4 Human platelet antigen (HPA) system

System	Antigen	Glycoproteins	Antigen frequency (%)
HPA-1	HPA-1a	GPIIIa	98
	HPA-1b		29
HPA-2	HPA-2a	GPIb	97
	HPA-2b		15
HPA-3	HPA-3a	GPIIb	88
	HPA-3b		54
HPA-4	HPA-4a	GPIIIa	>99
	HPA-4b		<1
HPA-5	HPA-5a	GPIa	98
	HPA-5b		21

- Management of severe thrombocytopenia involves transfusion of antigen-negative platelets. Frequently, washed (to remove antibodies) maternal platelets (negative for implicated antigens) are used for transfusion support

38.2.3 Human Neutrophil Antigen System

- The number of well-characterized neutrophil-specific antigen systems is limited (Table 38.5). Genotyping of human neutrophil antigen (HNA) is available using allele-specific PCR amplification
- The detection of neutrophil antibodies is less well established
- Clinical significance of HNA
  - Neutrophil alloantibodies are known to cause neonatal alloimmune neutropenia, immune neutropenia after hematopoietic stem cell transplantation, immune refractoriness to granulocyte transfusions, febrile nonhemolytic transfusion reactions, and transfusion-related acute lung injury

Table 38.5 Human neutrophil antigen (HNA) system

Antigen system	Antigen	Acronym
HNA-1	HNA-1a	NA1
	HNA-1b	NA2
	HNA-1c	SH
HNA-2	HNA-2	NB1
HNA-3	HNA-3a	5b
HNA-4	HNA-4a	MART
HNA-5	HNA-5a	OND

38.2.4 Blood Donor Screening for Infectious Diseases

- The risk of transmitting infection to transfusion recipients has been drastically reduced due to improved donor selection (exclusion of individuals with high-risk behaviors) and screening tests for antibodies developed as a result of infections and viral nucleic acids
- Because viremia precedes seroconversion (development of antibodies) by several days to weeks, nucleic acid amplification testing (NAT) to detect viral nucleic acids is more sensitive than antibody tests in early phase of infection and reduces the window of detection by as much as 60 days for hepatitis C virus and 11 days for human immunodeficiency virus (HIV) infection
- NAT screening of volunteer donor blood for HIV, hepatitis C virus, and West Nile virus is currently performed by commercially available assays

38.3 Posttransplant Chimerism Study

- Frequent monitoring of posttransplant hematopoietic chimerism is important to assess the successful engraftment of hematopoietic stem cells
- Complete hematopoietic reconstitution of the donor origin following an allogeneic stem cell transplant is referred to as complete chimerism. Coexistence of donor and recipient blood cells is called mixed chimerism. In the case of graft failure, there will be no donor origin hematopoiesis. Malignancy relapse is often heralded by the progressive reappearance of recipient's blood cells
- Molecular methods of quantification of donor's and recipient's nucleated blood cells after allogeneic transplantation are analyses of PCR-amplified polymorphic DNA markers, such as variable number of tandem repeats (VNTR) or short tandem repeats (STR)
- In a sex-mismatched transplant, fluorescent in situ hybridization (FISH) with probes specific for chromosomes X and Y can be applied
- If HLA-mismatched transplant, HLA markers can be used, but less frequently
- The specimens can be peripheral blood or bone marrow. Peripheral blood cells are generally more useful than bone marrow cells for chimerism analysis

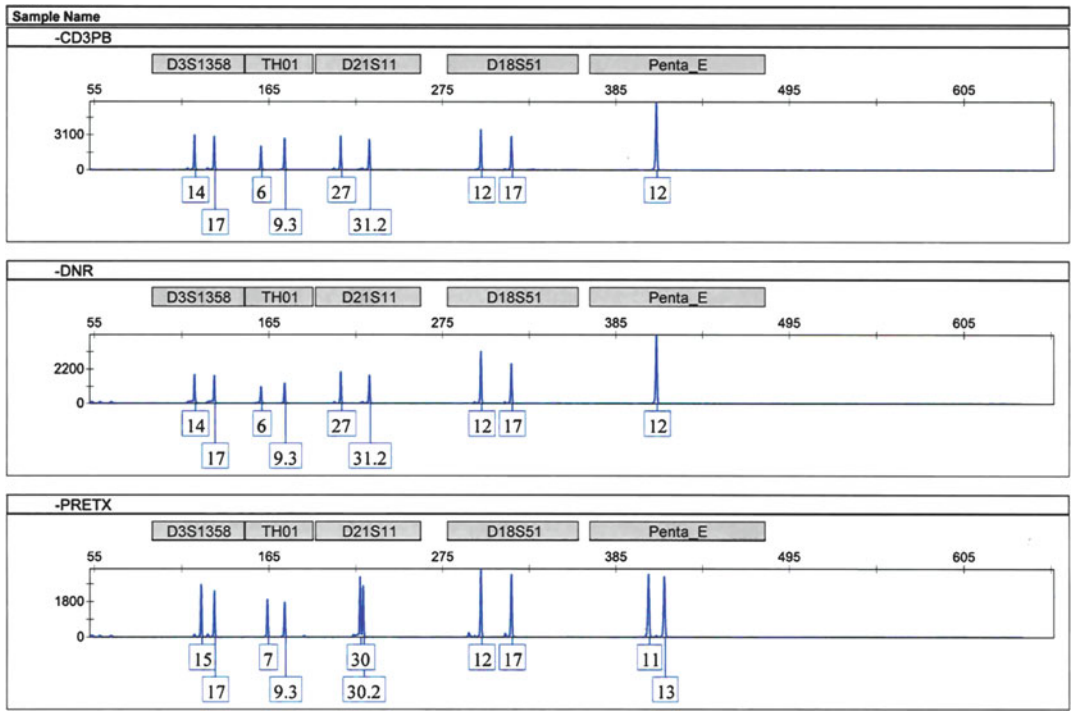


Fig. 38.6 Chromatogram of a set of five STR loci for engraftment study. The top panel is from peripheral blood CD3+ cells from a patient on day 28 posttransplant. The middle panel shows the alleles from the donor and the bottom panel from the recipient's pretransplant sample. Each peak is labeled with the name of allele. All of the four informative loci, except D18S51, show 100% donor-derived cells in the posttransplant sample (complete chimerism) (Case study provided by Anajane Smith and Chris McFarland, Clinical Immunogenetics Laboratory, Seattle Cancer Care Alliance, Seattle, WA)

- Genomic DNA is isolated from the donor and separate samples of the recipient collected before and at various intervals after transplantation to monitor the chimerism status (Fig. 38.6)

38.3.1 Quantification of Chimerism

- The marker loci are PCR-amplified with fluorescent primers followed by automated detection of fluorescently labeled PCR products separated by electrophoresis. For example, ALFexpress DNA sequencer (Pharmacia) or ABI 310 Genetic Analyzer (PE)
- The electrophoresis data are analyzed by software to calculate the amount of recipient's and donor's DNA from the informative markers distinguishing the two
- Quantification of donor's DNA is calculated using the following formula: percent of donor's DNA = (D1 + D2)/(D1 + D2 + R1 + R2) × 100, where D1 and D2 are peak areas of donor's alleles and R1 and R2 are peak areas of recipient's alleles. Only informative markers are used for the analysis. If donor and recipient are heterozygous but share one allele, only the nonshared alleles are considered for the calculation. Quantification of donor DNA is calculated for each informative STR locus, and the mean or median of all informative STR loci is reported as percent of donor's DNA (Fig. 38.7)
- The sensitivity depends on the size of alleles; the detection level is usually around 5% of patient's cells
- Although when using chimerism analysis, one cannot determine whether or not the population

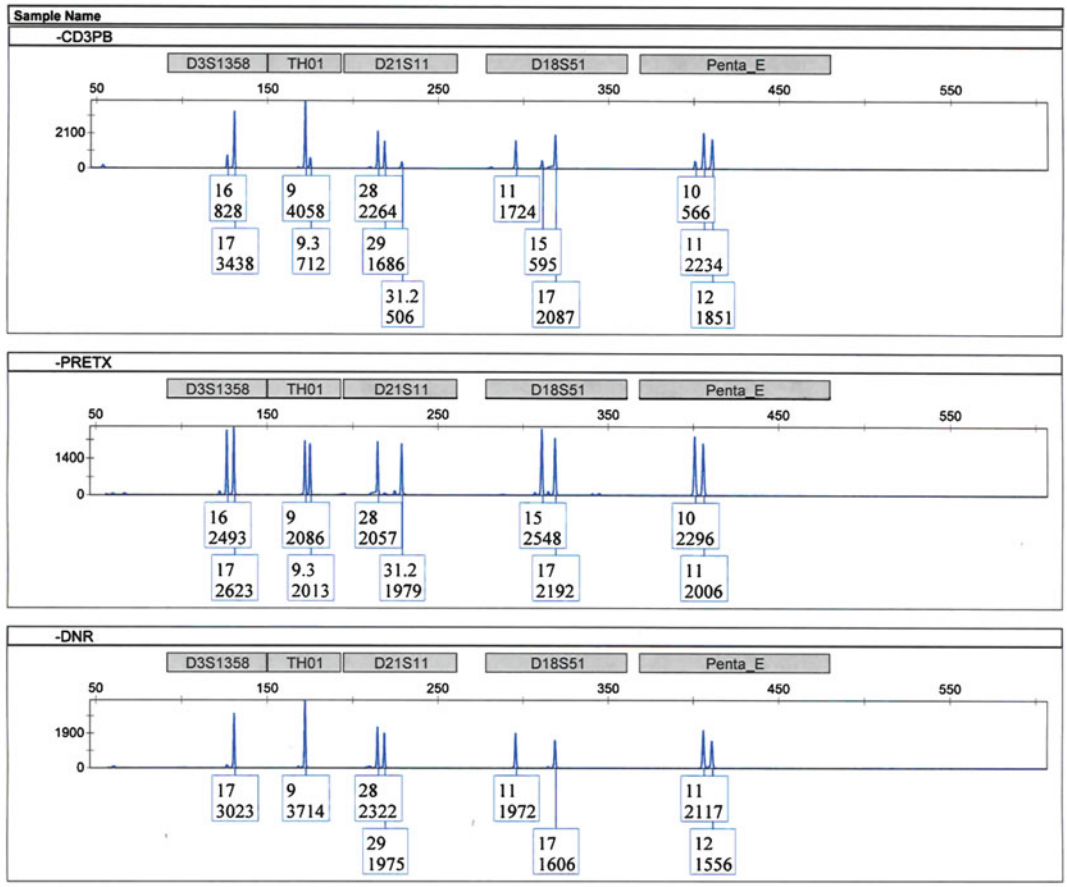


Fig. 38.7 Chromatogram showing posttransplant mixed chimerism. The top panel is from peripheral blood CD3+ cells from a patient approximately 1 year posttransplant. The middle panel is from the recipient's pretransplant sample and the bottom from the donor. Each peak is shown with the name of allele and its relative fluorescent unit value. Three STR loci (D21S11, D18S51, and Penta E) are informative to show mixed chimerism. The D21S11 locus shows 77% of donor origin (1,686/[1,686 + 506] × 100) in the posttransplant sample. Similarly, the D18S51 and Penta E show donor origin of 74% and 77%, respectively (Case study provided by Anajane Smith and Chris McFarland, Clinical Immunogenetics Laboratory, Seattle Cancer Care Alliance, Seattle, WA)

of recipient's nucleated cells contains leukemic cells; samples taken at various intervals can show if the expansion rate of the particular population is consistent with hematologic and clinical manifestations of the disease

Further Reading

Antin JH, Childs R, Filipovich AH, et al. Establishment of complete and mixed donor chimerism after allogeneic lymphohematopoietic transplantation: recommendations from a workshop at the 2001 Tandem meetings. Biol Blood Marrow Transplant. 2001;7:473–85.

Beck S, Trowsdale J. The human major histocompatibility complex: lessons from the DNA sequence. Annu Rev Genomics Hum Genet. 2000;1:117–37.

Bjorkman PJ, Parham P. Structure, function and diversity of class I major histocompatibility complex molecules. Annu Rev Biochem. 1990;59:253–88.

Bray RA, Hurley CK, Kamani NR, et al. National Marrow Donor Program HLA matching guidelines for unrelated adult donor hematopoietic cell transplants. Biol Blood Marrow Transplant. 2008;14:45–53.

Bux J. Human neutrophil alloantigens. Vox Sang. 2008;94:277–85.

Daniels GL, Fletcher A, Garratty G, et al. Blood group terminology 2004: from the International Society of Blood Transfusion committee on terminology for red cell surface antigens. Vox Sang. 2004;87:304–16.

Dickinson AM, Charron D. Non-HLA immunogenetics in hematopoietic stem cell transplantation. Curr Opin Immunol. 2005;17:517–25.

Feder JN, Gnirke A, Thomas W, et al. A novel MHC class I-like gene is mutated in patients with hereditary haemochromatosis. Nat Genet. 1996;13:399–408.

Goulmy E. Human minor histocompatibility antigens. Curr Opin Immunol. 1996;8:75–81.

Kristt D, Israeli M, Narinski R, et al. Hematopoietic chimerism monitoring based on STRs: quantitative platform performance on sequential samples. J Biomol Tech. 2005;16:380–91.

Mallal S, Phillips E, Carosi G, et al. HLA-B*5701 screening for hypersensitivity to abacavir. N Engl J Med. 2008;358:568–79.

Marsh SGE, Albert ED, Bodmer WF, et al. Nomenclature for factors of the HLA system, 2010. Tissue Antigens. 2010;75:291–455.

McCann SR, Crampe M, Molloy K, et al. Hemapoietic chimerism following stem cell transplantation. Transfus Apher Sci. 2005;32:55–61.

Metcalfe P, Watkins NA, Ouwehand WH, et al. Nomenclature of human platelet antigens. Vox Sang. 2003;85:240–5.

Scharf SJ, Smith AG, Hansen JA, et al. Quantitative determination of bone marrow transplant engraftment using fluorescent polymerase chain reaction primers for human identity markers. Blood. 1995;85:1954–63.

Shiina T, Hosomichi K, Inoko H, et al. The HLA genomic loci map: expression, interaction, diversity and disease. J Hum Genet. 2009;54:15–39.

Smith AG, Martin PJ. Analysis of amplified variable number tandem repeat loci for evaluation of engraftment after hematopoietic stem cell transplantation. Rev Immunogenet. 1999;1:255–64.

Clinical Pharmacogenomics

39

Catalina López-Correa and Lawrence M. Gelbert

Contents

39.1 Introduction 1010

39.2 Pharmacogenetic/Pharmacogenomic Patient Stratification 1010

39.3 The Evolution of Pharmacogenetic and Pharmacogenomic Tests 1011

39.4 Glossary 1012

39.5 Pharmacokinetics Versus Pharmacodynamics 1013

39.6 Genetic Variants 1013
39.6.1 Single-Nucleotide Polymorphisms 1013
39.6.2 Microsatellites 1014
39.6.3 Variable Number of Tandem Repeats 1014
39.6.4 Copy-Number Variants 1014
39.6.5 Next-Generation Sequencing 1014

39.7 Genetic Markers for Association and Linkage Studies 1015

39.8 Pharmacogenomic Association Studies 1015
39.8.1 Candidate Gene Search 1015
39.8.2 Genome-wide Association Studies 1015
39.8.3 Candidate Pathway Approach 1016

39.9 Genetic Testing 1016

39.10 Regulatory Issues in Pharmacogenomics 1017

39.11 Pharmacogenetics: Drug-Metabolizing Enzymes 1017

39.12 Warfarin (Coumadin): An Example of Drug-Metabolizing Enzymes and Target Variation Affecting Safety and Efficacy (Therapeutic Index) 1019

39.13 Pharmacogenomics of Asthma Treatment 1021
39.13.1 B_2-Agonists 1022
39.13.2 Leukotrienes 1022

39.14 Pharmacogenetics of Cardiovascular Disorders 1023
39.14.1 Hypertension 1023
39.14.2 Angiotensin-Converting Enzyme 1024
39.14.3 Angiotensinogen Gene 1024
39.14.4 β-Blockers 1024
39.14.5 Pharmacogenomics of Lipid-Lowering Agents 1024
39.14.6 Emerging Trends in Cardiovascular Pharmacogenetics 1025

39.15 Pharmacogenomics of Neuropsychiatric Diseases 1025
39.15.1 Alzheimer Disease 1025
39.15.2 Parkinson Disease 1026
39.15.3 Major Depression 1026
39.15.4 Schizophrenia 1026

39.16 Cancer Pharmacogenetics 1027
39.16.1 Monogenic Cancer Pharmacogenetics: Thiopurine S-Methyltransferase 1027
39.16.2 Multigene Cancer Pharmacogenomics: Microarray Expression Profiling 1028

39.17 Web Resources 1029

Further Reading 1030

C. López-Correa, MD, PhD
Affaires scientifiques/CSO, Génome Québec, Montréal (Québec)

L.M. Gelbert, PhD (✉)
Herman B Wells Center for Pediatric Research, Indiana University School of Medicine, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_39,

39.1 Introduction

- -OME is from the Latin *-oma* meaning "mass." A genome is defined as the complete DNA sequence of an organism, and genomics (also known as functional genomics) is the comprehensive sanalysis of genomic structure and function. The development of genomics was greatly facilitated by the Human Genome Project and the publication of the draft sequence of the human genome in 2001. This and the development of several enabling technologies such as automated high-throughput DNA sequencing and genotyping, microarray technology, and bioinformatics now allow for complete views of human biological systems including drug activity. These developments have shifted research from gene discovery to the association of specific genes and genetic variants with disease susceptibility, response to therapy, and other phenotypes of importance in healthcare
- The role of genetic variability in response to drug therapy was first described in the 1950s. Pharmacogenetics is defined as the analysis of genetic factors influencing response to drug treatment. The term pharmacogenomics (PGx) is used to describe the expanded analysis of multiple genetic factors affecting drug efficacy and toxicity and disease susceptibility. Although pharmacogenetics and PGx are commonly used interchangeably, more recently, pharmacogenetics has been used to define the genetic factors specifically associated with drug-metabolizing enzymes (DMEs). The goal of both these fields is to identify genetic biomarkers to predict disease susceptibility/progression (disease biomarker) and response to therapy (drug activity biomarkers)
- In this chapter, we will provide a historical background to the development of pharmacogenetics/PGx and emerging trends and examples of the current state of PGx in several disease areas

39.2 Pharmacogenetic/ Pharmacogenomic Patient Stratification

- Drug therapy for many diseases has significantly reduced mortality, improved quality of life, and has a significant positive economic impact. However, because of genetic and environmental heterogeneity (e.g., diet and exercise), the response rate for most commonly prescribed drugs ranges from 25% for cancer treatments to 80% for analgesic COX2 inhibitors. Similarly, environmental and genetic factors impact drug safety, contributing to approximately 6.5% of the patient population having an adverse drug reaction (ADR) resulting in >100,000 deaths annually
- Figure 39.1 summarizes the use of PGx to stratify patient populations. Patients are first diagnosed with a disease (such as hypertension, Alzheimer, cancer, and asthma) for which there is a standard drug therapy. The goal of PGx is to enrich patient populations for responders (PGx efficacy tests) or to identify patients genetically predisposed for an ADR (PGx safety test)
- Complex diseases result from multiple genetic and environmental factors whose individual effects are small and overlap. Therefore, pharmacogenomic tests are not absolutely predictive, but allow for the enrichment of patient populations to improve safety and/or efficacy
- Factors driving the adoption of PGx include
  - Drug safety: the recent recall of several marketed drugs has increased safety concerns and regulatory expectations for new and existing drugs. Pharmacogenomic safety assays are viewed as a way to address these safety concerns
  - Drug cost: with the cost of healthcare rising, there is increasing pressure to contain costs, including those of prescription medicines. This is especially true for newer targeted cancer therapies and biologics, which tend to be more expensive.

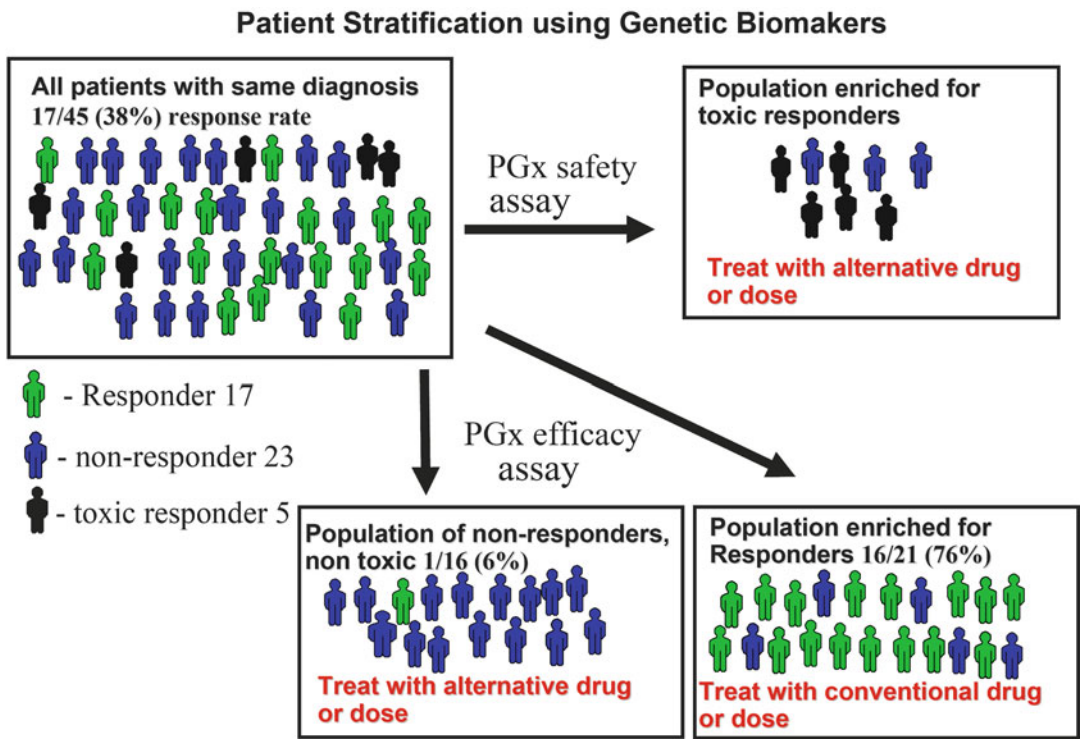


Fig. 39.1 Strategy to enrich patient populations to improve efficacy and reduce adverse drug responses using pharmacogenomic biomarkers. Safety assays include those measuring drug metabolism or on-target toxicities. Efficacy assays identify those patients who will respond to a treatment. Efficacy markers also include drug activity biomarkers that are used to determine the biologically optimum dose in clinical drug development (Adapted and reprinted, with permission, from the Ann Rev Gen Human Genet, Volume 2 © 2001 by Annual Reviews)

Targeting patients who respond is also of importance in developing countries, which have fewer financial resources for healthcare and will rely on genomic tests to maximize the benefit of prescription medicines and other forms of disease management

– Decreasing cost of drug discovery and development: the cost of developing a new drug now exceeds 800 million USD. The high cost of developing drugs results in great part from the attrition observed in clinical drug development. The overall success rate in clinical development is 11%, and reducing the cost of clinical drug development is recognized as a primary factor in controlling the cost of new drugs. The early development of pharmacogenomic markers will improve patient selection for clinical trials, reducing their size and complexity and allowing for a focus on efficacy testing, thus resulting in reduced attrition and lower costs

39.3 The Evolution of Pharmacogenetic and Pharmacogenomic Tests

- Historically, pharmacogenetics has been used to describe single-gene mutations or variants assayed with low-content/low-throughput techniques such as DNA restriction fragment

length polymorphisms (RFLPs), DNA sequencing, or differential electrophoresis methods such as single-stranded conformation polymorphism analysis. Such single-gene/low-content pharmacogenetic assays have been successful for diagnosing diseases or predicting drug response wherein a single gene with a large effect contributes to the phenotype (such as autosomal recessive familial disorders or genes regulating a rate-limiting step in drug metabolism). In the common complex diseases, multiple genes contribute to the phenotype, each with a small effect. To identify these genetic factors, pharmacogenomic approaches now include microarrays, high-throughput automated DNA sequencing and genotyping, informatics, and mass spectrometry. These new approaches are both more efficient and sensitive allowing for multiplexed analysis of mutations/variants in several genes and thus are more predictive for complex diseases wherein multiple genetic factors are involved

- Introduction of drugs with a defined molecular target (targeted agents) has led to the development of companion diagnostic tests. The FDA defines a companion diagnostic as an "in vitro diagnostic device that provides information that is essential for the safe and effective use of a corresponding therapeutic product (www.fda.gov)," and it is becoming increasingly common for drugs and companion diagnostics to be simultaneously approved. Recent examples include the BRAF inhibitor vemurafenib (Zelboraf®, Plexxikon/Roche), for patients with metastatic melanoma containing the $BRAF^{V600E}$ mutation, and the anaplastic lymphoma kinase (ALK) inhibitor crizotinib (Xalkori®, Pfizer), for the treatment of ALK-positive non-small cell lung cancer
- Figure 39.2 summarizes the pharmacogenomic biomarker discovery and development process. The process has several steps
  - Whole genome experiments (microarray gene expression or whole genome single-nucleotide polymorphism [SNP] analysis) drive the discovery phase when it is necessary to analyze a large number of genes in a small to moderate number of samples (patient specimens or in preclinical models)
  - Assay development and validation involves development of specific assays to measure a subset of genetic changes that are predictive of a specific clinical phenotype. This can include measurement of mRNA changes or genotyping using microarrays or by quantitative real-time polymerase chain reaction (PI-PCR) (TaqMan®, Applied Biosystems, Foster City, CA) and also includes approaches that measure proteins (e.g., Western blot, ELISA, and immunohistochemistry). Assay validation should be performed in both the original sample set as well as an expanded set of samples
  - Biomarker evaluation tests that the assay is robust enough to work in the appropriate biologic sample (fluid and tissue) under a variety of conditions observed in clinical applications (sample shipping, temperature variation, and so on)

39.4 Glossary

- *Haplotype*: a set of closely linked alleles (genes or DNA polymorphisms) inherited as a unit along a chromosome
- *Variable number of tandem repeat (VNTR) locus*: a region of DNA that is hypervariable because of tandemly repeated DNA sequences. Presumably variability is generated by unequal crossing over or slippage during replication. The repetitive sequence is present in different numbers in different individuals of a population or in the two different chromosome homologues in one diploid individual
- *SNP*: an SNP is a specific position in a stretch of DNA wherein there is a single-nucleotide substitution. Each alternate nucleotide is called an allele
- *Pharmacokinetic (PK)*: biologic properties related to altered drug uptake, distribution, metabolism, or excretion of the agent administered

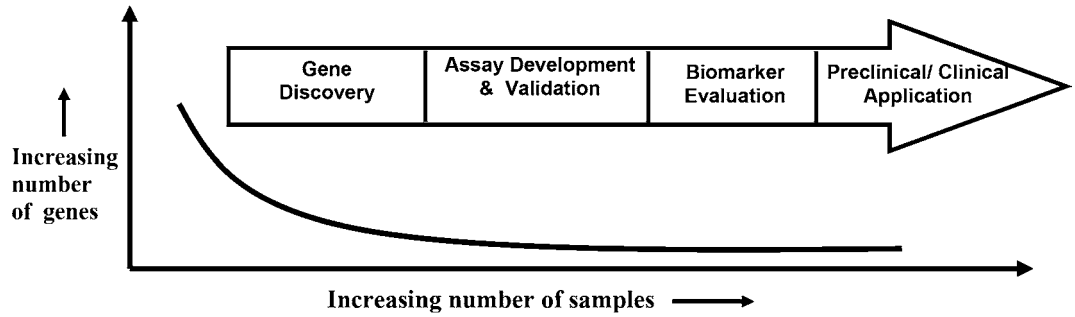


Fig. 39.2 Development flow scheme for pharmacogenomic assays

- *Pharmacodynamic (PD)*: biologic properties related to drug target modulation or in its pathway, leading to altered drug efficacy
- *Linkage disequilibrium (LD)*: LD is often termed "allelic association." When alleles at two distinctive loci occur in gametes more frequently than expected given the known allele frequencies and recombination fraction between the two loci, the alleles are said to be in LD. Evidence for LD can be helpful in mapping disease genes because it suggests that the two may be very close to one another

39.5 Pharmacokinetics Versus Pharmacodynamics

- Genetic factors can affect PK and PD activities. Polymorphisms in drug-metabolizing enzymes (DMEs) and transporters have their primary effect on drug PK properties, whereas those in a drug target or the target pathway primarily influence PD activities. Although these are separate biologic activities, it is ultimately the combination of PK/PD properties that is responsible for a drug to be both efficacious and safe
- The first pharmacogenetic studies were focused on PK changes and safety (polymorphisms in the metabolizing enzymes and transporters), but this has been shifting toward an increasing focus on PD variability (polymorphisms in the drug target and related pathways)

39.6 Genetic Variants

- Genetic variation in the human genome takes many forms, ranging from large, cytogenetic rearrangements (including chromosomal deletions, inversions, and duplications) to single-nucleotide changes (mutations, SNPs)
- Genetic variation can occur in the coding region of genes affecting the activity of the resulting protein or in regulatory regions of drug targets that can affect transcription and RNA splicing or stability, thereby increasing or decreasing the amount of protein present in a tissue

39.6.1 Single-Nucleotide Polymorphisms

- SNPs are the most abundant form of polymorphism in the human genome and have several advantages over previous genetic markers. The large number of SNPs and distribution across the genome provides a higher level of resolution for genetic studies than was previously possible; their power can be further enhanced by analyzing clusters of closely spaced SNPs called haplotypes, and because they are biallelic, their genotyping and analysis can be automated
  - *Synonymous SNP*: generates a codon that encodes the same amino acid
  - *Nonsynonymous SNP*: generates an amino acid change
  - *Nonsense SNP*: generates a premature stop codon

39.6.2 Microsatellites

- Microsatellites are also called simple sequence repeats or short tandem repeats. The usual repeat size is 1–5 bp
- In contrast to SNPs, microsatellites can have multiple alleles and are hypervariable
- Shorter overall length than VNTRs and are more amenable to PCR-based analysis
- Microsatellites have been largely used for family-based linkage studies

39.6.3 Variable Number of Tandem Repeats

- In humans, different types of repetitive sequences account for approximately 50% of the genome. Subsets of these are the tandem repeats, which are consecutively perfect or slightly imperfect repeats of DNA motifs
- VNTRs are defined as tandem repeats with a unit size of 6 bp or longer. Although extensively used as genetic markers in genetic linkage and forensic studies, only a handful of VNTRs have been reported to have functional consequences in gene expression or function
- Genes containing functional VNTRs that affect gene expression and the encoded protein product include insulin-like growth factor 2, the dopamine transporter solute carrier family 6, member 3, the dopamine transporter 1, dopamine receptor D4, and arachidonate 5-lipoxygenase (*ALOX5*)
- Technologies used for VNTR genotyping are
  - Southern blotting
  - DNA sequencing
  - PCR fragment sizing (gel electrophoresis)
  - Denaturing high-performance liquid chromatography

39.6.4 Copy-Number Variants

- Recent studies have discovered an abundance of submicroscopic copy-number variation of DNA segments ranging from kilobases (kb) to megabases (Mb) in size
- Deletions, insertions, duplications, and complex multisite variants, collectively termed copy-number variants (CNVs) or copy-number polymorphisms, are found in healthy and unrelated individuals
- Most of these CNV sequences may be functionally significant, but this has yet to be fully ascertained
- Different comparative genomic hybridization methods have been used to study copy-number variations
  - Comparative genomic hybridization (CGH)
  - SNP genotyping microarrays
  - Fiber fluorescent in situ hybridization (Fiber FISH)
- The chemokine (C–C motif) ligand 3-like 1 (*CCL3L1*) gene is an example of a CNV associated with a particular phenotype. Low *CCL3L1* copy number was found to be associated with enhanced susceptibility to HIV infection and AIDS progression

39.6.5 Next-Generation Sequencing

- In the decade since the completion of the first human genome sequence, technical improvements in DNA sequencing technology, known as next-generation sequencing or NGS, have dramatically reduced cost and increased throughput. This technology makes it possible to sequence an individual's entire genome or the genome of a patient's tumor, allowing for the identification of all DNA variants that may affect a patient's disease or response to treatment. NGS is expected to lead to new discoveries in pharmacogenomics and cancer biology
- NGS has also become a powerful new approach for identifying genes that underlie Mendelian disorders for which conventional approaches have failed. Examples of diseases where new mutations have been discovered using the new sequencing technologies are Charcot–Marie–Tooth disease, Bartter syndrome, and X-linked lymphoproliferative syndrome type 2 (XLP2)

Table 39.1 Features of SNP and microsatellite genetic markers

SNPs	Microsatellites
Biallelic markers	Multiallelic markers
Genetically more stable	More unstable markers
Lower mutation rate ($\sim 10^{-9}$)	Higher mutation rate ($\sim 10^{-3}$)
Low degree of heterozygosity (<20%)	High degree of heterozygosity (~70%)
Shorter range of LD (<20 kb)	Larger range of LD (~100 kb)
Need to analyze millions of SNPs to perform GWA studies	Few thousand for GWA
Capture more ancient events	Capture more recent events
More rapid and highly automated genotyping, lower overall cost and sample requirements	Expensive and time-consuming genotyping and analysis, requires more DNA to analyze
Generally used for case and controls association studies	Generally used for familial-based linkage studies

39.7 Genetic Markers for Association and Linkage Studies

- Until recently, microsatellites have been the predominant marker for family-based linkage analyses. They are abundant, well dispersed throughout the genome, and highly polymorphic/highly informative. The increased availability of SNPs now provides an opportunity for alternative automated approaches. SNPs are more abundant than microsatellites and are also well dispersed throughout the genome, but they are less informative than microsatellites (because they are biallelic). To overcome this last feature, a considerably larger number of SNP are required to achieve an information content equal to a small number of microsatellites
- Each of those markers has advantages and disadvantages that should be carefully reviewed when designing genetic studies (Table 39.1)

39.8 Pharmacogenomic Association Studies

- Genetic factors other than those associated with known drug metabolism and molecular drug targets affect drug safety and efficacy. Genetic association studies are used to identify such genetic factors

39.8.1 Candidate Gene Search

- This approach directly tests the association of selected genes with particular phenotypes, drug response, or ADRs
- The main source of candidate genes is through hypothesis generation and literature surveys. This can include analysis of cDNA and expressed sequence tags databases that help identify new candidate genes by using sequences that are similar to genes known to be associated with a drug's activity. Whole genome expression profiling using microarrays is now routinely used to identify novel candidates that are differentially expressed in responder/nonresponder cell lines or individuals
- Genetic variants (e.g., SNPs and microsatellites) in candidate genes are analyzed in populations of patients treated with the drug of interest and tested for statistical association with drug response
- This approach tests only known genes and therefore misses genes for which no previous involvement in drug response or particular phenotypes is unknown

39.8.2 Genome-wide Association Studies

- The candidate gene approach has considerable intuitive appeal, but is limited to the analysis of genes with known function, and excludes those whose function is not yet been determined

- Increasingly robust and sensitive techniques that have been developed for the analysis of large sets of SNPs, spanning the entire genome, have now made genome-wide associated studies (GWAS) feasible
- Genome-wide analysis of random SNPs is based on LD mapping or statistical association between SNPs in proximity to each other. The number of SNPs used for a GWAS is determined by patterns and strength of LD in a given population. GWA studies rely on the assumption that LD enables one SNP to act as a marker for association to other sequence variants in that region
- Large sets of affected individuals (cases) and controls are needed to perform these studies. The typical outcome is presented as one or more SNPs associated with the cases (or with drug responders) but not with the controls. Those SNPs can be used as markers to predict a certain phenotype. These studies are in general laborious and costly but are increasingly being used as they are more relevant for the analysis of complex diseases

39.8.3 Candidate Pathway Approach

- An appealing intermediate between the candidate gene and the genome-wide approach is the analysis of well-characterized biologic networks/pathways including dozens or hundreds of interacting genes. Genetic markers located in those genes are used to test the large sets of patients and controls in order to identify genes that might be associated with particular disorders or phenotypes
- Increasingly sophisticated pathway analysis software and computational tools are helping to analyze how the behavior of a complex biologic system changes in response to the variation of an individual component of that network
- As for other areas of genomics, pathway and biologic network analysis is expected to positively impact PGx through the association of novel genes and their function with drug response

39.9 Genetic Testing

- An increasing number of genetic association studies are performed to discover new genetic variants that could be translated into tests of drug response and/or disease susceptibility
- Confirming initial pharmacogenomic associations is a challenging endeavor. Ideally different researchers in different settings and working with different populations should obtain similar results in order to validate the association of a genetic variant with a certain phenotype
- A multitude of pharmacogenomic testing services are available directly from diagnostic labs, but most of them remain at the exploratory level (sometimes called home brew assays)
- Only few pharmacogenomic associations have been extensively validated to date. Well known and often cited examples include
  - Thiopurine S-methyltransferase (TPMT) variants and toxicity to the anticancer drug 6-mercaptopurine (6-MP) in children with acute lymphoblastic leukemia
  - *CYP2C9* and VKORC1 variants and dose requirements of the blood thinner drug warfarin in patients with clotting disorders
  - UGT1A1 variants and toxicity to the anticolon cancer drug irinotecan
- Genetic tests can evaluate different types of genetic changes
  - Heritable changes (constitutional) such as the genetic variants in the cytochrome P450 DMEs
  - Nonheritable changes (somatic) such as the genetic variants observed in cancer tissues and detected by gene expression profiling or gene amplification analysis
- Currently there are only a handful of marketed pharmacogenomic test kits
  - Herceptin® (trastuzumab therapy) for breast cancer treatment
  - AmpliChip® (Roche Molecular Systems Inc., Branchburg, NJ) test for two of the cytochrome P450 DMEs (*CYP2D6* and *CYP2C19*)

  - Oncotype Dx™ (Genomic Health Inc., Redwood City, CA) and MammaPrint Dx® (Agendia, Amsterdam, the Netherlands) gene expression assays in adjuvant breast cancer chemotherapy

39.10 Regulatory Issues in Pharmacogenomics

- As part of the Food and Drug Administration's (FDA) strategic plan, the FDA is developing standards to apply emerging technologies (e.g., PGx and other biomarkers) to provide effective translation of new scientific discoveries into safe and effective medical products
- The FDA published initial guidelines for genomic data submission in 2004. This guidance was intended to encourage voluntary genomic data submission by sponsors using PGx in exploratory research during drug development
- The FDA is encouraging the application of pharmacogenomic approaches and data into the evaluation of patient variability during clinical drug development. This data can now be submitted to the FDA

39.11 Pharmacogenetics: Drug-Metabolizing Enzymes

- DME is a term used to define genes whose activities are involved in the metabolic activation and/or inactivation of xenobiotics (environmental chemicals) including drugs and were the initial focus of clinical pharmacogenetics. DMEs are functionally divided into phases I and II DMEs. Phase I DMEs convert xenobiotics to reactive intermediates through oxidation, reduction, or hydrolysis reactions, whereas phase II DMEs conjugate the reactive intermediate to a small functional group, which makes the compound more easily excreted (Fig. 39.3a)
- The impact of genetic factors on drug response was first noted by Arno Moltulsky in a 1957 report. In the 1970s, Robert L Smith and colleagues studied the metabolism and PKs of debrisoquine (an antihypertensive drug) by administering the drug to themselves and monitoring its metabolism. In response to a standard dose, Smith soon became dizzy and hypotensive for 2 days. Drug metabolite analysis showed that Smith eliminated debrisoquine almost completely as the parent compound (poor metabolizer), rather than the 4-hydroxy metabolite observed in individuals who did not exhibit any adverse effects. Subsequent studies in larger populations showed that approximately 6–10% of Caucasians were poor metabolizers of debrisoquine, which was later shown to result from genetic variants in the *CYP2D6* gene. Genotype–phenotype studies of *CYP2D6* variants are now routinely performed using debrisoquine or dextromethorphan as surrogate substrates
- The large number of DME genes and variants in the human genome are believed to result from the selective advantage of these variants provided during the coevolution of plants and animals. The molecular mechanisms for the functional variation of DME genes include
  - Splice site mutations resulting in exon skipping (e.g., *DPD, CYP2C19*, and *CYP3A5*)
  - Microsatellite repeats (e.g., *CYP2D6*)
  - Gene duplications (e.g., *CYP2D6*)
  - Point mutations resulting in early stop codons (e.g., *CYP2D6*)
  - Enhanced proteinolysis (e.g., TPMT)
  - Altered promoter functions (e.g., *CYP2A5* and *UGTIA1*)
  - SNPs causing amino acid substitutions (e.g., *NAT2, CYP2D6, CYP2C9*, and *CYP2C19*)
  - Large gene deletions (e.g., *GSTT1, GSTM1*, and *CYP2D6*)
- There are over 70 alleles of *CYP2D6*, many of which alter the enzymatic activity of the encoded protein. The level of enzymatic activity varies from no activity (null), poor, intermediate (average), extensive metabolizers, and ultrametabolizers. *CYP2D6* is responsible for the metabolism of approximately 25% of all drugs (Table 39.2), and genetic variation in this gene is responsible for a significant amount of ADRs or a lack of efficacy

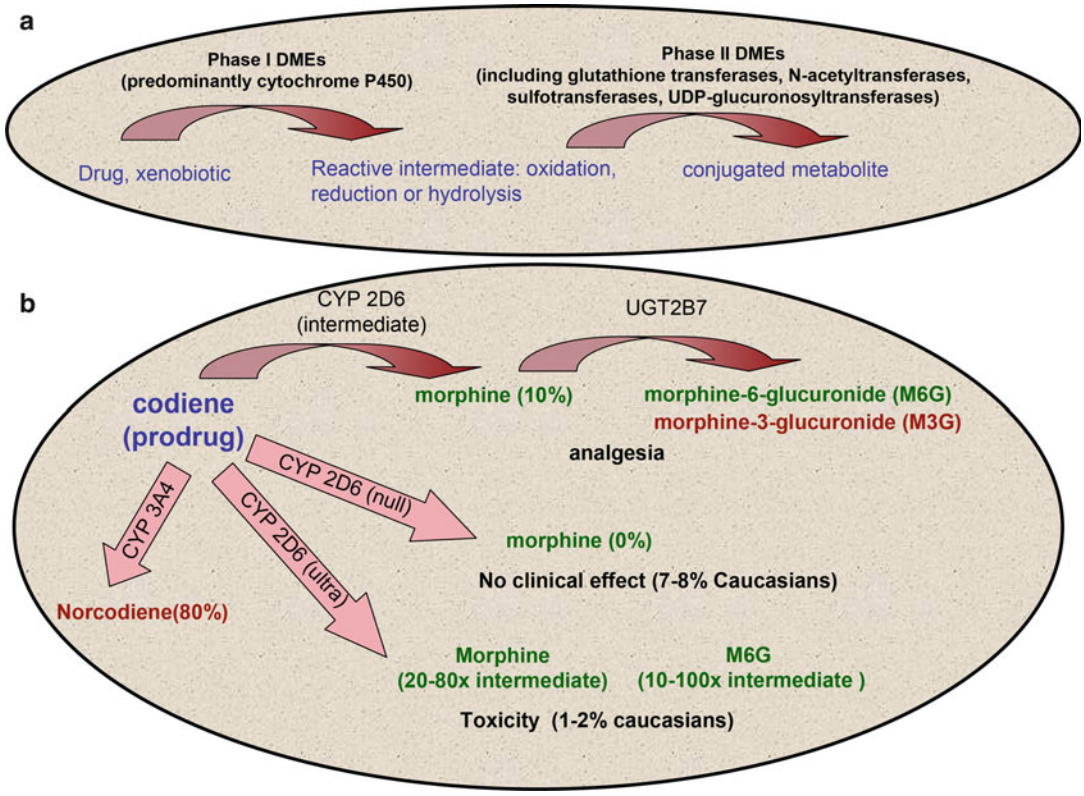


Fig. 39.3 Pharmacogenetics of drug metabolism. (**a**) Overview of phases I and II biotransformation of drugs. Phase I DMEs generate a polar functional group through oxidation, reduction, or hydrolysis that may sometimes be reactive. In drug metabolism, these newly revealed functional groups constitute an intermediate metabolite, which is subsequently conjugated by a phase II enzyme to a polar endogenous compound like a sugar that facilitates excretion. The balance of phases I and II activity is important to prevent the overproduction of intermediates and end products that cause oxidative stress, may be toxins or carcinogens. Naturally occurring variants in the genes encoding DMEs result in imbalance of the system, as do environmental factors (drug–drug and drug–diet interactions). (**b**) Example of drug metabolism and impact of pharmacogenetics. Codeine is a prodrug that possesses little analgesic activity. Green text indicates metabolites with analgesic activities and red inactive metabolites. The majority of codeine is biotransformed by *CYP3A4* to norcodeine. Biotransformation by *CYP2D6* to morphine and subsequently by the phase II enzyme glucuronosyltransferase UGT2B7 to M6G and morphine-3-glucuronide (M3G). Morphine and M6G have analgesic activity, whereas norcodeine and M3G are inactive. Genetic variation in *CYP2D6*, UGT2B7, and *CYP3A4* results in different ratios of the active metabolites of codeine, resulting in different clinical outcomes including no analgesia (*CYP2D6* null) and codeine toxicity (*CYP2D6* ultrametabolizers) (Relative levels of morphine and M6G estimated from data reported by Gasche et al. 2004)

- Although *CYP2D6* poor metabolizers are responsible for most of the variability in drug efficacy and ADRs, variants that increase enzymatic activity and result in an ultrarapid metabolism phenotype are also clinically important. Codeine is an example of a drug wherein clinical efficacy and safety are affected by rapid drug metabolism. Codeine is a prodrug that is metabolically activated to morphine and morphine-6-glucuronide (M6G) by the *CYP2D6* and UGT2B7. Codeine clearance is modulated by another phase 1 *DME*, *CYP3A4*. Genetic variation in *CYP2D6* has been shown to be associated with lack of efficacy and toxicity (Fig. 39.3b)
- The frequency of phenotypic variation in DMEs varies in different populations, as shown for *CYP2D6* and *CYP2C19* poor

Table 39.2 Common drugs that are metabolized by *CYP2D6*, prodrugs activated by *CYP2D6* are indicated by tamoxifen and codeine

Drug	Class
S-metoprolol	β-blocker
Propafenone	β-blocker
Timolol	β-blocker
Dextromethorphan	Antitussive
Paroxetine (Paxil)	Antidepressant
Haloperidol	Antipsychotic
Risperidone	Antipsychotic
Thioridazine	Antipsychotic
Tamoxifen	Antiestrogen – breast cancer prevention
Codeine	Analgesic

Table 39.3 Prevalence of *CYP2D6* and *CYP2C19* poor metabolizers in different ethnic groups

Population	PM phenotype (%) Median
CYP2D6	
African	3.4
Asian	0.5
Caucasian	7.2
CYP2C19	
African (sub-Saharan)	4
Asian	15.7
Caucasian	2.9
Middle East/North Africa	2

metabolizers in Table 39.3. The variation in metabolic frequencies between different ethnic populations and the large number of variants that affect metabolism highlights the essential need for a cost-effective and accurate genotyping platform to make clinical pharmacogenetics a reality

- Initial DME genotyping assays utilized RFLP or PCR technology, but were technologically cumbersome, time consuming, and error prone. Recent advances in microarray technology now allow for the genotyping of a large number of different alleles on a single assay platform. The AmpliChip *CYP450* is a microarray for genotyping clinically relevant *CYP2D6* and *CYP2C19* variants (Figs. 39.4 and 39.5). The use of microarray technology allows for a single standardized platform, reduced cost, and more robust data, which will speed the transition of DME genotyping from the laboratory to the patient
- The extensive knowledge on genetic variation in DMEs has led to rational adjustment of dose or dosing interval and to appropriate warnings or precautions. For example, the labeling of atomoxetine (Strattera, Eli Lilly and Company, Indianapolis, IN), thioridazine (Mellaril, Novartis Pharmaceuticals, Basel, Switzerland), voriconazole (Vfend, Pfizer, New York, NY), and irinotecan (Camptosar, Pfizer, New York, NY, Pharmacia and Upjohn) now contains information about the genetics of metabolizing enzymes

39.12 Warfarin (Coumadin): An Example of Drug-Metabolizing Enzymes and Target Variation Affecting Safety and Efficacy (Therapeutic Index)

- Warfarin is a commonly prescribed anticoagulant drug for the prevention and treatment of venous and arterial thromboembolic disorders. However, warfarin therapy is difficult to manage because of the drug's narrow therapeutic index and the wide interindividual variability in both drug metabolism and efficacy (e.g., resistance)
- Bleeding is a severe side effect of warfarin; both gastrointestinal and cerebral bleeds are of particular concern; therefore, the usual protocol is to start on a low dose and "titrate" to achieve the desired degree of anticoagulation. However, dose titration has been proven ineffective in many cases and carries inherent risk for adverse effects
- *CYP2C9* is associated with warfarin toxicity as a result of drug metabolism. This gene encodes the enzyme responsible for metabolizing more potent S-enantiomer form of warfarin to inactive metabolites
- Since the initial cloning of *CYP2C9*, a large number of allelic variants of the gene have

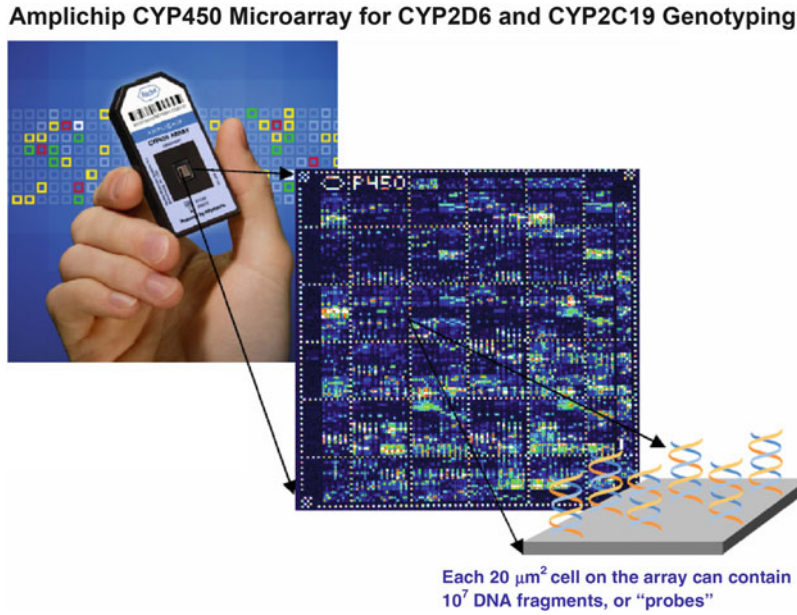


Fig. 39.4 Microarray assay platform for the comprehensive analysis of DME polymorphisms. The Amplichip *CYP450* is an oligonucleotide array that can detect alleles of *CYP2D6* and *CYP2C19* that affect drug metabolism. Each array contains over 15,000 different oligonucleotide probes, approximately 240 for each polymorphism. Genomic DNA is amplified in two multiplexed PCR reactions and hybridized to the array. The hybridization to specific sets of oligonucleotides representing specific alleles is detected by fluorescent staining of a biotin label incorporated into the sample. System software is then used to detect and call the genotype. The microarray format allows for the incorporation of a large number of controls and a standardized assay system which increases the accuracy of the data. In addition to this FDA-approved platform, other microarray assays have been developed that include phase II DMEs and drug transporters in which genetic variants are known to influence a patients response to a drug (DMET microarray from Affymetrix Inc., Santa Clara, CA)

been associated with drug response to warfarin. The two commonest are *CYP2C9*2*, in which cysteine substitutes for arginine at amino acid 144 (R144L), and *CYP2C9*3*, in which leucine substitutes for isoleucine at residue 359 (1359L). Both are well known to result in decreased S-warfarin metabolism. Patients with these variant alleles require significantly lower doses of warfarin than patients with the wild-type gene. The frequency of carriers for one or more of these variants is roughly 30% of the general population

- In 2004, coding-region mutations in vitamin K epoxide reductase complex subunit 1 (VKORC1, warfarin's pharmacologic target) were found to cause a rare syndrome of warfarin resistance. This discovery led to the investigation of the association of other *VKORC1* genetic variants with drug response to warfarin. This is an example of variable drug response related to variation of the target gene
- Common variants in *VKORC1* have now been found to account for a much greater fraction of variability in warfarin response (21%) than do variants in *CYP2C9* (6%). The –1639 G>A polymorphism in the transcriptional promoter of *VKORC1* has a major effect on warfarin dosing as the A allele results in lower expression of the target for encoded protein. The development of multiplexed assays to genotype patients for *CYP2C9* and *VKORC1*

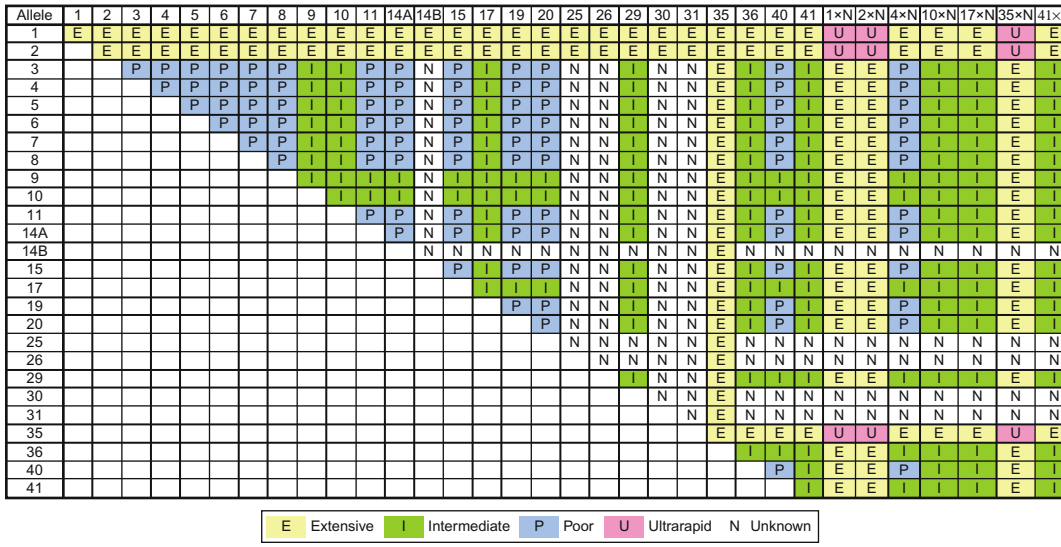


Fig. 39.5 Genotyping results from Amplichip *CYP450* array for *CYP2D6* gene. Allele number indicates change nucleotide sequence of the gene (http://www.cypalleles.ki.se/cyp2d6.htm), and color indicates predicted phenotype. The use of a single assay platform improves the accuracy of genotyping. Full details can be found in the package insert for the product (http://www.amplichip.us/presslounge/)

variants affecting warfarin activity, along with dosing algorithms to incorporate the pharmacogenetic data, now allows for adjustment of warfarin dosing to compensate for genetic differences in patients and is being integrated into clinical practice

- DME variants are known to affect the efficacy of another widely used antithrombotic agent, clopidogrel (Plavix®, Sanofi-Aventis/Bristol-Myers Squibb). Clopidogrel is an antiplatelet prodrug metabolized to the active metabolite by *CYP2C19*. Pharmacogenetic studies have previously shown that patients with reduced function variants in *CYP2C19* have an increased rate of cardiovascular adverse events, resulting in inclusion of an FDA "boxed warning." However, recent reanalysis of the data indicates there is not a significant association between *CYP2C19* genotype and cardiovascular events, highlighting the need for caution in interpreting such results and the need for additional clinical studies to incorporate pharmacogenetics into clinical practice

39.13 Pharmacogenomics of Asthma Treatment

- Asthma affects an estimated 300 million individuals worldwide, resulting in substantial morbidity, mortality, and healthcare utilization. Asthma is a complex respiratory disease characterized by airway inflammation and reversible airflow obstruction. The prevalence of the disease has increased dramatically over the past two decades, accentuating the need for effective pharmacologic treatments. Over the past few years, substantial effort has been made to explore how a patient's genotype determines asthma drug efficacy and ADRs
- Response to the three major classes of asthma therapy, β-agonists, leukotriene antagonists, and inhaled corticosteroids, demonstrates wide interindividual variability, including a significant number of nonresponders. Available data suggests that genetics

Table 39.4 Summary of polymorphisms, in two different genes that affect asthma drug treatment

Pathway	Gene	Location	Polymorphism	Associated phenotype
β_2-Adrenoceptor	B_2AR	5q31.32	16 (Arg → Gly)	Decreased response to β_2-agonists
			27 (Gln → Glu)	Reduced bronchial responsiveness
			164 (Thr → Ile)	Potentially decreased response to β_2-agonists
Leukotriene synthesis	ALOX5	10q11.12	Number of tandem (VNTR) Sp-1 and Egr-1 binding sites in promoter region	Decreased response to the *ALOX5* inhibitors

These examples must be considered in the light that until recently PGx studies were limited in scope and considered only a very small subset of variations in a few genes. These examples show the wide range of genomic variants that might influence drug response

contributes as much as 60–80% to the interindividual variability in treatment response. Here we present two different examples of genetic variants that have been shown to affect drug response and could eventually help in asthma patient stratification to improve drug response

39.13.1 B_2-Agonists

- B_2-adrenergic agonists (e.g., albuterol) are the first-line therapy for bronchodilatation in asthma patients. They act through binding to B_2-adrenergic receptors (B_2AR), a cell surface G protein-coupled receptor. The *B_2AR* gene contains numerous SNPs and some of these polymorphisms may act as disease modifiers in asthma and may be responsible for part of the known interindividual variation in the bronchodilating response to B_2-agonists
- The *B_2AR* variants Arg16Gly, Gln27Glu, and Thr164Ile appear to be functionally relevant based on data generated using cell gene transfection systems, from ex vivo studies from individuals with known genotype and from population studies (Table 39.4). These "nonsynonymous" variants result in amino acid substitutions altering receptor structure and therefore appear to be the most clinically relevant
- These polymorphisms are common, with allele frequencies as high as 50% in some cases, and have been associated with secondary effects related to the chronic use of long-acting β-agonists and also with general response to B_2-agonists in asthmatic patients
- In addition to studies of single base pair changes, there have also been studies focused on the analysis of haplotypes and their association with drug response in asthma patients. One example is a haplotype of 13 SNPs in the promoter and coding region of *B_2AR* that suggest altered receptor function correlated with clinical phenotype
- Despite these developments for B_2-agonist PGx, significant gaps in understanding the genetics remain and need to be answered before clinical pharmacogenomic assays can be developed. It is still unknown if other polymorphisms that occur in the noncoding and flanking regions of the *B_2AR* gene may be important in regulating *B_2AR* expression, thereby affecting individual drug responses. Little is known about other *B_2AR*-related genes (same pathway or transduction cascade) that may play a role in asthma drug response
- Owing to the complexity of the action of B_2-agonists and its therapeutic response, a broader genomic/genetic analysis (including other genes in the same pathway and other – noncoding – regions of the *B_2AR* gene) will help in optimizing therapy for the individual patient

39.13.2 Leukotrienes

- Leukotrienes are important mediators of asthma as they induce bronchoconstriction,

tissue edema, and airway secretion. Therefore, inhibition of leukotriene activity has therapeutic benefit in asthma patients. Leukotrienes are synthesized from arachidonic acid by 5-lipoxygenase (*ALOX5*), and several pharmacologic studies for leukotriene inhibition have focused on genes from the *ALOX5* pathway, mainly 5-lipoxygenase (*ALOX5*) and 5-lipoxygenase-activating protein (*ALOX5AP*)

- Studies have associated clinical response for a selective inhibitor of *ALOX5* with VNTR variations in the promoter region of the *ALOX5* gene, which naturally varies in the general population. The number of repeat copies is associated with changes in gene expression and promoter activity. Patients carrying the common length VNTR (5 copies) have normal gene expression levels whereas patients carrying the less common or mutant variants of the VNTR (3, 4, 6, and 7 copies) show diminished gene transcription with decreased *ALOX5* production and a diminished clinical response to treatments with drugs targeting this pathway. Individuals homozygous for any of the mutant alleles (3, 4, 6, or 7) demonstrated significantly decreased response as measured by lung function tests when compared with individuals heterozygous or homozygous for the common or wild-type allele. However, there is some discrepancy in the data published so far with some articles indicating that the mutant VNTR is associated with increased gene expression and increased leukotriene B4 production and inflammation
- The distribution of the VNTR rare alleles varies; it is present in 6% of the North American population and has been shown to vary across different ethnic groups with higher prevalence among Asians (19.4%), blacks (24%) than among Hispanic (3.6%), and non-Hispanic whites (3.1%). This variability in allele distribution indicates that response to *ALOX5* inhibitors will not be the same in all the human populations
- This example highlights the importance of the variation in the regulatory (promoter) region of genes, in addition to the coding regions of the gene itself, and of the differences in human populations
- The complexity in understanding drug response to leukotriene inhibitors is illustrated by the fact that only 6% of asthma patients carry the rare VNTR alleles at the *ALOX5* promoter locus, but >6% of asthmatic patients do not respond to leukotriene inhibitors. Thus, there are likely other genetic variants in this pathway, yet to be identified, that play a role in leukotriene regulation

39.14 Pharmacogenetics of Cardiovascular Disorders

- Cardiovascular disease is one of the leading causes of morbidity and mortality, and drug therapy is a major modality to attenuate its burden. At present, conditions such as hypertension, lipid disorders, and heart failure are pharmacologically managed with an empirical trial and error approach. However, this approach fails to adequately address the therapeutic needs of many patients, and pharmacogenetics has been explored as a tool to enhance patient-specific drug therapy
- There are a number of studies in the published literature that provide proof of concept that genetic variation contributes to the variable response that is observed on administration of cardiovascular drugs. Some of these examples are discussed here, mainly focusing on the PD aspects (genetic variants in drug targets)

39.14.1 Hypertension

- Hypertension is the most common chronic disease in the western world. Even with the large number of drugs from which to choose therapy, only 34% of North Americans respond to these treatments
- The renin–angiotensin system plays an important role in blood pressure (BP) regulation, and previous studies have reported that response to

antihypertensive medications is influenced by genetic variation in the renin–angiotensin–aldosterone system

39.14.2 Angiotensin-Converting Enzyme

- Angiotensin-converting enzyme (ACE) is essential for the production of angiotensin II and for the degradation of bradykinin, two peptides involved in vascular physiology and regulating BP. ACE inhibitors reduce peripheral vascular resistance and therefore reduce BP
- Several studies have associated an insertion/deletion (indel) polymorphism of the ACE gene with the occurrence of cardiovascular diseases, myocardial infarction, and drug response to ACE inhibitors. The ACE I/D polymorphism results from the presence or absence of a 287-bp DNA fragment in intron 16 of the ACE gene on chromosome 17. This intronic variation shows tight LD with the clinical phenotype, but the causative (functional) variant has not yet been determined
- The D allele has a frequency of approximately 0.53 in Caucasian populations and has been associated with higher levels of ACE activity. Several studies have shown that individuals homozygous for the deletion (del/del) appear to have the highest response to ACE inhibitors whereas individuals homozygous for the insertion (in/in) genotype show a lower response. Differences in plasma ACE activity associated with the ACE genotype that affect the therapeutic response of ACE inhibitors explain in the interindividual variability in cardiovascular or renal response to equivalent doses of ACE inhibitors
- However, clear patterns of association with the ACE indel polymorphism have failed to emerge, with as many studies pointing to the increased response of I allele carriers as studies pointing to D allele carriers. Other studies have found no pharmacogenetic effect of the indel polymorphism and ACE inhibitors on BP and related outcomes. These divergent results suggest that other genetic and environmental factors are associated with response to ACE inhibitors

39.14.3 Angiotensinogen Gene

- A polymorphism in the angiotensinogen (*AGT*) gene, M235T, wherein the T allele is associated with higher plasma *AGT* has been linked to elevated blood pressure (BP) and myocardial infarction. However, neither the M235T nor the T174M polymorphisms, both in exon 2, seem to affect function, secretion, or metabolism of *AGT*

39.14.4 β-Blockers

- β-Blockers (e.g., atenolol and bisoprolol fumarate) are a first-line treatment for hypertension. Several studies have found significant pharmacogenetic effects with β-blockers, including studies that reported an association in the β1-adrenergic receptor Arg389 variant and response to β-blockers. Other polymorphisms have been described, but most of them have not been confirmed in independent populations
- Pharmacogenetic studies of β-blockers have been complicated by relatively small sample size and had variable durations of treatment. However, in aggregate they suggest that variants in the β1-adrenergic receptor gene may play a future role in determining response to β-blockers

39.14.5 Pharmacogenomics of Lipid-Lowering Agents

- The prevention of cardiovascular disease has been greatly facilitated by lipid-lowering therapy including 3-hydroxymethyl-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibitors (statin drugs, including pravastatin and atorvastatin) and cholesterol absorption inhibitors. These drugs are

generally well tolerated but severe adverse effects occur in a small number of patients, and there is also a subset of patients that do not respond
- Studies have shown that SNPs in HMG-CoA are significantly associated with reduced efficacy of pravastatin. Individuals heterozygous for variants in the HMG-CoA reductase gene may experience significantly smaller reductions of cholesterol (22% lower reduction of total cholesterol and 19% lower reduction of low-density lipoprotein) when treated with pravastatin. However, these initial results have not yet been confirmed in other populations

39.14.6 Emerging Trends in Cardiovascular Pharmacogenetics

- Although numerous polymorphisms in several genes have been associated with drug response to β-blockers, ACE inhibitors, and statins, additional studies are required to translate the PGx of cardiovascular disorders into clinical practice. Future studies are required that use adequately sized patient cohorts, reduce variability in methodology, and assess complexity of the larger biologic organization (e.g., gene networks and pathways)
- Additional challenges include the applicability of pharmacogenetic findings across specific population groups as a number of reports show differences in cardiovascular disorders and severity among different racial categories

39.15 Pharmacogenomics of Neuropsychiatric Diseases

- Impressive advances have been made in the genetics of neuropsychiatric diseases. Synergies between genetic studies, elaboration of intermediate phenotypes, and novel applications in neuroimaging are revealing the effects of positively associated disease alleles on aspects of neurologic function
- Genetic and pharmacogenomic studies suggest that the subcategorization of individuals based on various sets of susceptibility alleles could make the treatment of neuropsychiatric more predictable and effective
- Much of the pharmacogenomic and pharmacogenetic data available today has come from psychiatric disorders as well as from neurologic conditions often resulting in psychiatric sequelae (e.g., Alzheimer and Parkinson diseases), and these areas will be the focus of this chapter

39.15.1 Alzheimer Disease

- Genetic polymorphisms in the apolipoprotein E (APOE) gene are associated with predicting response to therapy for Alzheimer disease as well as for lipid-lowering drugs
- There are numerous allelic variants of the human APOE gene (e.g., *APOEϵ2, APOEϵ3*, and *APOEϵ4*), which contain one or more SNPs that alter the amino acid sequence of the encoded protein
- Studies with tacrine (a cholinesterase inhibitor) to treat Alzheimer disease have shown that 83% of patients lacking the *APOEϵ4* allele had a positive response to tacrine as compared with 40% of the patients with at least one *APOEϵ4* allele. However, the greatest improvement was observed in a patient with a single *APOEϵ4* allele, the unfavorable genotype, illustrating that a single gene will not always predict the response to a given treatment
- Followup studies indicate that the prognostic value of APOE genotype for tacrine treatment was strongest in female patients, suggesting that additional genes and other factors may be involved in the response to tacrine treatment
- Other studies have suggested that patients not carrying the *APOEϵ4* allele respond to PPARγ agonist rosiglitazone (cognitive and functional improvement), whereas *APOEϵ4* allele carriers showed no improvement
- Analysis of APOE has shown association between APOE genotype and susceptibility

to Alzheimer and response to treatment. However, prospective clinical evaluation with robust clinical end points and sufficient sample size are needed to define better the usefulness of the clinical implementation of an APOE pharmacogenetics test

39.15.2 Parkinson Disease

- More than 50% of Parkinson disease (PD) patients treated with L-DOPA develop L-DOPA-induced dyskinesias (LIDs). Some patients exhibit severe dyskinesia soon after starting low doses of L-DOPA, whereas other patients remain free of this disabling complication despite long-term treatment. Avoiding or delaying the appearance of LIDs is a major issue in the management of PD
- Some studies have associated several genetic polymorphisms with the risk of developing LIDs, including variation in the dopamine receptors 2, 3, and 5
- Genetic predisposition to LIDs is likely to involve several distinct genes (or multiple alleles), each producing a small effect that might increase the risk of developing LIDs by two- to threefold

39.15.3 Major Depression

- Serotonergic activity is thought to play an important role in the regulation of mood, motor activity, and sleep patterns. Serotonin reuptake is controlled by the serotonin transporter gene (*SERT* or *SLC6A4*)
- *SERT* displays an I/D polymorphism in its promoter region (5-HTTLPR), which is the presence or absence of a 44-bp insertion. This variant results in a biallelic polymorphism designated long (l) and short (s). The shorter variant of the promoter is associated with reduction in the transcriptional efficiency of the gene resulting in decreased gene expression
- Antidepressant efficacy of selective serotonin reuptake inhibitors (SSRIs) has been shown to depend on this functional promoter polymorphism. It has been reported that carriers of the short allele have a poor outcome after treatment with SSRIs and a higher rate of adverse effects whereas individuals homozygous for the long variant have two times more efficient response to SSRIs and might have a better long-term outcome when treated with antidepressants
- However, contradictory results that might be explained by interethnic differences, or differences in haplotypes, have been reported. Therefore, sufficiently large and well-planned, controlled randomized studies are needed to finally prove that antidepressant therapy might benefit from *SERT* genotype diagnosis

39.15.4 Schizophrenia

- There is substantial unexplained interindividual variability in the drug treatment of schizophrenia with an important proportion of patients that respond inadequately to antipsychotic drugs, and many experience limiting side effects. Converging data suggest that the identification of the molecular variants that influence antipsychotic drug response and adverse effects may soon be feasible
- For the most part, the pharmacogenomic studies in schizophrenia have focused on the secondary effects of the new atypical antipsychotic agents. One of the most studied of such secondary effects is weight gain associated with those agents. Weight gain appears to be a serious side effect encountered during treatment with many antipsychotic drugs. Although the propensities of inducing weight gain vary considerably between agents, this adverse effect is mostly observed with administration of atypical antipsychotic drugs
- A specific SNP in the 5-hydroxytryptamine 2C (*5-HT2C*) receptor has been associated with weight gain across diverse samples.

Several recent reports have linked a –759C/T polymorphism of the *5-HT2C* receptor gene with chlorpromazine-, risperidone-, and clozapine-induced weight gain

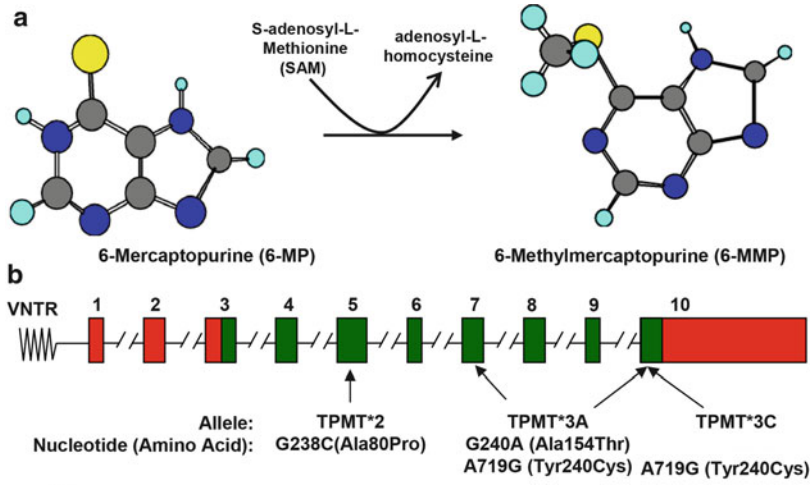
- Aside from adverse effects, some studies have suggested that variation in the gene that codes for the dopamine D2 receptor may significantly influence the clinical efficacy of a number of typical and atypical antipsychotic drugs. Some of these polymorphisms are the –141C indel polymorphism and the *TaqI* polymorphism. Although several advances have been made, particularly in understanding the pharmacogenetics of some limiting side effects, genetic prediction of drug response remains elusive and more studies with larger and better-defined populations are needed

39.16 Cancer Pharmacogenetics

- Tumorigenesis is a multistep process wherein cells acquire somatic genetic alterations followed by clonal expansion of those carrying alterations providing a growth advantage. This results in a tumor that has a genotype unique from the normal tissue which it was derived. Thus, cancer is a genetic disease, which makes it ideally suited for the use of genetic assays for tumor classification/outcomes and PGx
- The use of PGx in cancer is further driven by the narrow margin of safety (MOS – the ratio of the toxic dose to therapeutic doses) for cancer cytotoxic drugs and the emergence of targeted therapies such as trastuzumab (Herceptin®), imatinib (Gleevec®), and gefitinib (Iressa®) that have efficacy only in tumor harboring specific genetic alterations. The first widespread use of a pharmacogenomic companion diagnostic to assess drug target status was for trastuzumab
- Detailed descriptions of cancer PGx are provided in other chapters. Here, we will provide a brief historical overview and frame the current status of pharmacogenomic applications

39.16.1 Monogenic Cancer Pharmacogenetics: Thiopurine S-Methyltransferase

- Thiopurine drugs are used to treat leukemia (acute myelogenous leukemia and acute lymphoblastic leukemia), autoimmune diseases (inflammatory bowel disease, systemic lupus erythematosus, and rheumatoid arthritis), and organ transplant patients. This class of drugs includes 6-MP, 6-thioguanine, and azathioprine (6-MP prodrug). Myelosuppression is the primary toxicity observed with this class of drugs
- TPMT is a cytosolic enzyme that catalyzes the S-methylation of these drugs using S-adenosyl-L-methionine (SAM) as a methyl donor (Fig. 39.6). Inherited variation in TPMT was first reported in 1980 by Weinshilboum and Sladek. Measuring TPMT enzymatic activity in red blood cells of unrelated adults and in families, they showed a distribution of TPMT that conformed to Hardy–Weinberg prediction for codominant autosomal alleles for high and low activity and suggested that inherited variation TPMT was responsible in part for the clinical response/adverse events of thiopurines. Approximately 89% Caucasian and African-Americans have high activity, 11% intermediate, and $<1\%$ (1:300) no detectable activity
- The cloning of the *TPMT* cDNA in 1993 facilitated the molecular analysis of *TPMT* deficiency, which is now known to be because of sequence variation. Genomic characterization of the *TPMT* gene has shown that the gene is located on chromosome 6p22.3 and consists of 10 exons. Twenty-one *TPMT* polymorphisms have been identified, 18 of which are nonsynonymous SNPs (alleles *2, *3A, *3B, *3C, *5–14, and *16–*19), two of which alter mRNA splice sites (alleles *4 and *15), one premature stop codon (*3D), and one changing the translational start codon (*14)



c

Genotype	Caucasian Frequency	TPMT Phenotype
WT/WT	89.8	High
WT/*2	0.5	Intermediate
WT/*3A	8.0	Intermediate
WT/*3C	0.6	Intermediate
*3A/*3A	0.4	Low
*3A/*3C	0.2%	Low

Fig. 39.6 Polymorphisms in the *TPMT* gene. (**a**) illustrates the enzymatic activity of TPMT. Using SAM as a methyl donor, *TPMT* catalyzes the S-methylation of 6-MP to 6-methylmercaptopurine. Yellow indicates sulfur atom, blue – nitrogen, gray – carbon, and cyan – hydrogen. (**b**) Structure of the *TPMT* gene. There are a total of ten exons, eight of which encode the protein. Coding regions are indicated by green, and noncoding by red. Location of the three most common genetic variants is shown below. Allele *3A contains two nonsynonymous SNPs resulting in amino acid changes in exons 7 and 10. A VNTR located in the promoter region has been shown to be polymorphic, but the significance of this VNTR on TPMT activity is not yet been determined. (**c**) Genotype–phenotype correlation in Caucasians for the three most common alleles (*2, *3A, and *3C). Phenotyping is performed in erythrocytes measuring enzymatic conversion of 6-MP to 6-MMP using radio-labeled SAM

- The most common alleles for low activity in Caucasians are *2, *3A, and *3C all which result in a protein with essentially no enzymatic activity. The proteins encoded by these alleles are targeted for rapid proteasome degradation or aggresome formation
- The pharmacogenetic analysis of TPMT allows for strategies for eliminating thiopurine toxicity and improving efficacy by adjusting dose based on TMPT genotype. Several genotyping and phenotyping methods are used including RFLP-PCR, direct DNA sequencing, single-stranded conformation polymorphism, denaturing high-performance liquid chromatography analysis, and TPMT enzymatic/metabolite measurement in erythrocytes
- Significant dose adjustment (6–10% standard dose) is recommended for patients with low TMPT activity. Although more controversial, studies suggest that dose adjustment (~65% standard dose) reduces ADRs in intermediate patients

39.16.2 Multigene Cancer Pharmacogenomics: Microarray Expression Profiling

- The first whole genome analysis of genomic changes in human cancers showed that the number and complexity of genomic changes

Table 39.5 TPMT genotypes and phenotype

Genotype	Caucasian frequency	TPMT phenotype
WT/WT	89.8	High
WT/*2	0.5	Intermediate
WT/*3A	8.0	Intermediate
WT/*3C	0.6	Intermediate
*3A/*3A	0.4	Low
*3A/*3C	0.2%	Low

was significantly different from what was predicted from previous studies. The first large-scale cancer genome resequencing projects focused on the analysis of genes that were known to be mutated in cancer (i.e., kinases). Stephens et al. resequenced essentially all human kinases ($n = 518$) in a small set of 25 breast cancers. This was followed with larger studies in different tumor types and with a larger set of genes. In general, although these studies found mutations in genes known to be altered in cancer, the number of novel alterations was larger and the distribution and frequency of known cancer genes more complex than previously thought (Table 39.5). The results of these and other studies have led to the development of the National Institutes of Health (NIH)-funded project, the Cancer Genome Anatomy (http://cancergenome.nih.gov). Currently the Cancer Genome Anatomy is funding a 100 million USD pilot to globally assess genomic changes in cancer, including somatic mutations, transcriptional changes, and epigenetic alterations such as DNA methylation and histone modifications

- One genomic approach that has already been exploited to analyze clinical PGx of cancer is whole genome expression profiling using microarrays. Microarrays are manufactured and automated tools that allow for the analysis of a large number of nucleic acid sequences on a small solid support. Current microarray technology is based on technology first described by Ed Southern in 1975 for the immobilization of nucleic acids on a solid support. Currently, there are two predominant microarray technologies, cDNA and oligonucleotide arrays, and since the first description of cDNA microarrays in 1995 and oligonucleotide arrays in 1996, their use in cancer research has rapidly expanded and is now an established method of classifying tumors and to predict clinical outcome (survival, response to treatment)
- The rapid uptake of microarray technology in cancer research has resulted in the development of the first disease-specific multigene pharmacogenomic assays: the MammaPrint for breast cancer patient stratification (risk of recurrence) and Oncotype Dx which is used to assess the need for adjuvant tamoxifen treatment in estrogen receptor-positive/node-negative breast cancer. These assays are a significant advancement in clinical PGx and are described in greater detail in Chap. 10

39.17 Web Resources

- NCBI – http://www.ncbi.nlm.nih.gov/
- PharmGKB – http://www.pharmgkb.org/
- Genetests – http://www.genetests.org/
- NIH biomarkers – http://ospp.od.nih.gov/biomarkers/
- FDA – http://www.fda.gov/cder/genomics/default.htm
- NIH clinical trials registry – http://www.clinicaltrials.gov
- PGx for every nation – http://pgeni.unc.edu/
- National coalition for health professional education in genetics – http://www.nchpeg.org/
- Professional education for genetic assessment and screening – http://www.pegasus.nhs.uk/
- Comprehensive research on expressed alleles in the therapeutic evaluation – http://pharmacogenetics.wustl.edu/
- Personalized medicine coalition – http://www.personalizedmedicinecoalition.org/
- Drug interaction website – http://medicine.iupui.edu/flockhart/

Further Reading

Basile VS, Masellis M, De Luca V, et al. 759C/T genetic variation of 5HT(2C) receptor and clozapine-induced weight gain. Lancet. 2002;360:1790–1.

Cheok MH, Evans WE. Acute lymphoblastic leukaemia; a model for the pharmacogenomics of cancer therapy. Nat Rev Cancer. 2006;6:117.

Davies H, Hunter C, Smith R, et al. Somatic mutations of the protein kinase gene family in human lung cancer. Cancer Res. 2005;65:7591–5.

Drysdale CM, McGraw DW, Stack CB, et al. Complex promoter and coding region beta 2-adrenergic receptor haplotypes alter receptor expression and predict in vivo responsiveness. Proc Natl Acad Sci USA. 2000;97:10483–8.

Dwyer JH, Allayee H, Dwyer KM, et al. Arachidonate 5-lipoxygenase promoter genotype, dietary arachidonic acid, and atherosclerosis. N Engl J Med. 2004;350:29–37.

Evans WE, Hon YY, Bomgaars L, et al. Preponderance of thiopurine *S*-methyltransferase deficiency and heterozygosity among patients intolerant to mercaptopurine or azathioprine. J Clin Oncol. 2001;19:2293–301.

Farlow MR, Lahiri DK, Poirier J, et al. Apolipoprotein E genotype and gender influence response to tacrine therapy. Ann N Y Acad Sci. 1996;802:101–10.

Gasche Y, Daali Y, Fathi M, et al. Codeine intoxication associated with ultrarapid CYP2D6 metabolism. N Engl J Med. 2004;351(27):2827–31.

Heils A, Teufel A, Petri S, et al. Allelic variation of human serotonin transporter gene expression. J Neurochem. 1996;66:2621–4.

Holmes MV, Perel P, Shah T, Hingorani AD, Casas JP. CYP2C19 genotype, clopidogrel metabolism, platelet function, and cardiovascular events: a systematic review and meta-analysis. J Am Med Assoc. 2011;306(24):2704–14.

Israel E, Chinchilli VM, Ford JG, et al. Use of regularly scheduled albuterol treatment in asthma: genotype-stratified, randomised, placebo-controlled cross-over trial. Lancet. 2004;364:1505–12.

Kalayci O, Birben E, Sackesen C, et al. ALOX5 promoter genotype, asthma severity and LTC production by eosinophils. Allergy. 2006;61:97–103.

Kirchheiner J, Fuhr U, Brockmoller J. Pharmacogenetics-based therapeutic recommendations – ready for clinical practice? Nat Rev Drug Discov. 2005;4:639–47.

Krynetski EY, Tai HL, Yates CR, et al. Genetic polymorphism of thiopurine *S*-methyltransferase: clinical importance and molecular mechanisms. Pharmacogenetics. 1996;6:279–90.

Liggett SB, Mialet-Perez J, Thaneemit-Chen S, et al. A polymorphism within a conserved beta(1)-adrenergic receptor motif alters cardiac function and beta-blocker response in human heart failure. Proc Natl Acad Sci USA. 2006;103:11288–93.

Linazasoro G. New ideas on the origin of L-dopa-induced dyskinesias: age, genes, and neural plasticity. Trends Pharmacol Sci. 2005;26(8):391–7.

Mardis ER. A decade's perspective on DNA sequencing technology. Nature. 2011;470(7333):198–203.

Marshall E. Preventing toxicity with a gene test. Science. 2003;302:588–90.

Motulsky AG. Drug reactions enzymes, and biochemical genetics. J Am Med Assoc. 1957;165:835–7.

Nebert DW. Polymorphisms in drug-metabolizing enzymes: what is their clinical relevance and why do they exist? Am J Hum Genet. 1997;60:265–71.

Palmer CN, Lipworth BJ, Lee S, et al. Arginine-16 beta2 adrenoceptor genotype predisposes to exacerbations in young asthmatics taking regular salmeterol. Thorax. 2006;61:940–4.

Poirier J, Delisle MC, Quirion R, et al. Apolipoprotein E4 allele as a predictor of cholinergic deficits and treatment outcome in Alzheimer disease. Proc Natl Acad Sci USA. 1995;92:12260–4.

Rieder MJ, Reiner AP, Gage BF, et al. Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose. N Engl J Med. 2005;352:2285–93.

Risner ME, Saunders AM, Altman JF, et al. Efficacy of rosiglitazone in a genetically defined population with mild-to-moderate Alzheimer's disease. Pharmacogenomics J. 2006;6:246–54.

Roper N, Storer B, Bona R, et al. Validation and comparison of pharmacogenetics-based warfarin dosing algorithms for application of pharmacogenetic testing. J Mol Diagn. 2010;12(3):283–91.

Schafer M, Rujescu D, Giegling I, et al. Association of short-term response to haloperidol treatment with a polymorphism in the dopamine D(2) receptor gene. Am J Psychiatry. 2001;158:802–4.

Sjoblom T, Jones S, Wood LD, et al. The consensus coding sequences of human breast and colorectal cancers. Science. 2006;314:268–74.

Smith RL. The Paton Prize Award. The discovery of the debrisoquine hydroxylation polymorphism: scientific and clinical impact and consequences. Toxicology. 2001;168(1):11–19.

Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol. 1975;98(3):503–17.

Stephens P, Edkins S, Davies H, et al. A screen of the complete protein kinase gene family identifies diverse patterns of somatic mutations in human breast cancer. Nat Genet. 2005;37:590–2.

Taylor DR. Pharmacogenetics of beta2-agonist drugs in asthma. Clin Rev Allergy Immunol. 2006;31:247–58.

Wang L, Weinshilboum R. Thiopurine *S*-methyltransferase pharmacogenetics: insights, challenges and future directions. Oncogene. 2006;25:1629.

Weinshilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: monogenic inheritance of erythrocyte thiopurine methyltransferase activity. Am J Hum Genet. 1980;32:651–62.

Wilffert B, Zaal R, Brouwers JR. Pharmacogenetics as a tool in the therapy of schizophrenia. Pharm World Sci. 2005;27:20–30.

Winkelmann BR, Russ AP, Nauck M, et al. Angiotensinogen M235T polymorphism is associated with plasma angiotensinogen and cardiovascular disease. Am Heart J. 1999;137:698–705.

Zhu Y, Shennan M, Reynolds KK, et al. Estimation of warfarin maintenance dose based on VKORC1 (−1639 G > A) and CYP2C9 genotypes. Clin Chem. 2007;53(7):1199–205.

Molecular Forensic Pathology

40

Stephen R. Dlouhy and P. Michael Conneally

Contents

40.1 Introduction to Forensic Molecular Analysis and Parentage Testing 1033
40.1.1 Overview 1033

40.2 Forensic Molecular Analysis 1034
40.2.1 Overview 1034
40.2.2 Sample Collection and Preservation 1034
40.2.3 Forensic DNA Isolation 1035
40.2.4 Example Forensic DNA Isolation Protocols 1035
40.2.5 PCR Amplification 1036
40.2.6 STR Marker Typing 1037
40.2.7 Comparison of Results and Determination of Profile Frequencies 1041

40.3 Parentage Testing 1042
40.3.1 General 1042
40.3.2 Examples 1042

Further Reading 1047

40.1 Introduction to Forensic Molecular Analysis and Parentage Testing

40.1.1 Overview

- DNA methods have revolutionized the science of human identification
- DNA profiling has been particularly useful in the identification of criminals, especially in rape and murder cases
- Previous to DNA identification methods, blood groups and protein polymorphisms were used. In general, they were not very discriminatory. They also required a relatively large sample of blood and had extremely limited usefulness in tissue identification
- The original DNA typing markers were a class of restriction fragment length polymorphisms (RFLPs) known as variable number of tandem repeats (VNTRs). Typing was performed using the Southern blot assay and required microgram (μg) quantities of high quality DNA
- The first use of DNA polymorphisms was in Leicester, England, in the mid-1980s when Professor Alec Jeffries used multilocus VNTR probes to identify the murderer of two teenage girls
- The next major step was the polymerase chain reaction (PCR), where DNA from as little as a single cell can be amplified to produce a sufficient quantity of DNA, from the region of interest, to be tested

S.R. Dlouhy, PhD (✉) • P.M. Conneally, PhD
Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_40,

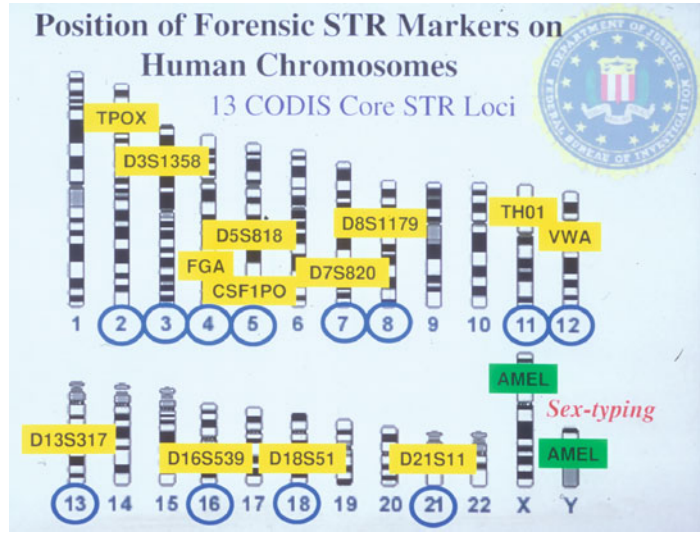


Fig. 40.1 Human STR markers. Individual locus names and approximate chromosomal positions are shown. The different chromosomal positions of the markers allow them to be used as independent loci in calculations used for forensic/paternity analysis

- The use of short tandem repeats (STRs), discussed in more detail later, has superseded the use of VNTRs in forensic typing. A standard set of combined DNA index system (CODIS) STR markers (usually 13, see Fig. 40.1) are now used for forensic typing, human identification, and parentage testing

40.2 Forensic Molecular Analysis

40.2.1 Overview

- There are five steps involved in forensic molecular analysis
  - Sample collection and preservation
  - DNA isolation
  - PCR amplification
  - STR marker typing
  - Comparison of results and determination of profile frequencies

40.2.2 Sample Collection and Preservation

- Individual sample collection is usually by venipuncture into acid citrate dextrose (ACD) or ethylenediamine tetra acetic acid tubes, a finger-prick blood drop on specially treated filter paper (fast technology analysis [FTA] paper) or a buccal swab
- Buccal swabs are the most popular since they are the least invasive method. However, contamination by oral bacteria is inevitable
- Long-term storage and preservation in the liquid state requires low temperature (−80 °C) freezers. Samples spotted on FTA paper can be stored for years at room temperature in the dry state
- For sample integrity, that is, no human DNA contamination, it is critical that powder-free disposable gloves be used during sample procurement and storage. The exquisite sensitivity of PCR amplification can allow even very minute amounts of a contaminant human DNA to show up on marker tests. For this reason, sterile swabs or FTA paper should never be touched with bare hands
- Forensic samples. Extreme care should be used in the collection of forensic samples since they are usually very limited in amount and cannot be recollected. This material may be any type of tissue which contains DNA, including any organ, teeth, bones, fingernail scrapings, hair roots, and so on. Care should be taken when examining crime scene samples to protect the examiner from potential toxic substances
- Chain of custody. All forensic samples must have a documented paper trail proving that

they have come from the person in question or the specific crime scene and have been properly sealed to prevent interference or contamination before the results of testing can be admitted into evidence. All individuals who have had access to the sample need to be documented. It should be clear that no tampering has occurred and that all seals are intact. Photographs of the evidence are very useful. A unique identifier should accompany each sample to avoid mix ups. This is especially true of crime scene samples, which are unique and irreplaceable

40.2.3 Forensic DNA Isolation

- DNA isolation/extraction from forensic material varies depending on the large variety of forensic material on which human cells have been deposited
- The type of DNA extraction may also depend on the type of DNA analysis to be performed
- Three widely used methods are Chelex extraction, FTA paper, and organic extraction. Other methods are also available
  - Chelex extraction. A major advantage is that the extraction is rapid, involves fewer steps, and thus decreases the likelihood of contamination. Another advantage of this method is that the chelating resin prevents DNA degradation by chelating metal ions. The disadvantage of this method is that it produces single-stranded DNA and is therefore only useful where PCR amplification is involved
  - FTA paper collection, storage, and DNA isolation. This is an absorbent cellulose paper with four chemical compounds, which protect DNA from degradation when stored at room temperature for several years. FTA paper is available from a number of commercial sources. A major advantage of this method is that the sample can be stored with other evidence since it does not require refrigeration
  - Organic extraction. This method is the oldest method and is also referred to as the phenol extraction method. This allows for high molecular weight DNA to be obtained and can be used for RFLP or PCR typing

40.2.4 Example Forensic DNA Isolation Protocols

40.2.4.1 Chelex Extraction

- Place 3 μL of blood sample and 1 mL of sterile water in a microcentrifuge tube and incubate 15–30 min at room temperature
- Microcentrifuge 3 min at 12,000*g* and discard the supernatant
- Add 200 μL of 5% Chelex suspension, keeping suspended on magnetic stirrer, and incubate 15–30 min at 56 °C
- Vortex for 10 s at maximum speed, then microcentrifuge for 3 min at 12,000*g* to pellet the Chelex resin
- The DNA can then be pipeted from the supernatant and used for analysis or concentrated or diluted if needed

40.2.4.2 DNA Isolation Using FTA Paper

- Place a 1–3-mm punch paper from the center of the blood spot into a thin-walled 0.5-mL PCR tube
- Washing: add 200 mL of FTA purification reagent
- Vortex tube for 1–2 s at low speed and incubate 5 min at room temperature
- Microcentrifuge 30 s at 2,000*g*. Discard supernatant
- Repeat the above washing steps two more times
- Add 200 μL of TE buffer pH 8.0 and repeat vortex and microcentrifuge steps above twice
- After air-drying punch for 1 h at room temperature, sample is ready for PCR amplification (may be done in original tube)

40.2.4.3 Organic Extraction (Also Known as Phenol Extraction)

- First, the sample is digested as follows: pipet 10–50 μL of whole blood into a microcentrifuge tube. Add the following

  - 467-μL protein lysis buffer
    - Protein lysis buffer
      - 10 mL of 1 M Tris-HCl, pH 7.4
      - 20 mL of 0.5 M ethylenediamine tetra acetic acid
      - 2 mL of 5 M NaCl
      - 968 mL of H_2O
      - Sterilize by autoclaving
      - 25 μL of 20% of sodium dodecyl sulfate
      - 7.5 μL of 10 mg/mL proteinase K
    - Proteinase K
      - 100 mg of lyophilized proteinase K (Sigma [St. Louis, MO])
      - 10 mL of sterile glass-distilled water
      - Incubate/digest a few hours to overnight at 37 °C
- Perform organic extraction as given next (take note of volume of digest). DNA should remain in the aqueous phase throughout
- Add approximately 1–1.5 volume of buffered phenol to the digest (Molecular grade buffered phenol solution is available from a number of commercial suppliers or can be made in-house. Phenol is a *hazardous* chemical and must be handled appropriately, utilizing personal protection equipment such as gloves and laboratory glasses and coat)
- Mix to form an emulsion
- Centrifuge 3–5 min at 12,000*g* at room temperature. Two layers should form. Denatured protein appears at the interphase between the layers
- Separate the aqueous and phenol phases (e.g., by pipeting), being careful to not contaminate the aqueous fraction with denatured protein from the interphase. Save the aqueous phase (for further extraction) and dispose of the phenol phase and any interphase protein (The phenol phase is usually the bottom layer and is usually darker)
- To the aqueous phase from the previous step, add approximately 1–1.5 volume of a solution containing a mixture of phenol, chloroform, and isoamyl alcohol. To make the mixture, first prepare a mixture chloroform and isoamyl alcohol (24:1), then mix that with an equal volume of buffered phenol (e.g., 10-mL chloroform/isoamyl alcohol and 10 mL of phenol)
- Mix to form an emulsion
- Centrifuge 3–5 min at 12,000*g* at room temperature. Two layers should form. The aqueous phase should again be the top layer
- Remove/save the aqueous phase and discard the organic phase. The aqueous phase should be extracted with approximately 1–1.5 volume of a 24:1 chloroform/isoamyl alcohol mixture. This step helps to remove any residual phenol that might inhibit subsequent PCR reactions
- Mix to form an emulsion
- Centrifuge 3–5 min at 12,000*g* at room temperature. Two layers should form. The aqueous phase should again be the top layer
- The aqueous phase may be used for PCR and/or further extracted with water-saturated butanol and/or concentrated if needed
- Concentrate DNA by Centricon
  - Add 1.5 mL of TE buffer, pH 8.0, to the upper chamber of a labeled Centricon 100 concentrator (Millipore [Billerica, MA]) and then transfer the DNA phase into the upper chamber of the concentrator as well
  - Centrifuge concentrator 15 min at 1,000*g*, 10 °C
  - Add 2 mL of TE buffer, pH 8.0, to the upper chamber of the concentrator
  - Repeat these steps two more times
  - Invert upper chamber of concentrator into a labeled conical collection tube (supplied with the concentrator). Centrifuge 5 min at 1,000*g*, 10 °C, to collect DNA
- Alternatively, if there is sufficient volume, the DNA may be concentrated by alcohol precipitation

40.2.5 PCR Amplification

40.2.5.1 General

- The goal of PCR is to enzymatically amplify the amount of a specific DNA sequence or sequences, such as the DNA-representing STR loci
- This amplification facilitates further analysis (e.g., allele typing) of the DNA
- While many of the general concepts of PCR amplification can be found elsewhere, it is

important to remember that amplification of specific sequences requires the design (in advance) and use of oligonucleotide primers that are specific for the target sequence(s)
- The design of primers used for PCR not only depends upon the markers selected for typing but also on the planned method of detection and analysis of the PCR product (as discussed further later)
  - PCR is used for the 13 standard CODIS STR marker loci already mentioned (and shown in Fig. 40.1). Additional markers are also available or may be developed by a particular laboratory
  - Kits for STR markers can be obtained from commercial suppliers such as Applied Biosystems (Foster City, CA) or Promega

40.2.5.2 Limitations/Considerations

- As with other applications that involve PCR, one must keep in mind a number of limitations of this technique. These may be of particular concern if the original sample/template DNA is only a small amount, as may be the case for forensic samples
- There is the possibility that PCR product(s) from contaminating DNA may interfere with the analysis and/or subsequent interpretation
- There are also a number of scenarios in which the PCR product may not accurately reflect the makeup of the original DNA, even if there is no contamination
  - Mutations may be introduced during the amplification process. If these happen early enough during the reaction sequence, "mutant/artifact" DNA may constitute a major portion of the final PCR product
  - Differential amplification of alleles. It is possible that there may have been two different alleles at a given STR locus in the original genomic DNA, but only a single allele is seen in the PCR product. This phenomenon is sometimes referred to as "allele drop out"
    - For example, a DNA profile obtained from a buccal sample taken at interrogation of the suspect may have two alleles at a given STR locus, but only a single allele is seen in the DNA profile of a forensic sample
    - It is also possible that there are "mutations" in the original DNA sample at the site at which a primer is supposed to bind. This can prevent (or impair) primer binding and the subsequent failure of amplification and detection of an allele

40.2.6 STR Marker Typing

40.2.6.1 General

- The underlying basis of STR markers is the existence in genomic DNA of short stretches of bases that are repeated in tandem a number of times (e.g., see Fig. 40.2). A single locus can have many different alleles with different numbers of repeats
- For a given STR marker, the number of bases that make up the repeated unit is expected to be the same (e.g., a two base, GT repeat is shown in Fig. 40.2). However, another STR marker may have a different number of bases within the repeat, for example, a three-base repeat unit
- When such repeats are found at specific chromosome locations and are flanked by unique sequences, they may be useful as markers. This is because the number of repeats of the unit can be different on different chromosomes, thus providing polymorphism useful for genetic analysis (e.g., see chromosomes A [6 repeats] and B [13 repeats] in Fig. 40.2)
- The flanking sequences can be used to design primers for PCR. Thus, while the repeat itself (e.g., a GT repeat) might be present many times in the genome, the association with specific flanking sequence enables specificity during PCR amplification
  - It is combination of the repeat and the specific flanking sequences that makes a particular marker locus unique
  - As a result of differences in tandem repeat number in the genomic DNA, the PCR products of different STR alleles (see Fig. 40.2) will differ in size

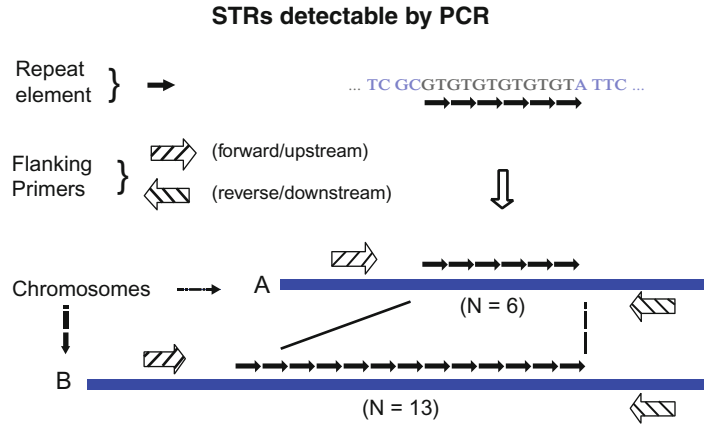


Fig. 40.2 Diagramatic representation of molecular basis of STR markers. The two homologous chromosomes shown (**a** and **b**) each have a different number of copies of the GT repeat unit

- Thus, STR locus products are typically analyzed by electrophoresis. There are a number of different systems that can be used such as acrylamide gel (discussed directly next and shown in Fig. 40.3) or capillary electrophoresis (as discussed and shown in later figures)
- In order to determine what alleles are present for a particular STR marker, it is best to include an allelic ladder as part of the electrophoretic analysis. The allelic ladder helps to ensure correct identification of specific alleles at each locus. Allelic ladders may be included in kits provided by various manufacturers and may be limited to contain only the more common alleles for each locus
- Important advantages of STR markers are that the allele designations have been standardized (e.g., across laboratories) and that specific population frequencies of alleles are available

40.2.6.2 Limitations

- Higher mutation rate than RFLP loci
- Markers may not be as informative as some RFLP systems

40.2.6.3 Example of Acrylamide Gel Electrophoresis

- With this system, the samples from different individuals or specimens can be electrophoresed concurrently and compared side by side. This allows for rapid visual comparison
- Figure 40.3 demonstrates results of STR analysis for four samples and allelic ladders that have been resolved by acrylamide gel electrophoresis. Samples A, B, and C (represented in **lanes 2, 3**, and **4**) are from three different individuals, and they each demonstrate distinct banding patterns
- Individuals may share some alleles in common, as can be seen for allele "19" of individuals A and B (**lanes 2** and **3**) in the figure. This may be by chance or because the individuals are closely related
- When two individuals have different banding patterns, it is possible to recognize contributions from both individuals in samples that are mixtures from different individuals (e.g., see **lane 5** of Fig. 40.3), such as may be encountered in forensic scenarios. In mixed samples, the relative amounts of DNA from each individual may vary. In this example, the relative contributions are nearly equivalent

40.2.6.4 Example of PCR and STR Analysis by Capillary Electrophoresis

- As a further illustration, we will briefly describe a protocol modeled after the Profiler Plus™ kit for PCR utilizing an ABI 310 Genetic Analyzer for capillary electrophoretic analysis (Applied Biosystems). Additional information can be found in the manufacturer's literature

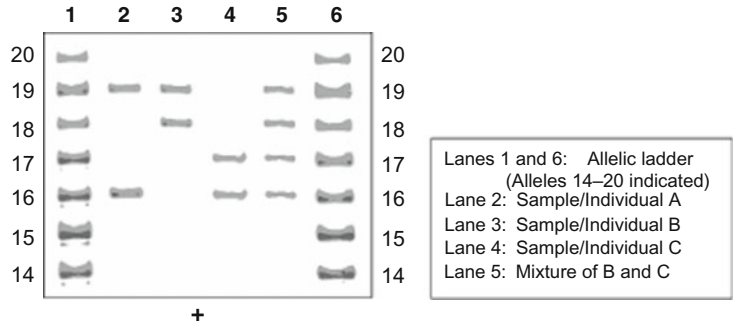


Fig. 40.3 STR markers resolved by gel electrophoresis. Samples were loaded at the top and electrophoresed toward the positive pole at the bottom. Bands were visualized by silver staining. Other methods of visualization are possible such as by direct staining of DNA with dyes like ethidium bromide or by use of x-ray film or phosphorimaging to detect fragments that have been radioactively labeled. Regardless of the method of detection, minor/weak background bands may sometimes be seen

- In this system, there are fluorescent dyes (Applied Biosystems) attached to the primers that allow detection of the PCR products. Thus, these reagents are light sensitive. The primer set should be protected from light while not in use. Amplified DNAs should also be protected from light
- The system is designed for multiplex analysis, that is, multiple loci are amplified simultaneously in one tube and can be analyzed simultaneously during electrophoresis. There are not only differences in PCR product sizes due to allelic variation, but also the PCR products from one STR locus differ in size from the products of another STR locus
- Example PCR reaction
  - Place the required number of 0.2-mL reaction tubes into a rack and label them
  - In order to provide better uniformity of reaction conditions from sample to sample, it is best to first make a master mix that contains all components except for DNA. This can then be distributed (in aliquots) into the reaction tubes, followed by DNA addition to each individual tube
  - Components for the master mix may be produced in-house or provided as stocks by a commercial supplier (e.g., PCR reaction mix, primer set, and DNA polymerase as supplied by ABI)
  - Prior to setting up the master mix, the stock tubes should be vortexed briefly to mix and then spun so that any liquid that may be adhering to the top or the sides of the tube can be collected in the bottom
  - Example preparation and dispensing of a master mix
    - In a 1.5-mL tube, combine
      - Number of samples × 10.5 μL of PCR reaction mix
      - Number of samples × 0.5 μL of DNA polymerase
      - Number of samples × 5.5 μL of primer set
    - Dispense 15 μL of master mix into each PCR tube
    - To individual tubes, add 10 μL (1 μg) of control DNA or 1 μL of undiluted DNA of unknown sample plus 9 μL of double-distilled water
    - The final volume in each PCR tube is 25 μL
- Place the PCR tubes in a thermocycler (Applied Biosystems) and start the program as follows
  - Hold 95 °C for 11 min
  - 28 cycles of
    - 94 °C for 1 min
    - 59 °C for 1 min
    - 72 °C for 1 min

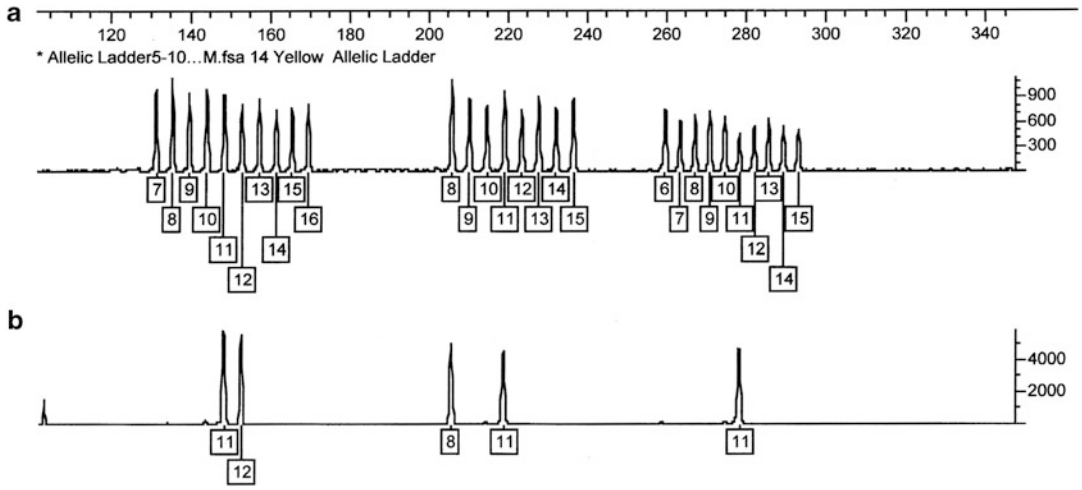


Fig. 40.4 Resolution of STR markers by capillary electrophoresis. The results for only 3 of the 13 CODIS STR marker loci are shown. Smaller fragments are to the *left* and larger fragments are to the *right*. The panel contains two plots (*lines*) of data. The top plot (*line A*) represents an allelic ladder that includes common alleles from three different STR marker loci (i.e., D5S818 on the left, D13S317 in the middle, and D7S820 on the *right*). The other plot (*line B*) is data from one individual. The tested individual has two different alleles (is heterozygous) at each of two loci (D5S818 has alleles 11 and 12; D13S317 has alleles 8 and 11). The individual has only one allele (i.e., they appear to be homozygous) at the third locus (D7S820 has allele 11). Note that small peaks are often seen to the *left* of larger peaks (e.g., see *arrow* for D7S820). These artifacts are referred to as "stutter" peaks that may result from slippage of *Taq* polymerase and which may or may not interfere with interpretation of results

  - Hold 60 °C for 45 min
  - Hold/store 4 °C
- Capillary electrophoresis
  - Before analyzing (running) the PCR product, the samples and capillary electrophoresis apparatus/analyzer must be prepared for the run according to manufacturer's specifications or as determined by the laboratory for a particular protocol. This includes installing or checking files for detection of the specific fluorescent dye(s) used in the application
  - Example preparation of allelic ladder and samples
    - Combine 12 μL of deionized formamide, 0.5 μL of size standard, and 1 μL of PCR product or 0.5 μL of allelic ladder
    - Mix by pipeting up and down and microcentrifuge briefly
    - Denature each sample 5 min at 95 °C
    - Chill tubes on ice for at least 3 min
    - Place tubes in sample tray
  - Place the sample tray in the capillary apparatus, launch the appropriate software, enter/list sample information, and initiate the run
  - After the run is complete, the data is analyzed (see Sect. 2.7 below)
- Determining genotypes may be done with the aid of software and/or by visual inspection. Genotypes are assigned by comparing the sizes obtained for unknown samples with the sizes obtained for the alleles in the allelic ladders (e.g., see Fig. 40.4). Genotypes, not sizes, are used for comparison of data between runs

40.2.6.5 Limitations/Considerations

- In addition to the primary peaks, other peaks (of varying size) may be seen on the tracings from capillary electrophoretic analysis (e.g., see "stutter" in Fig. 40.4)
- Such peaks may represent stutter peaks (as described above) or contaminating DNA

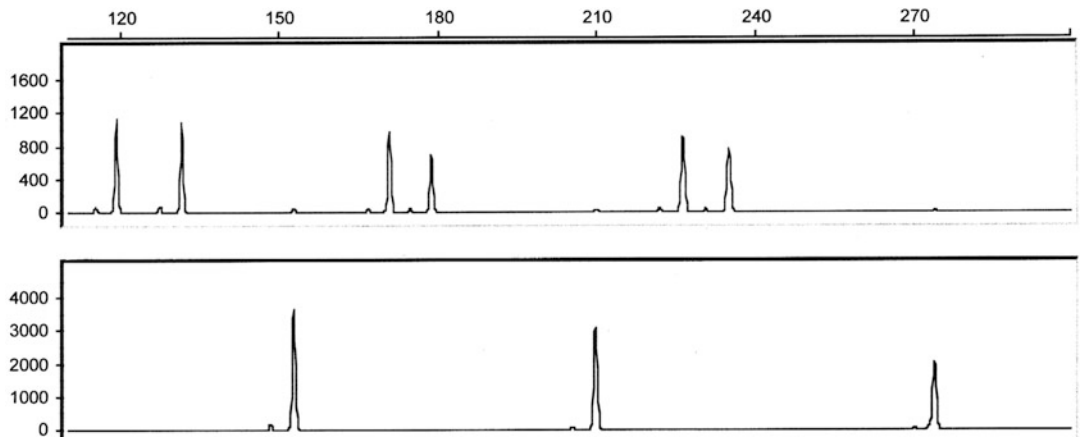


Fig. 40.5 Capillary electrophoretic analysis demonstrating pull-up artifacts. In addition to the stutter peaks that are slightly *left* off the main peaks in both plots, the *top plot* contains other minor peaks that correspond to the major peaks that appear in the *bottom plot*. Although relatively small in this example, these pull-up peaks can be of considerable height and might be misinterpreted as real peaks. This phenomenon occurs when the amount of PCR is particularly high for one set of markers (note the scale to the *left* of each plot)

- If the amount of initial template DNA is limiting and the overall amplitude of amplification is low, a number of "background" peaks may appear along the baseline. Although these may be randomly distributed, some can have mobility similar to that of fragments from real STR alleles. Thus, it may be difficult to unambiguously identify all alleles, particularly in mixed samples (e.g., in a sample that contains a low percentage contribution of suspect DNA)
- During capillary electrophoresis, fragments from different STR loci are labeled/tagged with different colored dyes and are electrophoresed and detected simultaneously in the same capillary. Thus, in a particular plot/line there may appear peaks that actually are caused by a type of spill over from the loci of a different plot. This phenomenon is sometimes referred to as "pull up" (see Fig. 40.5)

40.2.7 Comparison of Results and Determination of Profile Frequencies

40.2.7.1 General

- If the STR genotypes for the evidentiary and the suspect's samples match for all 13 CODIS loci, this strongly suggests that they originated from the same individual. Multiple loci are used to increase mathematic confidence in the results
- The next step is to determine the frequency of the profile. This requires a database of gene frequencies. Since gene frequencies differ across populations, databases are available for the major racial or ethnic groups such as Caucasian, African-American, and Hispanic
- The size of the database is an important consideration. Although one might wish to have frequencies based on a very large number of individuals in each population, sample procurement, and STR typing precludes such an endeavor. Typically, the sample size for most databases is approximately 200 individuals. The number of genes is twice the number of individuals
- At a given locus, an individual may be homozygous with one allele or heterozygous with two alleles. The frequency of homozygotes will be p^2 (or q^2) and the frequency of heterozygotes is $2pq$, where p and q denote the allele frequencies. The genotype frequency is calculated for each locus, and these frequencies are multiplied together to obtain the overall frequency of the profile

Table 40.1 Calculation of a profile frequency. Expected genotype frequencies, based on allele frequencies, are calculated for each locus. The genotype frequencies are multiplied to give the expected overall profile frequency

		Allele frequency		
Locus	Alleles 1,2	1	2	Genotype frequency
TPOX	8,10	0.5443	0.0369	0.0402
D3S1358	15,15	0.2463	–	0.0607
FGA	20,25	0.1454	0.0689	0.02
D5S818	11,12	0.4103	0.3538	0.2903
CSF1PO	10,11	0.2537	0.3005	0.1525
D7S820	10,11	0.2906	0.202	0.1174
D8S1179	13,15	0.3393	0.1097	0.0744
THO1	7,7	0.1724	–	0.0297
vWA	16,17	0.2015	0.2628	0.1059
D135317	10,12	0.051	0.3087	0.0315
D165539	12,13	0.3391	0.1634	0.1108
D18S51	12,14	0.1276	0.1735	0.0443
D21S11	30,31.2	0.2521	0.0995	0.0462

Overall frequency 4.2 E-16

40.2.7.2 Example

- Table 40.1 illustrates the calculation of a profile frequency
- Note that in this example, the sum of the allele frequencies at any one locus does not add up to one. That is because each locus has many other alleles that are not involved in the example
- The overall frequency of the profile is 4.2 E-16 or (taking the reciprocal) 1 in 2.4 E 15 or 1 in 2.4 quadrillion. Given that the world's population is more than six billion, this is clearly a highly significant number and is close to uniqueness. The frequencies used in this example are from a Caucasian database

40.3 Parentage Testing

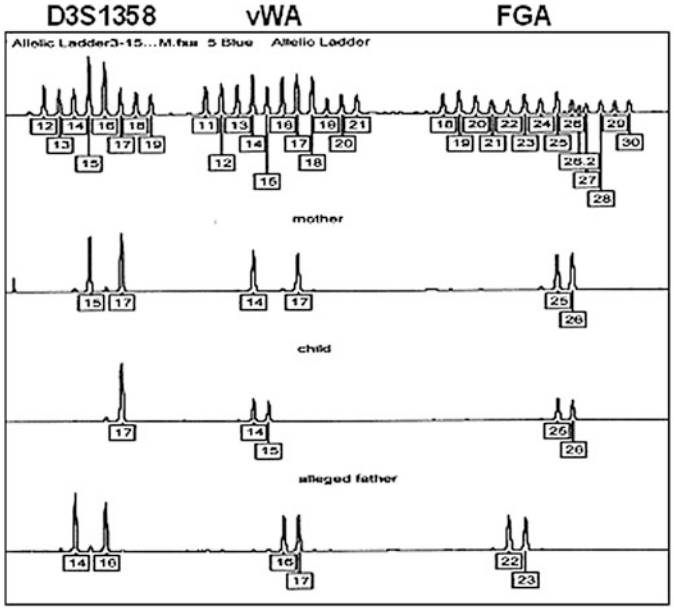
40.3.1 General

- In the vast majority of cases of parentage testing, the parent in question is the father (paternity). However, there are many variations of testing with related individuals. This is especially important in the case of mass tragedies where a relative of a missing individual may provide DNA information, which may be sufficient to identify the deceased
- In the case of paternity testing, where the father of the child but not the mother is in doubt, DNA samples will be available for all three individuals, putative father, mother, and child
- Essentially all paternity testing laboratories in the United States use the standard 13 CODIS STR markers, though some may add extra STRs
- In the case of paternity testing laboratories, their cases are usually to prove paternity for child support, while forensic laboratory cases usually involve rape, statutory rape, or incest
- Once DNA testing is complete, the outcome is either exclusion or inclusion. In the case of exclusion, this means that the child has one or more markers, which could not have been inherited from the putative father

40.3.2 Examples

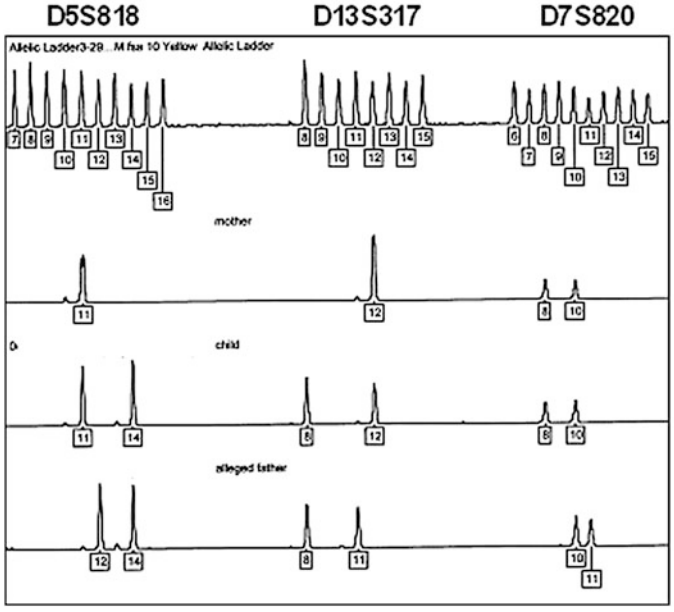
- As an example, let us assume that the child is type AB, the mother type AC, and the putative father type CD. Clearly the child inherited the A allele from its mother and, therefore, must have inherited the B allele from its father. This excludes the questioned man as being the father

- An exception to this explanation is mutation. It is possible that the allele contributed by the father has mutated prior to being passed on to the child. While mutations are usually very rare (<1 per 20,000 meioses) in the case of STR repeat alleles, the frequency is much higher (in the order of 1 in 1,000 meioses). Exclusion at 1 of the 13 loci is not unusual and does not necessarily exclude the putative father
- Data demonstrating exclusion at three loci (D3S1358, vWA, and FGA) is also shown in **panel A** of Fig. 40.6
- In some circumstances, samples may be available from more than one alleged father or more than one suspect. Clearly, as with any case, chain of custody and sample integrity are critical in such circumstances
- A case involving two alleged fathers is shown in Fig. 40.7
- Based on results for marker A (**top panel**, Fig. 40.7), either man could be the father since they both share the lower band with the child. The mother contributed the child's upper band
- Marker B (**bottom panel**, Fig. 40.7) shows exclusion for AF2. The child appears to be homozygous. Thus, the mother and father would each be expected to have contributed the same allele to the child, an allele that AF2 lacks (*Note*: Upon completion of this chapter, the astute reader may think of an alternative explanation of the findings for marker B)
- In the case presented in Fig. 40.7, AF2 was also excluded at number of other STR markers. However, it is not uncommon for an excluded man to share alleles with the child at more than one marker
- Typically, there should be evidence of exclusion with at least two different markers before considering ruling out a given man as the father (e.g., see discussion of mutation). It is important to test multiple markers in order to achieve a high degree of confidence
- If there are no exclusions, this means that the putative father is very likely to be the biological father of the child. Thus, a statistical calculation is necessary to understand the strength of the conclusion that the putative father is indeed the actual father. The statistic involved is known as the *paternity index (PI)*. The PI is a *likelihood ratio* or a *ratio of two probabilities*
  - The *numerator* is the probability of the child's genotype given that the man is the father. This will almost always be a Mendelian ratio either 0.25, 0.5, or 1
  - The *denominator* is the probability of the child's genotype given that another man is the father. This probability is composed of two parts, the probability that the mother transmitted a specific allele to her child which is 0.5 or 1 times the probability that the random man transmitted the other allele, which is the frequency of that allele in the population
- Examples in Table 40.2
- The gene frequencies are obtained from databases published by the FBI and others. Since there are differences in frequencies among various ethnic groups, there are separate gene frequency databases for Caucasians, African Americans, and Hispanics
- The PI is obtained for each of the STR loci, and the overall (combined) PI is simply the product of the individual PIs
- Table 40.3 shows an actual example of a PI calculation using 13 CODIS STR loci. The PI is 1,648,000. This means that the putative father is 1,648,000 times as likely to be the father of the child rather than a man chosen at random. This is very compelling evidence in favor of paternity
- Paternal mutation
  - As discussed earlier, an exclusion at only 1 of 13 tested loci does not necessarily exclude the putative father. In such a circumstance, it would be necessary to calculate a PI that takes into account the mutation frequency at that locus. Details of such analyses are beyond the scope of this chapter (e.g., *see* "Mutation in Paternity" in "Further Readings")
- Maternal contribution
  - Although a PI can be determined without a sample from the mother, it is best to have a maternal sample whenever possible



A: Exclusion

Paternity case demonstrating exclusion of the alleged father. For each marker locus, the alleged father does not have an allele that matches the child.



B: Inclusion

Paternity case demonstrating inclusion of the alleged father. For each marker locus, the alleged father has an allele that he could have contributed to the child.

Fig. 40.6 (*Panel A*) Exclusion: paternity case demonstrating exclusion of the alleged father. For each marker locus, the alleged father does not have an allele that matches the child. *(Panel B)* Inclusion: paternity case demonstrating inclusion of the alleged father. For each marker locus, the alleged father has an allele that he could have contributed to the child. Each panel contains four plots (*lines*) of data. The *top line* represents an allelic ladder that includes alleles from three different STR marker loci (i.e., the left, middle, and right groups). The other three plots/lines in each panel are data for three individuals, a mother, child, and putative (alleged) father

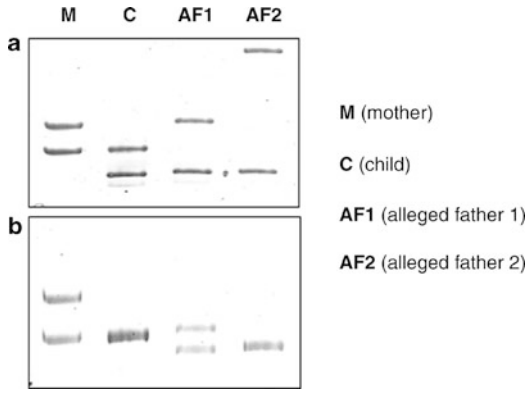


Fig. 40.7 Paternity case involving two alleged fathers. STR analysis for two markers, (**a**) (*top panel*) and (**b**) (*lower panel*), was done with a commercial kit for PCR, acrylamide gel electrophoresis to separate fragments, and silver staining to visualize bands

Table 40.2 Example formulae for calculation of a paternity index. The formula that is used for a particular marker can vary and depends on the marker results for each member of the trio

Putative father	Mother	Child	PI
AA	AA	AA	$1/p$
AB	AA	AB	$0.5/p = 1/2p$
AA	AB	AA	$0.5/0.5p = 1/p$
AB	BC	BB	$0.25/0.5p = 1/2p$
AB	AB	AB	$0.5/(0.5p + 0.5q) = 1/(p + q)$

PI paternity index

Table 40.3 Example paternity index calculation. A paternity index (PI) is calculated for each marker based on the results for each member of the trio, use of the appropriate formula (e.g., see Table 40.2) and allele frequencies (not shown) obtained from a database. The PIs for individual marker loci are multiplied to obtain the overall (combined) PI

Marker	Putative father	Mother	Child	PI
FGA	19,27	20,25	25,27	39.06
TPOX	8,11	8,10	8,10	0.86
D8	13,14	13,15	13,15	1.11
vWA	14,15	16,17	15,16	4.46
D18	12,14	12,14	14,14	2.88
D21	30,30	30,31	30,31	3.02
THO1	7,9.3	7,7	7,9.3	1.64
D3	14,17	15,15	15,17	2.36
CSF1PO	10,12	10,11	10,12	1.54
D16	12,14	12,13	12,14	15.53
D7	9,10	10,11	9,11	3.38
D13	10,12	10,12	10,12	2.78
D5	11,12	11,12	11,12	1.31

PI paternity index. Overall (combined) PI: 1,648,000

  - Having knowledge of the mother's genetic contribution to the child enables a better determination of exclusion or inclusion of the putative father
  - The mother is expected to have at least one allele in common with the child at each locus
    - Mismatches between the woman and child could indicate that the woman is not the mother
  - For prenatal testing, such as may be involved in a case of rape, it is particularly important to have a sample from the mother
    - This helps to establish that the results obtained for the "fetal" sample do not actually represent maternal contamination
    - This helps to validate the integrity/chain of custody of the fetal sample
- Example
  - A case of what appeared to be a mismatch between mother and child is shown in Fig. 40.8
  - When initially analyzed in routine fashion with the standard allelic ladder, the mother and child each appeared to be homozygous for different alleles (see **Panel A**)
  - This, however, was the only mismatch; they shared alleles at all 12 other loci. When data from the run was reanalyzed, additional peaks were seen, consistent with the mother and child sharing a rare allele
  - *Note*: a similar scenario of "apparent homozygosity" of parent and child for different alleles could occur if they actually shared an allele with a mutation in a primer binding site. Thus, the DNA from that allele might not be amplified by standard PCR and no peak (no PCR product) would be detected for either individual. Similar circumstances that could be misinterpreted as homozygosity for different alleles could also occur in a paternity assessment

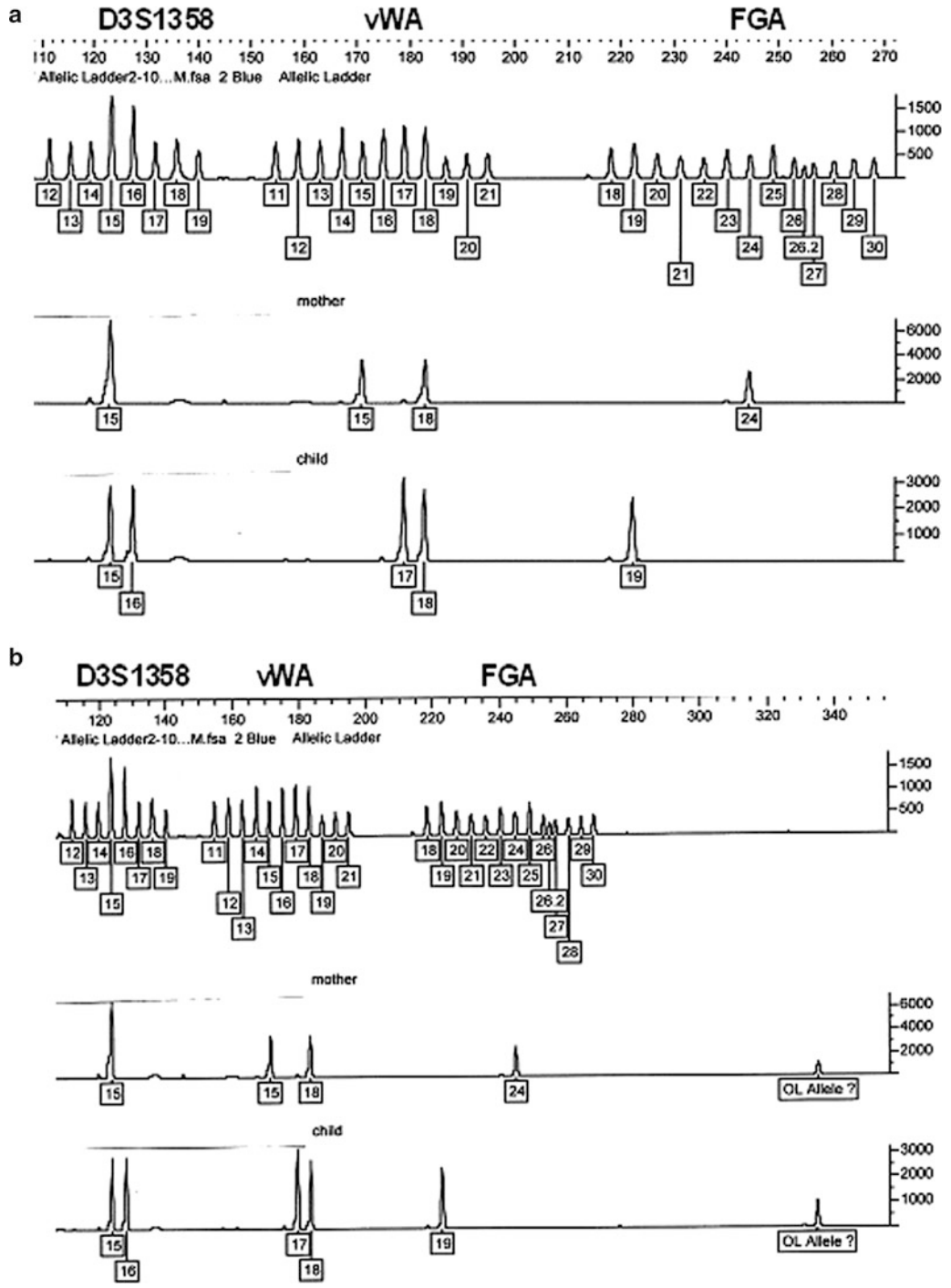


Fig. 40.8 Capillary electrophoresis STR results for a mother and child. *Panel A*: printout of standard analysis with three loci and allelic ladders of common alleles shown. The two individuals appear to differ at the FGA locus. *Panel B*: printout demonstrating unusual peak that is shared by mother and child. Since rare alleles in this size range are reported in the database(s) for this population, this finding was interpreted as a match between mother and child

Further Reading

Budowle B, Moretti RR, Niezgoda SJ, Brown BL. CODIS and PCR-based short tandem repeat loci: law enforcement tools. Second European Symposium on Human Identification. Madison: Promega Corporation; 1998.

Budowle B, Hebson DL, Smerick JB, Smith JAL. Low Copy Number – Consideration and Caution. Proceedings of the twelfth international symposium on human identification; 2001. http://www.promega.com/products/pm/genetic-identity/ishi-conference-proceedings/12th-ishi-oral-presentations/

Butler JM. Forensic DNA typing, second edition: biology technology and genetics of STR markers. Burlington: Elsevier Academic; 2005.

Butler JM, Reéder DJ. (1997, with continuous updates) Short tandem repeat DNA internet database. http://www.cstl.nist.gov/biotech/strbase/

Evett IW, Weir BS. Interpreting DNA evidence: statistical genetics for forensic scientists. Sunderland: Sinauer Associates; 1998.

Jeffreys AJ, Wilson V, Thein SL. Individual-specific fingerprints of human DNA. Nature. 1985;316:76–9.

Brenner C. Mutations in paternity. http://www.dna-view.com/mudisc.htm. July 9, 2007.

Standards for parentage testing laboratories, 6th edition, American Association of Blood Banks, Bethesda, Maryland. http://www.aabb.org. July 9, 2007.

User's manuals/guides as supplied by Applied Biosystems for ABI 310 and Profiler Plus. Also see, http://www.appliedbiosystems.com. July 9, 2007.

Specimen Identification and Bone Marrow Engraftment Analysis

41

Loren Joseph

Contents

41.1 Specimen Identification Overview 1049

41.2 Errors in Medicine 1050
41.2.1 The Institute of Medicine Study 1050
41.2.2 Identification Error in Anatomic Pathology 1050
41.2.3 Identification Error in Clinical Pathology 1051

41.3 Methods for Specimen Identification 1053
41.3.1 Preanalytical Considerations 1053
41.3.2 Analytical Targets and Methods 1053
41.3.3 Medicolegal Considerations 1058

41.4 Clinical DNA Chimerism Analysis 1058
41.4.1 Analysis for Hematopoietic Stem Cell Transplantation 1058
41.4.2 Maternal Cell Contamination 1060
41.4.3 Gestational Trophoblastic Disease 1062
41.4.4 Tumors Arising in Transplanted Organs 1063
41.4.5 Biobanking 1063
41.4.6 Cell Line Identification 1063

Further Reading 1063

41.1 Specimen Identification Overview

- Specimen misidentification is a persistent serious source of medical error
- DNA genotyping can resolve many specimen identification problems
- The dominant method is analysis of microsatellites ("DNA fingerprinting")
- Analysis of single nucleotide polymorphisms (SNPs) and short insertion–deletions (indels) are alternative but less well-standardized methods and at present little used
- Some new genetic technologies can be engineered to include sample identification
  - Whole genome/exome/transcriptome sequencing
  - Highly multiplexed parallel mutation testing
- DNA genotyping has an important role as a diagnostic test in which specimens are expected to be admixtures
  - Bone marrow engraftment analysis
  - Parentage testing (see Chap. 40)
  - Maternal cell contamination assessment in prenatal testing
  - Gestational trophoblastic disease classification
  - Identification of origin of tumors post transplantation
  - Biobanking quality assurance/control
  - Cell line identity confirmation (primarily for research)

L. Joseph, MD
Department of Pathology, University of Chicago Medical Center, Chicago, IL, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_41,

41.2 Errors in Medicine

41.2.1 The Institute of Medicine Study

- The Institute of Medicine Study "To Err is Human" helped catalyze the patient safety movement
- This study estimated 100,000 deaths/year attributable to medical error including identification errors
- The plausibility of this figure has been challenged but not the conclusion
- The study calls for reduced reliance on vigilance! The goal is to engineer systems to reduce opportunity for error. Methods like barcoding blood sample tubes are already widely implemented or gaining acceptance in the case of tissue cassettes

41.2.2 Identification Error in Anatomic Pathology

41.2.2.1 Incidence of Identification Errors

- Accessioning errors – CAP Q Probe. A survey of 1,004,115 cases from 417 institutions found significant identification errors in accessioning in 0.6%
- Extraneous tissue contaminants (ETCs). A combined prospective/retrospective Q Probe study of 321,757 slides surveyed the incidence of ETC (also known as "tissue floaters"), a form of identification error
  - ETCs were identified in 0.58% of slides in the prospective and in 2.9% in the retrospective study
  - The contaminant was judged to be a slide contaminant in 59.4%/72.9% of ETC cases
  - Of cases with ETC, 9.4%/1.1% presented a diagnostic problem

41.2.2.2 Laboratory Phase Errors

- A national reference laboratory survey of identification errors in the analytical phase (or "laboratory phase" – *after* accessioning) covered 29,479 cases/109,354 blocks/248,013 slides
  - 75 errors detected (0.25% of cases, 0.068% of blocks, 0.030% of slides)
  - 69% were attributed to error in the gross room, predominantly mislabeling of cassettes
  - 13/75 could have led to significant inappropriate clinical action
- Opportunities for error after receipt of a correctly identified specimen
  - Mislabeling cassettes
  - Acquisition of extraneous tissue – in the gross room or histology lab
  - Mislabeling slides
- Opportunity for error in molecular genetic testing of pathology specimens
  - Starting with DNA extraction, mislabeling or contamination can afflict every step
  - Extraneous tissue contaminants
    - If DNA is prepared from tissue sections on slides which are matched to a stained slide to guide macrodissection or microdissection, ETC should not be an issue (beyond any error made at diagnosis)
    - If DNA is prepared from tissue scrolls, ETC could be acquired which are NOT present in the block
      - With Sanger sequencing, as much as 20% contamination will probably not generate a false positive signal for a mutation
      - More sensitive mutation-specific tests could lead to a false positive
  - Labeling the DNA storage tube incorrectly
  - Errors in aliquoting during assay setup
  - Entering sample identification information into the analysis software
- Additional signs of unsuspected identification error
  - Diagnosis is unexpected based on clinical findings

- Subsequent clinical course is not consistent with the surgical pathology diagnosis. A simple example is the "vanishing" cancer identified on biopsy but not in the subsequent resection
- The posttransplant sample genotype matches *neither* the recipient *nor* the donor genotype(s)

41.2.2.3 Opportunities to Reduce Error

- Barcoding of containers, cassettes, and slides
- Radiofrequency identification (RFID) "tags" can be included in the sample collection container and maintained throughout the histology process
- Barcodes and RFID tags facilitate efficient tracking of specimens and save labor in addition to reducing error
- Inking – A low-tech "solution" uses a prescribed sequence of colored inks to stain consecutive biopsies. Permits error detection in all laboratory phases
  - The above methods can each significantly reduce error at various phases, but without further operational engineering, none can keep a pathologist or technologist from putting the wrong tissue into the "right" cassette or on a slide
- Include identification markers in the test method
  - Illustrated in one commercial microsatellite instability test kit (Fig. 41.1)
  - Not suitable for most current single target molecular testing
  - Could be engineered into highly multiplexed parallel assays such as microarray studies, multigene sequencing, and single nucleotide extension assays (e.g., MALDI-TOF, SnaPShot) by including highly informative microsatellites, SNPs, or indels among the targets
    - Requires genotyping of peripheral blood (or normal tissue) from the same patient
      - This increases cost and introduces possibility for additional errors, but ones which should be readily detected and resolved
- Automate assays (robotics) for DNA or RNA testing

41.2.3 Identification Error in Clinical Pathology

41.2.3.1 Incidence of Errors

- CAP Q Probe study of clinical laboratory identification errors
  - 120 institutions surveyed (90% used barcodes)
  - 6,705 errors for a rate of 324 errors/1,000,000 billable tests
  - 85% of errors were found before results were released
  - 345 errors held risk for adverse consequences loosely defined
  - 251 led to patient inconvenience, typically a redraw
  - 16 led to a treatment change, no change in outcome
  - 78 were of unknown consequence

41.2.3.2 Opportunities to Reduce Errors

- Barcoding of blood collection tubes, already in wide use
- For most tests, if an error is suspected, it is simplest to get another specimen
- If redraw is not feasible or if a medicolegal issue is involved, it is possible to extract adequate DNA from serum or plasma (and occasionally urine) to permit DNA genotyping
- Molecular genetics
  - DNA can be stored frozen indefinitely for retesting and DNA identification
  - Whole blood (or DNA) can be archived at room temperature on specific dry matrices; this permits archiving multiple aliquots of blood at the time of sample receipt
  - Testing of family members, where indicated, offers additional opportunity for patient identity confirmation
  - Automation (robotics) for DNA/preparation and testing

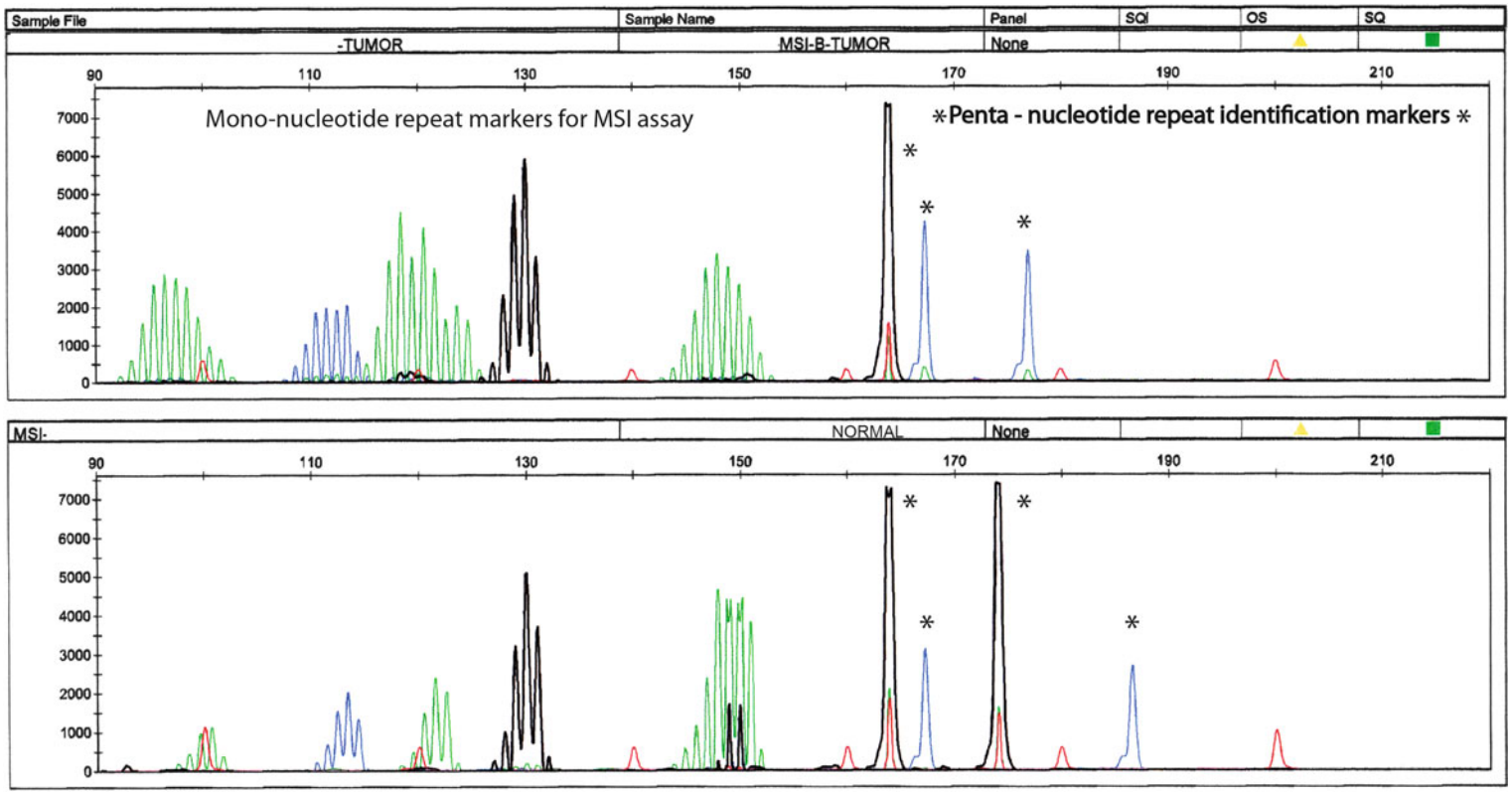


Fig. 41.1 Inclusion of genetic identity markers in a multiplex molecular genetics assay. A commercial (Promega) assay kit for microsatellite instability analysis uses a single tube multiplex PCR reaction which includes primer pairs for five quasimonomorphic mononucleotide microsatellite repeats *and* two primers pairs for pentanucleotide microsatellites. The pentanucleotide STRs are highly polymorphic and can be used to ensure that the DNA from the tumor tissue and the DNA from the normal tissue come from the same patient. (Although the markers are putatively monomorphic, this can still serve as a useful safeguard.) Each panel shows a chromatogram in which the *x*-axis shows the size of the amplicon in base pairs, and the *y*-axis shows the fluorescent signal intensity. In the *top panel*, the sample shows microinstability – high. In the *bottom panel*, DNA from the putative normal tissue does show microsatellite stability *but* the identification markers indicate it comes from a different source than the tumor DNA. Note the bell-shaped, sawtoothed profile of the microsatellites used for MSI assessment: these STR are mononucleotides; even in vitro the polymerase has difficulty making a faithful copy. The pentanucleotide repeats give a single amplicon size for each allele (figure courtesy of Julia Joseph)

41.3 Methods for Specimen Identification

41.3.1 Preanalytical Considerations

41.3.1.1 Specimen Types Suitable For DNA Identification Testing

- Paraffin blocks
  - DNA can be obtained from blocks up to several decades old
  - Some processing methods significantly inhibit recovery and/or PCR
    - Acid decalcification
    - Heavy metal fixatives
- Stained slides
  - Some stains inhibit DNA recovery/analysis. Reports vary concerning inhibition of PCR by H&E. This could reflect variation in the source of hematoxylin, a plant derivative
  - Alcohol-based stains generally permit excellent recovery of DNA
- Blood and bone marrow
  - DNA can be extracted successfully from samples refrigerated up to several months
- Serum/plasma/urine
- Prenatal specimens – includes amniotic cells and chorionic villi samples, analyzed directly or after culturing
- Buccal swabs
  - Useful for obtaining recipient genotypes *after* bone marrow transplantation

41.3.1.2 Amount of DNA Needed

- For unfixed specimens like blood or serum, microsatellite PCR can be performed with as little as 2 ng DNA (300 cells). For engraftment analysis, 20 ng (3,000 cells) provides more robust results for the minority component
- For paraffin sections, the minimum number of cells needed depends on variables which affect DNA recovery and quality like the age of specimen and duration of fixation. Consider a minimum of two sections, each 4 μ thick, with a minimum tissue dimension of 2 mm × 2 mm

41.3.1.3 Additional Prerequisites for Testing

- Specimen misidentification – To demonstrate that a specimen (is/is not) from patient X requires a "gold standard" genotype from patient X. Possible sources are as follows:
  - Concurrent peripheral blood
  - Prior surgical specimen
- Extraneous tissue contamination assessment – It is sufficient to demonstrate that the two tissue pieces show the same or different genotypes
- Genotyping a tissue contaminant could require macrodissection or microdissection

41.3.2 Analytical Targets and Methods

- Short tandem repeats (STRs or "microsatellites") in genomic DNA
- SNPs in genomic DNA
- Chromosomes – Usually limited to detection of X and Y sex-linked chromosomes
  - Phenotype and genotype can be discordant
- Genotype variations *not* in routine use for identification at present
  - (Structural) Copy number variants – heritable absence of large segments of the genome (kilobases to megabases)
  - Short indel polymorphisms (several nucleotides long)

41.3.2.1 Short Tandem Repeats (Microsatellites)

- The most widely established system for identification, typically implied by the term "DNA fingerprinting"
- A microsatellite is defined as a short tandem array of repeats of a simple sequence, typically 1–5 base pairs long
  - The microsatellites used for DNA fingerprinting typically have a 4-nucleotide repeat such as "AGAT"
  - Microsatellites used for "microsatellite instability" testing are typically "mononucleotide" repeats, a single nucleotide repeated many times
  - Repeats with a three nucleotide motif, especially in coding DNA, are

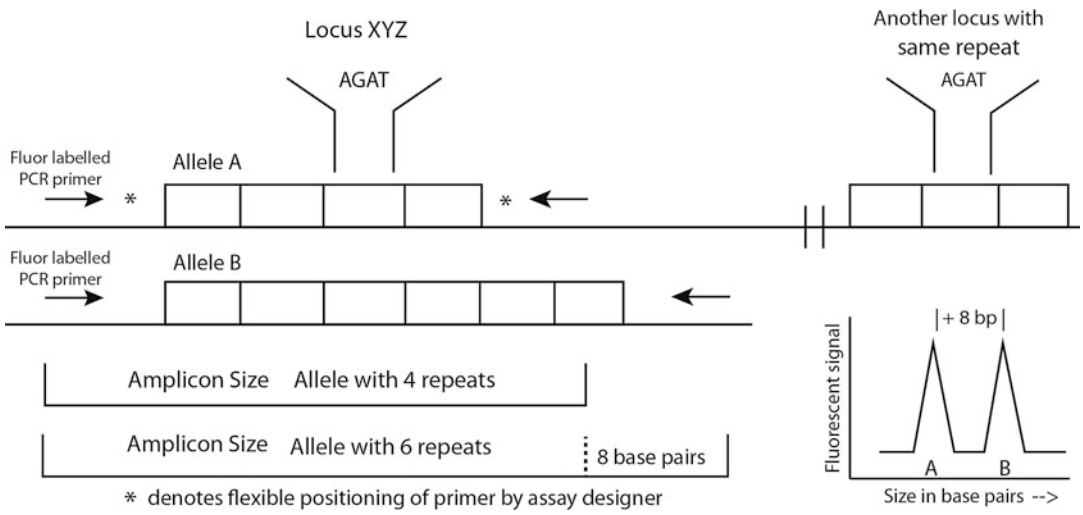


Fig. 41.2 Schematic for the design of a typical assay for identifying microsatellite alleles. Any given microsatellite repeat unit, here the tetranucleotide AGAT, occurs in many microsatellite loci throughout the genome. Here two AGAT microsatellites are illustrated along the same chromosome. The AGAT microsatellite at locus XYZ is specifically amplified by a single primer pair targeted at the unique sequence flanking the repeat. One primer carries a fluorescent tag. The XYZ locus on one chromosome has an allele in which there are four copies of the motif, the other chromosome has an allele with six copies. The PCR primers are the same distance from the ends of the repeat unit in both cases so the final amplicons will differ by the length of two repeats, here eight base pairs. The expected chromatogram is shown in the lower right hand quadrant of the figure (figure courtesy of Julia Joseph)

frequently associated with sudden expansions associated with distinct pathogenic processes

- Microsatellites with a given repeat motif occur at hundreds or thousands of sites throughout the genome. Although the repeat motif is the same, the surrounding DNA is often unique, permitting design of PCR primers to amplify a specific STR locus at a specific position on a chromosome (Fig. 41.2)
- The number of repeats at a given locus is usually stably inherited (see Chap. 40). The number of repeats for a given locus on corresponding chromosome can be the same (homozygous) or different (heterozygous)
- A given number or repeats constitutes an "allele" for that locus
- For a given microsatellite locus, there can be many different alleles in the population. The more possible alleles and the more uniform the distribution through the population, the more informative the locus (the lower the chance of two people sharing a given allele)
- Combined DNA Index System (CODIS) alleles
  - FBI labs identified the STR loci which are most informative across multiple ethnic groups
  - Developed multiplex PCR for up to 16 CODIS (including 13 cores) loci.
  - Common alleles for each locus have assigned numbers (Fig. 41.3)
    - This system permits exchange of information in the absence of sample or sample DNA and is independent of the assay platform
    - Example – A patient status post bone marrow transplant moved to another state. Pretransplant host and donor DNA samples are not available. Engraftment monitoring can continue by using the CODIS allelic profiles
- Also, see Chap. 40, "Molecular Forensic Pathology"
- Method
  - Multiplex amplification requires one primer for each amplicon to be fluorescently labeled

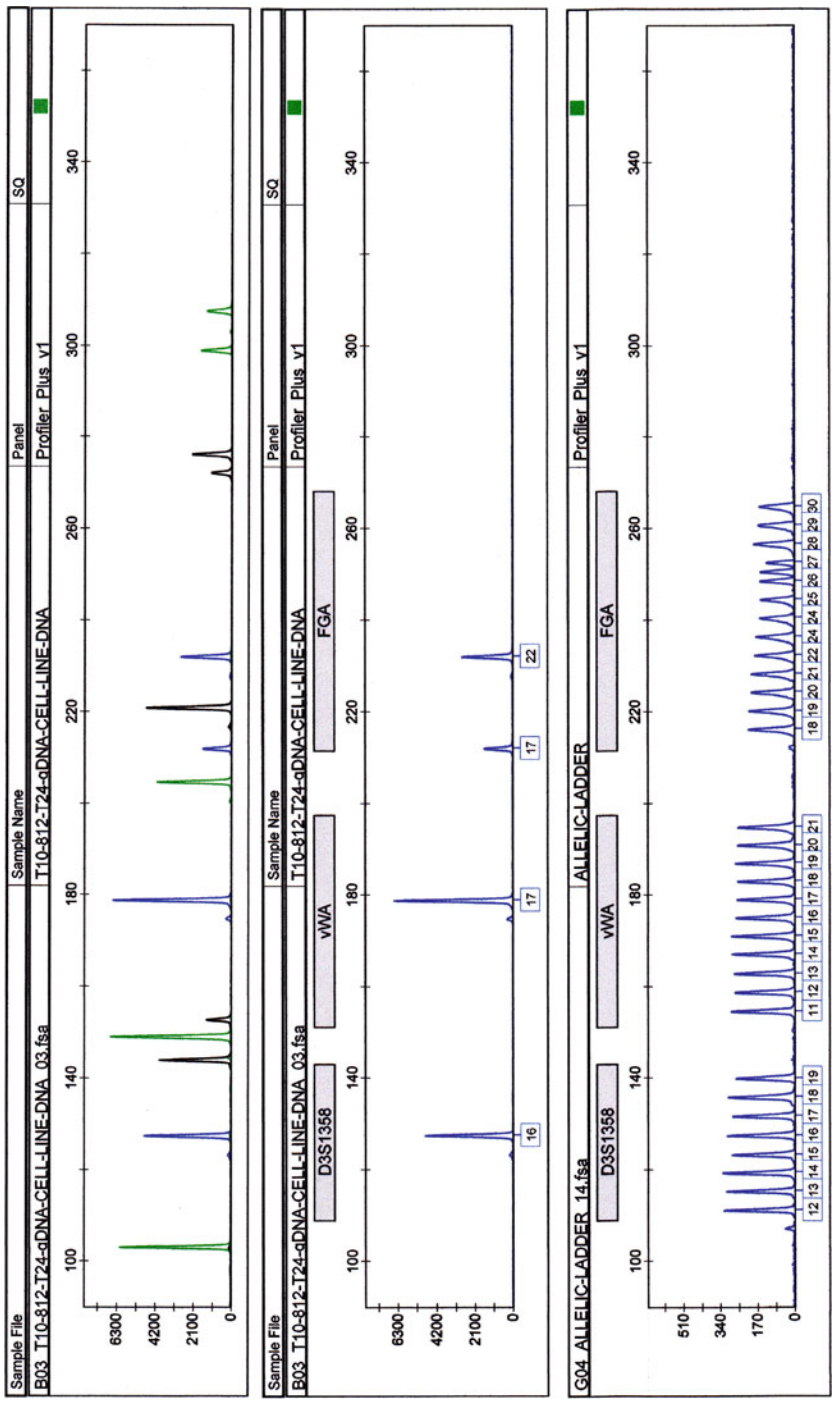


Fig. 41.3 Multiplex microsatellite analysis and assignment of CODIS alleles. As illustrated in Fig. 41.2, the assay designer has flexibility and the two primers are positioned relative to the repeat units thereby controlling the possible amplicon sizes. This allows same color fluorescent tags to be used for several primer pairs in one reaction, as long as the range of amplicon sizes (this is different from the number of repeats) do not overlap for the several loci. Amplicons with identical or similar sizes can be distinguished on capillary electrophoresis by using different fluorescent tags. (*Top panel*) This is a capillary electrophoresis chromatogram generated with the Profiler Plus™ kit (ABI) and GeneMapper™ (ABI) software to verify the identity of the T24 bladder cancer cell line. (*Middle panel*) A set of three loci marked by the same fluor is shown from the *top panel*. Loci names are at the top, the allelic assignments just below the peaks. (*Bottom panel*) This is part of the allelic ladder control showing the most frequent alleles for the three loci in the *middle panel* (figure courtesy of Julia Joseph)

  - The PCR amplicons are analyzed using a capillary electrophoresis analyzer (DNA sequencer) and fragment-size analysis software
  - Alleles and loci are distinguished by BOTH the size of product in base pairs and color of fluorescent label
  - Qualitative analysis of individual microsatellite loci can be done with unlabelled PCR primers and polyacrylamide gel electrophoresis
  - Commercial kits are available which implement single tube multiplex reactions (up to 16 CODIS loci)
  - Sequence information is available for all CODIS loci primers
  - Laboratory development of such a multiplex assay is not trivial
  - Each run should include an "allelic ladder," a kit-specific mixture of amplicons of the common alleles. This facilitates assignment of alleles to samples. The "rungs" of the ladder are PCR products, a potential source of contamination
- Interpretation qualitative comparison of alleles for sample identification should be sufficient, provided sufficient DNA is available to permit adequate amplification
- Pitfalls
  - Degraded DNA, especially from FFPE
    - DNA might be too degraded to amplify long amplicons
    - Possible solutions include:
      - Limit analysis to CODIS loci with short amplicon size ranges
      - Minifiler™ (short amplicon) STR panels are commercially available
      - SNP analysis –Because analysis does not depend on size of the amplicon, the amplicons are usually short
  - Low DNA concentrations – This can lead to random (stochastic) loss of alleles at any locus for DNA from fresh or fixed tissue
    - This can be mistaken for evidence of nonidentity
  - Loss of heterozygosity (loss of an allele) resulting from cytogenetic alterations
    - This can be mistaken for evidence of nonidentity in interpretation
  - In a specimen misidentification, the proportion of nonshared alleles should be so high that loss of one or two alleles will not preclude analysis
  - In bone marrow transplants from closely related donors, loss of even one informative allele might limit engraftment analysis

41.3.2.2 Single Nucleotide Polymorphisms

- Multiple methods are available for SNP detection (see Chaps. 8 and 10)
- Amplicons can be much shorter than for STR analysis
- Informativity
  - Usually only two alleles (choice of nucleotide) are known for any given SNP compared to 10–20 alleles at each STR locus so informativity of one SNP is much less than that of an STR locus
  - Estimates vary from 10 to 50 for number of SNPs needed for a "universal" panel. Bone marrow transplantation often involves related donors, and so could require a larger panel for discrimination
  - SNP panels have not been standardized to the same extent as the CODIS panels
- Methods – Real-time PCR SNP analysis
  - Numerous real-time PCR methods such as dual hydrolysis probes are suitable (Fig. 41.4) and can show sensitivity down to the 2% allelic level
  - Non–real-time methods such as PCR followed by restriction enzyme digestion then agarose gel electrophoresis, but these are impractical
  - Some allele-specific amplification methods adapted to real-time PCR show sensitivity to 0.1% or 0.01% but this is unnecessary for analysis of misidentified specimens and not of proven clinical utility for engraftment analysis

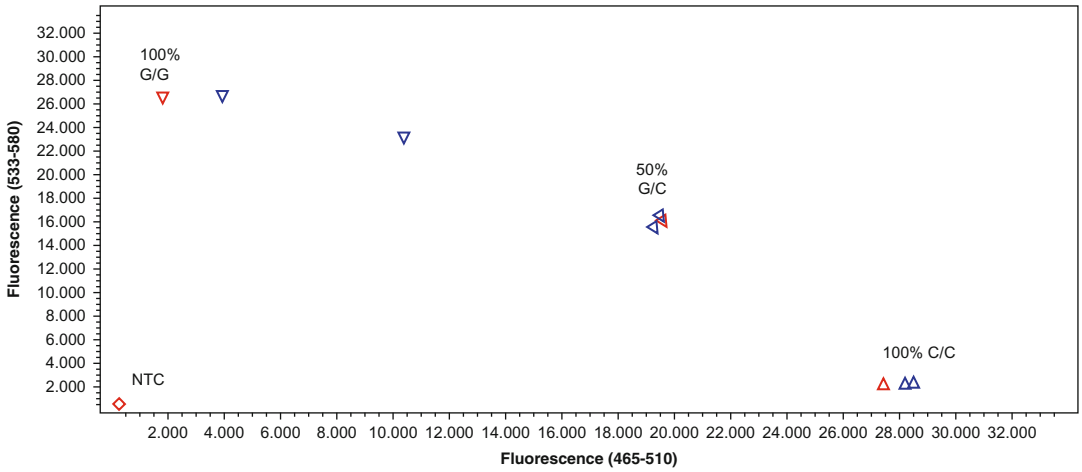


Fig. 41.4 Two-color real-time PCR SNP assay. The assay shown uses a dual probe hydrolysis format. The probe detecting the "G" allele (Y axis) and the probe detecting the "C" allele (X axis) carry different fluors. Analysis is by endpoint fluorescence at cycle 40. This does *not* take full advantage of real-time PCR but because it is a competitive assay with both probes present in the same reaction throughout, meaningful quantitative results can be obtained from the "plateau phase." *NTC* no template control (figure courtesy of Julia Joseph)

  - Advantages
    - Does *not* require expensive capillary electrophoresis analyzer
    - Assay faster than microsatellite analysis
    - More successful than STR analysis with degraded samples
  - Limitations
    - Cannot multiplex in a single tube
    - SNP genotypes can change as a result of mutation or cytogenetic change
  - Pitfalls
    - If two or more individual SNP assays are run individually, mislabeling of the reactions could be overlooked in the results since typically every assay uses the same fluors, unlike STR analysis in which each allele/locus has a distinct size and fluor combination
    - Running a fixed panel, preferably in a plate format, for every engraftment sample will reduce error (and throughput)
- Methods – SNPs by single base extension (SBE)
  - PCR amplify a region surrounding the SNP of interest. After purification, the product is denatured and hybridized with an "extension" primer which stops just before the SNP site. The extension reaction mix uses dideoxynucleotides but *no* deoxynucleotides so the polymerase can only extend the primer by a single base, which identifies the SNP
  - Mutations can be analyzed in a highly multiplexed assay
  - Each extension primer in a multiplex reaction is designed to have a unique length
  - If both tumor and normal tissue from a patient are analyzed, inclusion of reactions for several informative SNPs would permit confirmation that results for the tumor and the normal specimen belong to the same patient
  - SBE analysis by capillary electrophoresis:
    - Multiplex PCR of at least 30 SNP assays
    - Targets distinguished by length of extension primer
    - Dideoxynucleotides are fluorescently labeled
  - SBE analysis by MALDI-TOF mass spectrometry
    - MALDI-TOF can determine which base was added because each of the four dideoxynucleotides has a distinct molecular weight
    - Targets are distinguished by total mass of each extension primer

- Future methods – Next generation sequencing (NGS)
  - SNP and CODIS STR data "fall out" of whole genome sequencing
  - SNP data also result from exome and transcriptome sequencing
  - Comparison with standard STR or SNP analysis from additional (normal) tissue or peripheral blood could confirm identification

41.3.3 Medicolegal Considerations

- If DNA fingerprinting analysis of a potential mixup is requested it is advisable to initiate testing only after consultation with the medicolegal or risk management group
- An "informal" conclusion can become part of subsequent fact-finding
- If tissue is limited, processing might interfere with subsequent investigation

41.4 Clinical DNA Chimerism Analysis

41.4.1 Analysis for Hematopoietic Stem Cell Transplantation

- Role of bone marrow/ hematopoietic stem cell (HSCT) transplantation
  - Treatment of hematopoietic malignancy
  - Treatment of genetic diseases such as immunodeficiencies
- Outcomes of bone marrow/hematopoietic stem cell transplantation
  - "Chimerism" – Complete replacement of recipient blood cells by donor cells
  - "Partial chimerism" or "incomplete chimerism" – An admixture, sometimes maintained for years, of recipient and donor cells. This is worrisome after leukemia/ lymphoma because of recurrence risk
  - "Split chimerism" – The extent of replacement can vary among cell lineages
  - A low level of recipient cells for a short period after transplantation for malignancy can be beneficial. This is hypothesized to stimulate a "graft v leukemia" effect
- Donor considerations
  - It is important to verify that informative alleles are present before transplantation
  - The use of related donors can make engraftment analysis difficult because many alleles are shared (noninformative)
  - Cord blood donor cells are associated with lower rates of graft versus host reaction than are adult donor cells *but* cord donors often provide too few cells to protect adult recipient from infection soon after transplantation
  - "Haplo-cord" transplantation – Donor stem cells from an adult donor and from one or more cord blood donors are infused together. The numerous adult cells protect from infection while the cord blood donor cells expand and eventually replace both remaining recipient cells and adult donor cells
  - The use of two or more donors reduces the number of informative alleles
- Role of engraftment analysis
  - Hematopathological analysis cannot distinguish donor from recipient cells; consequently, it cannot assess progress of the transplantation barring recurrent disease,
  - In multidonor transplantation, quantify each donor contribution – important for donor selection for lymphocyte infusion and/or subsequent retransplantation
  - Hematopoietic subsets engraft at different rates, significance is controversial
    - Subsets can be analyzed after positive or negative selection
    - Single-marker selection (immunomagnetic beads, rosettes) on a small scale is feasible in a molecular laboratory
    - Selection methods (e.g., some large commonly used magnetic beads) complicate confirmation of purity by flow cytometry

    - Subset analysis can be important in specific settings, for example,
        - Erythroid lineage s/p transplantation for sickle cell disease
        - Assessing success of donor lymphoid or NK cell infusion
- Sample considerations
    - Pretransplant considerations
        - Ideally, samples for recipient and donor genotyping should be provided to the laboratory prior to transplantation
        - Recipient and donor samples should not be drawn on same day (and certainly not processed on the same day), unless gender disparate
        - Recipient and family members with the same name (especially parent–child combinations) *do* occur
        - Twins – Phenotypically identical twins are not necessarily genotypically identical so initial genotyping is appropriate
    - Sample type – The proportion of donor in peripheral blood and in bone marrow are usually similar but results should be tracked separately
- Methods and data analysis
    - STR analysis of CODIS loci is standard
    - Multiple commercial CODIS panels are available
    - Linearity over the range 5–95% (detection of the minor component)
    - Limit of detection in the range of 2–3% is usually clinically sufficient
    - SNP or indel profiles would "work," but are not standardized among labs
    - From the recipient and donor genotypes, identify which alleles are informative (unique) for each source
    - For a given locus at which the signal is *not* saturating, the signal peak height (or peak area) is proportional to the contribution of that allele at that locus
    - Calculate % donor (or recipient) for each locus separately
    - For a given locus, sum the signal peak heights (or areas) for all informative donor alleles (OR recipient), then divide by the total peak height (area) from all informative allele (donor AND recipient) sources to give the proportion of donor (recipient), being careful to use the *same* number of informative alleles from each source
    - Multiple patterns are possible (Fig. 41.5)
    - Calculate percent of each donor (recipient) for every informative locus
    - Report the average donor (or recipient) over all informative loci (Fig. 41.6)
    - Consult the clinical service for preference with respect to reporting % donor or % recipient, reporting both risks confusion
- Pitfalls
    - Stutter peaks
        - An amplicon one repeat unit shorter than the main amplicon formed as a reproducible "artifact" of polymerase slippage (Fig. 41.5)
        - Stutter signal is typically 2–10% of that of the main amplicon
        - If it overlies an informative allele, it limits the sensitivity of detection but the % stutter is reproducible so an estimated background can be judiciously subtracted if other supporting loci are available
    - Donor genotype available, recipient genotype not available
        - Tenative analysis is usually feasible if the donor is unrelated, alleles absent in the donor may be imputed to the recipient
        - Family members can differ by as little as one allele; consequently even if the posttransplant genotype matches that of the donor, a low level of an informative recipient allele(s) might be obscured by a stutter peak
        - Request a buccal swab. A buccal swab can include a significant donor component from inflammatory cells but recipient alleles *should* be dominant
    - Donor sample not available, recipient genotype available
        - Similar considerations as in the preceding discussion

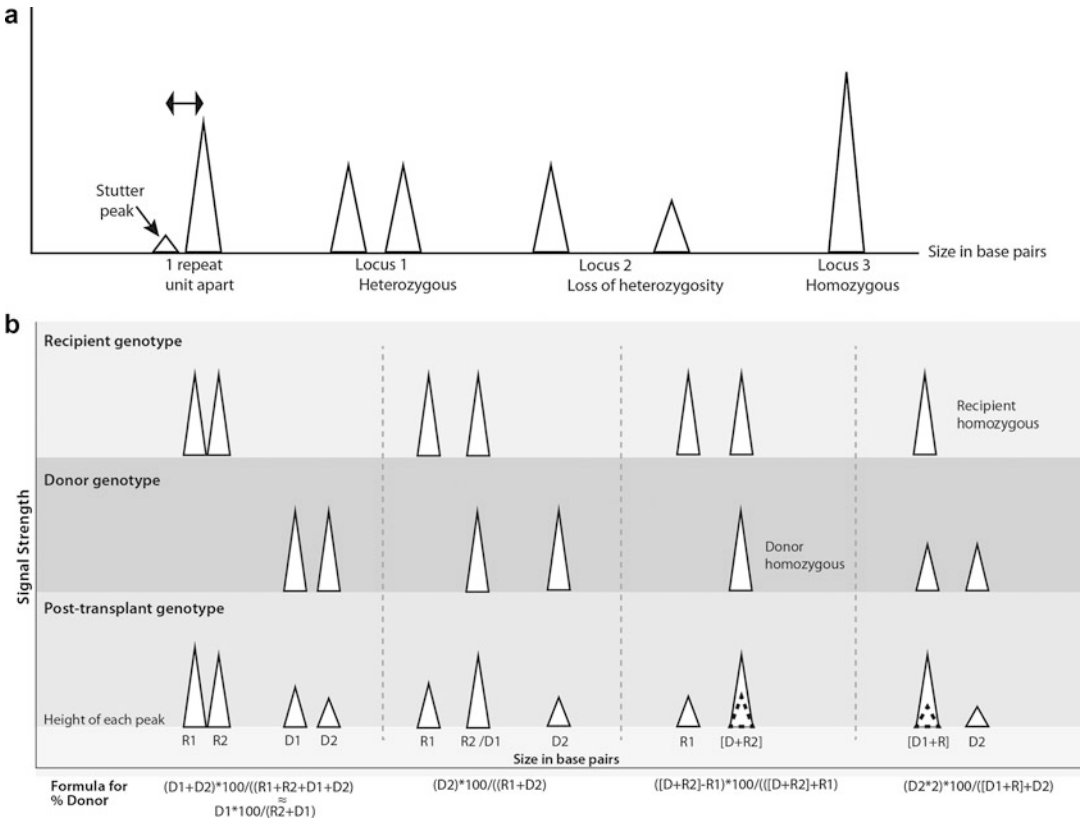


Fig. 41.5 Representative patterns of microsatellite amplicons. (*Top panel*) There is modest variation in signal intensity as a result of technical factors, beyond the presence of one, two, or zero copies of an allele. In general, the longer the amplicon, the lower the signal intensity. An SNP in a primer binding site can weaken *or* abolish the signal and simulate loss of heterozygosity. If a single amplicon is present and the signal strength is closer to that of alleles for the nearest heterozygous locus, rather than double the signal, it suggests the "homozygous" locus reflects loss or amplification failure of the other allele. (*Bottom panel*) This illustrates *some* of the possible combinations of recipient and donor alleles for a given locus and suggests formulae for calculating % donor. In combinations in which the recipient and donor share an allele, it can be helpful to visualize the component of the minority partner (shown here as a *triangle* with a *dashed* border). It is sometimes necessary to estimate the signal for an obscured (shared) allele with the value of the clearly displayed allele from the same source (figure courtesy of Julia Joseph)

- At least two informative loci should be analyzed
  - Because recurrent leukemia/lymphoma can show cytogenetic changes (usually loss) abolishing a CODIS locus, reliance on a single locus could lead to erroneous overestimation of the donor component
  - Example: Four loci showed an average of 95% donor. A fifth locus, D7S820, showed 90% donor. FISH analysis showed monosomy 7 consistent with recurrent AML

41.4.2 Maternal Cell Contamination

- Prenatal genetic diagnosis usually analyzes DNA from cells in amniotic fluid or chorionic villi
- Maternal cells are frequently present and can limit sensitivity of detection of an abnormality. If not recognized, maternal cell contamination (MCC) can lead to an incorrect diagnosis
- Cells can be analyzed for MCC directly or after culture
- MCC is usually lower in cultured samples; most maternal cell types do not proliferate in cell culture

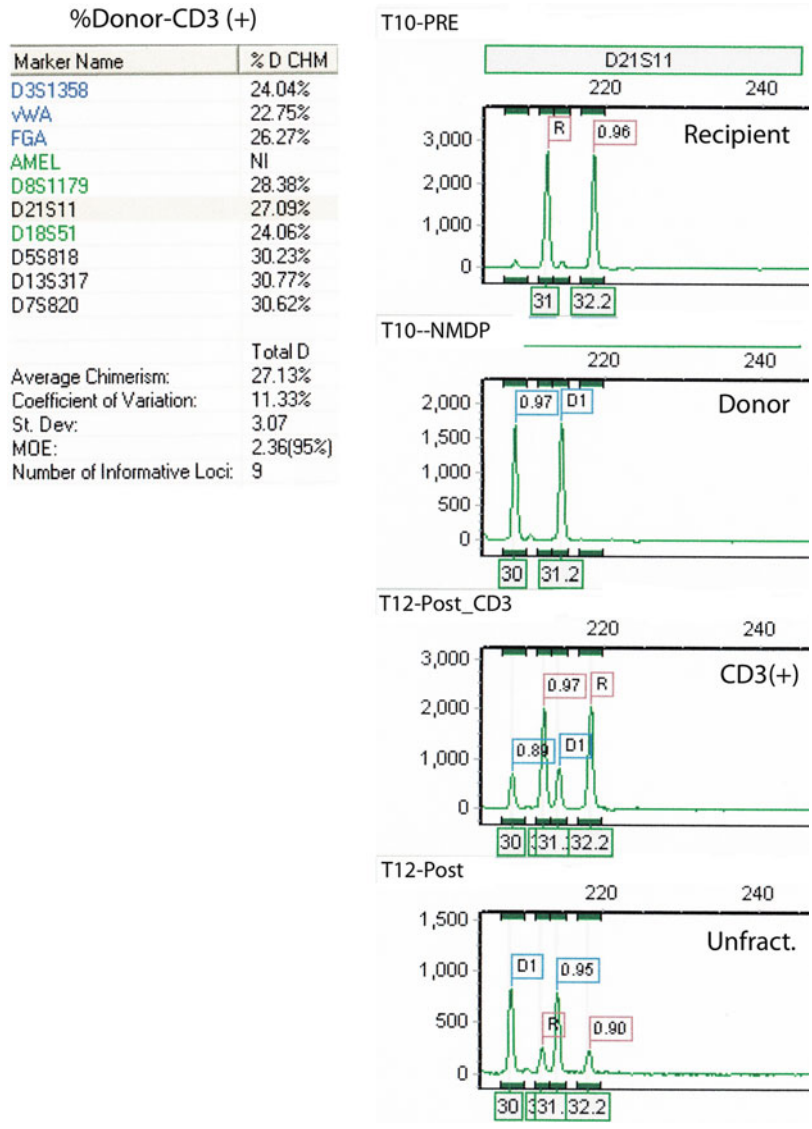


Fig. 41.6 Example of chimerism analysis in a posthematopoietic stem cell transplantation patient. The *top panel* shows the recipient genotype, the *second panel* shows the donor genotype, the *third panel* shows the mixed genotypes in the CD3(+) cells of a posttransplant bone marrow, the *bottom panels* show the mixed genotypes in the unfractionated bone marrow. Only the D21S11 locus is shown; displaying an entire DNA "fingerprint" raises a possible HIPAA issue, especially if the patient has a profile in a nonclinical database. Analysis was done automatically using ChimeraMarker ™(Softgenetics), the results for all loci in the unfractionated preparation (*bottom panel*) are shown, 70% for the D21S11 locus. Analysis of the CD3(+) prep shows 27% donor. The software will also store results by patient for trend analysis

- Methods/analysis
  - Microsatellite analysis is the dominant approach
  - Maternal specimen genotype *required* for comparison
  - Paternal specimen *usually* not required
  - Guidelines exist for preanalytical, technical, interpretive, and reporting stages; recommendations include the following:
  - All prenatal samples for genetic testing should have MCC tested
  - At least two informative loci should be analyzed; this typically requires testing 7–10 loci
  - The level of acceptable MCC (% of cells) depends on the intended test
- Pitfalls
  - If multiple passes are used to acquire amniotic fluid, each for separate tests, each sample should be analyzed for MCC or pooled prior to any test
  - Multiple gestations – Analyze each gestation separately for MCC

  - Surrogacy – If surrogacy status is not communicated, the fetal genotype could display a puzzling lack of the expected maternal alleles
  - Spontaneous change of an allele in meiotic transmission is uncommon but well described (See Chaps. 5 and 40)
  - Cytogenetic aberrations
    - Loss of a chromosome which carries a locus in the STR panel could appear as loss of an allele
    - Gain of a chromosome including an STR locus will appear as allelic signal imbalance, *not* as an additional allele
- Future directions
  - Microarrays
    - Cytogenetic microarrays which include SNP probes can detect at least a moderate level of MCC
  - Arrays including SNP probes have been reported to detect fetal chromosomal abnormalities in the presence of MCC, at the risk of increased false negatives
  - Next-generation sequencing
    - Preparations of fetal DNA circulating in maternal plasma *always* has a contaminating maternal component
    - The primary proposed application is for chromosome copy number analysis
    - Detection of translocations and point mutations could become feasible

41.4.3 Gestational Trophoblastic Disease

- Different subclasses can show different courses and require different therapies
- Best distinguished with combined clinical, histopathologic, *and* genetic information
  - DNA fingerprinting is needed to determine maternal and paternal genetic contributions
  - Cytogenetic or FISH analysis of the X and Y chromosomes will not identify the origin of the X chromosome(s) even when a Y (clearly paternal) chromosome is present
  - Maternal genotype profile at a minimum is needed for comparison
  - Paternal/androgenetic source might be uncertain, especially in choriocarcinoma, which can show latency of more than 10 years

41.4.3.1 Hydatidiform Mole: Proliferation of Villous Trophoblasts

- Complete hydatidiform mole (CHM)
  - Usual pattern – Diploid or tetraploid, androgenic (paternal) genotype only
  - Duplication of one sperm genome (90% of cases), or simultaneous fertilization by two sperms
  - Mosaicism underappreciated could account for uncommon cases with apparently biparental diploid genotypes
  - Even more rare familial forms of diploid biparental CHM described
  - Significantly increased risk of choriocarcinoma, after latency of years
- Partial hydatidiform mole (PHM) – Uncommon
  - Usual pattern – Triploid (Diandric monogynic = two paternal chromosome sets, one maternal chromosome set) resulting from fertilization of one ovum by two sperms (dispermic) or from one sperm (monospermic) followed by duplication of the paternal genome
  - No increase in risk of choriocarcinoma
  - Nonmolar pregnancies can show "triploid" DNA content but are digynic
- Invasive hydatidiform mole
  - Can follow either CHM or PHM

41.4.3.2 Choriocarcinoma and Other Placental Tumors

- Choriocarcinoma – Gestational
  - Most preceded by CHM (50%), but also can follow pregnancy or abortion
  - High cure rate
- Choriocarcinoma – Nongestational
  - Worse prognosis than gestational choriocarcinoma
  - Genotype is diploid and consistent with the patient's normal genotype

- Placental site trophoblastic tumor (PSTT)
  - Majority benign
  - Usually diploid BUT requires paternal X chromosome and absence of Y
- Epithelioid trophoblastic tumor (ETT)
  - Diandric, monogynic. Genotype can be important in histologic distinction from carcinoma

41.4.4 Tumors Arising in Transplanted Organs

- Uncommonly a tumor will arise in a transplanted tissue or marrow (e.g., hepatocellular carcinoma in a liver transplanted into a patient treated for hepatocellular carcinoma)
- This could represent reseeding by long dormant disseminated tumor cells or a de novo donor neoplasm
- Years after transplantation, the organ will not only contain recipient hematopoietic cells but may also have recipient cells providing the reticuloendothelial framework
- The methods of analysis described above can be applied but the fraction of non–liver recipient cells in the tissue specimen can be hard to estimate

41.4.5 Biobanking

- Traceability of samples is essential (even for anonymized research samples)
- Confirmation of random sample identity can be instituted as a QA/QC measure
- DNA identification by analysis of STRs or SNPs is suitable
- Traceability requires an aliquot of blood stored at time of collection

41.4.6 Cell Line Identification

- Misidentification of cell lines in research is a persisting problem
- Some journals require confirmation of cell line identity using the CODIS alleles
- Many CODIS fingerprints are available online (e.g., www.atcc.org)

Further Reading

Beadling C, Heinrich MC, Warrick A, et al. Multiplex mutation screening by mass spectrometry evaluation of 820 cases from a personalized cancer medicine registry. J Mol Diagn. 2011;13:504–13.

Buza N, Hui P. Gestational trophoblastic disease: histopathological diagnosis in the molecular era. Diagn Histopathol. 2010;16:526–37.

Cao D, Hafez M, Berg K, Murphy K, Epstein JI. Little or no residual prostate cancer at radical prostatectomy: vanishing cancer or switched specimen? A microsatellite analysis of specimen identity. Am J Surg Pathol. 2005;29:467–73.

Cardoso S, Valverde L, Odriozola A, Elcoroaristizabal X, de Pancorbo MM. Quality standards in biobanking: authentication by genetic profiling of blood spots from donor's original sample. Eur J Hum Genet. 2010;18:848–51.

Fordyce SL, Avila-Arcos MC, Rockenbauer E, et al. High-throughput sequencing of core STR loci for forensic genetic investigations using the Roche Genome Sequencer FLX platform. Biotechniques. 2011;51:127–33.

Francis DL, Prabhakar S, Sanderson SO. A quality initiative to decrease pathology specimen–labeling errors using radiofrequency identification in a high-volume endoscopy center. Am J Gastroenterol. 2009;104:972–5.

Gephardt GN, Zarbo RJ. Extraneous tissue in surgical pathology: a College of American Pathologists Q-probes study of 275 laboratories. Arch Pathol Lab Med. 1996;120:1009–14.

Gineikiene E, Stoskus M, Griskevicius L. Single nucleotide polymorphism-based system improves the applicability of quantitative PCR for chimerism monitoring. J Mol Diagn. 2009;11:66–74.

Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. J Am Med Assoc. 2001;286:415–20.

Huijsmans R, Damen J, van der Linden H, Hermans M. Single nucleotide polymorphism profiling assay to confirm the identity of human tissues. J Mol Diagn. 2007;9:205–13.

Kohn JT, Corrigan JM, Donaldson MS. To err is human: building a safer health system. Washington, DC: National Academy Press; 1999.

Layfield LJ, Anderson GM. Specimen labeling errors in surgical pathology: an 18-month experience. Am J Clin Pathol. 2010;134:466–70.

Lichter P, Allgayer H, Bartsch H, et al. Obligation for cell line authentication: appeal for concerted action. Int J Cancer. 2010;126:1.

Lou JJ, Andrechak G, Riben M, Yong WH. A review of radio frequency identification technology for the anatomic pathology or biorepository laboratory: much promise, some progress, and more work needed. J Pathol Inform. 2011;2:34.

Makary M, Epstein J, Pronovost P, Millman E, Hartmann E, Freischlag J. Surgical specimen identification errors: a new measure of quality in surgical care. Surgery. 2007;141:450–5.

Mulero JJ, Chang CW, Lagacé RE, et al. Development and validation of the AmpFISTR MiniFilerTM PCR amplification kit: a MiniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. J Forensic Sci. 2008;53:838–52.

Nagan N, Faulkner NE, Curtis C, Schrijver I. Laboratory guidelines for detection, interpretation, and reporting of maternal cell contamination in prenatal analyses a report of the association for molecular pathology. J Mol Diagn. 2011;13:7–11.

Nakhleh RE, Zarbo RJ. Surgical pathology specimen identification and accessioning: a College of American Pathologists Q-Probes study of 1,004,115 cases from 417 institutions. Arch Pathol Lab Med. 1996;120:227–33.

Pfeifer JD, Zehnbauer B, Payton J. The changing spectrum of DNA-based specimen provenance testing in surgical pathology. Am J Clin Pathol. 2011;135: 132–8.

Raff LJ, Engel G, Beck KR, O'Brien AS, Bauer ME. The effectiveness of inking needle core prostate biopsies for preventing patient specimen identification errors: a technique to address Joint Commission patient safety goals in specialty laboratories. Arch Pathol Lab Med. 2009;133:295–7.

Renshaw A, Kish R, Gould E. The value of inking breast cores to reduce specimen mix-up. Am J Clin Pathol. 2007;127:271–2.

Schrijver I, Cherny SC, Zehnder JL. Testing for maternal cell contamination in prenatal samples. J Mol Diagn. 2007;9:394–400.

Srebniak M, Boter M, Oudesluijs G, et al. Application of SNP array for rapid prenatal diagnosis: implementation, genetic counselling and diagnostic flow. Eur J Hum Genet. 2011;19:1230–7.

Circulating Tumor Cells: Liquid Biopsy for Molecular Pathology

42

Jeffrey S. Ross

Contents

42.1 Overview of Circulating and Disseminated Tumor Cells 1066

42.2 History of CTC Discovery and Development 1066

42.3 CTC/DTC Detection Methods 1066
42.3.1 CellSearch® 1069
42.3.2 The AdnaTest 1070
42.3.3 Membrane Microfilter Assay 1071
42.3.4 Microfluidic CTC Chip 1072
42.3.5 Fiberoptic Array Scanning Technology (FAST) 1072
42.3.6 Epithelial Immunospot (EPISPOT) 1072
42.3.7 Laser Scanning Cytometry (LSC) 1072
42.3.8 Differential Centrifugation 1073
42.3.9 Other Homebrew Techniques 1074
42.3.10 Morphologic Versus Molecular Detection of CTC 1074

42.4 CTC Platform Comparisons 1075

42.5 CTC Versus DTC 1075

42.6 Prognostic and Predictive Significance of CTC 1076

42.7 CTC/DTC Versus Positron Emission Scanning (PET-CT) 1076

42.8 Molecular Assessment of Captured Cells 1077

42.9 CTC and Next-Generation Sequencing (NGS) 1078

42.10 Summary and Future Clinical Implications 1078

Further Reading 1079

J.S. Ross, MD
Department of Pathology and Laboratory Medicine,
Albany Medical College, Albany, NY, USA

Foundation Medicine, Inc., Cambridge, MA, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_42,

42.1 Overview of Circulating and Disseminated Tumor Cells

- Cancer is the leading cause of death worldwide, with lung, stomach, liver, colon, and breast cancers accounting for the majority of mortalities
- Despite advances in early detection and treatment, there are still patients that die of disease even if presenting at an early stage most often caused by the presence of minimal residual disease (MRD) surviving after completion of primary treatment
- Recent advances in immunocytochemical and molecular techniques have enabled the detection of circulating tumor cells (CTC) in bloodstream or disseminated tumor cells (DTC) in bone marrow in all stages of cancer
- It is now widely held that continuing advancement in CTC/DTC detection techniques will enable the performance of sophisticated molecular testing on captured cells that will serve as a major guideline to the selection of targeted therapies for many cancer patients in the future

42.2 History of CTC Discovery and Development

- CTC were first described in the scientific literature by Ashworth in 1869
- In 1889, Paget described his theory of "seed and soil" based on his observations of metastatic breast cancer
- In the mid-1990s, the scientists were restricted by the sole utilization of CTC morphology as a selection criterion which greatly restricted progress and clinical interest in CTC
- In 1991, Smith et al. detected tyrosinase mRNA in peripheral blood of melanoma patients using RT-PCR
- The introduction of IHC staining simultaneously enabled detecting micrometastases in bone marrow and lymph nodes in breast cancer patients which was correlated with therapy response and disease outcome
- The discovery of the immunomagnetic cell capture system then led to the clinical assessment blood-based detection of CTC of numerous solid tumors, including breast, lung, colorectal, gastric, esophageal, pancreatic, prostatic, gallbladder, head and neck, bladder, ovarian, and cervical cancers
  - The peripheral blood and bone marrow have long been considered the ideal matrix to search for CTC, as they are the natural route of spread of malignant neoplasms and are easily accessible by routine techniques
  - Although the peripheral blood and bone marrow normally do not contain significant numbers of epithelial cells enhancing the ease of detection of carcinomas in these samples, the extremely low concentration of malignant epithelial cells at about 1 in 10^6–10^7 total nucleated cells, even in patients with widespread metastatic disease and heavy tumor cell burdens, greatly impacts the sensitivity of the various techniques developed to detect them

42.3 CTC/DTC Detection Methods

- A variety of methods have been developed to detect CTC in blood and DTC in bone marrow of breast cancer patients (Table 42.1)
- In order to enhance the identification of CTC/DTC present in very low concentrations, enrichment methods, including differential centrifugation, ficoll enrichment, and cell separation by immunomagnetic techniques, have been the cornerstone of both commercially developed and early-stage technologies
  - A major issue confronting all approaches is the problem of malignant cell loss due the inherent fragility of both CTC and DTC
- The positive detection of CTC/DTC have utilized a variety of methods including
  - Immunohistochemistry (IHC)
  - Immunofluorescence (IF)
  - Fluorescence in situ hybridization (FISH)
  - Flow cytometry (FACS)
  - Southern blot
  - Northern blot

Table 42.1 Methods of detection of CTC and DTC in breast cancer

Test name	Capture technology	Identification method	Test result	Commercial or academic source	FDA approval status	Commercial status	Biomarker detection strategy
CellSearch®	Immunomagnetic Bead Cell Capture using anti-EpCAM selection	EpCAM for capture Cytokeratins for positive identification CD45-negative selection	Microscopic Cell Count	Veridex, LLC. (Ortho Clinical Dx, Inc, Raritan, NJ)	Fully Approved at the Premarket Approval Level (PMA) for breast, colorectal and prostate cancer	On the market as both decentralized locally provided test and centralized in a US National Laboratory (Quest Diagnostics, Inc.)	HER2 status has been reported using FISH to measure HER2 gene amplification
AdnaTest BreastCancerSelect	Immunomagnetic Bead separation using a proprietary mixture of 3 antibodies to epithelial surface antigens RT-PCR	Epithelial antigen GA733-2 (mRNA) HER2 (mRNA) MUC1 (mRNA) Actin for platelet-negative selection	Relative CTC-associated mRNA levels	AdnaGen AG, Langenhagen, DE (Marketed in the US by Onco-Vista, Inc., San Antonio, CA)	Review process started in May, 2008	In development	HER2 status has been reported by RT-PCR determination of HER2 relative mRNA levels.
Membrane Microfilter Assay	Membrane filter capture on-filter electrolysis RT-PCR	To be determined	Relative CTC-associated mRNA levels	University of Southern California, Los Angeles, CA	Not currently under review	In development	Not developed to date

(*continued*)

Table 42.1 (continued)

Test name	Capture technology	Identification method	Test result	Commercial or academic source	FDA approval status	Commercial status	Biomarker detection strategy
Microfluidic CTC Chip	Antibody coated microposts Laminar microfluidics	EpCAM for capture Fluorescent nuclear stain and cytokeratin for positive identification CD45 for-negative selection	Microscopic Cell Count	Massachusetts General Hospital, Harvard Medical School, Boston, MA	Not currently under review	In development	Not developed to date
FAST	Fiber optic array scanning Immunofluorescence	Cytokeratins	Microscopic Cell Count	Scripps-PARC Institute for Advanced Biomedical Science, La Jolla, CA	Not currently under review	In development	Not developed to date
EPISPOT	Enzyme linked immunospot assay CTC cell culture	Muc1 Cytokeratin 19 CD45 for-negative selection	CTC epithelial-associated protein production in culture	University Medical Center Hamburg-Eppendorf, DE	Not currently under review	In development	Can profile proteins secreted by cultured CTC
LSC	Laser scanning cytometry Immunofluorescence	EpCAM CD45 for-negative selection	Microscopic Cell Count	Friedrich Schiller University, Jena, DE	Not currently under review	In development	Not developed to date

Table reprinted with permission from the *American Journal of Clinical Pathology*

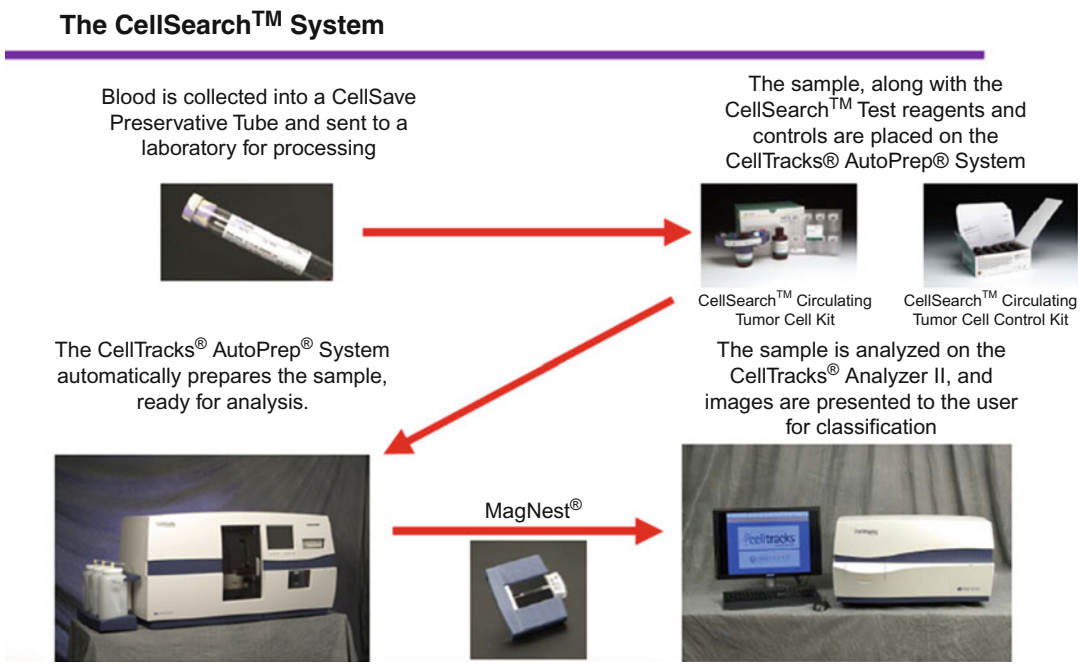


Fig. 42.1 The CellSearch system. The figure outlines the CellSearch procedure for determining CTC (Reprinted with permission from *Am J Clin Pathol*)

  - PCR
  - RT-PCR
  - Multiplex RT-PCR
  - Cell culture
  - Proteomic techniques
- For breast cancer, a variety of individual and multiplex biomarkers for CTC/DTC identification have been used including cytokeratins (CK), MUC1, GA733-2, mammaglobin, maspin, and CEA
- In addition, several studies and approaches have also used biomarkers such as CD45 to exclude circulating leukocytes and actin to exclude platelets

42.3.1 CellSearch®

- The CellSearch system (Veridex, LLC, Raritan, NJ) is a semiautomated device for detection of circulating tumor cells expressing the epithelial cell adhesion molecule (EpCAM) with antibody-coated magnetic beads as an enrichment method (Fig. 42.1)
  - The cells are then labeled with fluorescent monoclonal antibodies specific for
    - Leukocytes (CD45)
    - Epithelial cells (CK8, CK18, CK19)
  - Slides are then analyzed for fluorescently labeled CTC using an automated microscope
  - CTC are defined as CK+/CD45–
  - Nucleated cells are subsequently counted using the CellTracks® Analyzer
  - This method is now FDA approved for detection of CTC in patients with metastatic breast, colon, and prostate cancers
- The initial study in a large breast cancer cohort revealed that a high CTC count at the time of diagnosis was a significant adverse stand-alone prognosis factor
  - In this study, if selected chemotherapy given to a patient with a positive CTC assay in the metastatic breast cancer setting failed to reduce the number of CTC below five cells per 7.5 mL of blood, the disease would undoubtedly progress while under the selected treatment and the patient's prognosis was grave

  - Recently, the detection of five or more CTC in 7.5-mL blood in 43 of 83 (52%) in a series of patients with stage IV metastatic breast cancer prior to the initiation of first-line chemotherapy confirmed the original observations that CellSearch was highly predictive for progression-free and overall survival and provided more helpful information than conventional imaging procedures
  - A number of additional institutions have confirmed these results which has led to a randomized prospective clinical trial designed to validate that changing therapy in the face of a failure to reduce CTC levels below five cells per 7.5 mL of blood could result in an improved patient outcome
- The Southwest Oncology Group (SWOG) 0500 Trial Design
  - This trial is now accruing metastatic breast cancer patients into one of three groups
    - Group 1 participants will have fewer than five CTC at the start of the trial and a low risk of early progression. They will continue to receive regular chemotherapy treatment
      - Group 2 participants will have fewer than five CTC after completing one course of chemotherapy and a moderate risk of early progression
      - Their CTC levels will have decreased during treatment
      - They will continue to receive their current chemotherapy without change
    - Group 3 participants will have a CTC level greater than five with a high risk of early progression
      - They will be divided based on HER2 status and disease type (metastasis to the bone versus measurable metastasis to other areas of the body) and will be randomly assigned to one of two treatment arms
        - Arm 1 participants will continue to receive their current chemotherapy (the standard treatment)
        - Arm 2 participants will receive a new chemotherapy (the treatment under study)
  - The CellSearch technique features a high intraobserver and interobserver as well as interinstrument accordance
  - In the CellSearch clinical trials, a few apparently healthy individuals who had standard CTC determinations were found to have a positive CTC test. It is not known how many of these individuals subsequently were diagnosed with malignant disease
  - Using the CellSearch method, CTC have been detected in peripheral blood samples of patients suffering from a wide variety of additional metastatic epithelial malignancies including endocrine, gastrointestinal, pancreatobiliary, and genitourinary malignancies
  - The sensitivity of the CellSearch technique has been a subject of concern for some investigators
  - A recent study that has caused some concern for the CellSearch technique reported that the so-called normal molecular class of invasive breast cancers which accounts for about approximately 10% of all cases is typically negative for EpCAM expression and may thus be a cause of false-negative CellSearch-based CTC detection

42.3.2 The AdnaTest

- The AdnaTest CancerSelect (AdnaGen Diagnostics, Langenhagen, DE) is a CTC test which commences with a CTC enrichment procedure featuring a proprietary mixture of immunomagnetic beads coated with 1 of 3 antibodies to epithelial surface antigens (Fig. 42.2)

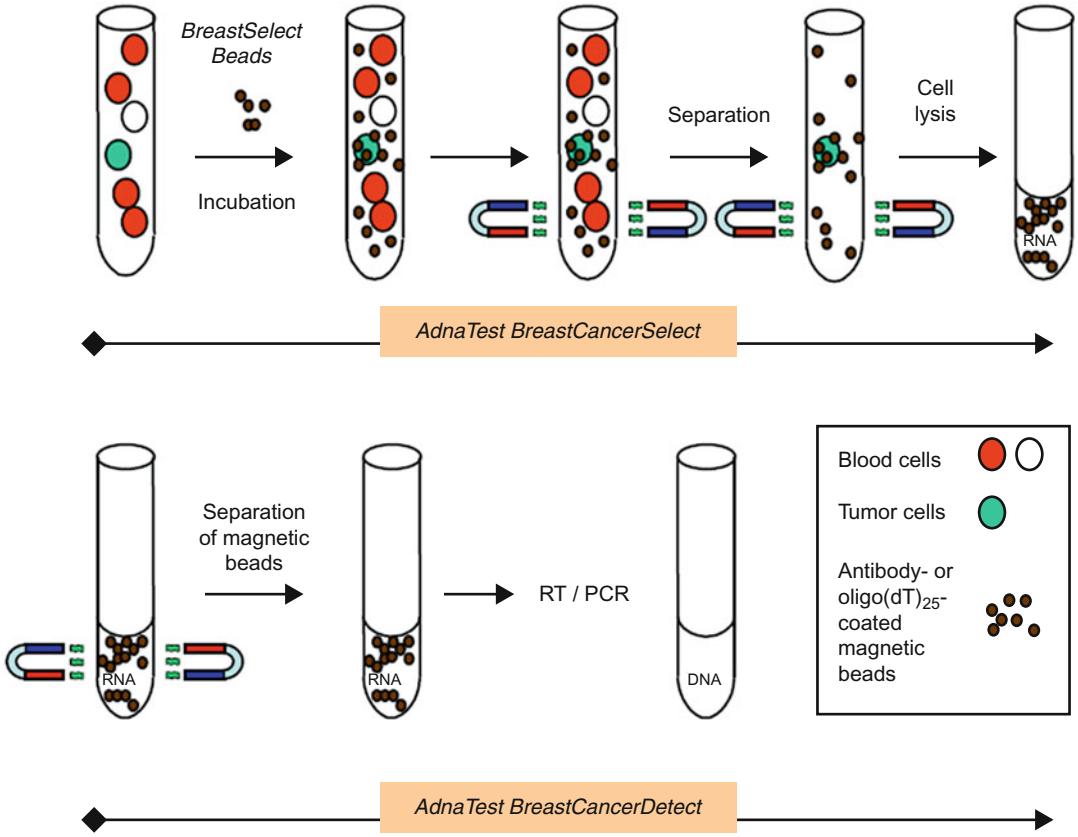


Fig. 42.2 The AdnaTest breast cancer select system. The figure outlines the AdnaTest procedure for determining CTC (Reprinted with permission from *Am J Clin Pathol*)

- After mRNA extraction with RNA quantity and integrity is confirmed (Agilent Bioanalyzer, Agilent Technologies, Inc., Santa Clara, CA), the "number" of CTC is indirectly determined by a semiquantitative RT-PCR technique using probes for three epithelial cell-associated mRNAs
  - MUC1
  - HER2
  - GA733-2
- The amount of "CTC-derived" mRNA is compared to the amount of mRNA extracted from platelets which is estimated by determining the relative actin mRNA level
- The AdnaTest claims to have a higher sensitivity than the whole cell detection-based CellSearch assay
- The AdnaTest, to date, has not received regulatory approval
- This test is commercialized in Europe for breast, prostate, and colorectal cancer, and marketing in the United States is expected to commence in 2011

42.3.3 Membrane Microfilter Assay

- Polycarbonate filters were an early method to concentrate the larger CTC for microscopic identification by pathologists and separate them from circulating leukocytes and platelets
- Thus, membrane filters became a relatively easy, efficient, and inexpensive technique for CTC enrichment

- More recently, a microfabrication technique using parylene-C has produced a specialized microfilter for separation of circulating CTC
- In the membrane microfilter assay for CTC, RT-PCR is performed on the capturing membrane using electrolysis
  - Although there are presently no published clinical trials using this method, it has significant promise given the apparent high efficiency of the technique
  - It has not, to date, been commercialized

42.3.4 Microfluidic CTC Chip

- The CTC chip is a microchip technology that employs a microfluidic device designed to improved CTC capture efficiency (Fig. 42.3)
  - The CTC chip consists of an array of microposts coated with anti-EpCAM antibodies
  - A laminar flow system passes whole blood from the patient through the coated posts
  - The captured CTC are subsequently identified by fluorescence microscopy after CTC staining with a nuclear marker for DNA content and an anticytokeratin epithelial marker
  - To prevent counting of leukocytes, negative selection with an anti-CD45 antibody is used
- In its initial study, the CTC chip achieved a high yield (>99%) and a large number of isolated CTC with high purity (>47%)
  - This was attributed to the use of a single-step technology from whole blood without the typical preparatory steps of centrifugation, washing, and incubation that result in loss and/or destruction of a significant proportion of CTC in other methods
  - The CTC chip results were positive in cases of metastatic lung, prostate, pancreatic, breast, and colon cancer in 115 of 116 (99%) samples, with a range of 5–1,281 CTC per mL and approximately 50% purity
  - CTC were also identified in 7/7 (100%) of patients with early-stage prostate cancer
- In another study of a small cohort of patients with metastatic cancer, the CTC chip results appeared to correlate with the clinical course of disease as measured by standard radiographic methods
- The CTC chip has recently been commercialized by On-Q-ity (On-Q-ity, Waltham, MA)

42.3.5 Fiberoptic Array Scanning Technology (FAST)

- A fiberoptic array scanning technology (FAST) employing fluorescence cytometry combined with an automated digital microscopy imaging system featuring laser-printing optics that scan 300,000 cells per second has been used to detect immunofluorescent-labeled CTC on a glass slide
- This method is in an early stage of development, and large-scale clinical trials have not been reported

42.3.6 Epithelial Immunospot (EPISPOT)

- The EPISPOT (epithelial immunospot) approach detects viable tumor cells by taking advantage of their ability to produce and secrete proteins after 48-h culture (Fig. 42.4)
  - CD45 is used to negatively select leukocytes
  - Dying or dead CTC that do not produce or secrete epithelial-associated proteins are not identified
  - The EPISPOT assay has successfully identified metastatic breast cancer cells cultured from both blood and bone marrow using MUC1 and CK19 as the identifying proteins
  - The test has not, to date, been evaluated in large-scale clinical trials or undergone formal commercial development

42.3.7 Laser Scanning Cytometry (LSC)

- This technique identifies CTC in peripheral blood by use of an automated laser scanning cytometer

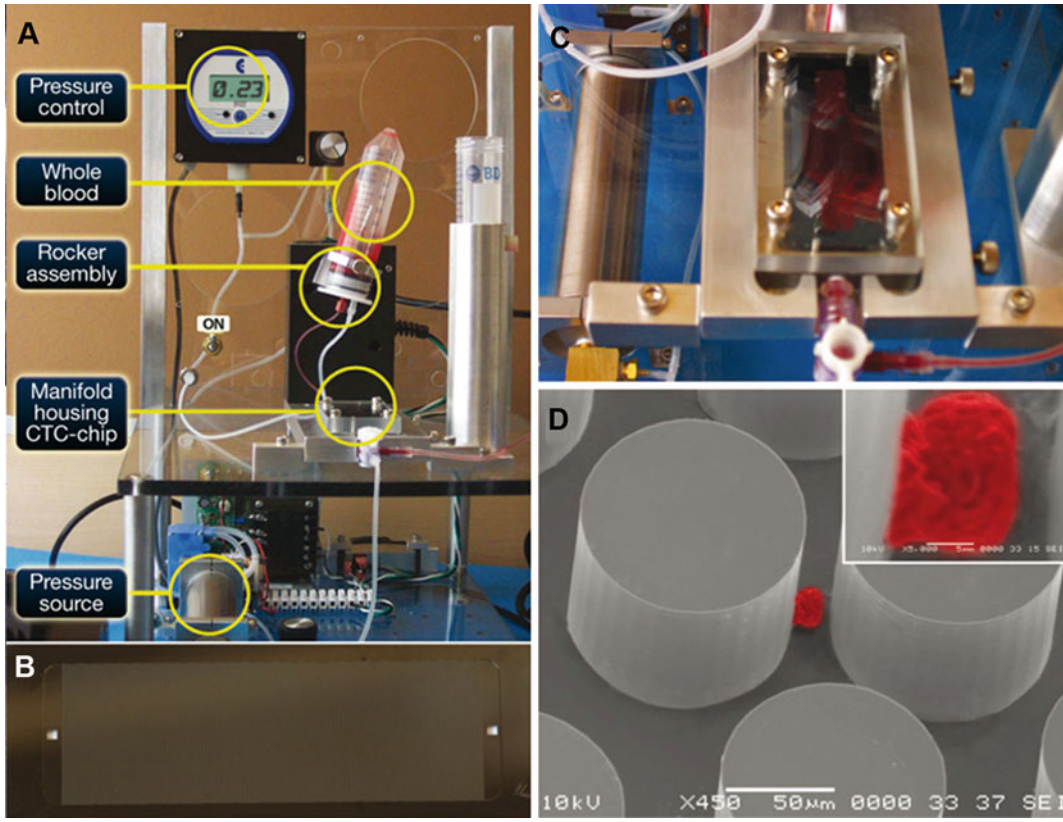


Fig. 42.3 The microfluidic CTC chip. (**a**) The workstation setup for CTC separation. The sample is continually mixed on a rocker and pumped through the chip using a pneumatic-pressure-regulated pump. (**b**) The CTC chip with microposts etched in silicon. (**c**) Whole blood flowing through the microfluidic device. (**d**) Scanning electron microscope image of a captured NCI-H1650 lung cancer cell spiked into blood (pseudocolored *red*). The *inset* shows a high-magnification view of the cell (Ross et al. 2009; Reprinted with permission from *Am J Clin Pathol*)

- Positive CTC identification is achieved using fluorescent-labeled anti-EPCAM
- Negative selection is performed with an anti-CD45 antibody
- In patients who received neoadjuvant treatment for their breast cancer, nearly all subjects were positive for CTC by LSC at the start of therapy
- CTC count correlated with tumor size
- In addition, an initial decrease in cell numbers highly correlated with the final tumor size reduction
- To date, the LSC approach has not been formally commercialized for CTC detection

42.3.8 Differential Centrifugation

- OncoQuick® is a system for the enrichment of CTC from blood that has been developed in cooperation between Hexal Gentech and Greiner Bio-One (Greiner Bio-One, Frickenhousen, DE)
  - The system combines the advantages of the cell separation by density gradient centrifugation with recovery rates that are comparable with immunobead methods
  - OncoQuick consists of a sterile 50-mL polypropylene tube with a porous barrier inserted on top of the unique separation medium

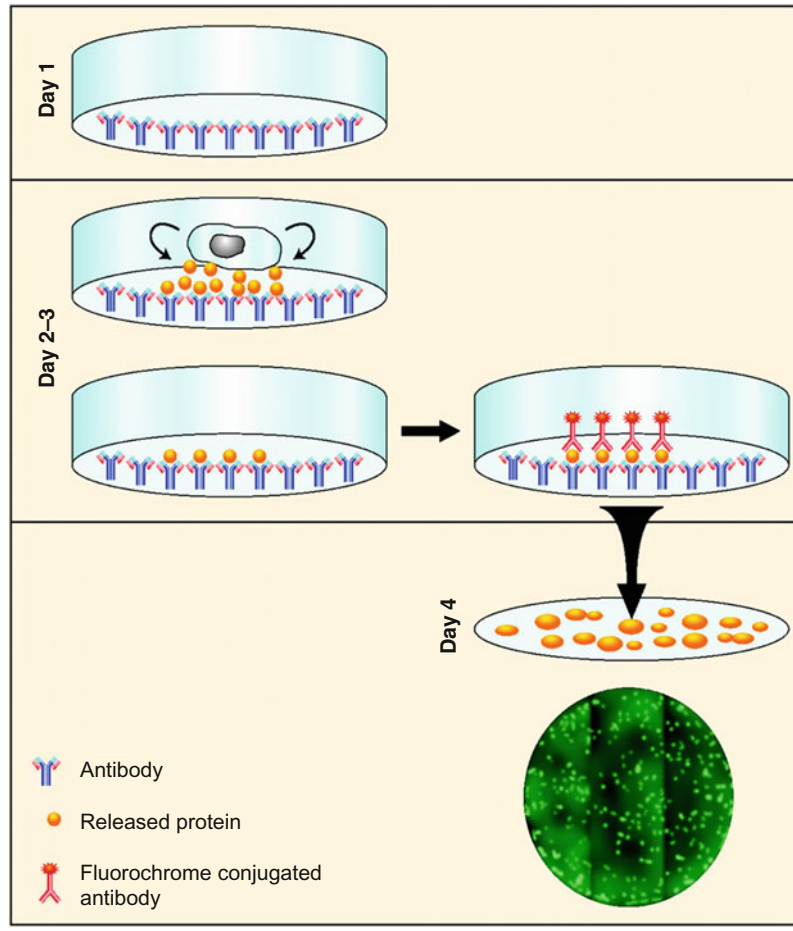


Fig. 42.4 The EPISPOT assay procedure. Day 1, the membranes of the EPISPOT plates are coated with a specific antibody. Days 2–3, the cells are seeded in each well and cultured for 48 h. During this incubation period, the released specific proteins are directly immunocaptured by the immobilized antibody on the bottom of the well. Plates are then washed, and cells are removed. The presence of the released protein is revealed by the addition of a fluorochrome-conjugated antibody. Day 4, fluorescent immunospots are counted with an automated reader. One immunospot corresponds to the fingerprint left only by one viable cell releasing the marker protein (Ross et al. 2009; Reprinted with permission from *Am J Clin Pathol*)

42.3.9 Other Homebrew Techniques

- Flow cytometry has been implemented in detecting CTC/DTC in patients with metastatic cancer, although it has a low sensitivity
- Routine (nonfluorescence) microscopy using immunohistochemistry for epithelial markers after enrichment techniques featuring differential centrifugation approaches has been utilized with variable results

42.3.10 Morphologic Versus Molecular Detection of CTC

- Morphology-driven slide-based cell counting features high specificity, but a number of investigators believe that this method lacks sensitivity in comparison with quantitative mRNA techniques
  - This likely reflects the fragility of CTC in general
  - Increasing the detection rate of CTC would allow the test to be more useful in that more patients would be positive at the start of their treatment, and then, in addition to obtaining prognostic information, they could subsequently be monitored for therapy response
- Molecular methods that utilize RT-PCR to amplify the target mRNAs of tumor cells can also distinguish them from circulating leukocytes and platelets
  - A wide variety of mRNA epithelial cell and more specific breast cancer cell biomarkers have been used

    - CK19
    - CK20
    - MUC1
    - Mammaglobin
  - Multimarker assays appear to offer greater sensitivity than single-probe assays
  - A main advantages of molecular methods is their enhanced sensitivity and large assortment of available primers
  - However, the extraction, storage, and preparation of the relatively unstable mRNA can lead to loss of sensitivity as well as misidentification of tumor-specific markers
  - Other issues with PCR-based tests include
  - The false-positive values which may be obtained due to low-level expression by noncancerous cells
  - Heterogeneity in the expression levels of particular target transcripts that cannot be predicted
  - False-negative values due to downregulation of the expression of a single target gene (often overcome when multimarker approaches are used)
- Finally, PCR can detect both viable and nonviable CTC which may be important when the CTC test is being used as a monitor of the response to therapy

42.4 CTC Platform Comparisons

- CellSearch versus AdnaTest
  - In a recent comparison between the CellSearch assay and a molecular test, the AdnaTest BreastCancer, concordant results regarding HER2 positivity were obtained in 50% of the patients
    - The study concluded that a universal internal and external QC system for both CTC detection and enumeration is urgently needed before their application in the clinic
  - In another study, the two techniques had an 81% concordance in breast cancer patients
- CellSearch versus OncoQuick
  - In this study, the CellSearch system was more accurate and analytically sensitive than OncoQuick for enumeration of CTC
- Immunodetection and separation systems
  - One study found that the success of all types of immunodetection systems was highly dependent on the level of EpCAM expression in the target CTC

42.5 CTC Versus DTC

- The presence of DTC in bone marrow at early stages of breast cancer has been extensively analyzed as an adverse prognostic factor for the disease
- Although, the clinical significance of the detection of DTC in bone marrow in breast cancer is well studied, the impact on the wide variety of other solid tumors that can also feature bone marrow DTC either early or late in the clinical course of these diseases is not well understood
  - In an analysis of pooled data from several prospective studies, the detection of bone marrow DTC in breast cancer was associated with a statistically significantly higher risk of recurrence and disease-specific death
    - In these studies, a variety of detection techniques were used mostly depending on immunohistochemistry to identify the rare malignant epithelial cells among the hematopoietic and stromal elements
  - In the 2007 recommendations update from the American Society of Clinical Oncology on the use of tumor markers in breast cancer, it was concluded that data were insufficient to recommend assessment of bone marrow micrometastases for management of patients with breast cancer
  - Although the ASCO group concluded that in contrast with the significant impact of bone marrow DTC positivity for patients treated with systemic adjuvant chemotherapy, the presence of bone marrow DTC in patients who did not receive

adjuvant systemic therapy predicted for a statistically significantly higher risk of relapse, but the difference in distant disease-free survival between those patients who had DTC versus those who did not was very small
  - The group further concluded that bone marrow DTC in patients with small, low-grade, node-negative breast cancers did not forecast a sufficiently worse prognosis such that it could be used to adjust recommendations for adjuvant therapy

42.6 Prognostic and Predictive Significance of CTC

- Prognostic
  - By the end of 2010, more than 700 published manuscripts employing a wide variety of techniques had considered the potential prognostic significance of the presence of CTC on the outcome of epithelial malignancies
  - The CellSearch assay is currently FDA approved in the US as a prognostic test for breast, colon, and prostate cancers
  - The negative prognostic impact associated with the detection of CTC during the course of breast cancer treatment is now widely accepted
- Predictive
  - It is generally accepted that the future expanded use of CTC testing for breast cancer lies in the use of the test to predict therapy efficacy and resistance and serve as a monitor of treatment response
  - However, despite the validated data for its prognostic significance, the ASCO tumor markers group concluded that, as of 2007, the measurement of CTC was not to be used to influence any treatment decisions in patients with breast cancer
    - The group further held that the CellSearch could not be recommended for use until additional validation confirms its clinical value
    - Given this resistance to large-scale clinical adoption based on original data approved by the FDA, investigators have recommended that prospective randomized breast cancer clinical trials be developed to verify the cost to benefit ratio of the test
    - The Southwest Oncology Group (SWOG) and the Breast Cancer Intergroup of North America are now conducting a prospective trial in which patients with metastatic breast cancer who have a positive CTC count using the CellSearch system after one cycle of first-line chemotherapy will be randomly assigned to either remaining on that therapy until clinical and/or radiographic evidence signals progression, or switching therapy at that time point to a different chemotherapeutic agent (see above)
    - As this trial and others that included CTC measurements in their biomarker protocols proceed in various clinical settings, the critical data needed to confirm the clinical utility of CTC for the ongoing management of breast cancer will be obtained, and the future of CTC testing will be decided

42.7 CTC/DTC Versus Positron Emission Scanning (PET-CT)

- The introduction of positron emission tomography-based scanning (PET) for the detection of recurrent and metastatic breast cancer created a competitive environment where functional imaging would be compared with in vitro molecular diagnostics
- Published studies designed to coevaluate PET scanning with CTC testing are currently limited
  - In one study, a significant correlation among PET scan, serum CA27.29 levels, and CTC counts was found

Table 42.2 Potential CTC-based cancer biomarker detection to select targeted therapies

Genetic event	Disease	Drug	Current test platform	Regulatory status
HER2 gene amplification	Breast and upper GI cancer	Trastuzumab/ Lapatinib	IHC, FISH	FDA approved
BCR–ABL translocation	CML	Imatinib/Dasatinib	Cytogenetics, FISH, PCR	FDA approved
RARA–PML translocation t (15;17)	APL	ATRA	Cytogenetics, FISH, PCR	FDA approved
KRAS mutation	CRC	Cetuximab/ Panitumumab	Traditional sequencing	FDA approved
BRAF mutation	CRC	Cetuximab/ Panitumumab	Traditional sequencing	Clinical trials
EGFR mutation	NSCLC	Gefitinib/Erlotinib	Traditional sequencing	FDA approved
BRAF mutation	Melanoma/thyroid cancer	Vemurafenib	Traditional sequencing	FDA approved
EML4–ALK translocation	NSCLC	Crizotinib	FISH	FDA approved
CKIT mutation	GIST/melanoma	Imatinib/Sunitinib	IHC	FDA approved
MET	NSCLC	Tivantinib, MetMab	IHC, FISH	Clinical trials

42.8 Molecular Assessment of Captured Cells

- Beyond their potential use as general prognostic and predictive markers is the application of CTC capture to targeted therapy selection (Table 42.2)
- This has led to the concept of the liquid biopsy where treatment selection will be based on CTC biomarker testing at the start of the disease or at the time of disease progression
- For breast cancer, biomarker testing on CTC has focused on HER2 assessment
  - A number of investigators have reported preliminary findings concerning the measurement of HER2 status on CTC and DTC and differences between these results and those for the patient's primary breast cancers
  - Mostly using the CellSearch method and FISH-based determination of *HER2* gene amplification status, several studies have found that CTC maintain the same HER2 status as the primary tumor, whereas other reports have claimed that CTC may be HER2-positive in cases where the primary tumor was originally HER2-negative
  - The methodological differences in assessing HER2 status in the primary tumor versus in the CTC may at least partially account for these discrepant results
    - The different CTC techniques have influenced the capability of performing HER2 testing with the CellSearch method requiring a slide-based HER2 test such as FISH or immunocytochemistry whereas the RT-PCR-based techniques, with or without immunomagnetic-based cellular enrichment, claiming an enhanced sensitivity based on their relative HER2 mRNA measurements
    - Recent studies have shown that HER2 overexpression on CTC/DTC was predictive of a poor clinical outcome in stage I–III breast cancer patients
    - The overexpression of HER2 in the primary breast cancer specimen is also predictive of the presence of CTC in early-stage breast cancer patients
    - In another study, it was concluded that HER2-positive DTC might identify additional patients who can benefit from anti-HER2-targeted therapy

- Note: there has been very limited consideration of estrogen and progesterone receptor determination on CTC and DTC

42.9 CTC and Next-Generation Sequencing (NGS)

- Among the potential additional uses of CTC are their use in discovering targets for therapy by having the DNA extracted from them and fully sequenced in a whole genome analysis test
  - The ability to detect mutations, insertions, deletions, rearrangements, and copy number alterations in a single NGS test could enable significant progress in CTC biomarker testing
  - The DNA extracted from CTC could be "combined" with circulating free DNA from tumor cells to achieve a meaningful NGS result

42.10 Summary and Future Clinical Implications

- Since the introduction of CTC testing, substantial progress has been made in the field including the FDA approval of the CellSearch assay
- Nonetheless, numerous significant technical and commercial challenges remain to be overcome before CTC testing achieves widespread acceptance and becomes incorporated into routine cancer management
  - The frequency of CTC-positive patients eligible for clinical monitoring must be increased by increasing CTC sensitivity
  - The goal to capture just a single circulating malignant cell in a 7.5-mL volume of whole blood, which may feature more than 10 billion benign blood cells, is a major challenge
  - This challenge must be approached by creating improvements in the cell capture technology with increased efficiency of CTC recovery and their more effective separation away from contaminating blood cells and bone marrow elements
- The competition between tests that count captured cells such as the CellSearch technique and tests that feature target gene amplification such as by RT-PCR must also be played out to verify which approach is the most accurate and clinically useful
  - The commercial development of CTC testing has also been controversial
    - The initial commercial launch of the CellSearch assay featured a decentralized approach where hospital-based and physician office laboratories were to acquire the testing equipment and necessary reagents and perform the CTC test themselves
    - More recently, commercial laboratory-based centralized testing has been introduced
- CTC/DTC testing appears to have significant future potential value in the clinical management of breast cancer including the identification of patients at high risk for relapse, the stratification of patients to specific adjuvant therapies, and the monitoring of response to treatment in both the metastatic and neoadjuvant settings
- Applying routine biomarker testing to captured cells such as HER2 IHC and FISH testing and more sophisticated methods such as NGS detection of therapy targets are major future developments for CTC evaluation in cancer patient management with great potential
- With further optimization and standardization of CTC/DTC detection techniques, improvements in test sensitivity, specificity, and reproducibility, and, most importantly the demonstration of a significant impact on patient outcome in prospective randomized trials, the inclusion of CTC/DTC detection into daily clinical practice could become a reality

Further Reading

Allard WJ, Matera J, Miller MC, et al. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res. 2004;10:6897–904.

Braun S, Schlimok G, Heumos I, et al. ErbB2 overexpression on occult metastatic cells in bone marrow predicts poor clinical outcome of stage I-III breast cancer patients. Cancer Res. 2001;61:1890–5.

Budd GT, Cristofanilli M, Ellis MJ, et al. Circulating tumor cells versus imaging-predicting overall survival in metastatic breast cancer. Clin Cancer Res. 2006;12:6403–9.

Christofanilli M, Budd T, Ellis MJ, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med. 2004;351:781–91.

Fehm T, Braun S, Muller V, et al. A concept for the standardized detection of disseminated tumor cells in bone marrow from patients with primary breast cancer and its clinical implementation. Cancer. 2006;107:885–92.

Gerges N, Rak J, Jabado N. New technologies for the detection of circulating tumour cells. Br Med Bull. 2010;94:49–64.

Hayes DF, Smerage JB. Circulating tumor cells. Prog Mol Biol Transl Sci. 2010;95:95–112.

Ignatiadis M, Kallergi G, Ntoulia M, et al. Prognostic value of the molecular detection of circulating tumor cells using a multimarker reverse transcription-PCR assay for cytokeratin 19, mammaglobin A, and HER2 in early breast cancer. Clin Cancer Res. 2008;14:2593–600.

Jacob K, Sollier C, Jabado N. Circulating tumor cells: detection, molecular profiling and future prospects. Expert Rev Proteomics. 2007;4:741–56.

Krebs MG, Hou JM, Ward TH, et al. Circulating tumour cells: their utility in cancer management and predicting outcomes. Ther Adv Med Oncol. 2010;2:351–65.

Nagrath S, Sequist LV, Maheswaran S, et al. Isolation of rare circulating tumour cells in cancer patients by microchip technology. Nature. 2007;450:1235–9.

Pantel K, Alix-Panabieres C. The clinical significance of circulating tumor cells. Nat Clin Pract Oncol. 2007;4:62–3.

Riethdorf S, Fritsche H, Muller V, et al. Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: a validation study of the cell search system. Clin Cancer Res. 2007;13:920–8.

Riethdorf S, Pantel K. Disseminated tumor cells in bone marrow and circulating tumor cells in blood of breast cancer patients: current state of detection and characterization. Pathobiology. 2008;75:140–8.

Riethdorf S, Wikman H, Pantel K. Review: biological relevance of disseminated tumor cells in cancer patients. Int J Cancer. 2008;123:1991–2006.

Ross JS, Cronin M. Whole cancer genome sequencing by next-generation methods. Am J Clin Pathol. 2011;136:527–39.

Ross JS, Slodkowska EA. Circulating and disseminated tumor cells in the management of breast cancer. Am J Clin Pathol. 2009;132:237–45.

Saad A, Kanate A, Sehbai A, et al. Correlation among [^{18}F]fluorodeoxyglucose positron emission tomography/computed tomography, cancer antigen 27.29, and circulating tumor cell testing in metastatic breast cancer. Clin Breast Cancer. 2008;8:357–61.

Sequist LV, Nagrath S, Toner M, et al. The CTC-chip: an exciting new tool to detect circulating tumor cells in lung cancer patients. J Thorac Oncol. 2009;4:281–3.

Smerage JB, Hayes DF. The prognostic implications of circulating tumor cells in patients with breast cancer. Cancer Invest. 2008;26:109–14.

Vincent-Salomon A, Pierga JY, Couturier J, et al. HER2 status of bone marrow micrometastasis and their corresponding primary tumours in a pilot study of 27 cases: a possible tool for anti-HER2 therapy management? Br J Cancer. 2007;96:654–9.

Wulfing P, Borchard J, Buerger H, et al. HER2-positive circulating tumor cells indicate poor clinical outcome in stage I to III breast cancer patients. Clin Cancer Res. 2006;12:1715–20.

Yu M, Stott S, Toner M, et al. Circulating tumor cells: approaches to isolation and characterization. J Cell Biol. 2011;192:373–82.

Zheng S, Lin H, Liu JQ, et al. Membrane microfilter device for selective capture, electrolysis and genomic analysis of human circulating tumor cells. J Chromatogr. 2007;1162:154–61.

Zippelius A, Pantel K. RT-PCR-based detection of occult disseminated tumor cells in peripheral blood and bone marrow of patients with solid tumors. An overview. Ann N Y Acad Sci. 2000;906:110–23.

Section 7

Laboratory Management in Molecular Genetic Pathology

Molecular Testing: Ethical and Legal Issues

43

Kimberly A. Quaid

Contents

43.1 Ethics Issues in Molecular Testing 1084
43.1.1 Ethics Defined 1084
43.1.2 Ethical Terms 1084
43.1.3 Ethical Theories 1084

43.2 General Ethical Principles 1084
43.2.1 Respect for Individuals 1084
43.2.2 Nonmaleficence 1085
43.2.3 Beneficence 1085
43.2.4 Justice 1085

43.3 Rules Developed in Light of These Principles 1085
43.3.1 Informed Consent 1085
43.3.2 Confidentiality/Privacy 1085
43.3.3 Limits of Confidentiality 1086
43.3.4 Duty to Disclose 1086
43.3.5 Duty to Recontact 1086

43.4 Genetic Testing of Children 1086
43.4.1 Genetic Testing of Children for Late-Onset Conditions 1086
43.4.2 Genetic Testing of Children in the Context of Adoption 1086
43.4.3 Special Issues Associated with the Genetic Testing of Children 1087

43.5 Legal Issues in Molecular Testing 1087
43.5.1 Standard of Care for Genetic Counseling 1087
43.5.2 Possible Legal Exceptions to General Rule of Confidentiality 1087
43.5.3 Wrongful Birth 1087
43.5.4 Wrongful Life 1088
43.5.5 Duty to Warn 1088
43.5.6 Employment Discrimination 1088
43.5.7 Insurance Discrimination 1089
43.5.8 Genetic Information Nondiscrimination Act (GINA) 1089

Further Reading 1089

K.A. Quaid, PhD
Indiana University School of Medicine, Indiana University Center for Bioethics, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_43,

43.1 Ethics Issues in Molecular Testing

43.1.1 Ethics Defined

- The establishment of a set of guidelines for morally acceptable conduct within a theoretical framework

43.1.2 Ethical Terms

- Principles – Sources of guidelines for moral behavior
- Values – Priorities that are considered good, desirable, and important
- Rules – Specific statements of what should or should not be done
- Duties – Behaviors that are defined by a person's role in society
- Virtues – Characteristics of an individual that are morally desirable
- Rights – Justified claims that individuals or groups can make on others or on society

43.1.3 Ethical Theories

43.1.3.1 Utilitarian or Consequence-Based Theory

- The primary focus of this theory is the promotion of happiness
- Actions that maximize good and promote the greatest amount of happiness over pain are considered right or acceptable actions
- An ethical dilemma can be resolved by looking at the consequences of doing or not doing an action

43.1.3.2 Deontological or Principle-Based Theory

- The primary focus of principle-based ethics is on the role of moral reasoning and analysis in making ethical decisions
- The core principles of autonomy, beneficence, nonmaleficence, and justice are used to clarify moral duties and obligations
- An ethical dilemma can be resolved by weighing competing principles, duties, and values

43.1.3.3 Virtue Ethics

- The primary focus of virtue ethics is on the character traits or virtues a good person should have
- A person with such traits is considered to naturally act in a morally acceptable way
- An ethical dilemma is resolved by asking how a virtuous person would act in that particular situation

43.1.3.4 Ethic of Care

- The primary focus of care ethics is the maintenance and enhancement of caring while also preserving the traditional values of other ethical theories
- Care ethics is focused on humanistic virtues and the characteristics values in interactive and intimate relationships
- Ethical dilemmas are resolved by promoting respect for equality while at the same time recognizing and valuing differences

43.2 General Ethical Principles

- General ethical principles to be used as "general guides that leave considerable room for judgment in specific cases and provide substantive guidance for the development of more detailed rules and policies" (Beauchamp and Childress 2008)

43.2.1 Respect for Individuals

- Principle from which informed consent derives its importance
- Implies the autonomy of competent people and the protection of those incapable of autonomy
- Requires that a medical professional's actions never impinge upon an individual's personal autonomy

43.2.2 Nonmaleficence

- Associated with the tenet embodied in the Hippocratic oath "above all first do no harm"
- Imposes an obligation not to inflict harm intentionally

43.2.3 Beneficence

- Requires that one takes positive steps to act for the benefit of the patient and not simply avoid harming them
- Requires prevention and/or removal of conditions that may be harmful to the patient

43.2.4 Justice

- Ensures that benefits and burdens are shared equally
- Patients are to be treated equitably, and vulnerable groups (pregnant women, fetuses, children, prisoners, the cognitively impaired) are protected

43.3 Rules Developed in Light of These Principles

43.3.1 Informed Consent

- Moral aspect of informed consent ensures that the patient's autonomy is respected and that the patient has an understanding of pertinent information and is free of controlling influences
- Informed consent implies the right to self-determination, to dictate what will be done to one's own body
- Elements for consent to be informed
  - Competence
    - Capacity to make a rational choice
    - Ability to understand information provided
    - Ability to understand consequences of choices
  - Ability to communicate a choice
  - Amount and accuracy of information
    - Risks and benefits
    - Available alternatives
  - Patient understanding
    - Barriers to understanding include illness, fear, denial, cultural beliefs, language, and lack of education
  - Voluntariness
    - Absence of control by others
    - Absence of coercion by others or by circumstances
  - Authorization
    - Must be an active choice by the client
- Informed consent for testing should encompass the following
  - Alternatives to testing
  - Risks and benefits
  - Rates of false-positive, false-negative, and inconclusive results
  - Potential effects of results on self-image, family relationships, employment, insurance coverage, and possible negative emotional burden
  - Costs of testing
  - Length of time to get results
  - How results will be given
  - Followup recommendations and/or treatment options
- Informed refusal
  - Right of clients to obtain all pertinent information before refusing genetic testing

43.3.2 Confidentiality/Privacy

- Protecting the confidentiality of information is essential for all uses of genetic tests
- Results should be released only to those individuals for whom the test recipient has given consent for information release
- The means of transmitting information should be chosen in order to minimize the likelihood that the results will become available to unauthorized persons or organizations
- Results with identifiers should not be provided to any third parties, including employers, insurers, or government agencies without the

express and written permission of the person tested
- In general, healthcare providers have an obligation to the person being tested not to inform other family members of the test results without the permission of the person tested

43.3.3 Limits of Confidentiality

- Genetic testing may reveal information about health risks faced not only by the patient but by their family members as well
- Duty to protect patient confidentiality is not absolute
- President's Commission for the Study of Ethical Problems in Medical and Biomedical and Behavioral Research and the American Society of Human Genetics suggest that genetic information can be released to relatives under certain conditions
  - All reasonable efforts have been unsuccessful in obtaining consent for release
  - There is a high probability of irreversible harm to a third party, such as a relative
  - The release of the information has high probability of preventing the harm
  - Only the information necessary to prevent the harm is released
- Each state has laws covering legal disclosures permissible by health practitioners, and health professionals need to be aware of the laws that apply

43.3.4 Duty to Disclose

- Health professionals may owe patients particular duties of awareness and disclosure of susceptibility to genetic disorders
- Physicians who are aware, or who by the standard of care reasonably should be aware, of a patient's genetic risk have a duty to inform the patient without being asked
- There is a related duty to discuss the availability of genetic testing if tests are reliable and accessible

43.3.5 Duty to Recontact

- New genetic technologies are continuously adding to our fund of knowledge
- May be a continuing obligation to recontact clients when new information becomes available that may have an impact on the client's decision making (Pelias 1991)
- Expanded duty to disclosure based on the recognized duty of physicians to recontact patients when new information regarding past medication or therapy is discovered
- In genetic testing, this new duty may apply to information regarding changes in diagnostic availability of new tests and new interpretations of prior test results
- Documentation by the healthcare professional that includes a request for the client to keep in touch with the clinic if individual circumstances such as a change of address occur is helpful

43.4 Genetic Testing of Children

43.4.1 Genetic Testing of Children for Late-Onset Conditions

- Current recommendations state that genetic testing is appropriate if the child will receive an immediate medical benefit, such as early surveillance and/or treatment
- Beyond the possibility of immediate medical benefit, parents and providers should exercise caution in the genetic testing of children in order to minimize harm to the child
- Testing children for adult-onset disease is not recommended unless direct medical benefit will accrue to the child, and this benefit would be lost if by waiting until the child had reached adulthood

43.4.2 Genetic Testing of Children in the Context of Adoption

- Genetic testing of newborns and children in the adoption process should be consistent with tests performed on all children of a similar age

for the purposes of diagnosis or identification of appropriate preventive strategies
- Primary justification for testing should be a timely medical benefit for the child and should be limited to conditions that manifest themselves in childhood or for which preventive strategies may be undertaken in childhood
- Children and newborns should not be tested for the purpose of detecting genetic variations of or predispositions to physical, mental, or behavioral traits within the normal range

43.4.3 Special Issues Associated with the Genetic Testing of Children

- Important to consider whether it is the needs of the parents, the child, or both that are being met
- Possible benefits to testing a child
  - Timely adoption of medical and lifestyle practices that may prevent or improve the disease process
  - Reduction in anxiety
  - Reduced need for potentially painful and/or expensive vigilant medical surveillance
- Possible harms from testing a child
  - Psychological harm in conditions where there is no cure
  - Disruption of parent/child or sibling/sibling relationship
  - Negative changes in self-concept
  - Interference with future autonomy
  - Discrimination by third parties

43.5 Legal Issues in Molecular Testing

43.5.1 Standard of Care for Genetic Counseling

- Professional proficiency in clinical and medical genetics
- Counseling and communication skills
- The ability to promote client decision making in an unbiased and noncoercive manner
- The ability to protect client privacy and confidentiality
- The ability to promote informed consent
- Sensitivity to cultural differences
- Healthcare professionals who provide genetic counseling must be well informed about
  - Nature of condition
  - The social and psychological implications of genetic testing
  - Assessment of familial genetic risk
  - Proper interpretation of test results

43.5.2 Possible Legal Exceptions to General Rule of Confidentiality

- When information is already available to others
- To protect the interests of patients
- When the patient waives the physician–patient privilege
- In the interest of justice in criminal proceedings
- To ensure quality medical treatment
- To protect the interests of a child
- To protect third parties
- To serve a substantial and valid interest of an employer
- To provide insurers with information
- Disclosure of medical information to spouses
- Disclosure to siblings
- To protect a criminal defendant's constitutional rights

43.5.3 Wrongful Birth

- Claim brought by parents for damages suffered from having a child born with birth defects or genetic diseases
- Based on health professional's failure to
  - Inform couples of an appropriate test
  - Properly diagnose a condition
  - Properly interpret test results
  - Deliver test results in time for couples to make decisions about whether or not to continue the pregnancy
- Vast majority of jurisdictions recognize claims for wrongful birth

43.5.4 Wrongful Life

- Claim brought by child based on the theory that, but for the negligence of the healthcare provider, the child's birth defects would have been detected and the parents would have terminated the pregnancy, thus preventing the plaintiff from being born and suffering from the illness
- Most courts reject these causes of action for the following reasons
  - The difficulty of comparing the value of not having been born with the harm of being born to suffer serious health ailments
  - Reluctance to suggest that the plaintiff's life is not worth living
  - The belief that the only harm is that suffered by the parents
  - A few jurisdictions have recognized claims for wrongful life under the theory that the plaintiff is burdened with extraordinary costs of care for which he or she should be compensated

43.5.5 Duty to Warn

- *Pate v. Threlkel* (FL Super. 1995)
  - Plaintiff was receiving treatment for medullary thyroid carcinoma
  - Sued physicians who had previously treated her mother for the same condition but with whom the plaintiff had no doctor–patient relationship
  - Highest state court in Florida held that a physician has a duty to warn a third party about a genetically inherited disorder
  - Duty can be satisfied by warning the patient about the genetic ramifications of a particular disease
- *Safer v. Estate of Pack* (NJ Super. 1996)
  - Plaintiff sued the estate of a physician who had treated the plaintiff's father for multiple polyposis with adenocarcinoma of the colon more than 30 years prior
  - At the time of the father's death of metastatic cancer, the plaintiff was 10 years of age
  - At 36, the plaintiff was diagnosed with cancerous blockage because of multiple polyposis of the colon with evidence of metastatic disease
  - Cause of action against the physician was for professional negligence alleging that multiple polyposis is a hereditary condition that if left undiscovered or untreated invariably leads to metastatic colorectal cancer
  - An intermediate appellate court in New Jersey ruled unanimously that the physician's duty to warn those known to be at risk of avoidable harm from a genetically transmissible condition extends to members of the immediate family
  - Courts would likely balance the interests at stake
    - How dire the risk in terms of magnitude and likelihood
    - Whether there are equally good alternatives to warning
    - Whether direct warning will effectively prevent the harm

43.5.6 Employment Discrimination

- Employers have a legitimate financial interest in information regarding the risk for developing genetic disease in their employees
  - Those at genetic risk may generate expenses such as sick days and higher insurance premiums
  - Information about genetic risk may affect decisions with regard to hiring, promotion, the provision of additional training, disability pay, and pensions
- Fear of employers using genetic information is a major concern
  - The Equal Employment and Opportunities Commission (EEOC) charged with enforcing employment laws like the Americans with Disabilities Act (ADA) has interpreted the ADA to apply to individuals who are discriminated against due to genetic information
  - In April of 2001, the EEOC settled a case that enjoined Burlington Northern Santa Fe

railroad from directly or indirectly requiring employees to submit to genetic testing or to use any information obtained from genetic testing
  - This settlement bolsters the case for the applicability of the ADA in cases of genetic discrimination but is not a binding legal precedent

43.5.7 Insurance Discrimination

- Fear of losing insurance is a major reason for avoiding genetic testing
  - In insurance law, those who seek coverage must volunteer full disclosure of relevant matters particularly within their own knowledge
  - Insurers expect that applicants who have genetic test results will disclose them
  - Failure to disclose may void the insurance contract
  - Health Insurance Portability and Accountability Act of 1996 (HIPAA) provides some protection
    - HIPAA Privacy Rule established strict confidentiality guidelines for the storage and transmission of health information and places
    - Narrow and precise conditions under which health professionals may disclose such personal information
    - HIPAA prohibits the classification of a genetic predisposition to disease as a preexisting condition which can be used to deny group insurance coverage unless the individual has already been diagnosed with that condition
    - HIPAA does not place any restrictions on health insurance rate setting

43.5.8 Genetic Information Nondiscrimination Act (GINA)

- Enacted May 21, 2008, GINA is a new federal law designed to prohibit the improper use of genetic information in both health insurance and employment
- Sets a minimum standard of protection that must be met in all states
- Parts of law relating to health insurers took effect in May 2009, those relating to employers took effect in November of 2009
- Prohibits employers from using an individual's genetic information when making hiring, firing, job placement or promotion decisions
- Prohibits group health plans and health insurers from denying coverage to a healthy individual or charging that person higher insurance premiums based solely on a genetic predisposition to developing a disease in the future
- The law does not cover life insurance, disability insurance, or long-term care insurance
- Does not prohibit medical underwriting based on current health status
- Does not apply to members of the military

Further Reading

American Society of Human Genetics Board of Directors (ASHG/ACMG). Report: points to consider: ethical, legal and psychosocial implications of genetic testing in children and adolescents. Am J Human Genet. 1995;57:1233.

American Society of Human Genetics Social Issues Committee and The American College of Medical Genetics Social. Ethical and Legal Issues Committee. ASHG/ACMG Statement. Genetic testing in adoption. Am J Human Genet. 2000;66:761–7.

American Society of Human Genetics Social Issues Subcommittee on Familial Disclosure. ASHG statement: professional disclosure of genetic information. Am J Human Genet. 1998;62:474–83.

Beauchamp TL, Childress JF. Principles of biomedical ethics. 6th ed. New York: Oxford University Press; 2008.

Faden RR, Beauchamp TL. A history and theory of informed consent. New York: Oxford University Press; 1986.

Florida. Supreme Court. Pate v. Threlkel. Wests South Report. 1995:661:278–82.

Holtzman NA, Watson MS, editors. Promoting safe and effective genetic testing in the United States: final report of the task force on genetic testing. Washington, DC: US Government Printing Office; 1997.

http://www.genome.gov. Website of the National Human Genome Research Institute.
Lowrey KM. Legal and ethical issues in cancer genetics nursing. Semin Oncol Nurs. 2004;20:203–8.
McAbee GN, Sherman J, Davidoff–Feldman B. Physician's duty to warn third parties about the risk of genetic diseases. Pediatrics. 1998;102:140–1.
New Jersey. Superior Court, Appellate Division. Safer v. Estate of Pack. Atl Report. 1996;677:1188–93.
Pelias MK. Duty to disclose in medical genetics: a legal perspective. Am J Med Genet. 1991;39:347–54.
Quaid KA. Presymptomatic genetic testing in children. In: Miller G, editor. Pediatric bioethics. New York: Cambridge University Press; 2010.
United States Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. Screening and counseling for genetic conditions. A report on the ethical, social and legal implications of genetic screening, counseling and education programs. Washington, DC: US Government Printing Office; 1983. p. 53.

Molecular Testing: Regulatory Issues

44

Frank S. Ong, Wayne W. Grody, and Kingshuk Das

Contents

44.1 Introduction 1092

44.2 Federal Regulation: Clinical Laboratory Improvement Act 1092

44.2.1 Chronology of Clinical Laboratory Improvement Act Regulations 1092

44.2.2 Laboratories Regulated by Clinical Laboratory Improvement Act 1092

44.2.3 Categories of Testing 1093

44.2.4 Clinical Laboratory Improvement Act Certification 1094

44.2.5 Proficiency Testing 1096

44.2.6 Facilities 1097

44.2.7 Quality Control and Assurance 1098

44.2.8 Personnel 1102

44.2.9 Inspection 1103

44.2.10 Enforcement of Regulations 1104

44.3 Other Governmental Regulations and Regulatory Agencies 1104

44.3.1 State Departments of Health 1105

44.3.2 Office of the Inspector General 1105

44.3.3 Centers for Medicare and Medicaid Services 1105

44.3.4 United States Food and Drug Administration, Office of In Vitro Diagnostic Device Evaluation and Safety 1105

44.3.5 Equal Employment Opportunity Commission 1106

44.3.6 Occupational Safety and Health Administration 1106

44.3.7 Environmental Protection Agency 1106

44.3.8 Nuclear Regulatory Commission 1106

44.3.9 Department of Transportation 1106

44.3.10 Health Insurance Portability and Accountability Act 1107

44.3.11 Ban on Physician Self-referral (Stark Law) 1107

44.3.12 Three-Day Rule 1107

44.4 Nongovernmental Agencies 1107

44.4.1 College of American Pathologists 1107

44.4.2 Commission on Office Laboratory Accreditation 1108

44.4.3 Joint Commission on Accreditation of Healthcare Organizations 1108

Further Reading 1108

F.S. Ong, MD
Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, USA

W.W. Grody, MD, PhD
Divisions of Medical Genetics and Molecular Pathology, Departments of Pathology and Laboratory Medicine, Pediatrics, and Human Genetics, UCLA School of Medicine, Los Angeles, CA, USA

K. Das, MD (✉)
Department of Pathology and Laboratory Medicine, UCLA David Geffen School of Medicine, Los Angeles, CA, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_44,

44.1 Introduction

- In response to reports of erroneous Pap smear testing, subsequent public outcry, and skepticism of quality standards in clinical laboratories as a whole, in 1967, the Clinical Laboratory Improvement Act (CLIA) was passed, and the underpinnings of the first consistent set of federal laboratory regulations were laid, with the following consequences
  - CLIA determined standards for all clinical laboratories
  - Legislation provided necessary oversight of clinical laboratories (molecular diagnostic laboratories comprising a minor numerical subset) to ensure patient safety through quality laboratory operations
  - Clinical laboratories have become among the most intensely regulated parts of the entire system of healthcare in the United States
  - It is paramount to understand not only these federal regulations but also state and local
  - Regulations in order to establish and operate a modern, compliant molecular diagnostic laboratory
- The following chapter will provide an overview of the various regulatory issues facing such laboratories. It is important to note that the regulatory issues presented here are common to most clinical laboratories; however, those issues most specific and pertinent to molecular testing will be highlighted

44.2 Federal Regulation: Clinical Laboratory Improvement Act

44.2.1 Chronology of Clinical Laboratory Improvement Act Regulations

- CLIA 1967: CLIA was passed in 1967 in response to questionable clinical laboratory practices involving Pap smear testing. However, several issues were identified with this Act
  - It was most applicable to only large independent laboratories
  - Oversight and enforcement of standards was inconsistent
  - Proficiency testing was found to be fairly ineffective
  - Many smaller laboratories, including physician-operated laboratories, were increasing in popularity, though largely immune to such regulation
- CLIA ’88: In order to address these issues, in 1988, Congress enacted Public Law 100–578, a revision of Section 353 of the Public Health Service Act (42 U.S.C. 263a) that amended the Clinical Laboratory Improvement Act of 1967 (also known as Clinical Laboratory Improvement Amendments of 1988, or CLIA ’88) requiring the Department of Health and Human Services (HHS) to establish regulations to ensure quality and reliable clinical laboratory testing
  - In 1990, rules were proposed
  - In 1992, HHS published the final rules (with comment), establishing CLIA ’88 regulations that describe requirements for all laboratories performing clinical testing
  - Since the 1992 final rules, there have been numerous notices and regulations published, revising CLIA ’88

44.2.2 Laboratories Regulated by Clinical Laboratory Improvement Act

- There are two types of laboratories regulated by CLIA, those meeting either definition listed under 42 Code of Federal Regulations (CFR), Section 493.2
  - . . .a facility for the biological, microbiological, serological, chemical, immunohematological, hematological, biophysical, cytological, pathological, or other examination of materials derived from the human body for the purpose of providing information for the diagnosis, prevention, or treatment of any disease or impairment of, or the assessment of the health of, human beings. These examinations also include procedures to determine, measure, or otherwise describe the presence or absence of various substances or organisms in the body.

  - Laboratories seeking payment under the Medicare and Medicaid Programs
- Virtually all clinical laboratories, including molecular diagnostic labs, will fit into at least one, if not both categories, and therefore fall under CLIA regulation
- Includes any laboratory falling under federal jurisdiction (although regulations may be modified for these laboratories)
- Notable exceptions to CLIA regulation are
  - Laboratories performing only forensic testing, including molecular testing laboratories performing only forensic DNA analysis
  - Laboratories testing human specimens that do not meet the aforementioned criteria in 42 CFR, Section 493.2; typically academic or private sector research laboratories, many of which employ molecular genetic methodologies
  - Laboratories certified by Substance Abuse and Mental Health Services Administration (SAMHSA), typically providing drug testing under SAMHSA regulations
    - However, all other testing performed by such laboratories is subject to CLIA regulation, as necessary
    - With regard to molecular testing, any pharmacogen(etic/omic) testing, if used for diagnostic purposes and/or seeking Medicare/Medicaid reimbursement for such testing, may be subject to CLIA regulation

44.2.3 Categories of Testing

- CLIA regulations recognize three categories of laboratory testing (waived, moderate complexity, and high complexity) and enforce regulations specific to these categories

44.2.3.1 Waived Testing

- This category consists of tests simple enough that they require minimal regulatory oversight by CLIA. According to 42 CFR Sec. 493.15, these tests are either
  - "Cleared by FDA for home use" or
  - "Employ methodologies that are so simple and accurate as to render the likelihood of erroneous results negligible" or
  - "Pose no reasonable risk of harm to the patient if the test is performed incorrectly"
- Common examples of such testing include many urine pregnancy tests (visual interpretation) and blood glucose monitoring devices cleared by the FDA for home use
  - The list of waived testing is determined and maintained by HHS
  - At the time of this writing, there are no waived molecular genetic tests available, although methodologies exist that may make such testing possible

44.2.3.2 Moderate and High Complexity Testing

- All other testing is divided into tests of moderate and high complexity, according to grading based on seven criteria (see 42 CFR Sec. 493.17)
  - Knowledge required to perform test
  - Training and experience required to perform test
  - Reagents and material preparation
  - Characteristics of operational steps
  - Calibration, quality control, and proficiency testing materials
  - Test system troubleshooting and equipment maintenance
  - Interpretation and judgment
- A subcategory of moderate complexity testing is known as provider-performed microscopy procedures (PPM). To be considered PPM procedure, test must meet several criteria (42 CFR Sec. 493.19)
  - Performed by either a physician or other professional healthcare provider, on own patient or patient of group practice or on specimen obtained during visit
  - Procedure must be moderate complexity
  - Test is performed with microscope (bright field or phase contrast)
  - Specimen is labile, or delay in performing test could compromise accuracy
  - Control materials not available to monitor entire testing process

  - Limited specimen handling/processing required
  - Testing limited to list determined by HHS (examples include direct wet mount microbiology procedures, urine sediment examination, qualitative semen analysis)
- At the time of this writing, only 12 tests listed as PPM procedures, with no molecular diagnostic assay listed; however, above criteria do not preclude PPM approval for future molecular diagnostic testing

44.2.3.3 Determination of Testing Complexity

- Testing manufacturers submit supporting data to the US Food and Drug Administration (FDA) for determination based upon above criteria
- For noncommercial systems, laboratory or group involved submits such application for determination by Centers for Disease Control and Prevention (CDC)
- If a test is not listed in the Federal Register notices, it is automatically categorized **high complexity**, until applicable determination process completed
  - All molecular diagnostic testing at the time of this writing categorized as moderate or high complexity
  - Nature of molecular diagnostic testing, as an evolution of basic/translational research, generally obviates new testing be categorized high complexity
- A clinical laboratory may perform any of three types of testing, or any combination thereof, if CLIA certificate permits (see below)

44.2.4 Clinical Laboratory Improvement Act Certification

- Any laboratory regulated by CLIA must have current (unrevoked or unsuspended) CLIA certificate
- There are five types of CLIA certificates
  - Certificate of waiver
  - Certificate for PPM procedures
  - Certificate of registration
  - Certificate of compliance
  - Certificate of accreditation
- Applications for CLIA certificate submitted on behalf of every laboratory location, with following exceptions
  - If laboratory is not at fixed location (such as a mobile lab or temporary facility), single application may be filed on behalf of primary site
  - Nonprofit, federal, state, or local government laboratories offering limited menu of moderate complexity and/or waived tests ($\leq$15 total tests), for public health purposes, can file one application
  - Multiple laboratories within hospital system, located in contiguous buildings on same campus or same physical location/street address, under common direction can file single application

44.2.4.1 Certificate of Waiver

- Laboratories performing only waived testing must carry such certificate (as noted earlier, no molecular diagnostic tests currently categorized as waived)
- However, when waived molecular diagnostic testing does become available, a laboratory operating under this certificate must notify HHS (or designee) before
  - Performing and reporting any nonwaived testing
  - Within 30 days of any change in owner, name, location, or director
- Certificate of waiver valid for maximum of 2 years

44.2.4.2 Certificate for Provider-Performed Microscopy Procedures

- Certificate for PPM procedures required for
  - Laboratories performing PPM procedures
  - Certificate of waiver laboratories intending to perform PPM procedures
- As already noted, currently no molecular diagnostic testing falls under category of PPM procedures

- Laboratories must notify HHS (or designee) before
  - Performing and reporting results for moderate and/or high complexity testing
  - Within 30 days of any change in owner, name, location, or director
- Certificate for PPM procedures valid for maximum of 2 years

44.2.4.3 Certificate of Registration

- Required of
  - Laboratories performing moderate and/or high complexity testing
  - Laboratories issued certificate of waiver or certificate for PPM procedures intending to perform moderate and/or high complexity testing
- Pending inspection by HHS to demonstrate compliance with regulations governing moderate and/or high complexity testing
- Certificate of registration valid for maximum of 2 years (or until inspection conducted, whichever is shorter)
- Laboratory must notify HHS (or designee) within 30 days of any change in owner, name, location, director, or technical supervisor (high complexity testing only)

44.2.4.4 Certificate of Compliance

- Certificate of compliance may be issued after successful HHS inspection of laboratory, authorizing waived, PPM procedures, moderate complexity, high complexity, or any combination of testing thereof
- Laboratories operating under this certificate must notify HHS (or designee) within 30 days of any changes in owner, name, location, director, or technical supervisor (high complexity testing only) and
  - Within 6 months after performing testing not included on certificate
  - Within 6 months after any deletions or change in test methodology for any test on certificate
- Certificate of compliance valid for maximum of 2 years

44.2.4.5 Certificate of Accreditation

- To perform waived, PPM procedures, moderate complexity, high complexity, or any combination of testing thereof, a laboratory can pursue a certificate of accreditation (in lieu of or in addition to certificate of compliance)
- Lab must secure a certificate of registration (valid for a period of no more than 2 years), carrying valid certificate of compliance, and then
  - Become accredited by private nonprofit entity approved by HHS for this purpose
  - Provide HHS with proof of accreditation (within 11 months of receiving certificate of registration or, if certificate of compliance is held, prior to expiration of that certificate)
- Lab must notify HHS and HHS-approved accreditation program within 30 days of change in owner, name, location, director, or supervisor (high complexity testing only) and
  - Notify the HHS-approved accreditation program no later than 6 months after deletion or change of testing methodologies
  - Notify HHS-approved accreditation program no later than 6 months after performing testing for which laboratory is not accredited (for reevaluation of compliance)
  - Maintain HHS-approved accreditation, including allowing accreditation program to release relevant data to HHS for scrutiny, and allow HHS validations and inspections as well
- Certificate of accreditation is valid for no more than 2 years

44.2.4.6 Department of Health and Human Services-Approved Laboratory Accreditation Programs and Clinical Laboratory Improvement Act Exemption

- The Centers for Medicare and Medicaid Services (CMS) oversee the CLIA program on behalf of HHS. The aforementioned certificate of accreditation process may be thought of as a CMS-approved laboratory accreditation

program "deeming" a laboratory compliant with relevant CLIA regulations
  - Such CMS-approved nonprofit laboratory accreditation organizations are described as having CLIA "deeming authority"
- Laboratory can also be exempted from CLIA requirements by CMS, if licensed or approved in certain states. These states are described as having been granted an "exemption" that allows licensed/approved laboratories in those states exemption from CLIA requirements, as long as certain criteria are met
- For private nonprofit laboratory accreditation organizations to be granted "deeming authority" or state licensure/approval programs to be granted "exemptions," CMS must approve their respective regulations to be at least as stringent as CLIA regulations. Such authority given by CMS cannot exceed 6 years. A list of such organizations made publically available by CMS. At the time of this writing, the following states have exempt status
  - New York
  - State of Washington
- The following private nonprofit accreditation organizations have "deeming authority"
  - AABB (formerly the American Association of Blood Banks)
  - American Osteopathic Association
  - American Society for Histocompatibility and Immunogenetics
  - College of American Pathologists (CAP)
  - Joint Commission on Accreditation of Healthcare Organizations (JCAHO)
  - COLA (Commission on Office Laboratory Accreditation)

44.2.5 Proficiency Testing

- Every laboratory conducting nonwaived testing must engage in proficiency testing (PT) program approved by CMS for each CLIA regulated specialty/subspecialty and each CLIA regulated analyte/test for which it is certified or seeks certification
- The following are specialties/subspecialties regulated by CLIA, which encompass analytes/tests regulated by CLIA and thus have specific CMS-prescribed PT requirements
  - Microbiology (specialty): bacteriology, mycobacteriology, mycology, parasitology, and virology
  - Diagnostic immunology: syphilis serology and general immunology
  - Chemistry: routine chemistry, endocrinology, and toxicology
  - Hematology: no subspecialties
  - Pathology: cytology (limited to gynecologic examinations)
  - Immunohematology: ABO group and D (Rho) typing, unexpected antibody detection, compatibility testing, and antibody identification
- Molecular testing is not recognized as specialty/subspecialty by CLIA; therefore, there are no specific CLIA requirements for molecular PT; however
  - Laboratories performing nonwaived testing not regulated by CLIA must adhere to alternative performance assessment for such tests/analytes, including molecular testing
  - Minimum guidelines require alternative assessment twice per year for each such assay

44.2.5.1 Proficiency Testing Guidelines

- Proficiency testing samples must be treated in same manner as patient specimens
  - Within regular clinical patient workload
  - By routine testing personnel
  - Using routine laboratory methods
  - Without communicating with any other laboratory regarding PT samples (until after PT event due date)
  - If laboratory has multiple sites performing test, no communication is allowed regarding PT results (until after PT event due date)
  - Must not send samples or aliquots to any other laboratory for analysis that laboratory is certified to conduct (violation can result in revocation of certificate)
    - Any laboratory receiving PT samples from another laboratory must report to CMS

  - All records must be kept for at least 2 years from PT event
- If laboratory has certificate suspended or limited due to PT performance failure on at least one specialty/subspecialty and analyte/test
  - Must perform satisfactorily on two consecutive PT events before certificate reinstatement and Medicare/Medicaid approval addressed
  - Suspension/limitation of certificate and/or Medicare/Medicaid penalties are at least 6 months in duration

44.2.5.2 Alternative Performance Assessment

- If specialty/subspecialty or analyte/test is not listed within CLIA guidelines, then alternative performance assessment must be performed in lieu of CMS-approved PT programs. This is the case for all molecular testing at the time of this writing
  - Recommended to participate in other PT programs, whenever available
  - Programs available for limited number of molecular genetic analytes (CAP PT programs)
    - Molecular PT samples often do not challenge entire testing process (i.e., sample may be purified DNA; therefore, it does not interrogate nucleic acid extraction)
    - Rare genetic disorders present further challenge: adequate samples may be difficult to procure; few laboratories will perform such testing
- Recommended strategies for alternative performance assessment in molecular testing
  - Because external quality assessments avoid bias of internally based assessments, recommend
    - Available PT programs
    - Interlaboratory exchange
    - Acquire externally derived PT materials
  - However, if such options are unavailable
    - Repeat testing of blinded (internal) samples
    - Blinded testing of known external samples
    - Sample exchange with research facility or foreign laboratory
    - Split samples for method comparison with different instrument or method
    - Interlaboratory data comparison

44.2.6 Facilities

- Laboratories performing nonwaived testing must meet certain criteria for facility quality
  - Adequate space, ventilation, and utilities
  - Procedures and controls to minimize contamination of specimens, equipment, and reagents
    - For molecular testing
      - This includes unidirectional workflow for nucleic acid amplification procedures (not contained in closed systems)
      - Separate areas for specimen preparation, reagent preparation, amplification, and detection
  - Appropriate equipment, reagents, and supplies
  - Compliance with applicable federal, state, local regulations
  - Adequate and accessible safety procedures to protect from physical, chemical, biochemical, electrical hazards, and biohazards
  - Records, slides, blocks, and tissues stored under proper conditions for preservation
  - Test requisitions and authorizations retained for at least 2 years
  - Test procedures retained for at least 2 years after procedure discontinued
  - Retain quality control and patient result records for at least 2 years
  - Retain analytical system performance records for lifetime of use (no less than 2 years total)
  - Retain proficiency testing records for at least 2 years
  - Retain quality assurance records for at least 2 years

- Retain (or be able to retrieve) test reports for at least 2 years after reporting (pathology reports for at least 10 years)
- If slides/blocks/tissues are retained by the molecular diagnostic laboratory, cytology slides must be retained for at least 5 years from examination, histopathology slides for at least 10 years from examination, blocks for at least 2 years from examination, and tissue until a diagnosis is made
  - Note that several professional practice guidelines recommend, although not required by CLIA, that molecular testing specimens should be retained as long as they are stable, and at least until next PT event, to allow retesting of specimens in the event assay error is identified
- If laboratory closes operation, all retained samples and records must be maintained for the required time frames listed

44.2.7 Quality Control and Assurance

- Laboratories conducting nonwaived testing must develop and document quality control and assurance procedures that monitor the preanalytic, analytic, and postanalytic phases of testing, in addition to general laboratory practices, and provide for continuous quality improvement

44.2.7.1 General Laboratory Quality

- The laboratory must
  - Ensure confidentiality of patient information
  - Ensure specimen identification and integrity
  - Document problems and complaints and necessary investigations
  - Identify communication issues between laboratory and end users
  - Develop policies on personnel competency
  - Identify tests/analytes not evaluated by a CMS-approved PT program (at least twice per year) or for which CMS-approved PT program is unavailable
  - Document and review any problems, corrective actions, effectiveness of corrective actions, and any necessary revisions to policies/procedures
- All of the above guidelines are useful for any clinical laboratory, including those conducting molecular testing

44.2.7.2 Preanalytic Phase

- All laboratories performing nonwaived testing must
  - Receive written/electronic request for patient testing (from individual authorized to request such testing)
    - Laboratory may accept verbal test requests if written/electronic authorization received within 30 days (or documentation of efforts to obtain)
    - Test requisition must list
      - Name and other suitable identifiers of requestor, including (as necessary) contact person for reporting alert value(s)
      - Patient name or unique patient identifier
      - Sex and age or date of birth of patient
      - Test(s) to be performed
      - Source, as necessary
      - Date (and time, if necessary) of specimen collection
      - Additional necessary information for interpretation
      - Note that patient chart or medical record may suffice as test requisition, provided it is available to laboratory at time of test, and CMS (upon request)
      - Information transcribed accurately
  - Establish specimen submission, handling, and referral policies
  - Establish and maintain preanalytic quality control and assurance procedures
- For molecular testing, the majority of laboratory errors occur in preanalytic phase
  - Major culprit thought to be inappropriate test ordering, likely resulting from inadequate information of those ordering testing (including clerks, nursing staff, and clinicians)

    - Users may not understand limitations/specifications of test, risking misinterpretation
    - Users may incorrectly order test, due to inappropriate indication or method
  - Therefore, although not required by CLIA, it is common professional practice that laboratory convey relevant test information clearly and accurately to users, including
    - Intended use of test
    - Indications for testing
    - Test method
    - Analytical and clinical specifications
    - Limitations of test
  - Also, although not required by CLIA, it is common professional practice for laboratory to request information relevant to test analysis and interpretation, particularly for molecular genetic testing, for example
    - Gender
    - Ethnicity
    - Selected clinical/laboratory information
    - Pedigree or family history
    - Informed consent and/or pretest genetic counseling (recommended prior to testing for many heritable conditions, required by law in many states)

44.2.7.3 Analytic Phase

- All laboratories conducting nonwaived testing must
  - Maintain a written procedure manual for each assay performed
  - Select appropriate test systems and establish and maintain criteria for proper operation of the test systems
  - Establish criteria and maintain proper storage of reagents and specimens
- Error rates in analytic phase of testing for molecular assays reported to range between 0.06% and 0.12% (comparable to nongenetic testing)
  - Errors reportedly encompass all phases of analytic process, including specimen handling
  - For instance, mutations or polymorphisms leading to false-positive or false-negative results have been reported for some common molecular genetic tests (such as cystic fibrosis mutations and *HFE*-associated hereditary hemochromatosis testing)
  - Such interferences are difficult to predict or evaluate during method evaluation, as they are specific by definition from specimen to specimen, and exert effects on assay specific to each variant

Analytical Test Validation or Verification

- All laboratories must establish or verify performance specifications for test systems implemented after April 24, 2003
  - To verify and document performance specifications of unmodified, FDA-cleared/approved test systems, a laboratory must demonstrate comparable performance specifications to those obtained by the manufacturer for the following parameters
    - Accuracy
    - Precision
    - Reportable range
    - Reference intervals – verify that manufacturer's stated intervals appropriate for laboratory-specific patient population
  - To establish and document performance specifications (for a modified FDA-cleared/approved test; or test system not subject to FDA clearance or approval, such as a laboratory-developed test, or test system without provided specifications), the laboratory must determine the following parameters, as applicable
    - Accuracy
      - Note that for a molecular test evaluating multiple targets, all variants should be included
      - For rare variants alternative samples such as cell lines or synthetic nucleic acid constructs may be necessary
        - Precision
        - Analytical sensitivity
        - Analytical specificity (including interfering substances)
          - Note that in addition to common exogenous or endogenous

interfering substances, other alleles, homologous sequences, and contaminants may need to be assessed on a case-by-case basis for positive/negative interference
    - Reportable range
    - Reference intervals; for molecular genetic testing, this may constitute the reference sequence
    - Any other performance characteristic required for test performance (such as clinical sensitivity/specificity)
- Test system maintenance must be performed and documented at least according to manufacturer's recommendations
  - If the test system is laboratory-developed or such protocols are not available, the laboratory must establish such maintenance protocols to ensure reliable test results

Calibration Procedures

- Calibration and calibration verification are necessary to ensure and document accuracy of testing and should interrogate the entire reportable range of the test system. The laboratory must
  - Follow the manufacturer's recommendations using provided materials (if possible) or
  - Use laboratory-established acceptance criteria and frequency (based on method verification or establishment) using traceable (to reference method or material) calibration materials
    - Include at least a minimal, a midpoint, and maximum value near upper limit of the reportable range
  - Must occur at least every 6 months and whenever any of the following occur
    - Complete change of reagents
    - Major preventive maintenance or repair of test system
    - Control materials detect potential failure mode, and other measures have not identified a source
    - More frequent calibration verification required

Quality Control Procedures

- Each test system must utilize control procedures that monitor entire analytical process, to detect real-time failure modes, as well as developing trends that may signal future failure. These control procedures should be developed in accordance with each test's verification or establishment of performance characteristics. Quality control (QC) should also determine long-term accuracy and precision
- QCs should be performed
  - In same manner as patient specimens
  - According to specifications verified or established in laboratory
  - At least once each day, patient testing is performed
    - For quantitative testing, two control materials of different concentrations
    - For qualitative testing, a negative and positive control
  - For testing involving extractions, at least one control should interrogate the extraction procedure
  - After complete change of reagents
  - After major preventive maintenance or repair
  - With different lot of calibrator material, if calibrators used as control material
  - All control procedures should be documented
- Specific QC issues for molecular testing
  - For molecular testing involving amplification
    - At least two controls are required
    - If reaction inhibition is of concern, include control to sensitive to inhibition (spike-in, or endogenous)
  - Although not required by CLIA, it is common professional practice
    - To minimize or detect contamination by nucleic acid target or amplicon
      - At least one no template control (NTC) sample should be tested, starting with amplification step
      - Preferable to test NTC through entire analytical process (i.e., starting with extraction step)

- If multiple genotypes to be assayed, rotate positive controls over time, to avoid excessive number of controls per run
- For rare variants with no natural controls available, use artificial or cell line controls
- For sequencing assays, bidirectional sequencing can be confirmatory
- If using pre-extracted DNA control (thus not controlling for extraction), can assay surrogate for DNA extraction in each sample such as exogenous spike-in control or internal control gene(s)
- If laboratory develops nucleic acid controls, must avoid cross-contamination of patient specimens and reagents

Comparison of Test Results

- Between laboratory methods: laboratories performing same test using different methods or same method at multiple sites must perform comparison testing twice per year to establish and document relationship
- Between results: laboratories must identify potential testing inconsistencies, based on criteria such as patient age, sex, diagnosis, and/or other test parameters or previous results

44.2.7.4 Postanalytic Phase

- All laboratories that perform nonwaived testing must monitor their postanalytic systems, including
  - Test reporting
    - Accuracy and timeliness
    - Reports must indicate patient identification, name and address of laboratory location, date reported, date performed, specimen source, and result and allow access to reference intervals
    - Reports can only be released to authorized persons, person responsible for using test result (if applicable), and laboratory that requested test
    - There must be a mechanism for reporting alert values to responsible entities
    - Reports must be maintained and readily accessible
- The laboratory must have a procedure for monitoring, assessing, and correcting postanalytic systems and documenting these activities

Postanalytic Issues Specific to Molecular Testing

- Most frequently encountered error is misinterpretation of test results by healthcare providers, often due to misunderstanding of testing limitations
  - As mentioned earlier, this can be remedied in preanalytical arena
- Clinical validity and utility difficult to assess because of rapid evolution of testing
  - Validity challenging with regard to genotype–phenotype correlations because novel/emerging targets are less characterized, in addition to the presence of variants of uncertain significance
  - Clinical utility difficult to determine as clinical impact of emerging diagnostics often requires more time and sample numbers, as opposed to analytical validation of assays
- Because of the implications of test results from molecular testing, laboratories should consider including additional information not required by CLIA, in clear, concise, and informative fashion, including
  - Test indication
  - Test method
  - Test specifications and limitations
  - Current standard nomenclature, as well as commonly used terms (to avoid confusion regarding results)
  - Detailed interpretation
  - Genetic counseling recommendation, when applicable
  - Implications of test results for family members
  - Updated or amended reports as new information becomes available
    - Previously reported deleterious mutations may be reclassified as benign variants/polymorphisms, or vice versa
    - Laboratories may consider placing a system to amend reports and contact patients and providers, in the event of

such changes, and to maintain a database of variants to identify those patients more readily
- Retain reports for at least 25 years, as opposed to minimal CLIA guidelines
  - Molecular genetic diagnoses are essentially a lifetime diagnosis that may need to be referenced and also may affect future relatives of the proband; also, new knowledge may surface that will alter original diagnosis

Release of Molecular Testing Reports

- As required by CLIA, laboratories may only release test results to authorized persons, the person responsible for using the test results (if applicable), and the laboratory that initially requested the test (42 CFR 493.1291)
  - Note that molecular genetic test results may be necessary for the healthcare of family members of patients
    - Healthcare provider of family member would need to receive authorization from patient to request patient test information
    - Laboratory should request and confirm patient's authorization before releasing results

44.2.8 Personnel

- Laboratories performing nonwaived testing are regulated with respect to all personnel involved in testing

44.2.8.1 Provider-Performed Microscopy Procedures

- The laboratory must have
  - Qualified director (42 CFR Sec. 493.1357)
    - Qualified to manage and direct the personnel and perform PPM procedures
    - Carrying current license as laboratory director issued by state, if required
    - Either a physician, midlevel practitioner (authorized by the state to practice independently), or a dentist
    - Who directs no more than five laboratories
  - Qualified testing personnel (42 CFR Sec. 493.1363)
    - Who possess a current license issued by the state, if required
    - Either a physician, midlevel practitioner (authorized by the state to practice independently), or a dentist

44.2.8.2 Moderate Complexity Testing

- The laboratory must have
  - A qualified director (42 CFR Sec. 493.1405)
    - Responsible for overall operation and administration of laboratory, including employment of competent personnel
    - Accessible to laboratory for onsite, telephone, or electronic consultation
    - Who may direct no more than five laboratories
  - A qualified technical consultant (42 CFR Sec. 493.1411)
    - Responsible for technical and scientific oversight of laboratory
    - Available to provide onsite, telephone, or electronic consultation
  - A qualified clinical consultant (42 CFR Sec. 493.1417)
    - Who provides consultation regarding appropriateness of testing and interpretation of results
    - Available to provide clinical consultation to clients
    - Qualified testing personnel (42 CFR Sec. 493.1423)
    - Responsible for specimen processing, test performance, and reporting results

44.2.8.3 High Complexity Testing

- Laboratories performing high complexity testing must have
  - A qualified director (42 CFR Sec. 493.1443)
    - Responsible for overall operation of laboratory, including employment of competent personnel

    - If qualified, may perform duties of the technical supervisor, clinical consultant, general supervisor, and/or testing personnel or delegate these responsibilities
    - Must be accessible to laboratory to provide onsite, telephone, or electronic consultation
    - May direct no more than five laboratories
  - A qualified technical supervisor (42 CFR Sec. 493.1449)
    - Responsible for the technical and scientific oversight of the laboratory
    - Accessible to the laboratory to provide onsite, telephone, or electronic consultation
  - A qualified clinical consultant (42 CFR Sec. 493.1455)
    - Provides consultation regarding appropriateness of testing ordered and interpretation of results
    - Available to provide consultation to clients
  - A qualified general supervisor (42 CFR Sec. 493.1461)
    - Responsible for day-to-day supervision or oversight of the laboratory operation and personnel
    - Accessible to testing personnel at all times; testing is performed to provide onsite, telephone, or electronic consultation
  - Qualified testing personnel (42 CFR Sec. 493.1489)
    - Responsible for specimen processing, test performance, and results reporting

44.2.9 Inspection

- Basic inspections: all laboratories issued a CLIA certificate and CLIA-exempt laboratories must permit CMS or a CMS agent to conduct validation and/or compliance inspections, which may include
  - Testing of samples or performing procedures
  - Interviews of personnel
  - Observation of testing
  - Examination of all records

44.2.9.1 Laboratories Issued a Certificate of Waiver or Provider-Performed Microscopy

- Not subject to biennial inspections
- However, CMS or a CMS agent may conduct an inspection at any time to
  - Determine if laboratory constitutes "imminent and serious risk" to public health
  - Evaluate complaints
  - Determine whether laboratory performing tests outside scope of certificate
  - Collect information regarding appropriateness of waived or PPM procedures

44.2.9.2 Laboratories Issued a Certificate of Compliance

- Laboratory issued certificate of registration must permit initial inspection to assess laboratory compliance with applicable CLIA regulations
- CMS or a CMS agent may conduct subsequent inspections on biennial basis (or with other frequency, based on compliance history)

44.2.9.3 Inspection of Clinical Laboratory Improvement Act-Exempt Laboratories or Those Requesting or Issued a Certificate of Accreditation

- CMS or a CMS agent may conduct validation inspection of any accredited or CLIA-exempt laboratory at any time
  - CMS or a CMS agent may conduct complaint inspection of CLIA-exempt laboratory (or one requesting or issued certificate of accreditation) at any time upon receiving complaint
  - If validation or complaint inspection results in noncompliance
    - Laboratory-issued certificate of accreditation is subject to full review
    - CLIA-exempt laboratory is subject to appropriate enforcement actions under state licensure program

44.2.10 Enforcement of Regulations

- The Clinical Laboratory Improvement Act of 1967, amended by CLIA '88, set forth requirements for all laboratories performing diagnostic testing on human specimens and
  - Required federal certification of such laboratories
  - Granted the Secretary of HHS broad authority for enforcement, allowing
    - Intermediate sanctions
    - Suspension, limitation, or revocation of certification
    - Civil suit (for laboratory activity posing "significant hazard" to public health)
    - Imprisonment or fine (for conviction of intentional violation of CLIA regulations)

44.2.10.1 Basis of Sanctions

- Violations of CLIA regulations are detected by
  - Deficiencies found by CMS (or its agents)
  - Proficiency testing results
- One or more of the following sanctions imposed upon finding "condition level" violations (42 CFR Sec. 493.1806 for additional information), based on one or more following factors
  - Deficiencies pose "immediate jeopardy"
  - Nature, incidence, severity, and duration of deficiencies
  - Repeated offenses
  - Accuracy, extent, and availability of laboratory records, with regard to deficiency
  - Relationship of multiple deficiencies
  - Overall compliance history
  - Intended outcome of sanctions
  - Whether laboratory has taken opportunity to correct deficiencies
  - Recommendation by state agency as to appropriate sanction(s)

44.2.10.2 Types of Sanctions

- Principal sanctions
  - Three principal CLIA sanctions
    - Suspension
    - Limitation
    - Revocation of CLIA certificate
- Alternative sanctions
  - A directed plan of correction
  - State onsite monitoring
  - Civil money penalty
  - Civil suit
  - Criminal sanctions: individual convicted of intentionally violating CLIA regulation may be imprisoned or fined
- Additional sanctions
  - For laboratories approved for Medicare reimbursement
    - Principal sanction: cancellation of laboratory approval for Medicare reimbursement
    - Alternative sanctions
      - Suspension of payment for subset of tests performed on or after sanctioning date
      - Suspension of payment for all tests performed on or after sanctioning date
- Medicare consequences
  - Suspension or revocation of any CLIA certificate leads to CMS cancellation of laboratory approval for Medicare reimbursement
  - Limitation of any CLIA certificate leads to CMS limitation of Medicare approval to specialties/subspecialties still authorized by certificate
- Medicaid consequences
  - Reimbursement under state plans limited if CLIA certificate and/or state licensure/approval is limited

44.3 Other Governmental Regulations and Regulatory Agencies

- In addition to the Department of Health and Human Services (HHS) and its CMS branch for overseeing CLIA regulations, other branches of federal and state governments have regulatory impact on laboratories – and HHS and its branches (including CMS) have other regulatory oversight of clinical laboratories, in addition to CLIA regulations.

The following are selected federal and state regulatory agencies that have significant impact on the modern molecular diagnostic laboratory

44.3.1 State Departments of Health

- As described in Section III under CLIA certification, various state Departments of Health exert a spectrum of regulatory influence over clinical laboratories
- New York state and state of Washington have been granted "exemption" by CMS
  - Clinical laboratories licensed in the state of New York or Washington is eligible for CLIA-exempt status

44.3.2 Office of the Inspector General

- The Office of the Inspector General (OIG) is a division of the Department of Health and Human Services
- Established in 1976, involved in auditing and identifying inefficiency, fraud, and abuse of HHS programs, including Centers for Medicare and Medicaid Services
- Impact on clinical laboratories is to identify fraud and abuse of CMS with respect to clinical laboratory testing, particularly issues of reimbursement and medical necessity of testing
  - Molecular diagnostics is a rapidly evolving field, so medical necessity of this type of testing is often not well disseminated, and therefore reimbursements are at risk for denial without a certain amount of vigilance

44.3.3 Centers for Medicare and Medicaid Services

- A division of HHS, together the CMS are the largest healthcare entity in the United States
- The Medicare and Medicaid programs were enacted in 1965, as Title XVIII and Title XIX of the Social Security Act, targeting healthcare to individuals aged 65 or older, impoverished youths lacking parenting (including their caretakers), the blind, and those with disabilities
- Currently Medicare primarily covers the elderly; however, it also provides coverage for some persons with disability or end-stage renal disease
- Medicaid primarily covers low-income individuals through cooperation with state programs
- The impact of CMS through CLIA has already been presented; however, CMS also sets reimbursement rates for laboratory testing that not only impacts its own healthcare system but also is referenced by many other third party payers
  - Because of the reagent and labor costs of most molecular diagnostic assays, as well as difficulties with reimbursement of such leading edge technology and testing, CMS reimbursement rates are vital to the operation of the modern molecular diagnostic laboratory

44.3.4 United States Food and Drug Administration, Office of In Vitro Diagnostic Device Evaluation and Safety

- The Food and Drug Administration (FDA) is a division of HHS, and its Office of In Vitro Diagnostic Device Evaluation and Safety (OIVD) was founded in 2002 to consolidate regulation of in vitro diagnostic devices (IVDs) and test kits
- The OIVD regulates both in-home and laboratory IVDs
- As noted earlier in Section III, the OIVD also regulates commercial in vitro diagnostic testing complexity (waived vs. moderate/high complexity) for CMS
- Of the three divisions (Chemistry and Toxicology Devices, Immunology and Hematology Devices, and Microbiology Devices), it is the Division of Immunology and Hematology Devices (DIHD) that regulates the

majority of IVDs and test kits that would be encountered in a molecular diagnostic laboratory

44.3.5 Equal Employment Opportunity Commission

- Created by Title VII of the Civil Rights Act of 1964 (amended by Congress in the Equal Employment Opportunity Act of 1972), the Equal Employment Opportunity Commission (EEOC) enforces federal laws prohibiting discrimination against applicants or employees on the basis of race, religion, sex, age, disability, or genetic information
- Most employers with 15+ employees, labor unions, and employment agencies are covered by EEOC
- Hiring of personnel for molecular diagnostics laboratories is subject to the applicable EEOC regulations

44.3.6 Occupational Safety and Health Administration

- A division of the United States Department of Labor, the Occupational Safety and Health Administration (OSHA) was created by Congress with the Occupational Safety and Health Act of 1970
- This administration promotes and enforces healthy and safe workplace standards to protect employees
- As a place of employment, molecular diagnostic laboratories must adhere to OSHA standards to protect their employees, which are the single greatest asset to any laboratory

44.3.7 Environmental Protection Agency

- The Environmental Protection Agency (EPA) was established in 1970 because of concerns over pollution and the physical environment
- In order to protect the environment, research, monitoring, standards, and enforcement were enacted
- The molecular diagnostic laboratory must dispose of environmentally hazardous laboratory materials in accordance with EPA standards
  - Commonly occurring such materials include ethidium bromide, formalin, and xylene

44.3.8 Nuclear Regulatory Commission

- US Nuclear Regulatory Commission (NRC) was established in 1974 to create and enforce federal guidelines ensuring safe use of radioactive materials (for nonmilitary purposes only) and protect the public and environment
- NRC provides oversight through licensing, inspections, and enforcement of regulations of nuclear products such as nuclear power plants, nuclear medicines, and other radioactive materials (such as research reagents)
- Although most molecular diagnostic laboratories have replaced radioactive reagents such as Southern blot probes and radiolabeled nucleotides, with fluorescent or chemiluminescent analogs, if a radioactive material is employed in the laboratory, its use is governed by NRC guidelines

44.3.9 Department of Transportation

- The Department of Transportation (DOT) was established in 1966 to provide efficient, economical, and environmentally conscious national transportation system
- Primary agency in federal government regarding safety, adequacy, and efficiency of transportation system
- DOT is responsible for regulating transport of biohazard materials (including human specimens)
- All specimens received and referred by clinical laboratories must adhere to DOT guidelines

44.3.10 Health Insurance Portability and Accountability Act

- The Health Insurance Portability and Accountability Act (HIPAA) of 1996 (Public Law 104–191), an amendment to the Internal Revenue Code of 1986, aimed to improve portability and continuity of health insurance; decrease inefficiency, fraud, and abuse in health insurance and healthcare; promote healthcare savings accounts; improve accessibility to long-term care; and simplify administration of health insurance
- One benefit of this Act was a reform to healthcare information management known as the HIPAA Privacy Rule, safeguarding patient "protected health information (PHI)" from unauthorized written, oral, or electronic dissemination
- This Act is the basis for many current individual healthcare information standards and must be followed by all clinical laboratories

44.3.11 Ban on Physician Self-referral (Stark Law)

- A provision in the Omnibus Budget Reconciliation Act of 1989 (OBRA 1989), known as "Stark I" (named for Congressman Pete Stark, who sponsored bill), barred physicians from self-referring Medicare patients for laboratory services (effective 1992)
- There have been several additional provisions (known as Stark II and III), now among other changes governing self-referral for Medicare and Medicaid patients
- Directors of molecular diagnostic laboratories must understand these provisions, especially those who are also clinicians (and thus may refer testing to a laboratory they direct), in order to avoid unintentional wrongdoing

44.3.12 Three-Day Rule

- Omnibus Budget Reconciliation Act of 1990, OBRA 1990, Pub. L. 101–508; Preservation of Access to Care for Medicare Beneficiaries and Pension Relief Act of 2010, Pub. L. 111–192, Section 102
- These laws require that diagnostic services (including laboratory testing), or other services related to an inpatient admission, provided by the admitting hospital (or an entity owned/operated by the hospital) for admitted patient during the 3 days preceding date of admission is considered part of the inpatient stay (for admission to a hospital paid under inpatient prospective payment system, "IPPS"). The payment window for admission to a hospital not paid by IPPS (such as psychiatric, inpatient rehabilitation, long-term care, children's, and cancer hospitals) is 1 day preceding the date of admission
- These laws state that reimbursement for testing performed in a molecular diagnostic laboratory under any of the circumstances listed above would be included on a DRG system

44.4 Nongovernmental Agencies

- In addition to governmental organizations and regulations, there are many nongovernmental agencies that impact molecular diagnostic laboratories. An exhaustive list is beyond the scope of this writing; however, three significant organizations will be introduced

44.4.1 College of American Pathologists

- The CAP, a nonprofit organization, was formed in 1947 to promote pathology education, research, practice, and service for patients, physicians, hospitals, and the general public
- As noted earlier, CAP is a CLIA-deemed authority
- CAP accreditation recognized by JCAHO
- Offers the largest proficiency testing program in the United States
- Offers a wide range of proficiency testing programs for molecular diagnostic laboratories, including genetic and somatic mutation testing
- Offers an accreditation program tailored to molecular diagnostic laboratories

44.4.2 Commission on Office Laboratory Accreditation

- COLA is a nonprofit organization originally founded in 1988 to assist physician office laboratories (POLs) comply with CLIA
- COLA laboratory accreditation program granted CLIA deeming authority in 1993
- COLA accreditation recognized by JHACO
- Expanded program to include accreditation of hospital and independent laboratories
- Accredits following types of laboratories: physician office, community hospital, mobile clinic, Veterans Administration, and Department of Defense
- CMS-approved to accredit the following specialties: chemistry, hematology, microbiology, immunology, and immunohematology/transfusion services

44.4.3 Joint Commission on Accreditation of Healthcare Organizations

- Founded in 1951, JCAHO is an independent and nonprofit organization that evaluates and accredits healthcare organizations
- To maintain accreditation, a healthcare organization undergoes onsite surveys at least every 3 years, with laboratories surveyed at least every 2 years
- In surveying a healthcare organization, JCAHO typically surveys the laboratory as well
- Has accreditation programs for following healthcare organizations: hospital, home care, behavioral health, laboratory, ambulatory (including office-based surgery and primary care medical home), long-term care, and critical access hospitals
- Has CLIA-deemed authority, and is CMS-approved to accredit all CLIA specialties
- JCAHO accreditation also satisfies select state licensure and insurer requirements for laboratories

Further Reading

Centers for Medicare and Medicaid Services (http://www.cms.gov/)
Chronology of CLIA regulations (http://wwwn.cdc.gov/clia/chronol.aspx)
CLIA (http://wwwn.cdc.gov/clia/pdf/PHSA_353.pdf, http://wwwn.cdc.gov/clia/regs/toc.aspx)
CMS-deemed authorities and Exempt States (http://www.cms.gov/CLIA/13_Accreditation_Organizations_and_Exempt_States.asp#TopOfPage)
COLA (http://www.cola.org/)
College of American Pathologists (CAP, http://www.cap.org/apps/cap.portal)
Environmental Protection Agency (http://www.epa.gov/)
Good Laboratory Practices for molecular genetic testing for heritable diseases and conditions (http://www.cdc.gov/mmwr/pdf/rr/rr5806.pdf)
Health Insurance Portability and Accountability Act (HIPAA, http://www.hhs.gov/ocr/privacy/hipaa/understanding/index.html)
Joint Commission on Accreditation of Healthcare Organizations (JCAHO, http://www.jointcommission.org/)
List of CLIA waived testing (https://www.cms.gov/CLIA/downloads/waivetbl.pdf)
Listing of provider-performed microscopy procedures (https://www.cms.gov/CLIA/downloads/ppmp.list.pdf)
New York State Department of Health (http://www.health.ny.gov/)
Occupational Safety and Health Administration (http://www.osha.gov/index.html)
Office of Inspector General, United States Department of Health and Human Services (http://oig.hhs.gov/)
Searchable database of CLIA testing (http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCLIA/search.cfm)
Stark Rule Phase III (http://www.gpo.gov/fdsys/pkg/FR-2007-11-15/pdf/07-5655.pdf)
State of Washington Department of Health (http://www.doh.wa.gov/)
Three Day Rule (https://www.cms.gov/AcuteInpatientPPS/08a_Three_Day_Payment_Window.asp)
United States Department of Transportation (http://www.dot.gov/)
United States Equal Employment Opportunity Commission (http://www.eeoc.gov/)
United States Food and Drug Administration (http://www.fda.gov/)
United States Nuclear Regulatory Commission (http://www.nrc.gov/)

Molecular Pathology Reporting

45

Stacey S. O'Neill and Margaret L. Gulley

Contents

45.1 Goal of the Molecular Pathology Report 1109

45.2 Terminology 1109
45.2.1 Abbreviations 1109
45.2.2 Human Gene Nomenclature 1110
45.2.3 Basics of Reporting Sequence Variants 1110

45.3 Content of a Molecular Pathology Report 1111
45.3.1 Identifying Information 1111
45.3.2 Results 1111
45.3.3 Interpretation 1113
45.3.4 Comment 1114
45.3.5 Methods 1114
45.3.6 Billing Information 1114

45.4 General Considerations 1115

45.5 Special Considerations for Molecular Reporting 1115

45.6 Modifications to Reports 1115

45.7 Role of Other Healthcare Workers in Genetic Test Reporting 1116

Further Reading 1116

45.1 Goal of the Molecular Pathology Report

- A molecular pathology report communicates results of a molecular test and their significance in a concise manner to help guide patient management
- The report should be written such that a nonspecialist physician can understand and act on its content
- When appropriate, the report should provide an interpretation of results in the context of the setting and indications for testing in that particular patient

45.2 Terminology

45.2.1 Abbreviations

- Abbreviations are generally not acceptable due to possible misinterpretation or confusion by clinicians
- Exceptions include those abbreviations defined upon first use in the report and the following universally accepted abbreviations
  - DNA for deoxyribonucleic acid
  - RNA for ribonucleic acid
  - cDNA for complementary deoxyribonucleic acid
  - mRNA for messenger RNA
  - PCR for polymerase chain reaction
  - bp for base pair and kb for kilobase

S.S. O'Neill, MD, PhD (✉) • M.L. Gulley, MD
University of North Carolina, Chapel Hill, NC, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_45,

Table 45.1 Colloquial and standard nomenclature for common gene mutations

Colloquial nomenclature	Standard nomenclature
A1AT Z mutation	*SERPINA1* NM_001127705.1: c.1096G>A [p.Glu366Lys]
A1AT S mutation	*SERPINA1* NM_001127705.1: c.863A>T [p.Gul288Val]
BRAF V600E	*BRAF* NM_004333.4: c.1799T>A [p.Val600Glu]
CFTR F508del	*CFTR* NM_000492.3: c.1521_1523delCTT [p.Phe508del]
Factor V Leiden	*F5* NM_000130.4: c.1746G>A [p.Arg534Gln]
HFE C282Y	*HFE* NM_000410.3: c.845G>A [p.Cys282Tyr]
HFE H63D	*HFE* NM_000410.3: c.187C>G [p.His63Asp]
JAK2 V617F	*JAK2* NM_004972.3: c.1849G>T [p.Val617Phe]

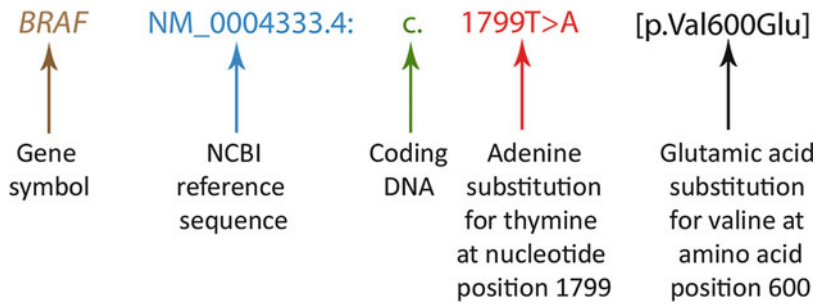


Fig. 45.1 Standardized gene nomenclature. The description begins with the gene symbol in *italics* followed by the NCBI reference sequencing number and a colon mark. The type of reference sequence is then indicated by a single letter followed by a *period*. The nucleotide position and change is indicated. The presumed amino acid change and involved codon is *bracketed*

- Abbreviations with multiple meanings should be clarified
  - For example, RT-PCR could be interpreted as real-time PCR or reverse transcription PCR

45.2.2 Human Gene Nomenclature

- Standardized gene nomenclature is critical so that the gene being tested is unambiguous
- The Human Genome Organization (HUGO) Gene Nomenclature Committee has established an online database of designated accepted gene symbols and their longer, more descriptive names (Table 45.1)
- There are ongoing efforts to implement widespread use of HUGO-designated gene symbols in molecular reporting, publishing, and marketing of probes
- During this transition, it is acceptable to include commonly used gene symbols in parentheses in the report to ensure the clinician understands which gene has been tested, e.g., *ERBB2 (HER2)*
- Microbial organism nomenclature follows conventions described in author guidelines of journals sponsored by the American Society for Microbiology

45.2.3 Basics of Reporting Sequence Variants

- Standardized gene nomenclatures is critical so that results of gene testing are unambiguous (Fig. 45.1)
- Gene symbols are italicized while the encoded protein is not (e.g., BRCA1 protein is encoded by the *BRCA1* gene)
- Sequence variants should be described relative to a reference sequence
  - The National Center for Biotechnology Information (NCBI) provides an online

collection of reference sequences called RefSeqGene that serves as an anchor by which to describe variants at the genomic, coding, or protein levels. See http://www.ncbi.nlm.nih.gov/refseq/rsg
    - Each coding (mRNA) sequence has an alphanumeric designation beginning with "NM"
    - Proteins are assigned an identification number beginning with "NP"
    - Predicted mRNA transcripts and proteins begin with "XM" and "XP," respectively
  - Designate the type of reference sequence used
    - "g" for genomic sequence
    - "c" for coding sequence
    - "r" for RNA sequence
    - "m" for mitochondrial sequence
    - "p" for amino acid (protein) sequence
- Designation of specific changes
  - Substitutions
    - Nucleotide: Nucleotide position followed by the reference sequence nucleotide ">" altered nucleotide (e.g., *BRAF* NM_4333.4:c.1799T>A designates a T to A nucleotide change at position 1,799 in the coding DNA sequence of the *BRAF* gene)
      - Nucleotides in DNA written in capital letters
      - Nucleotides in RNA written in lower case H-Amino acid: Use three-letter abbreviations for amino acids
      - Reference sequence amino acid followed by codon position then altered amino acid (e.g., p.Cys282Tyr, colloquially abbreviated C282Y)
  - Insertions: first and last position affected, separated by an underscore followed by "ins" and residue(s) inserted (e.g., c.76_77insG)
  - Deletions: position affected followed by "del" and residue(s) deleted (e.g., c.35delG)
  - Duplications: first and last position affected followed by "dup" and residue(s) duplicated (e.g., c.77_79dupCGG)
  - Two sequence changes in different alleles of the same gene are each listed in brackets separated by a semicolon (e.g., c.[35delG];[167delT])
  - It is recommended that a nucleotide change and the inferred amino acid change are reported, with the inferred change in brackets since it was not specifically tested
  - There are currently no consensus standards to report DNA methylation or other epigenetic changes. Use a practical approach of identifying the gene target or promoter, describing the molecular findings, and interpreting them in the context of the clinical setting

45.3 Content of a Molecular Pathology Report

- The College of American Pathologists (CAP) has identified key elements that should be included in every lab report, as outlined on the CAP Lab General Checklist and further described below
- Since molecular testing is rapidly evolving, a comprehensive report is recommended to include an interpretation, a comment that provides additional educational information, a brief description of methods, and billing information
- Model reports for heritable disease and cancer mutation testing are presented as Boxes 45.1 and 45.2, respectively

45.3.1 Identifying Information

- Patient identifying information including full name, medical record number, and date of birth
- Ordering physician's name and address
- Reporting laboratory name and address with optional additional contact information such as phone number, fax number, email address, and facility website
- Sample details including specimen type and status at receipt (e.g., fresh, frozen, paraffin-embedded)
- Interpreting physician's name

45.3.2 Results

- Concise, objective statement of the result based on analytic interpretation of raw data and controls

Box 45.1: A Model Report for Inherited Disease Testing

The Genetics Laboratory
1111 Laboratory Avenue
Nowhere, State 00839
(555) 867-5309
www.GenLab.org

Patient Name:	**Carol C. Patient**		
ID number:	**038907343**	**Race/Ethnicity:**	Caucasian
DOB:	**09/28/1981**	**Family history of cystic fibrosis:**	Negative
Gender:	Female	**Personal history:**	Negative
Ordering Physician:	Dr. Good		
Accession Number:	080520111155		

Test Performed:	***CFTR* mutation panel**
Indication for testing:	**Cystic fibrosis carrier testing in an asymptomatic patient**
Specimen Type:	Blood
Date of Collection:	11-5-2011

Result: **Heterozygous for *CFTR* delta F508 mutation**

Interpretation: **This patient is a carrier of a cystic fibrosis-associated mutation**

Comment:

- A carrier has one deleterious mutation and is not clinically affected by cystic fibrosis.
- Genetic consultation is recommended to inform the patient of their risk for having an affected child, and implications for other family members. To arrange for the patient to be seen by a geneticist or genetic counselor, call (555) 867-5309.
- Each of this patient's children has a 1/2 (50%) chance of inheriting this patient's mutation.The risk of having a child affected by cystic fibrosis is based upon the mutation status this patient's reproductive partner.
- Testing is indicated for blood relatives of this patient because of the ir increased risk of harboring a familial mutation.
- Cystic fibrosis is an autosomal recessive disease requiring two deleterious mutations, one inherited from the mother and the other from the father.
- Clinical pathology consultation about this test result is available by contacting the laboratory.General information is available at GeneTests (www.genetests.org) and at this laboratory's website. Patient/family information is available at Genetics Home Reference (http://ghr.nlm.nih.gov).
- Mutation nomenclature:*CFTR*NM_000492.3:c.1521_1523delCTT [p.Phe508del]

Method: The *CFTR* gene was tested for the presence of 23 mutations and variants by polymerase chain reaction and allele specific oligonucleotide ligation using the Cystic Fibrosis Genotyping Assay (Abbott Molecular, order # 06L20-01; http://www.abbottmolecular.com/us/products/genetics/sequencing/cystic-fibrosis.html).

Interpreted and electronically signed by John Doe, M.D.

Date reported:11/12/2011

Billing codes: 83890; 83900; 83901x14; 83914x16; 83909

Adapted from Lubin et al. (2009)

- Generally placed at the top of the report for ease of use
- Raw data such as bands on a gel or melt curves are not directly illustrated in the report
- The reference range or expected result in a normal individual should be provided for comparison to patient results
  - Few assays such as HLA typing and parentage testing do not have relevant reference ranges
  - In microbiology applications, it is helpful to indicate whether the identified agent is pathologic or commensal flora
- Assay performance characteristics do not need to be integrated into every report, particularly since performance characteristics can vary depending on the clinical setting and the indications for testing. United States (US) federal regulations require that a clinical consultant be available to discuss

Box 45.2: A model report for cancer mutation testing

The Genetics Laboratory
1111 Laboratory Avenue
Nowhere, State 00839
(555) 123-4567
www.GenLab.org

Patient: **Jim T. Patient**
MRN: **1234567-8**
DOB: 02/02/1951
Gender: Male
Ordering Physician: Dr. Clinician

Molecular Pathology Report

Accession Number: MO11-1234

***KRAS* MUTATION ANALYSIS**

SPECIMEN:
Paraffin sections labeled AB11-12345 A1 are received from Pathology Laboratory, with an original collection date of 11/2/2011 and a reported diagnosis of colorectal adenocarcinoma.

RESULT:
Positive for c.34 G>T [p.Gly12Cys] mutation in codon 12 of the *KRAS* gene
Gene nomenclature: *KRAS* NM_4985.3: c. 34G>T [p.Gly12Cys])

INTERPRETATION:
Presence of an activating *KRAS* mutation in this patient's tumor diminishes the likelihood of response to anti-EGFR chemotherapy. (www.nccn.org; Reference: PMID: 19386128)

METHOD:
Following tumor cell enrichment by macrodissection, pyrosequencing was performed on codons 12, 13, and 61 of *KRAS* DNA.

This test was developed and its performance characteristics determined by The Genetics Laboratory. It has not been approved by the US Food and Drug Administration. However, such approval is not required for clinical implementation, and test results have been shown to be clinically useful. This laboratory is CAP accredited and CLIA certified to perform high complexity testing.

Electronically Signed by: David T. Pathologist, MD Date: 11/16/2011

Billing Fee Code(s): 83890; 83907; 83892; 83898x2; 83904x2

results and their significance with clinicians upon request
- Results of positive and negative controls are not reported

45.3.3 Interpretation

- Practicing clinicians perceive the interpretation as a key component of a molecular report that helps apply the findings in clinical decision making
- There are two types of interpretation
  - Analytic interpretation processes the objective results into a conclusion that a clinician can use diagnostically. This reportable result is displayed in the Sect. 45.3.2 above

  - Clinical interpretation describes the results in the context of the clinical indications for testing and implications in patient care, thus describing the clinical significance of the results in general or in relation to the particular patient's medical information
- Interpretation should be done by a qualified individual having both technical and medical expertise
- Discuss whether a genetic change is a benign polymorphism, disease-causing mutation, or variant of unknown significance
- Describe major limiting factors impacting confidence in the results, if appropriate. Examples are
  - Analytic and clinical sensitivity of the assay
    - Analytic sensitivity reflects how well the assay is able to detect the specified analyte (e.g., no *BCR-ABL1* detected to a sensitivity of .001 international units)
    - Clinical sensitivity reflects how well the assay is able to detect the related disease (e.g., a diagnosis of chronic myelogenous leukemia is highly unlikely)
  - Quality of the specimen
    - Marginal specimen quantity or quality due to collection or handling variables. For example, scant tumor visualized by microscopy, or previously frozen specimen received thawed, or no organ donor specimen by which to compare posttransplant identity test result

45.3.4 Comment

- This is a flexible section of the report that can be used to describe pertinent information beyond the result and interpretation
- Recommendations for additional testing or referrals (i.e., genetic counseling)
- Educational resources and references

45.3.5 Methods

- Brief statement of the gene or locus tested and the methodology used since there are often multiple methodologies possible for molecular testing and each has different performance characteristics and billing codes
- Source of a commercial kit (specify version number) or analyte-specific reagent
- If a class I analyte-specific reagent is used in the US, the following disclaimer should appear
  - "This test was developed and its performance characteristics determined by (laboratory name). It has not been cleared or approved by the US Food and Drug Administration"
  - Example additional wording might include, "This test is used for clinical purposes. It should not be regarded as investigational or for research. This laboratory is certified under the Clinical Laboratory Improvement Amendments of 1988 (CLIA-88) to perform high complexity clinical laboratory testing"

45.3.6 Billing Information

- Current Procedural Technology (CPT) coding is a system maintained by the American Medical Association to identify medical, surgical, and diagnostic services to facilitate communication between providers, patients, and payers
  - CPT codes for molecular microbiology testing are mainly organism-based, leading to transparency for payers
  - CPT codes are being devised and implemented for selected heritable disease, HLA, and oncology applications
  - CPT codes for other applications use "stacking codes" based on the types of technical procedures performed, such as DNA extraction, nucleic acid amplification, sequencing, or melt curve analysis. Concerns have been raised by payers that stacking codes do not provide information about which gene is being tested
- International Classifications of Disease (ICD) coding is a system used to designate diagnoses and description of symptoms to unify communication between providers, patients, and payers

  - ICD codes are used to document the indication for testing and to justify the medical necessity of the test that is ordered
  - Codes are developed and monitored by the World Health Organization (WHO)
  - In the US, the Centers for Medicare and Medicaid Services regulates modifications to ICD codes

45.4 General Considerations

- Reports should generally be no more than one printed page
- Timely reporting
  - Target turnaround times should be established by the laboratory for each assay. Urgency of results for clinical management and complexity of methods and interpretation should factor into targeted turnaround times
  - A system for monitoring turnaround times should be in place
- US laboratories must have policies and procedures to protect reports as patient healthcare information, in compliance with the Health Information Portability and Accountability Act (HIPAA)
- Clinical laboratory reports should be retained for at least 2 years as recommended by CAP

45.5 Special Considerations for Molecular Reporting

- Patient characteristics can influence results and/or interpretation
  - In cystic fibrosis carrier testing, ethnicity affects the calculation of residual risk in a patient with a negative screening test
- Specimen factors can affect the quality of data and the confidence in results that may warrant explanation in the report
  - A delay in processing samples for RNA extraction can decrease specimen quality
  - Type of fixative and prolonged fixation time can negatively impact recovery of amplifiable nucleic acid
  - Macrodissection of tumor from formalin-fixed paraffin-embedded tissue sections can enrich the proportion of malignant cells, thus increasing assay sensitivity
- Patient safety considerations mandate a clear, concise report stating results and their significance, thus facilitating appropriate action by the healthcare team
  - There is currently no consensus on reporting benign polymorphisms
    - Benign polymorphisms may later be found to be functionally or clinically important which argues for their inclusion in a molecular report
    - On the other hand, clinicians may misinterpret a sequence variant as deleterious, arguing against reporting benign polymorphisms. A long, cluttered report may cause important findings to be missed
- Clinical applications of research testing
  - Molecular pathology is a constantly evolving field with rapid advancements emanating from translational research bridging into clinical-grade testing
  - However, results of research testing should not be included in a clinical report; only results of validated assays should be reported to patients or their healthcare providers
  - If research testing was performed as part of an institutional review board-approved protocol, results should be handled as stated in that research protocol

45.6 Modifications to Reports

- Reports may require modification due to errors, or additional testing or information
- Any changes to a report need to be well documented, and the revised report should

clearly indicate the change from the previous version
- An addendum is issued when new information is added to the report
- An amended report is issued when the result or meaning of the result is changed
- The ordering physician should be notified of a modification that changes the result or interpretation
  - Documentation of this notification can be included in the revised report

45.7 Role of Other Healthcare Workers in Genetic Test Reporting

- Studies of test ordering practices have shown that nurses or clerks frequently fill out test requisition forms, sometimes leading to misordering or errors in additional information to be conveyed on the requisition form
  - Consider strategies promoting improved outcomes, such as electronic ordering by physicians
- Studies also reveal that test results are sometimes relayed to patients by individuals lacking adequate clinical expertise, such as clerks
  - Individuals who report results to patients or who act on results should
    - Understand results and their medical significance
    - Educate the patient and provide guidance on next steps in clinical evaluation and management, such as referrals for genetic counseling
    - Refer the patient to patient-oriented support resources

Further Reading

American Medical Association. Coding billing insurance; current procedural technology. 2011. Available at http://www.ama-assn.org/ama/pub/physician-resources/solutions-managing-your-practice/coding-billing-insurance/cpt.page. Accessed 8 Sept 2011.

Check W. Fleshing out molecular pathology reports. CAP Today. 2004;18:1–24.

College of American Pathologists. Laboratory general checklist. 2011. Retired version available free at http://www.cap.org/apps/docs/laboratory_accreditation/checklists/laboratory_general_sep07.pdf. Accessed 8 Sept 2011.

College of American Pathologists. Molecular pathology checklist. 2011. Retired version available free at http://www.cap.org/apps/docs/laboratory_accreditation/checklists/molecular_pathology_sep07.pdf. Accessed 8 Sept 2011.

den Dunnen JT, Antonaraki SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat. 2000;15:7–12.

Gulley ML, Braziel RM, Halling KC, et al. Clinical laboratory reports in molecular pathology. Arch Pathol Lab Med. 2007;131:852–63.

HUGO Gene Nomenclature Committee (HGNC), EMBL Outstation – Hinxton, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK. 2011. Available at www.genenames.org. Accessed 8 Sept 2011.

Human Genome Variation Society. Recommendations for the description of DNA sequence variants- v2.0. 2011. Available at www.hgvs.org/mutnomen/recs-DNA.html. Accessed 8 Sept 2011.

Lubin IM, Caggana M, Constantin C, et al. Ordering molecular genetic tests and reporting results. J Mol Diagn. 2008;10:459–68.

Lubin IM, McGovern MM, Gibson Z, et al. Clinician perspectives about molecular genetic testing for heritable conditions and development of a clinician-friendly laboratory report. J Mol Diagn. 2009;11:162–71.

Ogino S, Gulley ML, den Dunnen JT, et al. Standard mutation nomenclature in molecular diagnostics. J Mol Diagn. 2007;9:1–6.

Organization for Economic Cooperation and Development (OCED). OECD Guidelines for quality assurance in molecular genetic testing. Paris: OECD; 2007, pp 17–18, 33–35.

Seal RL, Gordon SM, Lush MJ, et al. Genenames.org: the HGNC resources in 2011. Nucleic Acids Res. 2011;39:D514–9.

Laboratory Management: Quality Assurance and Inspection

46

Kristin M. Post and Liang Cheng

Contents

46.1 Inspection 1118

46.2 Supervision and Personnel 1118

46.3 Test Validation 1119

46.4 Procedure Manual 1120

46.5 Proficiency Testing 1120

46.6 Quality Management Plan and Quality Control 1121

46.7 Preanalytic Phase of Testing 1121
46.7.1 Requisition 1121
46.7.2 Specimen Handling 1122

46.8 Analytic Phase of Testing 1122
46.8.1 Calibration 1122
46.8.2 Reagent Verification and Controls 1123
46.8.3 Extraction 1124
46.8.4 Restriction Endonuclease Digestion 1124
46.8.5 Sequencing, Electrophoresis Agarose, and Polyacrylamide, Capillary Electrophoresis 1124
46.8.6 Real-Time Polymerase Chain Reaction 1126
46.8.7 Arrays 1126
46.8.8 Fluorescence In Situ Hybridization and Brightfield In Situ Hybridization 1126

46.9 Postanalytic Phase of Testing 1126

46.10 Equipment 1127

46.11 Safety 1128

Further Reading 1128

K.M. Post, MT (✉) • L. Cheng, MD
Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

L. Cheng, D.Y. Zhang, J.N. Eble (eds.), *Molecular Genetic Pathology*,
DOI 10.1007/978-1-4614-4800-6_46,

46.1 Inspection

- The molecular pathology laboratory has two main inspecting bodies and a laboratory must always be in a state of inspection readiness. These inspecting bodies are the individual state boards of health and the federal Clinical Laboratory Improvement Amendments of 1988 (CLIA) program. High complexity testing certification for laboratories from CLIA is obtained from inspections conducted by the College of American Pathology (CAP) and the Joint Commission. The information discussed regarding requirements for a molecular pathology laboratory is taken from the CLIA regulations and CAP checklist requirements
- CAP has designated checklists based on specialty. The CAP defines molecular testing as a laboratory performing DNA or RNA probe hybridization or amplification. The molecular pathology checklist covers clinical molecular genetics testing including
  - Oncology
  - Hematology
  - Inherited disease
  - HLA typing
  - Forensics
  - Parentage
- Laboratories performing molecular testing for infectious disease for both FDA-cleared and non-FDA-cleared tests may use the microbiology checklist for inspection
- The cytogenetic checklist may be used to inspect fluorescence in situ hybridization (FISH) in a cytogenetics laboratory division
- The anatomic pathology checklist may be used to inspect FISH testing or in situ hybridization (ISH) testing in an anatomic pathology laboratory section
- In 2011, due to each department or laboratory-defined section unit during a CAP inspection being inspected from multiple checklists, the CAP has started to eliminate redundancies in the checklists including creating an "all common checklist" to cover
  - Proficiency testing
  - Procedure manuals
  - Results reporting
  - Method performance specifications
  - Reference intervals
- CAP requires a self-inspection be conducted during the year when CAP is not on site for inspection

46.2 Supervision and Personnel

- CLIA regulations states there is a single laboratory director per CLIA site number
- In larger laboratories, the medical director of a molecular pathology laboratory may be considered as a technical supervisor under CLIA
- The CAP director (technical supervisor) must have one of the following qualifications
  - Pathologist
  - Board-certified physician in specialty other than pathology
  - Doctoral scientist in a chemical, physical, or biologic science with training or experience in molecular pathology
- The individual involved in technical operations oversight must have one of the following qualifications
  - Qualifies as a director
  - MB(ASCP), BS, BA, or MLS (ASCP)/MT (ASCP) with at least 4 years of experience; at least 1 year in molecular pathology under a qualified director
- The individuals involved in molecular pathology testing must have one of the following qualifications
  - Experience in the field under direct supervision by a qualified director; and for laboratories under the United States regulation, qualified to perform high complexity testing
  - MT(ASCP) certified or equivalent
  - BA or BS degree in biologic sciences and experience with molecular methods
- The supervisors of the laboratory must ensure all personnel working in the laboratory have a diploma of highest level of education on file
- CLIA requires a laboratory to have an up-to-date CMS-209 form on file at all times signed by the laboratory director and having all personnel job titles listed

- A new technologist in the molecular laboratory must have adequate training along with a training checklist for each method trained to perform and documentation must be on file
- Technologist must have access to continuing education programs
- Technologist must be assessed on yearly competency for methods they are currently performing and documentation must be on file
- Six elements must be assessed for yearly competency of nonwaived test
  - Direct observation of routine patient testing
  - Monitoring the reporting of test results
  - Review of intermediate test results or worksheets, quality control records, proficiency testing results, and preventative maintenance records
  - Direct observation of performance of instrument maintenance and function checks
  - Assessment of test performance through testing of previously analyzed specimens or external proficiency testing samples
  - Evaluation of problem-solving skills

46.3 Test Validation

- Validations must be performed prior to assays being used for diagnostic testing and documentation of the validation must be available
- Manufacturers such as Life Technologies have developed software like the EZValidation™ to help laboratories design validations and conduct validations
  - Laboratories must validate the software before implementing the software in a clinical laboratory
- For FDA-cleared/FDA-approved assays, a laboratory must verify the manufacturers claims on
  - Analytic accuracy
  - Precision
  - Reportable range
  - Limit of detection
  - Linearity
  - Qualitative test, a comparison with another comparable method
- A laboratory may validate specimen types and collection devices not listed in a package insert. The validation study must document that the specimen type or collection device does not affect the performance of the assay
- During a validation of a laboratory-developed or modified FDA-cleared/FDA-approved assay, a laboratory must evaluate the following as applicable
  - Accuracy
  - Precision
  - Linearity
  - Analytical sensitivity = true positive/(true positive + false negative)
  - Analytical specificity = true negative/(false positive + true negative)
  - Interferences
  - Reference range
  - Specimen stability
  - Carryover
  - Correlation with a reference or other defined method
  - Clinical/diagnostic sensitivity and specificity should be determined either by the laboratory to the best of their ability and cite literature that addresses these
- If the laboratory makes clinical claims about a laboratory-developed test, all the claims must be validated by the laboratory, such as claims about diagnostic sensitivity, specificity, or clinical usefulness
- The validation study must include a reasonable number of samples determined by the medical director (technical supervisor)
- A summary of the validation must be filed in the laboratory before patient testing is performed. The summary must include a statement that "the validation study has been reviewed and the performance of the assay is considered acceptable for patient testing" then signed by the laboratory director
- Validation documentation must be retained by the laboratory while testing of the method is being performed and for at least 2 years after testing is discontinued
- CAP requires a document listing a test not FDA approved/FDA cleared being performed in the last 2 years for review of validation data

46.4 Procedure Manual

- A procedure manual must be available to be used at the workbench
- Package inserts from manufacturers are not acceptable for a procedure manual but is acceptable as a component of the overall procedure
- Card files or work cards that contain key reference information from the procedure may be used at the workbench if a complete manual is available for reference and the director has signed off on these materials for use
- An electronic version of the procedure is acceptable and must be readily available to the technologist when needed
- Each procedure must include
  - Principle and clinical significance
  - Patient preparation
  - Specimen collection
  - Labeling
  - Storage
  - Preservation
  - Transportation
  - Processing
  - Referral
  - Criteria for acceptability and rejection
  - Step by step performance of procedure
  - Analytical measurement range, if applicable
  - Control procedures
  - Calibration and calibration verification procedures
  - Corrective action when calibration or controls fail
  - Limitations
  - Reference intervals
  - Critical results, if applicable
  - Pertinent literature references
  - Laboratory's system for entering results
  - Description of action to take if test is inoperable
- Policies and procedures should be reviewed and approved by the CLIA-defined laboratory director before testing is performed on patient specimens
- Policies and procedures must be reviewed annually by the technical supervisor
- A laboratory must develop a system to show personnel are knowledgeable of the contents of a procedure such as signing off on the procedure before performing patient testing
- If there is a change in directors, the new director must review all procedures in a timely manner
- Policies and procedures must be kept on-site for at least 2 years after being retired

46.5 Proficiency Testing

- Proficiency testing is a peer group comparison of the technical performance of an assay to help monitor accuracy of an assay
- Proficiency testing must be performed at least twice per year for each analyte tested by the laboratory for clinical testing
- A laboratory must have a written procedure in place for handling and testing of proficiency specimens along with a protocol for investigating and correcting problems if proficiency testing results are unacceptable
- The CAP activity menu for a laboratory must be updated as testing changes to reflect current testing being performed
- The proficiency testing attestation statement must be signed by the director or designee along with the technologist performing the testing
- If CAP-approved proficiency testing is not available, an alternative performance testing must be performed at least twice per year. Examples of alternative proficiency testing
  - Split sample testing with other laboratories
  - Split sample testing with another in-house method
  - Assayed material
  - Regional pools
  - Clinical validation by chart review
- Proficiency testing samples must be incorporated in routine testing and handled in the same manner as patient samples
- CLIA regulations prohibit interlaboratory communications about proficiency testing and a policy must be in place

- CLIA regulations prohibit proficiency testing referrals to another laboratory, and a policy must be in place

46.6 Quality Management Plan and Quality Control

- A written quality management plan must be in place to evaluate and monitor the quality of the laboratory systems
- The quality management plan incorporates the quality assurance (QA) plan and the quality assurance assessment
- The quality assurance plan must include
  - Preanalytic systems
  - Analytic systems
  - Postanalytic systems
- The quality assurance assessment includes quality indicators that the laboratory determines a process or outcome measure to be used to determine the quality achieved by the laboratory
- Quality indicators determined by the laboratory may include monitoring
  - Patient/specimen identification errors
  - Accuracy of correctly ordering test into the computer system
  - Test turnaround time
  - Documentation of notification of critical values
  - Customer satisfaction
  - Percentage of specimens acceptable for testing
  - Corrected reports
- Quality indicators must be monitored monthly and reviewed by the medical director/technical supervisor
- Laboratory must have in place a documentation system to detect or correct clerical and analytical errors in a timely manner
- Correct storage conditions and specimen handling is essential for some molecular testing, and the laboratory must have a procedure in place to describe specimen preservation and storage before testing
- Laboratory must have a written policy for monitoring turnaround times and must define appropriate turnaround times for each test
- The laboratory must have a procedure in place for calculating statistics including thresholds for some molecular test to monitor and take corrective action if needed
  - Percentages of normal and abnormal findings
  - Allele frequencies
  - Percent positivity rates for some infectious disease test such as *Chlamydia trachomatis* and *Neisseria gonorrhoeae*
- A written procedure is in place to prevent carryover in the molecular laboratory. This includes adequate separation of areas, unidirectional workflow, dedicated materials for each area, and decontaminating work areas
  - Ideally, a molecular laboratory should have three areas which include a clean area to prepare master mix, processing area, and an amplification/detection area
  - Amplicon contamination should be monitored by laboratories by incorporating swipe test into a QA plan
- Procedures should contain information about any probe or primer used in the assay unless information is considered proprietary
- A specimen must be properly identified through all phases of testing such as from specimen receipt to extraction to amplification to detection to reporting and storage by either
  - Text
  - Numeric
  - Bar coded

46.7 Preanalytic Phase of Testing

46.7.1 Requisition

- The laboratory must have a written or electronic request for testing by an authorized individual
  - Genetic testing may require informed consent before molecular testing can be performed. There are no federal regulations on this matter; however, the state regulations should be reviewed before testing if offered by a laboratory

  - A pedigree or racial/ethnicity may be required for interpretation of some molecular testing; if required, it should be included as a part of the requisition form
- The requisition must include, as applicable
  - Patient identification information such as name
  - Patient sex
  - Patient date of birth or age
  - Name and address of physician or person ordering test if different than laboratory
  - Test requested
  - For gynecologic specimens, last menstrual period
  - For some testing, time and date of specimen collection is necessary
  - For some testing, source of specimen is necessary
  - Some testing requires clinical information to be included

46.7.2 Specimen Handling

- For proper identification of the primary collection container, the specimen must contain at least two identifiers such as
  - Patient name
  - Date of birth
  - Hospital number
  - Social security number
  - Requisition number
  - Accession number
  - Unique random number
- A written procedure must be in place describing specimen acceptability and patient preparation
- Laboratories should follow manufacturer guidelines on proper transportation conditions of specimens
- Written criteria should be established for rejecting specimens if unacceptable such as
  - Improperly labeled
  - Improper collection container
  - Inadequate volume
  - Improper specimen type
  - Possibly commingled specimen, such as a specimen received after the container had been entered by a sampling device
- Documentation must be kept of notification of rejection of specimen
- Some testing requires personnel to aliquot a sample before testing and a written procedure should be in place to prevent cross-contamination
  - An aliquot should never be returned to the original container
- Written procedure is in place to ensure specimens are processed and stored promptly to avoid nucleic acid degradation
- Proper storage conditions must be defined by specimen type and testing performed
- The laboratory should develop a system to easily retrieve stored specimens if further testing is requested
- When RNA or RNA-probes are used, a ribonuclease-free condition is maintained to prevent degradation
- Optimal laboratory design to prevent contamination (Fig. 46.1)

46.8 Analytic Phase of Testing

46.8.1 Calibration

- For quantitative assays, calibrators must have a matrix similar to what the laboratory is testing
  - Calibrators are used during the validation of the assay to verify the accuracy, linearity, limit of detection, and limit of quantification
  - Two calibrator points must be run at least every six months and the laboratory must establish acceptable criteria for accepting the accuracy
  - Calibrations must be performed whenever major system components change or lot numbers of reagents change, unless a laboratory can prove patient results are not affected by these changes
  - If the calibration fails, the laboratory must recalibrate
- Calibrator material may include
  - Calibrators used to calibrate the system
  - Materials provided by the vendor of the system

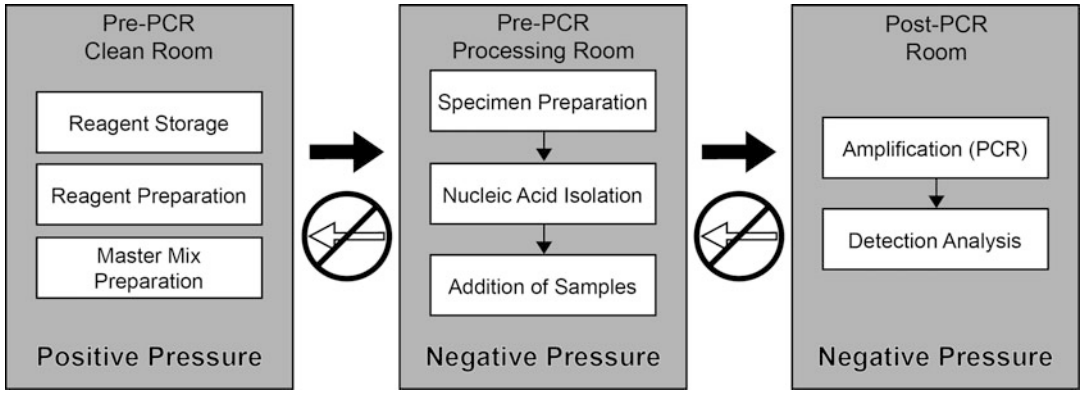


Fig. 46.1 Ideal molecular laboratory design to allow for unidirectional workflow and prevention of contamination with separate rooms

  - Previously tested patient specimens
  - Primary or secondary standards with appropriate matrix characteristics
  - Third party reference material if matrix is documented to be similar to patient samples
  - Proficiency testing material
  - Quality control (QC) material if appropriate matrix and has a mean established by peer group
- Calibrators should be separate from controls; however, if separate materials are not available, the calibrators and controls should be different lot numbers
- The analytical measurement range (AMR) must be checked at least every 6 months for quantitative assays
  - If the calibration performed includes points in the low, mid, and high range and passed established criteria by laboratory, the AMR is checked
  - If the range is not covered by the calibration or the laboratory extends the range from manufacturer recommendations, the AMR must be checked
  - AMR validation material used must be similar matrix to patient samples

46.8.2 Reagent Verification and Controls

- Reagents must be checked before placing in service for a new lot number or new shipment
- For qualitative assays, a positive and negative sample should be used to check reagent quality
- For qualitative assays, the reagent verification should include a negative and two levels of positives
- When possible multiplex assays should include each analyte detected to verify new lots and shipments
- To verify acceptability of new lot number, patient samples run previously with the old lot number should be tested
- QC material may be used to verify new shipments of the same lot number currently in service for patient testing. QC material may be used for verifying new lot number acceptability if a mean is established by a peer group; however, a laboratory should be aware of matrix interference
- Laboratories should have acceptable criteria to determine if reagent QC is acceptable and results should be documented
- Reagents should be properly labeled with
  - Content and quantity
  - Storage requirements
  - Date prepared or reconstituted
  - Expiration date
- Laboratories are not required to label reagents with a date received or open date
- Laboratories should not use reagents past the expiration date, and laboratories should assign an expiration date to reagents that the manufacturer has not assigned a date

- Qualitative assay should include a negative and positive control for each analyte in the run
  - For multiplex assays containing a large panel test, rotating the positive controls is acceptable
  - When a cutoff value is used to interpret a positive from a negative, the value must be verified at least every 6 months or change in lot number
- Quantitative assays should include a negative and two positives at different levels
- Control acceptability must be verified before patient results are reported and corrective action must be documented when the controls do not pass established criteria
- Controls must be run in the same manner as patient specimens
- Daily controls may be limited to electronic/procedural/built-in controls for tests meeting the following criteria established by CAP
  - Quantitative test - the system includes two levels of electronic/procedural/built-in internal controls that are run daily
  - Qualitative test - the system should include an electronic/procedural/built-in internal control that is run daily
  - The system should be FDA approved/FDA cleared and not modified by the laboratory
  - Laboratory must validate the accuracy of limiting daily QC to the electronic/procedural/built-in controls by a daily comparison of external controls to built-in controls for at least 20 days when patient samples are tested. Laboratory director must determine if acceptable before limiting daily QC to built-in controls
  - External surrogate sample controls must be run for new lot number, new shipment, major system maintenance, and software upgrades
  - External surrogate samples controls should be run at the frequency recommended by the manufacturer or at least every 30 days
- Controls for quantitative assays are monitored monthly for trending
- QC data must be reviewed by the director or designee monthly
- QC materials are available from several commercial companies (Table 46.1)

46.8.3 Extraction

- Commercially available kits or instruments for extraction of nucleic acids should be used or a laboratory must validate a method developed in-house
- A laboratory must measure nucleic acid concentration when the accuracy of an assay depends on the concentration
- RNA quality should be measured for human RNA targets due to the degradation of RNA
- Laboratories performing assays similar to KRAS and KIT using paraffin-embedded tumor should keep specimen documentation of the assessment of neoplastic cell content
- If the internal control does not go through extraction, then an extraction control must be included in each run
- If organic nucleic acid extractions are performed, a chemical safety cabinet should be available

46.8.4 Restriction Endonuclease Digestion

- Internal controls should be used in nucleic acid amplification assays to prevent release of false-negative results due to extraction issue or inhibition
- Laboratories using restriction endonuclease digestion should have a written procedure for application

46.8.5 Sequencing, Electrophoresis Agarose, and Polyacrylamide, Capillary Electrophoresis

- Laboratories performing sequencing should have
  - Literature documenting the wild type sequence and known mutations/polymorphisms
  - Sequence throughout the length of the target sequence should be readable

Table 46.1 Commercial sources for control materials

Company	Contact information	Products available	Analytes available
Life Technologies AcroMetrix	www.acrometrix.com	Infectious disease controls, validation kits, linearity panels	Adenovirus, Influenza H1N1, KRAS, Hepatitis B (HBV), Hepatitis C (HCV), HIV, BKV, Epstein-Barr virus (EBV), Cytomegalovirus (CMV), Herpes Simplex virus 1 & 2 (HSV), Varicella Zoster virus (VZV), Human papillomavirus (HPV), Enterovirus (EV), MRSA/MSSA, Group B *Streptococcus* (GBS)
Advanced Biotechnologies, Inc	www.abionline.com	Infectious disease native source DNA and RNA controls, purified virus/virus lysates, antigens, antibodies, native and recombinant viral proteins	Adenovirus, BK virus, CMV, *Chlamydia*, EBV, HCV, HIV, Human papillomavirus (HPV), *Helicobacter pylori*, Human herpes virus 6 (HHV6), HHV7, HHV8, HSV 1, HSV 2, HPV, Human T-lymphotropic virus, JC virus, *Mycobacterium tuberculosis*, *Mycoplasma pneumonia*, *Neisseria*, SV-40, VZV
American Type Culture Collection	www.atcc.org	Bacteria, bacteriophages, cell lines, hybridomas, filamentous fungi, yeast, tissue cultures, viruses, and so on	
Boston Biomedical, Inc	www.bbii.com	Genotype panels, linearity panels, qualification panels	*Chlamydia*, CMV, HBV, HCV, HIV, HPV, Parvovirus, WNV
Coriell Institute of Medical Research	www.coriell.org	Cell cultures and DNA derived from cell cultures for use as positive controls for many genetic disorders	Repositories included are National Institute of General Medical Sciences, National Institute on Aging, National Institute of Neurological Disorders and Stroke, American Diabetes Association, Autism Research Resource, U.S. Immunodeficiency Network, Center for Disease Control and Prevention, Leiomyosarcoma Cell and DNA Repository
Maine Molecular Quality Controls	www.mmqci.com	Synthetic DNA for inherited disease, infectious disease, pharmacogenetics	Cystic fibrosis (CF), Factor II G20210A, Factor V Leiden G1691A, MTHFR C677T, MTHFR A1298C, Warfarin, Hereditary hemochromatosis
SeraCare Life Sciences	www.seracare.com	Performance panels, verification panels, linearity panels	HBV, CMV, HCV, HIV, HPV, Cystic Fibrosis, *Chlamydia*, *Neisseria*, CYP2C9, VKORC1, EBV, HSV, Lyme Disease, MTHFR, *Toxoplasma*, West Nile virus
ZeptoMetrix Corporation	www.zeptometrix.com	Infectious disease controls, linearity panels	BKV, *Chlamydia*, *Clostridium difficile*, CMV, EBV, HSV 1 & 2, *Escherichia coli* O157:H7, HIV, HBV, HCV, HHV-6, 7 & 8, Norovirus, Rotovirus, Parvovirus, Influenza A H1, Influenza A H1N1 2009, Influenza B, *M. tuberculosis*, MRSA, MSSA, *Mycoplasma pneumonia*, *Neisseria*, Respiratory Viral Panel, VZV, West Nile Virus

  - Acceptance and interpretation of primary sequencing data should be established by the laboratory
  - Sequence of sense and antisense strands should be included in the procedure for heterozygous templates, rare alleles, or rare combinations of alleles
- Autographs and gel photographs should have quality resolution to accurately interpret
- Laboratories running agarose and polyacrylamide gel electrophoresis should
  - Load standard amounts of nucleic acid
  - Run molecular weight markers spanning the range of expected bands
  - Use visual/fluorescent bands to determine endpoint of electrophoresis
  - Establish objective criteria to interpret

46.8.6 Real-Time Polymerase Chain Reaction

- Real-time polymerase chain reaction (PCR) assays where results are interpreted based on a melting temperature, a temperature range must be defined and monitored
- Quantitative PCR calibrators should fall in the determined range for each run
- If multiple amplification runs do not include the extraction control, then an amplification control must be run

46.8.7 Arrays

- Array quality is verified with each change in lot

46.8.8 Fluorescence In Situ Hybridization and Brightfield In Situ Hybridization

- An anatomic or clinical pathologist must evaluate the corresponding hematoxylin and eosin slide for ISH testing to ensure invasive tumor cells are used in the study
- The laboratory must have a procedure in place for scoring of FISH results, and control loci must be included in each analysis
- *HER2* gene amplification by ISH must be validated by the laboratory on a minimum of 25 cases
- The laboratory procedure for *HER2* gene amplification by ISH includes the length of fixation and includes guidelines to report results using the ASCO/CAP scoring criteria
- Brightfield ISH for each sample should have a positive control probe against endogenous target to verify assay conditions and tissue pretreatment. For assays that detect RNA in target tissue or use an RNA probe, the laboratory must maintain a ribonuclease-free condition

46.9 Postanalytic Phase of Testing

- The laboratory should have a system in place to ensure test results are sent reliably from laboratory to final destination
- If a laboratory releases preliminary reports, they should be released in a reasonable amount of time, and discrepancies with the final report should be investigated
- Molecular reports should include the methodology, loci or mutations tested, and analytical interpretation
- Federal regulations and CAP require for laboratories testing using class I analyte-specific reagents (ASRs) that a disclaimer be attached to the report
  - "This test was developed and its performance characteristics determined by (laboratory name). It has not been cleared or approved by the US Food and Drug Administration. FDA does not require this test to go through premarket FDA review. This test is used for clinical purposes. It should not be regarded as investigational or for research. This laboratory is certified under the CLIA of 1988 as qualified to perform high complexity clinical laboratory testing."

- Patient confidentiality must be maintained during the release of molecular genetic test reports
- Reports must include standard nomenclature to designate genes and mutations; however, a laboratory may also choose to include the common name to help with clarification of testing
- Reports for ISH results must include interpretation of the ISH results
- Federal, state, and local regulations must be followed for retention of final report, result records, membranes, autoradiographs, gel photographs, and ISH slides
- Reports must be reviewed and approved every two years by the laboratory director
- Laboratory should maintain records of individual assays performed containing information of list of specimens, assay conditions, reagent lot numbers, serial number of instruments, and any other variable condition

46.10 Equipment

- When multiple instruments or methods are used to detect the same analyte, the laboratory must check the correlation of results at least twice a year when under a single CAP number
  - Manufacturer controls may be used to check correlation if the same control and reagent lot number are used on both instruments
- The laboratory should follow at least manufacturer guidelines for instrument maintenance and function checks
- Instrument maintenance and functions check should be well documented
- Temperatures should be recorded daily for refrigerators and freezers used to store reagents, controls, and patient specimens. The initials of the person checking the temperature must also be documented
  - A frost-free freezer may be used in a laboratory only if the maximum/minimum temperatures are recorded daily or it is equipped with a continuous monitoring system
- Temperatures should be checked and recorded daily if used for patient testing on those days for
  - Water baths
  - Heating blocks
  - Incubators and ovens where temperature control is necessary for a procedure
- Thermocycler wells should be checked for temperature accuracy before putting into use and at least yearly thereafter
  - Detected circuit breakers should be available to avoid electrical fluctuation
- Pipettes used for quantitative dispensing should be checked for accuracy and reproducibility before putting into use and at least annually thereafter
- Laboratory must check centrifuge speed at least annually
- Documentation of instrument maintenance/function checks and temperatures are reviewed monthly by the technical supervisor/director
- Procedures must be in place for personnel to accurately operate and set-up testing on instrumentation
- Procedures should describe minor troubleshooting and repairs of the instrument
- Laboratory should have maintenance, repair, and service records available to the personnel using the instrumentation
- A regular schedule should be established for checking critical operating functions of all instruments in use for patient testing
- When tolerance limits are established for instruments, these limits should be documented
- Laboratory should routinely service and repair film-processing equipment
- Proper protective equipment should be available to personnel when using ultraviolet light
- For scintillation counters, luminometers, and densitometers, the background level is compared each day of testing to established criteria
- Laboratory using instruments that measure multiple fluorochromes should have a protocol in place to identify and correct for bleed-through signal

- Spectrophotometer filters should be checked at least annually
- Spectrophotometer wavelength calibration is checked regularly following manufacturer criteria

46.11 Safety

- Laboratory must establish a policy for properly handling and processing samples
- A fume hood must be available for when using volatile chemicals
- A biological safety cabinet should be available and certified at least annually to ensure the filters function and airflow rates are within specifications
- Refrigerators should not contain improper materials such as food, externally contaminated specimens, or volatile materials
- For laboratories handling radionuclides
  - Radiation safety manual is up to date
  - Benches and sinks are decontaminated each day of use
  - Policy is in place for handling of radionuclides including authorization or restriction of personnel
  - Written procedure for notifying if damaged or leaking radionuclide shipment is received
  - Written procedure for proper storage of radionuclide
  - Documentation is kept for regular radiation surveys and wipe test
  - A sign is posted in areas where radioactive material is used or stored
  - Personnel have proper training documentation of decontamination, handling, and disposal
  - Documentation is maintained for proper disposal of radioactive waste, and waste is stored separate from normal trash

Further Reading

Association for Molecular Pathology Clinical Practice Committee. Molecular diagnostics assay validation. Bethesda: Association for Molecular Pathology Clinical Practice Committee; 2009.

Buckingham L. Quality assurance and quality control in the molecular laboratory. In: Buckingham L, Flaws ML, editors. Molecular diagnostics fundamentals, methods, and clinical applications. Philadelphia: F.A. Davis; 2007. p. 403–25.

Burd EM. Validation of laboratory-developed molecular assays for infectious diseases. Clin Microbiol Rev. 2010;23:550–76.

Centers for Medicare and Medicaid Services, Centers for Disease Control and Prevention. 42 C.F.R. Part 493. Medicare, Medicaid, and CLIA programs; laboratory requirements relating to quality systems and certain personnel qualifications. Final rule; 2003. p. 3640–714.

College of American Pathology. All common checklist. http://www.cap.org. Accessed 11 July 2011.

College of American Pathology. Laboratory general checklist. http://www.cap.org. Accessed 11 July 2011.

College of American Pathology. Microbiology checklist. http://www.cap.org. Accessed 11 July 2011.

College of American Pathology. Molecular pathology checklist. http://www.cap.org. Accessed 11 July 2011.

Ferreira-Gonzalez A, Wilkinson DS, Bluth MH. Establishing a molecular diagnostic laboratory. In: McPherson RA, Pincus MR, Henry JB, editors. Henry's clinical diagnosis and management by laboratory methods. Philadelphia: Elsevier Saunders; 2011. p. 1314–29.

Mifflin TE. Setting up a PCR laboratory. In: Dieffenbach CW, Dveksler GS, editors. PCR primer: a laboratory manual. New York: Cold Spring Harbor Laboratory Press; 2003. p. 5–14.

Morshed MG, Lee MK, Jorgensen D, Isaac-Renton JL. Molecular methods used in clinical laboratory: prospects and pitfalls. FEMS Immunol Med Microbiol. 2007;49:184–91.

National Committee for Clinical and Laboratory Standards. Assessment of laboratory tests when proficiency testing is not available; approved guidelines. NCCLS document GP29-A2. Wayne: NCCLS; 2008.

National Committee for Clinical and Laboratory Standards. Collection, transport, preparation, and storage of specimens for molecular diagnostic methods; approved guidelines. NCCLS document MM13-A. Wayne: NCCLS; 2005.

National Committee for Clinical and Laboratory Standards. Establishing molecular testing in clinical laboratory environments; approved guideline. NCCLS document MM19-A. Wayne: NCCLS; 2011.

National Committee for Clinical and Laboratory Standards. Quality management system: continual improvement; approved guideline. 3rd ed. NCCLS document GP22-A3. Wayne: NCCLS; 2011.

Seaton BL. Verification of molecular assays. In: Coleman WB, Tsongalis GJ, editors. Molecular diagnostics: for the clinical laboratorian. New York: Humana Press; 2006. p. 237–41.

Index

A
Abnormal maternal serum screening, 520
Accreditation, 1094–1096, 1103, 1107, 1108
Acute myeloid leukemia (AML), 866, 869, 870, 872–882, 884–885
Advanced maternal age, 520, 524
Alleged father, 1043–1045
Allele frequency, 1041, 1042, 1045
Alloimmunization, 1003–1005
Alpha-thalassemia, 977, 978, 982–983, 986, 989
Amino acid disorders, 576–581
AML. *See* Acute myeloid leukemia (AML)
Amniocentesis, 518
Amplification refractory mutation system (ARMS) PCR, 236–237
Analysis, 1114
Analytic phase of testing, 1122–1126
Aneuploidy, 47, 51–56, 73, 521–522, 524
Antibody
  panel, 304–305
  validation, 183–184
ARMS PCR. *See* Amplification refractory mutation system (ARMS) PCR
Array comparative genomic hybridization, 518, 525–526
Arrays, 1126
Automation, 415, 417, 418, 438, 441

B
Bacterial typing, 813–820
Bayesian analysis, 647, 649–651
B cell clonality, 825, 827–834, 839, 841, 851
bDNA assay. *See* Branch DNA (bDNA) assay
Beta-thalassemia, 977, 978, 983–986
Biobarcode, 385, 394, 395
Biochemical genetics, 561–590
Biomarker, 455–461
Biorepository, 483–496
Bone marrow engraftment chimerism, 1058–1063
Borrelia burgdorferi, 746–748
Branch DNA (bDNA) assay, 220–221
BRCA, 608, 610, 611, 613, 614, 617–620, 642
Breast cancer, 344–350, 353, 363
Brightfield *in situ* hybridization, 1126

C
Calibration, 1122–1123
Cancer, 1066, 1067, 1069–1078
  genetics, 608
  immunotherapy, 409–410
  stem cell, 79–84
  susceptibility, 608, 618, 633, 634
Capillary electrophoresis (CE), 225–227, 260–262, 442–443, 1038, 1040, 1041, 1046
Cardiac development, 592, 597–600
CDC. *See* Centers for Disease Control and Prevention (CDC)
CE. *See* Capillary electrophoresis (CE)
Cell
  culture, 467, 477, 479
  engineering, 408
  separation, 413–416
  sorting, 467, 477
Centers for Disease Control and Prevention (CDC), 1094
Centers for Medicare and Medicaid Services (CMS), 1095, 1105
Certification, 1094–1096
CGH. *See* Comparative genomic hybridization (CGH)
Chain of custody, 1034–1035, 1043
Challenges, 485–486, 495–496
CHD. *See* Congenital heart disease (CHD)
Chemotherapy response, 858, 862–863
Chimerism, 1001, 1005–1007
Chlamydia trachomatis, 734–739
Chorionic villus sampling (CVS), 519, 522–524
Chromosome, 4, 5, 7, 9, 11, 13, 14, 16, 18–21, 23–28, 36, 44–51, 53, 55–57, 59–67, 70–76
  instability, 88, 92–95, 107
  translocation, 859, 861
Chronic myelogenous leukemia (CML), 930–935
Circulating tumor cells (CTC), 1065–1078
CLIA ‘88. *See* Clinical Laboratory Improvement Amendment 1988 (CLIA ’88)
Clinical genomics, 335–366
Clinical Laboratory Improvement Amendment 1988 (CLIA ‘88), 1092, 1104
Clinical proteomics, 369–380
Clonality, 129–175
Clonality testing, 826, 831–855
Clopidogrel, 971–973
Clostridium difficile, 743–746

CML. *See* Chronic myelogenous leukemia (CML)
CMS. *See* Centers for Medicare and Medicaid Services (CMS)
CMV. *See* Cytomegalovirus (CMV)
Coamplification at lower denaturation temperature PCR (COLD-PCR), 237–239
CODIS. *See* Combined DNA index system (CODIS)
COLD-PCR. *See* Coamplification at lower denaturation temperature PCR (COLD-PCR)
Combined DNA index system (CODIS), 1034, 1037, 1040–1043
Communication, 1114
Comparative genomic hybridization (CGH), 595, 597, 598, 893–895, 902, 918, 920, 925, 927
Complex β-thalassemia, 978, 985
Complex sciences, 197
Congenital heart disease (CHD), 591–604
Contamination, 1034, 1035, 1037, 1040, 1045
Contiguous gene syndromes, 66–70
CTC. *See* Circulating tumor cells (CTC)
CVS. *See* Chorionic villus sampling (CVS)
Cytogenetic risk stratification, 870, 871, 873, 885
Cytogenetics, 43–76, 892, 901, 902, 904, 909, 916, 918, 921
Cytomegalovirus (CMV), 795–801, 806–808
Cytopathic effect, 699, 719

D
Data integration and mining, 197, 199, 205, 206
Deoxyribonucleic acid (DNA), 4–17, 23–30, 34–40, 468, 470, 473, 475–476, 478–480, 656–658, 664, 665, 667–670, 676, 677, 679–684, 686, 688–697, 699–703, 706–710, 717, 719, 720, 724, 726–730, 1109, 1111, 1113, 1114
  amplification, 222, 239, 242, 243
  binding dye, 235, 244–246
  copy number variations, 597
  detection, 455
  extraction, 213–214, 416, 1035
  fingerprinting, 1049, 1053, 1058, 1062
  isolation, 1034–1036
  methylation, 85–90, 95, 96
  polymorphisms, 1033
  profiling, 1033, 1037, 1041–1042
  quantification, 421, 422
  sequencer, 423, 426, 427
  testing, 538
  typing, 1033
Department of Health and Human Services (HHS), 1092–1095, 1104, 1105
DFA. *See* Direct fluorescent antibody (DFA)
Differential amplification, 1037
Direct fluorescent antibody (DFA), 701, 702, 712, 713, 715, 719, 720
Disseminated tumor cells (DTC), 1066
DMEs. *See* Drug metabolizing enzymes (DMEs)
DNA. *See* Deoxyribonucleic acid (DNA)
Drug metabolizing enzymes (DMEs), 1010, 1013, 1016–1021
DTC. *See* Disseminated tumor cells (DTC)

E
EBV. *See* Epstein-Barr virus (EBV)
Epigenetics, 80, 81, 84–88, 91, 95–97
Epstein-Barr virus (EBV), 795, 796, 798, 801–802, 946–950
Equipment, 1127–1128
Essential thrombocythemia (ET), 930, 932–933, 935–936
ET. *See* Essential thrombocythemia (ET)
Ethical issues, 1083–1089
Exclusion, 1042–1045
Expression profiling, 159, 171, 174, 175
Extraction, 1124

F
Factor V Leiden, 958–964
Familial cancer syndromes, 607–642
Fatty acid disorders, 564, 584–585
FDA. *See* Food and Drug Administration (FDA)
FISH. *See* Fluorescence *in situ* hybridization (FISH)
Flow cytometry, 301–333
Fluorescence *in situ* hybridization (FISH), 47–50, 523–525, 595, 596, 602, 892, 893, 897–899, 902–908, 910, 913–917, 924, 925, 1126
Fluorescence microscope, 450–452
Fluorescent probe, 235, 246
Food and Drug Administration (FDA), 1093, 1094, 1099, 1105–1106
Forensic, 1033–1046
Format, 118, 119, 205, 218, 221, 239, 242, 248, 271, 276, 283, 291, 357, 375, 430, 439–441, 505, 675, 676, 682, 775, 776, 867, 1057
Fungi, 787, 790–792

G
Gel
  electrophoresis, 1039, 1045
  imaging system, 449–450
Gene, 4, 6, 11, 13, 15–26, 28–30, 32, 33, 36–38, 40
  expression, 336, 337, 339–343, 345–353, 358, 359, 362, 900, 902, 912, 915, 918, 922, 926
  imprinting, 85, 95–98
  mapping, 118–123
  nomenclature, 1110, 1113
  rearrangement, 825–855
  therapy, 399–412
Genetic
  counseling, 645–651
  predisposition, 608, 609, 617, 621
  testing, 608–612, 616, 617, 619, 625, 627, 629–631, 633–636, 638, 640, 642, 1016–1017, 1085–1087, 1089
  variants, 1010, 1013, 1015–1017, 1020, 1022, 1023, 1028

Genotype, 1037–1043
Genotype frequency, 1041–1043
Gestational trophoblastic neoplasms, 1049, 1062
Graft *versus* host disease (GVHD), 999–1001
Growth factor receptor mutation, 866, 869, 870, 872–885
GVHD. *See* Graft *versus* host disease (GVHD)

H
Hardy-Weinberg equilibrium, 111–118
Hb. *See* Hemoglobin (Hb)
HbC. *See* Hemoglobin C (HbC)
HbCS. *See* Hemoglobin Constant Spring (HbCS)
HbD. *See* Hemoglobin D (HbD)
HbE. *See* Hemoglobin E (HbE)
HbS. *See* Hemoglobin S (HbS)
Hematological malignancies, 350
Hemoglobin (Hb), 976–977
Hemoglobin C (HbC), 976–978, 980–982, 986, 989
Hemoglobin Constant Spring (HbCS), 978, 982
Hemoglobin D (HbD), 976, 977, 981
Hemoglobin E (HbE), 976, 977, 981–982
Hemoglobinopathies, 975–991
Hemoglobin S (HbS), 978–980
Hemoglobin SC, 981
Hemolytic disease of the newborn, 1002–1004
Hemophilia, 967–969
HER2, 1067, 1070, 1071, 1075, 1077, 1078
Hereditary cancer, 610
Hereditary persistance of fetal hemoglobin (HPFH), 977, 985–986
HHS. *See* Department of Health and Human Services (HHS)
High resolution melting curve analysis, 245
High-throughput
  DNA sequencing, 442, 1010
  sequencing, 336, 337, 353–365
  technologies, 198, 201, 204
HLA. *See* Human leukocyte antigen (HLA)
HLA haplotype, 994–996, 999–1002, 1004
HNA. *See* Human neutrophil antigen (HNA)
HPA. *See* Human platelet antigen (HPA)
HPFH. *See* Hereditary persistance of fetal hemoglobin (HPFH)
Human leukocyte antigen (HLA), 993–1007
Human neutrophil antigen (HNA), 1005
Human platelet antigen (HPA), 1004–1005
Human tissues, 483–496
Hydatidiform mole, 167–169
Hypercoagulable state, 958, 959

I
IFA. *See* Immunofluorescent assay (IFA)
Immunodeficiency syndrome, 946, 949
Immunofluorescent assay (IFA), 699, 712, 719, 720
Immunoglobulin, 825–855
Immunohistochemistry, 181–193, 455, 461
Immunophenotyping, 302, 308, 311, 312, 316, 317, 320, 321, 323, 325–330
Inclusion, 1042–1045
Informatics, 499–513
Inspection, 1117–1128
Integrative pathology, 198, 205, 206

K
Karyotype, 518, 521, 524–526
Kidney, 800, 803, 807
Kinase inhibitors, 862

L
Laboratory information systems (LIS), 500, 502–505, 512–513
Lapatinib, 1077
Laser capture microdissection, 465–480
LCM. *See* Laser capture microdissection
Legal issues, 1083–1089
Legionella, 759–761
Linkage disequilibrium, 995
LIS. *See* Laboratory information systems (LIS)
Liver, 798, 801–803, 805–809
LOH. *See* Loss of heterozygosity (LOH)
Loss of heterozygosity (LOH), 144, 147–153, 157, 158, 161, 164
LS. *See* Lynch syndrome (LS)
LSD. *See* Lysosomal storage diseases (LSD)
Lymphocyte crossmatch, 997–999
Lymphoma, 946, 948–950
Lymphoproliferation, 945–950
Lynch syndrome (LS), 608, 611, 613, 617, 621–623
Lysosomal storage diseases (LSD), 575–576, 585–590

M
Major histocompatibility complex (MHC), 994–996, 1000
Malaria, 781–783
Malignant lymphoma, 914
Marker locus, 1037, 1044
Mass spectrometry (MS), 287, 292–299, 371, 377–380
Mathematical modeling, 197
MDS. *See* Myelodysplastic syndromes (MDS)
Medical interpretation, 1113–1114
Metabolic diseases, 562–565
Methicillin-resistant *Staphylococcus aureus*, 756–759
MHC. *See* Major histocompatibility complex (MHC)
MHC restriction, 996
Microarray analysis, 50–51, 69
Microarrays, 144, 159, 172–175, 336–343, 345–349, 351–353, 357, 359, 365, 423, 429, 435–443, 447, 1010, 1012, 1014, 1015, 1019, 1020, 1028–1029
Microfluidic device, 1072
MicroRNA (miRNA), 15, 17, 28, 30, 40, 85, 97–102, 159, 172–175
Microsatellite instability, 89–92
Minor histocompatibility antigens, 999, 1000

miRNA. *See* MicroRNA (miRNA)
Mitochondria DNA (mtDNA), 4, 6, 8, 15, 23, 28, 30, 35–40
MLPA. *See* Multiplex ligation-dependent probe amplification (MLPA)
MLST. *See* Multilocus sequence typing (MLST)
Molecular
  biology, 3–40
  diagnosis, 612, 617, 619, 622, 625–630, 633, 637, 638, 640, 641
  diagnostics, 40, 532–535, 547, 549, 558, 887–928
  diagnostic techniques, 783, 792
  genetics, 4, 7, 9, 11, 13, 20, 23, 28–30, 34–36, 39, 529–558, 930–934, 936–943
  information, 505–508
  pathology, 610, 619, 631, 632, 1109–1116
MPN. *See* Myeloproliferative neoplasm (MPN)
MS. *See* Mass spectrometry (MS)
mtDNA. *See* Mitochondria DNA (mtDNA)
Multilocus sequence typing (MLST), 813, 814, 820, 821
Multiplex ligation-dependent probe amplification (MLPA), 224–227
Mutation, 130, 132–134, 136, 141, 144, 147, 152–154, 171, 173, 1037, 1038, 1043, 1045
Mycobacterium tuberculosis, 761–769, 771–774
Myelodysplastic/myeloproliferative neoplasms, 929–944
Myelodysplastic syndromes (MDS), 869–873, 875–877, 879, 880, 884–885, 929–944
Myeloproliferative neoplasm (MPN), 869, 870, 876, 879–882, 884–885

N
NAIT. *See* Neonatal alloimmune thrombocytopenia (NAIT)
Nanofluidics, 384, 385, 396–398
Nanoparticle capture, 392
Nanoparticles, 384–395, 397, 398
Nanosensors, 385, 395
Nanotechnology, 383–398
NAT. *See* Nucleic acid amplification testing (NAT)
Neisseria gonorrhoeae, 736, 737, 739–743
Neonatal alloimmune thrombocytopenia (NAIT), 1004–1005
Newborn screening, 562–565, 579, 582, 583, 585
Next generation sequencing, 254, 264, 268, 276–279
Nucleic acid amplification testing (NAT), 1005

O
Operational issues, 484, 485, 488, 490, 491, 494
Organic acidurias, 583

P
Pancreas, 798, 800, 801
Parasites, 781, 783, 784
Parentage, 1033–1034, 1042–1046
Paroxysmal nocturnal hemoglobinuria (PNH), 331–332
Patch phenomenon, 135–136, 143
Paternity, 1034, 1042–1045
Paternity index (PI), 1043, 1045
Pathology
  image analysis, 512–513
  imaging, 506, 508–511
  informations, 499–513
Patient
  safety, 1050
  stratification, 1010–1011, 1022, 1029
Patterns of inheritance, 111–127
PCR. *See* Polymerase chain reaction (PCR)
Pharmacogenetics, 1010–1013, 1017–1019, 1021, 1023–1029
Pharmacogenomics, 863
PMF. *See* Primary myelofibrosis (PMF)
PNH. *See* Paroxysmal nocturnal hemoglobinuria (PNH)
Polycythemia vera (PV), 930, 932, 933, 935–936
Polymerase chain reaction (PCR), 212–215, 217, 219–221, 223–239, 244–251, 254, 256, 257, 262, 265, 267–271, 276–280, 416–419, 437–440, 443, 446, 447, 656, 657, 662, 664, 665, 668–671, 674–678, 681–683, 686–692, 695–697, 699–703, 707, 711–715, 717, 720, 722–730, 892, 893, 896, 897, 899, 902, 903, 905–909, 913, 917, 918, 924, 925, 997, 998, 1002, 1004–1006, 1034–1041, 1045
Postanalytic phase of testing, 1126–1127
Posttransplant lymphoproliferative disorder (PTLD), 946–948
Preanalytic phase of testing, 1121–1122
Precision medicine, 203
Predictive markers, 185, 192, 193
Primary and secondary immunodeficiency, 330–331
Primary myelofibrosis (PMF), 930, 932–933, 935–936, 943
Procedure manual, 1120
Proficiency testing, 1096–1097, 1120–1121
Profile frequency, 1042
Protein, 4, 8–12, 14–17, 21–24, 30–40
  pathway array, 291, 292
  targeting, 455, 456
Prothrombin G20210A, 959, 962–963
PTLD. *See* Posttransplant lymphoproliferative disorder (PTLD)
PV. *See* Polycythemia vera (PV)

Q
QDs. *See* Quantum dots (QDs)
Quality control, 1093, 1098–1102
Quality management plan and quality control, 1121
Quantum dots (QDs), 385–389, 391

R
Reagent verification and controls, 1123–1124
Real-time PCR, 229, 233–235, 244–246, 426–435, 1126
Regulatory issues, 493–494
Rejection, 994, 998–1000

Report, 1109–1116
Requisition, 1121–1122
Restriction endonuclease digestion, 1124
Restriction fragment length polymorphism (RFLP), 812–820
Reverse-phase protein microarrays, 373–376
RFLP. *See* Restriction fragment length polymorphism (RFLP)
Ribonucleic acid (RNA), 4, 6–9, 11–13, 15, 17, 22–25, 28–34, 36–38, 468, 470–473, 476, 478–480, 656, 658, 661, 665–679, 686, 688, 689, 695, 704, 706, 709–712, 714, 715, 717, 719, 720, 722–726, 1109, 1111, 1115
  detection, 455, 456
  extraction, 174, 215–219, 471
Risk estimation, 112, 115, 116, 118, 648–651
RNA. *See* Ribonucleic acid (RNA)
RNAi. *See* RNA interference (RNAi)
RNA interference (RNAi), 85, 95, 99, 102–108

S
Safety, 1128
Sample collection, 1034–1035
Sample preparation, 228, 229, 287–289, 295, 298, 418–421, 428, 433, 434, 438, 443–444
Sensitivity, 656–658, 661, 663, 665, 674, 675, 683, 685–689, 692, 697, 698, 700–702, 708, 709, 712, 713, 715, 720–729
Sequence-specific oligonucleotide probe (SSOP), 997
Sequence-specific primer (SSP), 998, 1004
Sequencing electrophoresis agarose and polyacrylamide, capillary electrophoresis, 1124, 1126
Serology, 656, 661–662, 664, 674, 680–681, 687, 694, 699, 702, 706, 712–713, 717, 719, 721–723, 725, 727, 729, 730
SERS. *See* Surface-enhanced Raman scattering (SERS)
Shell viral culture, 720
Short tandem repeats (STRs), 1005–1007
  locus, 1037–1039
  marker, 1034, 1037–1046
Single gene disorders, 530
Single nucleotide polymorphism (SNP), 1013
Single sequence polymorphism, 816, 819–821
SNP. *See* Single nucleotide polymorphism (SNP)
Specimen handling, 1122
Specimen identification, 1049–1063
Specimen mix-up, 1056, 1058, 1061
Spectrophotometer, 421–423
SSOP. *See* Sequence-specific oligonucleotide probe (SSOP)
SSP. *See* Sequence-specific primer (SSP)
Stem cell, 795–799, 801, 803, 807, 809
Streptococcus agalactiae, 749–752
Streptococcus pyogenes, 748–749
STRs. *See* Short tandem repeats (STRs)
Structural rearrangements, 56–66
Supervision and personnel, 1118–1119
Surface-enhanced Raman scattering (SERS), 385, 390–393
Systems pathology, 456

T
Targeted therapy, 181–193
Target sequence capture, 445–447
T cell clonality, 830–831, 833, 834, 838, 847–851
T cell receptor (TCR), 825–855
TCR. *See* T cell receptor (TCR)
Telomere, 93, 102–108
Test validation, 1119
Thermocycler, 423–429
Thrombophilia, 959, 961, 963
Tissue
  microarray, 455–461
  of origin, 152, 159, 172, 174, 175
  preservation, 380
TMA. *See* Transcription-mediated amplification (TMA)
Transcription-mediated amplification (TMA), 455–458, 460, 461
Translational research, 456
Transplant, 795–810
Trastuzumab, 1077
Tumor of unknown primary, 129–175

U
Ultrasound abnormalities, 525

V
Vancomycin-resistant *Enterococci* (VRE), 752–756
Variable number of tandem repeats (VNTR), 1005
Viral
  culture, 656, 661, 685–688, 698–702, 705, 713, 715, 717–721, 723, 725, 727, 729
  vectors, 399–412
Virus, 655–730
VNTR. *See* Variable number of tandem repeats (VNTR)
VRE. *See* Vancomycin-resistant *Enterococci* (VRE)

W
Warfarin, 962–964, 966, 969–971
Warfarin pharmacogenetics, 1019–1021
WHO classification, 312, 323, 328

X
X-chromosome inactivation, 75, 76, 135–146, 149