DEBRA G.B. LEONARD
EDITOR

Molecular Pathology in Clinical Practice

Springer

Molecular Pathology in Clinical Practice

Molecular Pathology in Clinical Practice

Editor

Debra G.B. Leonard, **MD, PhD**
Professor and Vice Chair for Laboratory Medicine, Department of Pathology and Laboratory Medicine, Weill Medical College of Cornell University and New York Presbyterian Hospital, New York, New York

Section Editors

Adam Bagg, **MD**
Director, Hematology, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

Angela M. Caliendo, **MD, PhD**
Vice Chair and Associate Professor, Department of Pathology and Laboratory Medicine, Emory University Hospital, Atlanta, Georgia

Karen L. Kaul, **MD, PhD**
Professor, Northwestern University Feinberg School of Medicine, Evanston, Illinois; Department of Pathology and Laboratory Medicine, Evanston Northwestern Healthcare, Evanston, Illinois

Vivianna M. Van Deerlin, **MD, PhD**
Director, Molecular Pathology Laboratory, Department of Pathology and Laboratory Medicine, University of Pennsylvania Health System, Philadelphia, Pennsylvania

Springer

Editor

Debra G.B. Leonard, MD, PhD
Professor and Vice Chair for Laboratory Medicine, Department of Pathology and Laboratory Medicine, Weill Medical College of Cornell University and New York Presbyterian Hospital, New York, NY 10021, USA

Senior Editors

Adam Bagg, MD
Director, Hematology, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Angela M. Caliendo, MD, PhD
Vice Chair and Associate Professor, Department of Pathology and Laboratory Medicine, Emory University Hospital, Atlanta, GA 30322, USA

Karen L. Kaul, MD, PhD
Professor, Northwestern University Feinberg School of Medicine, Evanston, IL 60201; Department of Pathology and Laboratory Medicine, Evanston Northwestern Healthcare, Evanston, IL 60201, USA

Vivianna M. Van Deerlin, MD, PhD
Director, Molecular Pathology Laboratory, Department of Pathology and Laboratory Medicine, University of Pennsylvania Health System, Philadelphia, PA 19104, USA

Library of Congress Control Number: 2006923893

ISBN-10: 0-387-33226-X e-ISBN-10: 0-387-33227-8
ISBN-13: 978-0-387-33226-0 e-ISBN-13: 978-0-387-33227-7

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

To Greg
With love and thanks

Preface

Molecular pathology now is recognized as an area of pathology specialty practice, with the first molecular genetic pathology board examination having been given by the American Board of Pathology and the American Board of Medical Genetics in November 2001. The anatomic and clinical pathology boards given by the American Board of Pathology also have an increasing emphasis on molecular pathology expertise in general pathology practice. The clinical practice of molecular pathology is becoming fundamental to almost every aspect of healthcare delivery, assisting with diagnosis of disease, therapeutic choice, therapeutic outcome monitoring, prognosis, prediction of disease risk, directing preventive strategies, beginning-of-life choices, patient and specimen identification, and clinical epidemiology. With this growing emphasis on molecular pathology, the publication of a textbook focused on the clinical practice of molecular pathology is very timely.

Molecular Pathology in Clinical Practice addresses all areas of molecular pathology clinical practice in a single textbook: molecular biology basics, genetics, infectious diseases, hematopathology, solid tumors, HLA typing, identity testing, and laboratory management. While other textbooks cover selected aspects of molecular pathology practice, or cover each of the areas of molecular pathology practice in a single chapter, each type of molecular pathology testing is not covered at a level of detail sufficient as a reference for molecular pathology practice. The purpose of this textbook is to provide a comprehensive reference for the practicing molecular pathologist, and to have some level of usefulness for practicing pathologists not specializing in molecular pathology, clinical colleagues, and trainees. This textbook is not meant to be a recipe book for molecular tests currently in clinical practice, simply because the specifics of testing change as new technologies emerge and are applied to clinical practice. Instead, the emphasis is on the molecular variations being detected for clinical use, how test results are used, and clinical and specific laboratory issues that require special attention.

Molecular Pathology in Clinical Practice aims to present the current state of our knowledge about this clinical practice. Because molecular pathology clearly is a rapidly changing field, with our knowledge of genomics increasing exponentially, the textbook likely will need regular updates. Some of the chapters present relatively nascent knowledge that is anticipated to be useful for clinical practice in the future, while other chapters present testing that is better established. My hope, as editor of this textbook, is that many editions will be needed in the future as we continue to apply the fruits of the Human Genome Project to the practice of medicine, resulting in improved outcomes for our patients.

Debra G.B. Leonard, MD, PhD

Contents

Contributors

Michael J. Ackerman, MD, PhD, FACC
Associate Professor, Department of Medicine, Pediatrics, and Molecular Pharmacology, Director, Long QT Syndrome Clinic and Sudden Death Genomics Laboratory, Consultant, Pediatric Cardiology, Mayo Clinic College of Medicine, Mayo Clinic, Rochester, MN 55905, USA

Jean A. Amos, PhD
Scientific Director, Department of Human Genetics, Focus Diagnostics, Cypress, CA 90630, USA

Daniel A. Arber, MD
Professor, Department of Pathology, Director of Clinical Hematology, Associate Director, Department of Pathology, Stanford University Medical Center, Stanford, CA 94305, USA

Mark Ayers
Bristol-Myers Squibb Company, Princeton, NJ 08543, USA

Adam Bagg, MD
Director, Hematology, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Daniel B. Bellissimo, PhD, FACMG
Director, Molecular Diagnostics Laboratory, Blood Center of Southeastern Wisconsin, Milwaukee, WI 53233, USA

Andrea Biondi, MD
Centro Ricerca M. Tettamanti, Università di Milano-Bicocca, H. San Gerardo Via Donizetti, 106, 20052 Monza, Italy

Rita M. Braziel, MD
Director, Hematopathology, Department of Pathology, Oregon Health and Sciences University, Portland, OR 97239, USA

Peter J. Bridge, PhD
Director, Molecular Genetics Laboratory, Alberta Children's Hospital, Calgary, Alberta T2T 5C7, Canada

Angela M. Caliendo, MD, PhD
Vice Chair and Associate Professor, Department of Pathology and Laboratory Medicine, Emory University Hospital, Atlanta, GA 30322, USA

Nancy J. Carpenter, PhD
Laboratory Director, Center for Genetic Testing, Saint Francis Health System, Tulsa, OK 74105, USA

Giovanni Cazzaniga, PhD
Centro Ricerca Tettamanti, Clinica Pediatrica Universitá di Milano Bicocca, Ospedale San Gerardo, Monza, Italy

Ethel Cesarman, MD, PhD
Associate Professor, Department of Pathology, Weill Medical College of Cornell University, New York, NY 10021, USA

Amy Chadburn, MD
Professor, Department of Pathology and Laboratory Medicine, Director, Immunopathology Laboratory, New York; Presbyterian Hospital/Weill Medical College of Cornell University, New York, NY 10021, USA

Edward Clark
Bristol-Myers Squibb Company, Princeton, NJ 08543, USA

Lisa Edelmann, PhD
Assistant Professor, Department of Human Genetics, Mount Sinai School of Medicine, New York, NY 10029, USA

Guang Fan, MD, PhD
Assistant Professor, Department of Pathology, Oregon Health and Sciences University, Portland, OR 97239, USA

Daniel H. Farkas PhD, HCLD
Director, Molecular Diagnostics, Department of Pathology, The Methodist Hospital, Houston, TX, 77030; Associate Professor, Department of Pathology and Laboratory Medicine, Weill Medical College of Cornell University, New York, NY 10021, USA

Jennifer M. Farmer, MS
Certified Genetic Counselor, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Andrea Ferreira-Gonzalez, PhD
Professor, Department of Pathology, Director, Department of Molecular Diagnostics Laboratory, Molecular Diagnostics, Virginia Commonwealth University, Richmond, VA 23298, USA

Sydney D. Finkelstein, MD
Cofounder and Scientific Director, Redpath Integrated Pathology, Pittsburgh, PA 15212, USA

Betty A. Forbes, PhD
Professor, Department of Pathology and Medicine, Virginia Common University Health Systems; Medical College of Virginia Campus, Richmond, VA 23298, USA

Julie D. Fox, BSc, PhD, MRCPath (UK)
Associate Professor, Microbiology and Infectious Diseases and Provincial Laboratory for Public Health (Microbiology), University of Calgary, Calgary, Alberta T2N 4W4, Canada

Christine C. Ginocchio, PhD, MT (ASCP)
Director, Microbiology/Virology and Molecular Diagnostics, North Shore Long Island Health System Laboratories, Lake Success, NY 11042, USA

David Grimwade, PhD, MRCP, MRCPath
Senior Research Fellow and Honorary Senior Lecturer in Haematology, Cancer Genetics Lab, Department of Genetics, Guy's, King's and St. Thomas' School of Medicine, Guy's Tower, Guy's Hospital, London, England

Kevin C. Halling, MD, PhD
Co-Director, Clinical Molecular Genetics and Molecular Cytology Laboratories, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA

Cyrus V. Hedvat, MD, PhD
Assistant Attending, Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10021, USA

John A. Heit, MD
Professor, Division of Hematology, Mayo Clinic, Rochester, MN 55905, USA

Gabriel N. Hortobagyi, MD
Professor, Department of Medicine, Chairman, Department of Breast Medical Oncology, University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA

Pei Hui, MD, PhD
Assistant Professor and Attending Pathologist, Medical Director, Molecular Diagnostics Laboratory, Department of Pathology, Yale University Schnool of Medicine, New Haven, CT 06520, USA

James C. Hyland, MD, PhD
Medical Director and CEO, Connective Tissue Gene Tests, Allentown, PA 18103; Hematopathologist, Health Network Laboratories and Lehigh Valley Hospital, Allentown, PA 18103, USA

Robert Jenkins, MD, PhD
Professor, Department of Pathology and Laboratory Medicine, Mayo Clinic, Rochester, MN 55905, USA

Jeanne A. Jordan, PhD
Associate Professor, Department of Pathology, University of Pittsburgh/Magee Women's Hospital, Pittsburgh, PA 15213, USA

Malek Kamoun, MD, PhD
Professor, Pathology and Laboratory Medicine, Philadelphia, PA 19104, USA

Karen L. Kaul, MD, PhD
Professor, Northwestern University Feinberg School of Medicine, Evanston, IL 60201; Department of Pathology and Laboratory Medicine, Evanston Northwestern Healthcare, Evanston, IL 60201, USA

Anant Khositseth, MD
Faculty of Medicine, Department of Pediatrics, Ramathibodi Hospital, Thung Phayathai, Bangkok, Thailand 10400

Ruth Kornreich, PhD
Assistant Research, Professor and Co-Director, Genetic Testing Laboratory, Department of Human Genetics, Mount Sinai School of Medicine, New York, NY 10029, USA

Debra G.B. Leonard, MD, PhD
Professor and Vice Chair for Laboratory Medicine, Department of Pathology and Laboratory Medicine, Weill Medical College of Cornell University and New York Presbyterian Hospital, New York, NY 10021, USA

Frank S. Lieberman, MD
Associate Professor, Department of Neurology, University of Pittsburgh, Pittsburgh, PA 15232, USA

Gerald P. Linette, MD, PhD
Assistant Professor, Department of Internal Medicine and Neurosurgery, Washington University School of Medicine, St Louis, MO 63110, USA

Ruth Ann Luna, MS
Assistant Director, Molecular Microbiology Laboratory, Department of Pathology, Baylor College of Medicine, Houston, TX 77030, USA

Amy L. Massey, PhD
Director of Curriculum and Student Affairs, Department of Biochemistry and Molecular Biology, University of Louisville School of Medicine, Louisville, KY 40292, USA

Kristin May, PhD
Laboratory Director, Genetic Diagnostic Laboratory, TC Thompson Children's Hospital, Chattanooga, TN 37403, USA

Holly L. Neibergs, PhD
Director, Hereditary Cancer Institute, Norton Healthcare, Louisville, KY 40232, USA

Frederick S. Nolte, PhD, D (ABMM), F (AAM)
Professor, Pathology and Laboratory Medicine, Emory University School of Medicine; Director, Clinical Microbiology and Molecular Diagnostic Laboratories, Emory Medical Laboratories, Atlanta, GA 30322, USA

Shuji Ogino
Assistant Professor, Department of Pathology, Brigham and Women's Hospital, Boston, MA; Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA; Harvard Medical School, Boston, MA 02115, USA

Deborah Ann Payne, PhD, ClSp(MB), CLDir(NCA)D
Associate Professor, Director of Molecular Diagnostics, Department of Pathology, University of Texas Southwestern Medical Center-Veripath, Dallas, TX 75390, USA

Herbert F. Polesky, MD
Former Director, Memorial Blood Centers of Minnesota; former Professor, Laboratory Medicine and Pathology, University of Minnesota School of Medicine, Healdsburg, CA 95448, USA

Nicholas T. Potter, PhD
Director, Molecular Diagnostics, Molecular Pathology Laboratory Network, Inc, Maryville, TN 38704, USA

Thomas W. Prior, PhD
Professor, Department of Pathology and Neurology, Director of Molecular Pathology, Columbus, OH 43210, USA

Rajiv K. Pruthi, MBBS
Department of Hematology and Internal Medicine, Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, MN 55905, USA

Joseph F. Pulliam, MD
Department of Pathology, Baylor College of Medicine, Methodist Hospital, Houston, TX 77030, USA

Lajos Pusztai, MD, DPhil
Associate Professor, Department of Medicine and Breast Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77230, USA

Hanna Rennert, PhD, FACMG
Director, Molecular Pathology Laboratory, Department of Pathology and Laboratory Medicine, New York Presbyterian Hospital, Weill Medical College of Cornell University, New York, NY 10021, USA

Andrew Ricci, Jr., MD
Department of Pathology, Hartford Hospital, Hartford, CT 06102, USA

Benjamin Roa, PhD
Assistant Professor, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA

Jeffrey S. Ross, MD
Cyrus Strong Merrill Professor and Chair, Department of Pathology and Laboratory Medicine, Albany Medical College, Albany, NY 12208, USA

Wendy S. Rubenstein, MD, PhD, FACMG
Medical Director, Evanston Northwestern Healthcare Center for Medical Genetics; Assistant Professor, Department of Medicine, Northwestern University Feinberg School of Medicine, Evanston, IL 60201, USA

Iris Schrijver, MD
Director, Molecular Pathology Laboratory, Stanford University Medical Center, Lucile Packard Children's Hospital, Medical Director, Stanford POCT, Director, Molecular Genetic Pathology Fellowship Program, Assistant Professor, Department of Pathology and (by courtesy) Pediatrics, Stanford University School of Medicine, Stanford, CA 94305, USA

Carlos Alberto Scrideli, MD, PhD
Department of Pediatrics, Ribeirão Preto Medical School, São Paulo University, Ribeirão Preto, São Paulo 14020-380, Brazil

Richard J.H. Smith, MD
Sterba Hearing Research Professor in Otolaryngology, Vice Chair, Department of Otolaryngology, Director, Molecular Otolaryngology Research Laboratories, Professor, Department of Internal Medicine, Division of Nephrology, University of Iowa, Iowa City, IA 52242, USA

Megan J. Smith-Zagone, MD
Pathologist, St Joseph Hospital, Eureka, CA 95501, USA

Karen Snow-Bailey, PhD, FACMG, FHGSA
Director, Department of Diagnostic Genetics, LabPlus, Auckland District Health Board, Auckland, New Zealand

James Stec
Abbott Labs, Chicago, IL 60064, USA

Karen Stephens, PhD
Research Professor, Department of Medicine and of Laboratory Medicine, Co-Director, Genetics Laboratory, University of Washington, Medical Genetics, Seattle, WA 98195, USA

Catherine Stolle, PhD
Adjunct Associate Professor, Department of Pathology and Laboratory Medicine, Director, Molecular Genetics Laboratory, Children's Hospital of Philadelphia, Abramson Research Center, Philadelphia, PA 19104, USA

Fraser Symmans
University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

Giovanni Tallini, MD
Professor, Department of Pathology and Attending Pathologist, Bologna University School of Medicine, Director, Diagnostic Molecular Pathology Laboratory, Ospedale Bellaria, Bologna, Italy

Jack Tarleton, PhD
Director, Fullerton Genetics Center Laboratory, Mission Hospitals, Asheville, NC 28803, USA

Stephen N. Thibodeau, PhD
Co-Director Molecular Genetics, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA

Erik C. Thorland, PhD
Co-Director, Clinical Cytogenetics Laboratory, Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA

Peter A. Tilley, MD, FRCPC
Medical Microbiologist, Department of Microbiology and Infectious Diseases and Provincial Laboratory for Public Health (Microbiology), University of Calgary, Calgary, Alberta T2N 4W4, Canada

Gregory J. Tsongalis, PhD
Director, Molecular Pathology Laboratory, Department of Pathology, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, USA

Guy Van Camp, PhD
Department of Medical Genetics, University of Antwerp, B2610 Antwerp, Belgium

Vivianna M. Van Deerlin, MD, PhD
Director, Molecular Pathology Laboratory, Department of Pathology and Laboratory Medicine, University of Pennsylvania Health System, Philadelphia, PA 19104, USA

Jo-Anne Vergilio, MD
Associate Pathologist, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

James Versalovic, MD, PhD
Assistant Professor, Department of Pathology and Molecular Virology and Microbiology, Baylor College of Medicine; Director, Microbiology Laboratories, Texas Children's Hospital, Houston, TX 77030, USA

Catherine Walsh Vockley, MS
Senior Genetic Counselor, Children's Hospital of Pittsburgh, Pittsburgh, PA 15213, USA

Victor W. Weedn, MD, JD
Visiting Professor, Bayer School of Natural and Environmental Sciences, School of Law, Duquesne University, Pittsburgh, PA 15282, USA

Erica Williams, MD
Assistant Chief Medical Examiner, Department of Legal Medicine and the Office of the Chief Medical Examiner, Virginia Commonwealth University School of Medicine, Richmond, VA 23219, USA

Thomas M. Williams, MD
Professor, Department of Pathology, University of New Mexico; Director of Genetics and Cytometry, TriCore Reference Laboratories, University of New Mexico, Albuquerque, NM 87131, USA

Robert B. Wilson, MD, PhD
Associate Professor, Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

Yaping Yang, PhD
Laboratory Director, Molecular Laboratory, Center for Medical Genetics, Houston, TX 77054, USA

Barbara Zehnbauer, PhD, FACMG
Director, Molecular Diagnostic Laboratory, Department of Pathology and Immunology, Washington University School of Medicine, St Louis, MO 63110, USA

Chapter 1

Basics of Molecular Biology

Deborah Ann Payne

Introduction

Molecular biology entails the analysis and study of the chemical organization of the cell. Molecules comprise the smallest chemical component capable of performing all the activities (structural or catalytic) of a substance. One or more atoms constitute each molecule. This chapter describes the physical organization of cells, cellular organelles, and molecules important in cell division, inheritance, and protein synthesis.

Organization of the Cell

The cell is a mass of protoplasm surrounded by a semipermeable membrane.[1] Cells constitute the smallest element of living matter capable of functioning independently; however, within complex organisms, cells may require interaction with other cells. To function independently, cells must produce nucleic acids, proteins, lipids, and energy. In complex organisms, these organic processes form and maintain tissues and the organism as a whole.

Genes consist of discrete regions of nucleic acids that encode proteins, and control the function of the cell. Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) comprise the two types of nucleic acids found in all cells. Chromosomes, made up of double-stranded DNA complexed with proteins, contain all the genes required for the cell to live and function.

Prokaryotic Cells

Prokaryotic cells are simple organisms lacking subcellular compartments, such as bacteria. The majority of prokaryotic nucleic acids form circular strands comprising approximately 1×10^6 base pairs (bp) (Table 1-1). Additional extrachromosomal genetic elements consist of circular plasmids also known as episomes and linear mobile genetic elements called transposons (30–40 bp). Plasmids range in size from 33 bp to 230 bp[2] and first gained notoriety in the 1950s by being associated with antibiotic resistance in bacteria.[3,4] Transposons also may confer antibiotic resistance on the host bacteria. All these genetic elements exist in direct contact with the bacteria's cytoplasm.

Eukaryotic Cells

Cytoplasm

In contrast to prokaryotic cells, eukaryotic cells are complex, highly compartmentalized structures. The cytoplasm contains multiple membrane-bound compartments known as organelles. The cellular membrane separates the cellular cytoplasm from the external environment. The membranes consist of hydrophobic lipid bilayers. The lipid bilayer contains proteins that serve as receptors and channels.

Nucleus and Nucleolus

The nucleus of the cell contains the cell's linear chromosomes and serves as the primary locus of inherited genetic material. Inner- and outer-pore-containing membranes define the nucleus and separate the chromosomes from the surrounding cytoplasm. Further partitioning occurs within the nucleus to generate the nucleolus, which functions as the ribosome-generating factory of the cell. Instead of additional membranes, fibrous protein complexes separate the nucleolus from the rest of the nucleus. In this structure, the nucleolus organizer (a specific part of a chromosome containing the genes that encode ribosomal RNA) interacts with other molecules to form immature large and small ribosomal subunits. Following processing, immature subunits exit the nucleolus and enter the nucleus. Eventually, mature ribosomal subunits and other molecules exit the nucleolus through the nuclear pores and enter the cytoplasm.

Table 1-1. Comparison of DNA Sizes of Various Genetic Elements

Genomic Element	Size in Base Pairs
Human chromosome	$1–3 \times 10^9$
Bacterial chromosome	$1–4 \times 10^6$
Mitochondrial chromosome	16,569
Bacteriophage	39,000
CAM plasmid	230
R388 plasmid	33
Transposons	30–40

Mitochondria

Mitochondria are membrane-bound organelles within the cytoplasm of cells that have several cellular functions. Inheritable genetic material, independent from the nuclear chromosomes, resides in mitochondria. These maternally derived organelles contain their own circular chromosomes (16,569 bp) and replicate independently from the cell and one another. As a result, not all mitochondria in a given cell have the same mitochondrial chromosomal sequence, resulting in genetic diversity of these organelles within and between different cells of the same organism, which is known as heteroplasmy. Mitochondrial genes encode mitochondria-specific transfer RNA molecules (tRNA). In addition, the mitochondrial chromosomes contain genes that encode proteins used in oxidative phosphorylation, including subunits of the cytochrome c oxidase, cytochrome b complex, some of the ATPase complex and various subunits of NAD dehydrogenase. Other components of the oxidative phosphorylation pathway are encoded by nuclear genes. For this reason, not all mitochondrial genetic diseases demonstrate maternal transmission. Analysis of mitochondrial DNA has applications for diagnosis of mitochondrial-inherited genetic diseases as well as for forensic purposes in the identification of severely decomposed bodies.

Other Cellular Organelles

Membranes not only segregate heritable genetic molecules into the nucleus and mitochondria, but also separate various cellular functions into distinct areas of the cell. The compartmentalization of cellular functions, such as molecular synthesis, modification, and catabolism, increases the local concentration of reactive molecules, thus improving the cell's biochemical efficiency. This partitioning also protects inappropriate molecules from becoming substrates for these processes. One example of this segregation is the endoplasmic reticulum (ER), which consists of a complex of membranous compartments where proteins are synthesized. Glycoproteins are synthesized by ribosome-ER complexes known as rough ER (RER), while lipids are produced in the smooth ER. The Golgi apparatus consists of numerous membrane-bound sacs where molecules generated in the ER become modified for transportation out of the cell. In addition, peroxisomes and lysosomes segregate digestive and reactive molecules from the remainder of the cellular contents to prevent damage to the cell's internal molecules and infrastructure.

Biological Molecules

Carbon can covalently bond to several biologically important atoms (i.e., oxygen, hydrogen, and nitrogen) and forms the scaffold for all biomolecules. Basic subunit biomolecules can combine to form more complex molecules such as carbohydrates, nucleic acids, and amino acids.

Carbohydrates

Carbohydrates serve as energy reservoirs and are a component of nucleic acids. In addition, carbohydrates also attach to lipids and proteins. The basic unit of a carbohydrate consists of the simple sugars or monosaccharides. These molecules have carbon, oxygen, and hydroxyl groups that most commonly form ringed structures. The oxygen can react with the hydroxyl group of another simple sugar to form a chain. As a result, the formula for a simple sugar is $(CH_2O)_n$, where n represents various numbers of these linked building block units.

Two pentose sugars, deoxyribose and ribose, comprise the sugar element of DNA and RNA molecules, respectively. As the name indicates, deoxyribose ("de-," a prefix meaning "off" and "oxy," meaning "oxygen") lacks one hydroxyl (OH) group compared to ribose.

Nucleic Acids

Nucleic acids are composed of chains of nucleotides. Each nucleotide is composed of a sugar (either ribose or deoxyribose), a phosphate ($-PO_4$) group, and a purine or pyrimidine base. The nucleotides are joined into a DNA or RNA strand by a sugar-phosphate-linked backbone with the bases attached to and extending from the first carbon of the sugar group. The purine and pyrimidine bases are weakly basic ring molecules, which form N-glycosidic bonds with ribose or deoxyribose sugar. Purines are comprised of two rings, a six-member ring and a five-member ring ($C_5H_4N_4$), while pyrimidines consist of a single six-member ring ($C_4H_2N_2$). Purines (guanine, G, and adenine, A) pair with pyrimidines (cytosine, C, and thymine, T) via hydrogen bonds between two DNA molecules (Figure 1-1). The additional hydrogen bond that forms between G and C base pairing (i.e., three hydrogen bonds) dramatically enhances the strength of this interaction compared to the two hydrogen bonds present between A and T nucleotides. This hydrogen-bonding capacity between G:C and A:T forms a pivotal molecular interaction for all nucleic acids and assures the passage of genetic information during

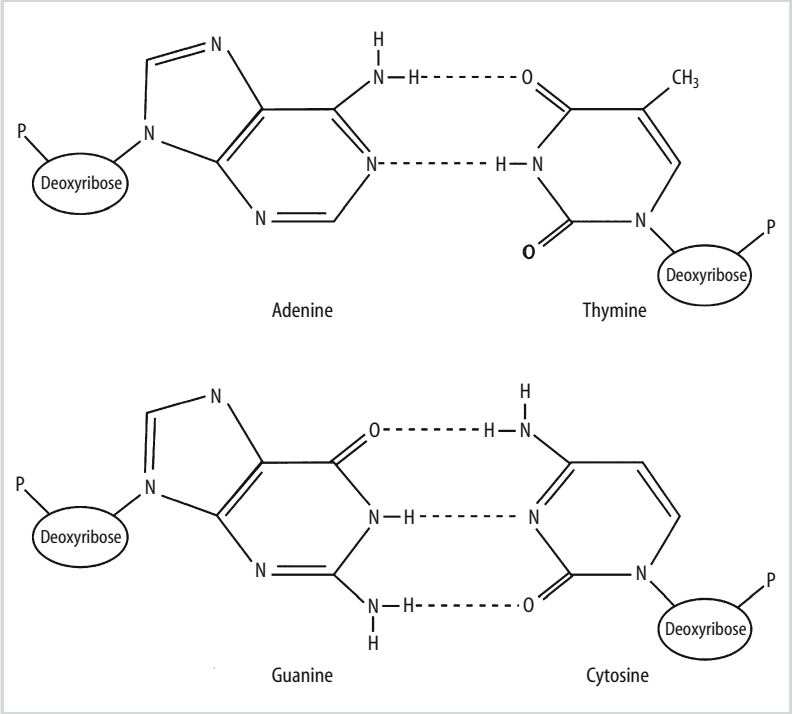


Figure 1-1. DNA base pairing. DNA nucleotides are composed of three moieties (e.g., sugar, base, and phosphate groups). The bases are either purine (adenine and guanine) or pyrimidine (thymine and cytosine). Note the difference in hydrogen bonds between adenine and thymine base pairs, with two hydrogen bonds, compared to cytosine and guanine base pairs, with three hydrogen bonds. (Reprinted from Leonard D. *Diagnostic Molecular Pathology*, copyright 2003, with permission from Elsevier.)

DNA replication, RNA synthesis from DNA (transcription), and the transfer of genetic information from nucleic acids to the amino acids of proteins.

Amino Acids

Amino acids are the building blocks of proteins. Amino acids linked together via peptide bonds form large, complex molecules. Amino acids consist of an amino group (NH_3), a carboxy group (COO–), an R group, and a central carbon atom. The R group can be a simple hydrogen, as found in glycine, or as complex as an imidazole ring, as found in histidine. Twenty different R groups exist; and determine whether an amino acid has a neutral, basic, or acidic charge (Table 1-2). The amino group of the polypeptides is considered the beginning of a protein (N-

Table 1-2. Amino Acids

Amino Acid	Amino Acid Symbols: Three Letter	Amino Acid Symbols: Single Letter	R Group
Alanine	ala	A	CH_3—$CH(NH_2)$—COOH
Arginine	arg	R	HN=$C(NH_2)$—NH—$(CH_2)_3$—$CH(NH_2)$—COOH
Asparagine	asn	N	H_2N—CO—CH_2—$CH(NH_2)$—COOH
Aspartic acid	asp	D	HOOC—CH_2—$CH(NH_2)$—COOH
Cysteine	cys	C	HS—CH_2—$CH(NH_2)$—COOH
Glutamine	glu	Q	H_2N—CO—$(CH_2)_2$—$CH(NH_2)$—COOH
Glutamic acid	gln	E	HOOC—$(CH_2)_2$—$CH(NH_2)$—COOH
Glycine	gly	G	NH_2—CH_2—COOH
Histidine	his	H	NH—CH=N—CH=C—CH_2—$CH(NH_2)$—COOH
Isoleucine	ile	I	CH_3—CH_2—$CH(CH_3)$—$CH(NH_2)$—COOH
Leucine	leu	L	$(CH_3)_2$—CH—CH_2—$CH(NH_2)$—COOH
Lysine	lys	K	H_2N—$(CH_2)_4$—$CH(NH_2)$—COOH
Methionine	met	M	CH_3—S—$(CH_2)_2$—$CH(NH_2)$—COOH
Phenylalanine	phe	F	Ph—CH_2—$CH(NH_2)$—COOH
Proline*	pro	P	NH—$(CH_2)_3$—CH—COOH
Serine	ser	S	HO—CH_2—$CH(NH_2)$—COOH
Threonine	thr	T	CH_3—CH(OH)—$CH(NH_2)$—COOH
Tryptophan	trp	W	Ph—NH—CH=C—CH_2—$CH(NH_2)$—COOH
Tyrosine	tyr	Y	HO—p—Ph—CH_2—$CH(NH_2)$—COOH
Valine	val	V	$(CH_3)_2$—CH—$CH(NH_2)$—COOH

*Proline has a ring shape arising from the covalent bond formed between the amino group and the central carbon.

terminus), while the carboxyl group is at the opposite end, providing directionality to the protein.

Genetic Molecules

Nucleic acids encode genetic information but also participate in additional physiological processes ranging from metabolism to energy transfer. Nucleotides constitute the monomeric units of nucleic acids (Figure 1-1). Nucleosides consist of two components (ribose or deoxyribose in RNA and DNA, respectively, and either a purine or pyrimidine base). A nucleotide is produced from a nucleoside by the addition of one to three phosphate groups through a covalent bond with the hydroxyl group of the 5′ carbon of the nucleoside's sugar ring.

Nucleic acids are formed by chains of nucleotides linked by phosphodiester bonds between the 3′ carbon of the first nucleotide's sugar ring and the 5′ carbon of the adjacent nucleotide's sugar ring. The phosphodiester linkages cause nucleic acids to have a 5′ to 3′ directionality. The alternating sugar-phosphate chain forms a continuous molecule with bases extending from the 1′ carbon of each sugar. For this reason, the sugar-phosphate chain is referred to as the backbone of nucleic acids (Figure 1-2). The phosphate groups give nucleic acids a negative charge that imparts important physiochemical properties to nucleic acids. The negative charge of DNA facilitates the binding of mammalian DNA to various proteins and allows separation of nucleic acid molecules by charge and size during gel or capillary electrophoresis.

Structure

In double-stranded DNA, the two DNA strands are held together by exact A : T and G : C hydrogen bonding between the bases of the two strands, in which case the two strands are said to be complementary. The two strands are oriented in opposite 5′ to 3′ directions, such that one strand is oriented 5′ to 3′ (↓) and the complementary strand is oriented 3′ to 5′ (↑) in an antiparallel fashion (see Figure 1-2). In this case, "anti-" refers to the head (or 5′ end) of one DNA strand being adjacent to the tail (or 3′ end) of the opposite strand.

The molecular curves of the two DNA strands form antiparallel helices known as the DNA double helix. This double helix form (the B form) has ten nucleotide pairs (base pairs) per turn, occupying 3.4 nm. Because the bonds between the sugar and the base are not perfectly symmetrical, the strands curve slightly. The slight curve of the offset glycosidic bonds results in major and minor grooves characteristic of the B form of the double helix.[5] Many molecular diagnostic assays target the minor groove of DNA with sequence-specific probes known as minor groove binding (MGB) probes. Two other forms of DNA exist as the Z and A forms. The Z form acquires a zigzag shape, while the A form has a very shallow and very deep groove.

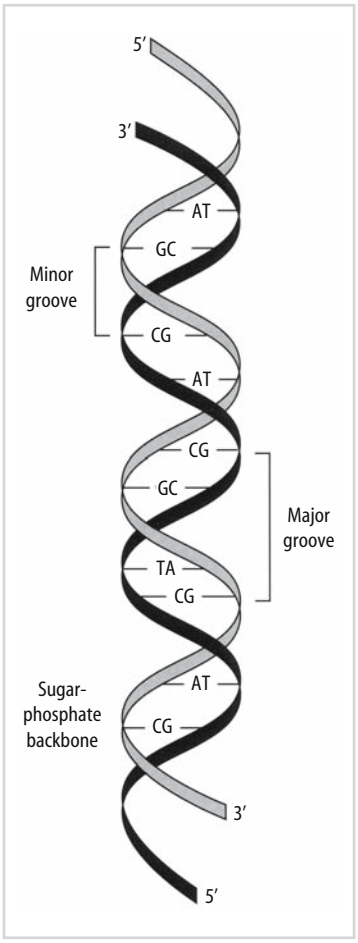


Figure 1-2. Double-stranded DNA. The two DNA strands are oriented in an antiparallel relationship, with asymmetric base pairing of two DNA strands that generates the minor and major grooves of the DNA double helix. (Reprinted from Leonard D. *Diagnostic Molecular Pathology*, copyright 2003, with permission from Elsevier.)

Thermodynamics of Nucleotide Base Pairing

Thermodynamics plays a major role in the structure and stability of nucleic acid molecules. The core mechanism of nucleic acid thermodynamics centers on the hydrogen-bonding capabilities of the nucleotides. The stability of these interactions not only influences the formation and stability of duplex nucleic acids but also impacts the structure and catalytic characteristics of single-stranded nucleic acids through intramolecular base pairing. In addition to these physiological functions, the phenomenon of complementary base pairing profoundly impacts clinical diagnostic assay development. Prior to the advent of clinical molecular diagnostic testing, many diagnostic tests required obtaining an antibody to identify or detect a target protein. The procedures for generating and validating diagnostic antibodies required extensive time and expense. The application of techniques utilizing the capability of two molecules to form a base pair as the basis for detection and characterization of target nucleic acids has greatly facilitated diagnostic test development. The formation of hydrogen bonding between two pieces of nucleic

acid is called hybridization, or annealing, and the disruption of the hydrogen bonds holding two nucleic acid molecules together is called denaturation, or melting. The fact that molecular diagnostic tests use hybridization techniques based on A : T and G : C base pairing underscores the necessity for understanding the thermodynamics of hydrogen base pairing of nucleic acids.

Short pieces of DNA or RNA called probes, or primers, that contain a specific sequence complementary to a disease-related region of DNA or RNA from a clinical specimen are frequently used in the molecular pathology laboratory. To achieve hybridization of a DNA or RNA probe to genomic DNA for a diagnostic test, the two genomic DNA strands must be separated, or denatured, prior to probe hybridization. Increasing the temperature of a DNA molecule is one mechanism for disrupting the hydrogen bonds between the DNA base pairs and denaturing double-stranded DNA into single-stranded form. The temperature at which 50% of the double-stranded DNA molecules separate into single-stranded form constitutes the melting temperature (T_m). The shorter the two complementary DNA molecules are, the easier it is to calculate the T_m. This primarily results from the decreased likelihood of nonspecific intramolecular annealing or base pairing compared to inter- and intramolecular base pairing. The simplest and least accurate formula for determining the T_m for short double-stranded DNA multiplies the sum of the G : C base pairs by 4 and multiplies the sum of the A : T base pairs by 2 and then adds these numbers together.

$$T_m = [4(G:C)] + [2(A:T)]$$

Although this is the least accurate method for calculation of the T_m of a double-stranded DNA molecule, it mathematically illustrates that G : C bonds are roughly twice the strength of A : T bonds. This formula works fairly well for short DNA molecules (i.e., ≤18 bp); however, as the length of the DNA molecule increases to 100 bp, the nearest neighbor T_m calculation for DNA and RNA is more accurate.[6,7]

$$T_m = \frac{\Delta H}{\Delta S + R \ln(Ct)} - 273.15$$

where

ΔH = enthalpy of the nucleic acid fragment
ΔS = entropy of the nucleic acid fragment
$R = 1.987\ \text{cal}\,K^{-1}\,\text{mol}^{-1}$
Ct = total strand concentration

For longer sequences (>100 bp), the most accurate formula for calculation of T_m is as follows:[8]

$$T_m = 81.5 + 16.6 \log [NA] + 0.41\ [\%G = \%C\} - 0.65 (\%\ \text{formamide}) - 675/\text{length} - \%\ \text{mismatch}$$

Table 1-3 demonstrates the effect of increasing the relative amounts of G : C base pairs on the T_m using these formulas.

Intramolecular base pairing also generates complex three-dimensional forms within single-stranded nucleic acid molecules. As a result, the single-stranded nature of eukaryotic RNA molecules affords great structural diversity via intramolecular base pairing. These conformations strain the linear RNA molecule and produce chemically reactive RNA forms. Catalytic RNA molecules play pivotal roles in cellular functions and in gene-targeting therapies.

Table 1-3. Melting-Temperature Calculations for Short Oligomers

Total Length	Number of G:C	Number of A:T	T_m*	%G:C†	A:T + G:C‡
30	30	0	106.2	100.0	100.0
30	25	5	101.2	93.2	100.0
30	20	10	89.5	79.5	90.0
30	10	20	83.4	72.7	80.0
30	0	30	71.6	59.0	60.0
20	20	0	90.4	88.8	80.0
20	10	10	72.7	65.1	60.0
20	0	20	55.9	47.8	40.0

*Nearest neighbor calculation of T_m.[6]
†T_m method for sequences over 100 bases.[8]
‡4(G + C) + 2(A + T) formula.

Intra- and intermolecular base pairing can negatively affect hybridizations. Dimers, bulge loops, and hairpin loops exemplify some of these interactions. Hairpins inhibit plasmid replication and attenuate bacterial gene expression.[2] These detrimental effects may also include initiation of spurious nonspecific polymerization, steric hindrance of hybridization of short stretches of nucleic acids (i.e., 10 to 30 base pieces of single-stranded nucleic acids, known as oligomers or primers), and depletion of probes or primers away from the specific target by either primer dimerization or other mechanisms. These interactions can result in poor sensitivity or poor specificity for diagnostic molecular tests.

Topology

The DNA and RNA molecules assume various geometric shapes or topologies that are independent of base pair interactions. Eukaryotic nucleic acids take on linear forms, in contrast to the circular forms of mitochondrial and bacterial chromosomal DNA. Viral genomes occur as different forms, ranging from segmented linear to circular. Although the conformation of RNA molecules can be complex via intramolecular base pairing, the topology of messenger RNA (mRNA) molecules is primarily linear. An organism's genomic topology influences the biochemical mechanisms used during replication and the number of replication cycles a given chromosome can undertake. In contrast to circular genomes, linear genomes limit the total number of possible replication cycles due to progressive shortening of the linear chromosome.

Mammalian Chromosomal Organization

The human genome contains approximately 10^9 base pairs of DNA. The total DNA is contained in 46 double-stranded

DNA pieces complexed with proteins to form chromosomes. The diploid human genome, therefore, is contained in 46 chromosomes: two of each of the 22 autosomal chromosomes, plus either two X chromosomes, in females, or one X and one Y chromosome, in males. Since the length of each helical turn of a double-stranded DNA molecule is 3.4 nm, consisting of ten bases, the length of the total genomic DNA in each cell measures approximately 1 m in length. For each cell to contain this long molecule, the double-stranded DNA must be compressed. A complex of eight basic histones (two copies of histone 2 [H2], H3, H4, and H5) package the DNA.[9] The histone complex contains positively charged amino acids that bind to 146 bases of negatively charged DNA. Histones fold the DNA either partially or tightly, resulting in compression of the DNA strand. Tight folding of the DNA condenses the DNA into heterochromatin. Following packaging and condensation, the nucleic acid strand widens from 2 nm to 1400 nm, with extensive overall shortening of the nucleic acid in the metaphase chromosome. Light microscopy easily permits the visualization of condensed metaphase chromosomes.

Less-condensed DNA binds histone 1 (H1) proteins or other sequence-specific DNA-binding molecules. Some of these DNA-binding molecules regulate gene expression (discussed later in this chapter). In contrast, tightly condensed chromosomes lack the "open spaces" for binding of regulatory proteins and prevent gene expression from tightly condensed DNA regions. These proteins may also prevent access to nucleic acid probes or primers for molecular diagnostic tests. As a result, many DNA extraction protocols include a protein-digestion step to liberate the DNA from these DNA-binding proteins. Removal of these proteins facilitates hybridization with short pieces of nucleic acid, such as primers or probes.

DNA Replication

Eukaryotic DNA Replication

The replication of DNA is a complex process requiring specific physiological temperatures and a host of proteins. As mentioned previously, molecular diagnostic techniques rely on the ability to denature or melt a double-stranded DNA template. Using chemical or physical conditions, separation of DNA strands can be accomplished with alkali or high temperatures (i.e., 95°C). Under physiological conditions, dissociation of DNA strands for replication is accomplished by numerous enzymes, such as helicases and topoisomerases. The region of transition from the double-stranded to separated single-stranded DNA is called the replication fork. The replication fork moves along the double-stranded DNA molecule as replication proceeds. At the replication fork, various primases, initiating proteins, and polymerases bind to the original or parental DNA strands and generate new daughter strands. Known collectively as a replisome, these enzymatic activities generate

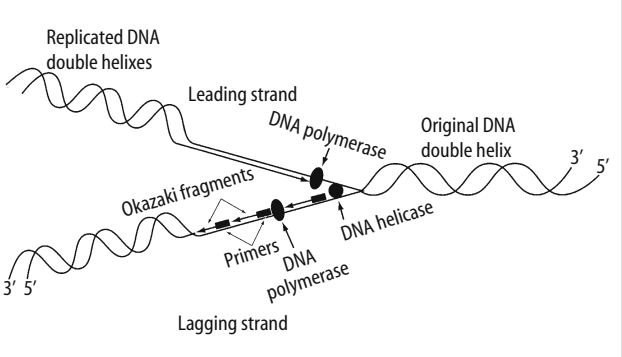


Figure 1-3. DNA replication. Replication fork depicting the leading and lagging strands and the numerous proteins and Okazaki fragments involved with replication. (Reprinted from Leonard D. *Diagnostic Molecular Pathology*, copyright 2003, with permission from Elsevier.)

two new nucleic acid strands that are complementary to and base paired with each of the original two template or parent DNA strands. This replication process is known as semiconservative because each resulting double-stranded DNA molecule consists of one new and one old DNA strand (Figure 1-3).

Polymerases function to synthesize new nucleic acid molecules from nucleotide building blocks. The sequence of the new strand is based on the sequence of an existing nucleic acid molecule, and the polymerase adds nucleotides according to the order of the bases of the parent strand, using G : C and A : T pairing. The new strand is antiparallel to the parent strand and is synthesized in a 5′ to 3′ direction. Of the two parent strands of genomic DNA, one strand (called the leading strand) can be read by the polymerase continuously in a 3′ to 5′ direction, with the new strand generated in a continuous 5′ to 3′ direction. In contrast, the opposite strand (known as the lagging strand) cannot be read continuously by the polymerase. The replication fork moves along the lagging strand in a 5′ to 3′ direction, and polymerases synthesize only by reading the parent strand in a 3′ to 5′ direction while synthesizing the new strand in a 5′ to 3′ direction. Therefore, synthesis cannot proceed continuously along the lagging strand, which must be copied in short stretches primed from RNA primers and forming short DNA fragments known as Okazaki fragments. The new strand complementary to the lagging strand is formed by removal of the RNA primer regions and ligation of the short DNA fragments into a continuous daughter strand complementary to the lagging strand. Discontinuous 3′ to 5′ replication results in the progressive loss of ends of the chromosomes known as telomeres in normal cells. Some malignant cells retain telomerase activity that permits the addition of these terminal telomeric sequences to the chromosomes.

While replication requires many proteins, the polymerase determines the speed and accuracy of new strand synthesis. The rate that the four nucleotides are polymerized into a nucleic acid chain defines the processivity of the enzyme. The processivity of most polymerases approximates 1000 bases per minute.

Table 1-4. Fidelity of Various Polymerases

Polymerase	Error Rate
pol β*	8×10^{-4}
pol α*	1×10^{-4}
pol θ*	$1.7–4 \times 10^{-5}$
Pfu†	1.3×10^{-6}
Deep vent†	2.7×10^{-6}
Vent†	2.8×10^{-6}
Taq†	8×10^{-6}
UITma†	55×10^{-6}
Klenow‡	$1–10 \times 10^{-7}$
HIV reverse transcriptase	$6–30 \times 10^{-4}$

*Reference 37.
†Reference 38.
‡Reference 39.

The fidelity of the polymerase refers to the accuracy of the enzyme to incorporate the correct complementary bases in the newly synthesized DNA. Incorporation of incorrect bases or other replication errors can result in cell death or oncogenesis. The error rate of polymerases varies widely from 1 in 1200 to 1 in 1,000,000 bases (Table 1-4). To correct the erroneous incorporation of bases or other replication errors, protein complexes proofread and correct synthesis errors. In normal cells, the cell cycle pauses to facilitate error repair in the G2 phase of the cell cycle (Figure 1-4). Malignant cells may not pause to allow for error correction, resulting in the accumulation of damaged or mutated DNA.

The complexity of the biochemical reactions necessary for replicating eukaryotic nuclear DNA demonstrates a high degree of regulation for generating two strands from one replication fork. In addition to these complexities, replication in eukaryotic cells occurs at multiple origins of replication (Ori). These multiple sites grow progressively until the newly generated strands join to form complete chromosomal-length DNA.

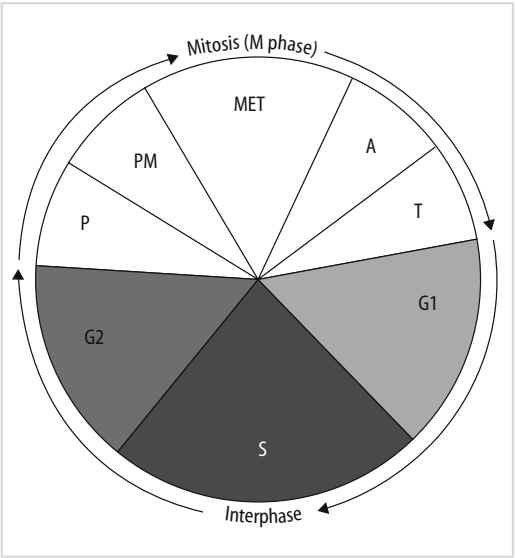


Figure 1-4. Cell cycle. The clear panels are the ordered phases of mitosis (M phase), while the gray and black panels are the ordered stages of interphase. P, prophase; PM, prometaphase; MET, metaphase; A, anaphase; T, telophase; G1, gap 1; S, DNA synthesis; G2, gap 2.

Bacterial and Mitochondrial Replication

The relatively small chromosomes of bacteria (~10^6 base pairs) utilize a simpler mechanism than eukaryotic replication. A single origin of replication initiates the duplication of the bacterial chromosome, and replication occurs simultaneously on both strands in opposite directions from the origin of replication. This efficient replication process depends on the circular topology of the bacterial genome.

Another unique feature of prokaryotic chromosomal replication is the mechanism bacteria have evolved to protect their chromosomes. The lack of a protective nuclear membrane in bacteria makes the chromosome susceptible to attack by viruses (bacteriophages). As a result, many bacteria produce restriction enzymes that degrade foreign nucleic acids. These restriction enzymes recognize specific short sequences and cleave the DNA at those sites. However, methylation of the recognition sequences in the bacterial chromosomal DNA prevents most restriction enzymes from digesting the chromosomal DNA of the bacteria. In this way, methylating enzymes add methyl groups to the replicated bacterial chromosome, preventing chromosomal degradation by its own restriction enzymes. This methylation and restriction process functions as a primitive immune system by destroying foreign bacteriophage DNA before it can usurp the bacteria's replication system. Bacterial restriction enzymes are used to specifically cleave DNA in molecular diagnostic tests and are useful for identifying genetic variations.

Additional types of replication occur in some viruses and bacteria. The rolling-circle mechanism of replication proceeds with an initial single-strand cut or nick in double-stranded circular genomes, followed by replication proceeding from the nick in a 5′ to 3′ direction. The new strand displaces the old strand. RNA viral genomes use the enzyme transcriptase for replication. In the case of retroviruses, a reverse transcriptase generates an intermediate DNA molecule, which integrates into the host chromosome and then is used for generation of progeny RNA molecules. The high error rate of human immunodeficiency virus (HIV) reverse transcriptase produces numerous mutations in the viral genome.[10,11] Some of these mutations confer resistance to antiretroviral therapies and can be identified by clinical molecular tests.

Cell Division and Cell Cycle

In eukaryotic cells, the cell cycle refers to the entire process of generating two daughter cells from one original cell, with chromosomal replication as one of the steps. The two parts of the cell cycle are called interphase and mitosis.

DNA synthesis occurs during interphase and consists of three stages: gap 1 (G1), synthesis (S), and gap 2 (G2) (Figure 1-4). Regulation of cell division depends on specific cell-cycle-dependent proteins known as cyclins and growth factors. Some of these factors cause the cycle to progress while others stop the cycle at certain stages. Checkpoints, or times when the cycle may be paused, exist at the G1/S and G2/mitosis interfaces and allow the cell time to repair any DNA damage that may be present in the cell before and after replication of the DNA, respectively.

Growth factors initiate the G1 phase via cell surface receptors. Several molecular events such as the dephosphorylation of the retinoblastoma protein and cyclin binding to cyclin-dependent kinases (Cdk) transition the cell toward the G1/S checkpoint. The amount of cellular P53 protein determines whether the cell progresses beyond this checkpoint, with higher levels preventing cell cycle progression. Because various DNA-damaging events, such as ultraviolet light, radiation, carcinogens, and double-stranded DNA breaks, induce production of P53 protein, this molecule serves as a sentinel for mutated DNA. The functional failure of P53 removes this sentinel from the cell cycle process and results in the accumulation of genetic errors. Therefore, inactivation of P53 facilitates oncogenesis.

Once DNA repairs have taken place during G1 prior to replication of the DNA, the cell proceeds to S phase. DNA synthesis occurs in the S phase, followed by the G2 phase. Replication errors occurring during the S phase are corrected in the G2 phase, the G2/M checkpoint. This final checkpoint marks the end of interphase.

Mitosis, the process of physical division of the parent cell into two daughter cells, occurs during the mitosis or M phase of the cell cycle. During mitosis, the duplicated chromosomes are physically separated so that each daughter cell receives the correct number of chromosomes. Mitosis consists of five phases: prophase, prometaphase, metaphase, anaphase, and telophase. The duplicated chromosomes condense during prophase. A structural element known as the mitotic spindle originates from two structures called centrioles, which move to opposite poles of the cell and the spindle forms between the centrioles. The nuclear membrance dissipates, proteins form kinetochores on the chromosmes, and microtubules attach to the kinetochores during prometaphase. The duplicated chromosome pairs attach at central points along the spindles. The arrangement of the highly condensed chromosome pairs along an equatorial cell plane denotes metaphase. As previously discussed, highly condensed chromosomes cannot bind proteins necessary for gene expression. As a result, the cell's internal machinery focuses solely on cell division during metaphase. The centriole-derived spindle guidelines pull the duplicate chromosomes apart and drag them toward each centriole during anaphase. With the separation of the daughter chromosomes (chromatids) into opposite poles of the cell and the reformation of nuclear membranes around the two daughter sets of chromosomes, telophase begins. Cytokinesis, or the division of the cytoplasm, is the last step in cell division. During cytokinesis, the mitochondria are randomly and potentially unevenly distributed in the daughter cells. The cell cycle can then be reinitiated by one or both of the daughter cells to generate additional cells. Alternatively, some cells become quiescent in a G0 phase (between telephase and G1) and either have a prolonged delay before initiating replication again or no longer divide.

Cell division to generate gametes is called meiosis and consists of two divisions, meiosis I and meiosis II. Like mitosis, this process begins with the duplication of chromosomes in prophase I. During metaphase I, the maternal and paternal homologous chromosomes pair (i.e., pairing occurs between each of the pairs of the 22 autosomal chromosomes, the two X chromosomes in females, and the X and Y chromosomes in males). Each pair attaches to the spindle apparatus along the equatorial plane of the cell spindle. DNA may be exchanged between the paired chromosomes by either crossing-over or recombination mechanisms during this pairing stage of meiosis I. During anaphase I, homologous chromosomes separate into daughter cells, resulting in 23 duplicated chromosomes in each daughter cell. A second cell-division cycle, meiosis II, separates the duplicated chromosomes, resulting in haploid cells, egg or sperm, containing only one copy of each of the 22 chromosomes plus an X (egg or sperm) or Y (only sperm) chromosome.

From Gene to Protein

The genomic DNA content is the same in all cells of the same person and encodes all the genetic information for cellular function. Encoded in the DNA are the blueprints for all the RNA and protein molecules present in any type of cell. Different parts of the genetic information are used by different types of cells to accomplish each cell's specific function. DNA is used to produce RNA and protein molecules by processes called transcription and translation, respectively. The regions of DNA that encode RNA and protein molecules are called genes.

Replication requires an increase in building materials for the duplicated daughter cells. Highly condensed metaphase chromatin cannot produce gene products because proteins that initiate gene expression cannot bind to the chromosomes at this phase of replication. In contrast, partially condensed or unfolded chromatin permits the binding of specific proteins (e.g., RNA polymerases) that synthesize mRNA and tRNA. Ultimately, these molecules facilitate the production of gene products, specifically proteins.

RNA molecules function as the mediators between DNA and protein. These molecules essentially speak the same language as DNA because, as nucleic acids, they can base pair with complementary DNA sequences. Like transferring spoken language to a written form, this process of copying information from DNA to RNA is referred to as transcription. The transcription complex of proteins must unwind the double-stranded DNA at the specific gene site to be copied, locate the polymerase binding site on one of

the DNA strands, and generate a primary (1°) transcript, which is one component of heterogeneous nuclear RNA (hnRNA) by reading the DNA strand in a 3′ to 5′ direction, with RNA synthesis proceeding in a 5′ to 3′ direction. The 1° RNA transcript is processed into mRNA, and finally the DNA in the region of the gene becomes double-stranded again. Numerous DNA sequences bind proteins that regulate and coordinate gene expression. These sequences can be used to identify the locations of genes within the entire human genome sequence. Since the generation of the first draft of the human genome, the interest in understanding gene structure has increased with the goal of identifying disease-associated genes.[12–14]

Gene Structure

Promoting Transcription

Sequences that bind RNA polymerases in combination with transcription factors drive and regulate the production of 1° RNA transcript (Table 1-5). Proteins and transcription factors bind to sequences located 5′, or upstream, of the gene to be expressed and are collectively called the promoter region of a gene. Negative numbering denotes the location of these sequences upstream of the first protein-encoding base. The promoter sequence initiates (or promotes) transcription of the downstream gene and harbors conserved sequences that are recognized by the transcription complex of enzymes.

The complexity and organization of the transcription regulatory sequences of genes differ between prokaryotic and eukaryotic cells. Prokaryotes contain a simple gene structure with sequences for polymerase binding occurring at –35 and –10 for each gene. The –10 sequence contains a consensus sequence of TATAAT, while the –35 region consists of TTGACA. Variations of these sequences as well as the sequences located adjacent to the gene determine the strength of the promoter's transcriptional activity. For example, small differences such as having a TATATA sequence rather than the consensus sequence at the –10 position will decrease the promoter's ability to bind to the RNA polymerase and result in decreased production of mRNA for that gene. In bacteria, operons regulate expression of multiple genes with related functions from the same promoter.

Table 1-5. Examples of Nucleic Acid Motifs

Name	Sequence
AP1 binding site	TGASTCAG
AP2 binding site	CCCCAGGC
AP3 binding site	GGGTGGGAAAG
AP4 binding site	YCAGCTGYGG
C/EBP	TGTGGAAAG
CCAAT box	CCAAT
CP1 binding site	YN(6)RRCCAATCA
CP2 binding site	YAGYN(3)RRCCAATC
CREB	TGACGTCA
CTF/NF1 binding sites	GCCAAT
GCN4 target site	ATGASTCAT
Glucocorticoid receptor	GGTACAN(3)TGTTCT
Homeobox protein-binding site	TCAATTAAAT
HSTF	CNNGAANNTTCNNG
INF-stimulated response	RGGAANNGAAACT
Lariat consensus sequence	YNYTRAY
MALT box	GGAKGGA
NF-1	TTGGMN(5)GCCAAT
Octamer sequence	ATTTGCAT
Poly A signal	AATAAA
Splice acceptor	Y(11)NYAGG
Splice donor	MAGGTRAGT
TATA box	TATA
Translational initiation sequence	RNNMTGG

R = A/G; Y = C/T; M = C/A; W = A/T; N = A/T/C/G.

In eukaryotic genes, various promoter sequences bind multiple proteins, which catalytically modify and activate other bound proteins. Enhancer sequences increase the production of mRNA but are far removed from the gene. One of the pivotal proximally located sequences comprises a TATA box (TATAAA) located at –25 (Figure 1-5). These bases initiate binding of a TATA-binding protein (TBP) within the transcription factor D complex. Following this binding, transcription factors B, H, and E bind to and open the DNA strands downstream from the promoter. Finally, transcription factor F and RNA polymerase II bind to the transcription complex. The close proximity of these proteins to RNA polymerase II permits phosphorylation of the polymerase and initiation of transcription. In eukaryotic cells, variations in the recognition sequences alter the efficiency of transcription. These variations may be base pair changes or base modifications. For example, promoter sequences that are highly methylated do not bind well to the transcription factors or polymerase. As a result, a gene

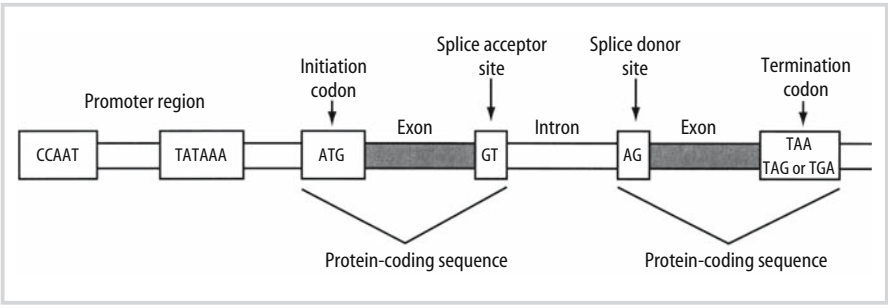


Figure 1-5. Gene structure. Gene structure depicting coding and noncoding regions of the eukaryotic gene. (Reprinted from Leonard D. *Diagnostic Molecular Pathology*, copyright 2003, with permission from Elsevier.)

may appear to be unaltered or intact but may be transcriptionally silent due to methylated bases in the promoter region.

Elongation and Termination of the mRNA

Once the RNA polymerase binds to the promoter, transcription begins at position +1. The polymerase reads the DNA in a 3′ to 5′ direction, while synthesizing the 1° RNA transcript in a 5′ to 3′ direction. In bacteria, the complete transcript serves as the template for translation. Transcription ends with termination. The mRNA must be terminated in bacteria; termination of the transcript can result from attenuation or the formation of hairpin structures. Termination occurs at several sites beyond the polyadenylation signal in eukaryotic cells and is dependent on bases near the stop codon.[15] Because the eukaryotic cell transcripts are polyadenylated, a termination of the transcription process by a process similar to attenuation is not necessary to regulate gene expression. Specifically, transcripts produced after the polyadenylation signal lack a 5′ cap, resulting in rapid degradation.[16,17]

In eukaryotic cells, once the 1° RNA transcript has been produced in the nucleus, this transcript is processed to form an mRNA by splicing to remove the non-protein-coding introns (intervening sequences) and join the protein-coding exons. Introns are located between sequences called exons, which encode the protein sequence and are translated during protein synthesis. Splicing involves a complex of ribonucleoproteins known as a spliceosome, which recognizes consensus sequences at the 5′ and 3′ ends of the intron. Genetic changes to these splice donor (A/C AG G U A/G AGU) and splice acceptor ($[U/C]_{11}$ N C/U AG G/A) consensus sequences may prevent the spliceosome from recognizing and catalyzing the splicing event.[18,19] Autoantibodies directed to or alterations in the steady-state level of the spliceosome may play a role in some diseases.[20–22] Alternate splicing may generate multiple distinct transcripts from a single gene. That is, some exons may be spliced out in one mRNA molecule but retained in another. As a result, alternate splicing generates different proteins from the same gene and 1° RNA transcript.[23–24]

An additional mechanism of generating diversity from 1° RNA transcripts entails trans-splicing (initially identified in *Drosophila* cells). Essentially, two separate, unrelated transcripts form a hybrid molecule by using the splice donor from the first transcripts and the splice acceptor from the second transcripts. Complementary intronic sequences in both transcripts facilitate the generation of the chimeric mRNA. This process has not been demonstrated in other eukaryotic cells. However, when the process is used for gene therapy applications, normal gene function has been restored from defective genes using trans-splicing.[25–26] Other therapeutic applications for catalytic RNA molecules involve innovative treatments for HIV-infected patients. In this application, synthetic ribozymes cleave drug-resistant variants of HIV.[25,27–29]

Additional modifications of the 1° RNA eukaryotic transcript enhance the stability and transport of the mRNA. One such modification occurs immediately on the generation of the 1° transcript and involves addition of a 7-methyl guanosine linked in an unusual 5′ to 5′ linkage to the triphosphate at the 5′ end of the transcript, also known as the 5′ cap. This cap protects the transcript from degradation. Another 1° transcript modification is cleavage at a polyadenylation signal (AAUAA) near the 3′ end of the transcript, followed by the addition of 100 to 200 adenosine residues (poly-A tail) by polyadenylate polymerase. The poly-A tail facilitates transportation of the mature mRNA into the cytoplasm and protection of the transcript from degradation by exonucleases. A given gene may have several polyadenylation signals, providing another level of variation for a single gene.[30–32]

Translation

Translation is the next step in using information from the DNA gene to produce a functional protein. This process changes the genetic information from a nucleic-acid-based language into an amino-acid-based language of polypeptides and proteins. For these reasons, the term "translation" describes this complex cascade of events.

Following transportation of the mRNA into the cytoplasm, translation begins with the mRNA binding to a ribosome and requires additional nucleic acids, specifically protein-associated RNA molecules (Figure 1-6). A ribosome is a complex of about 50 different proteins associated

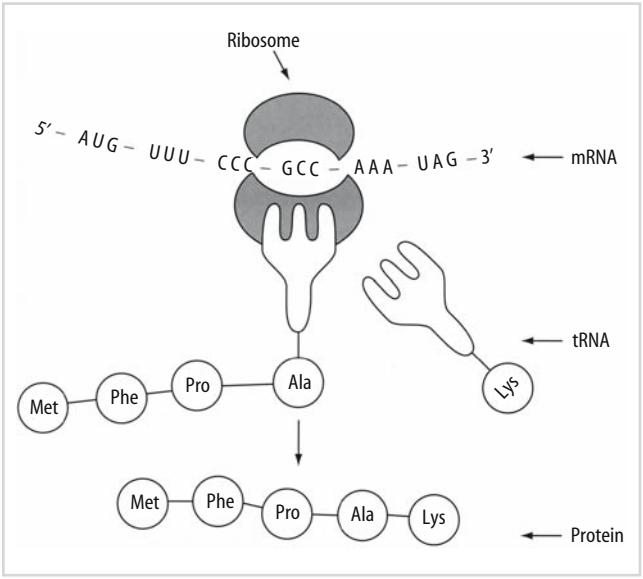


Figure 1-6. RNA translation. RNA is translated through binding events between the mRNA, a ribosome, tRNA, and amino acids, resulting in the production of a protein polypeptide chain. (Reprinted from Leonard D. *Diagnostic Molecular Pathology*, copyright 2003, with permission from Elsevier.)

Table 1-6. The Human Genetic Code

FIRST BASE OF CODON	SECOND BASE OF CODON: U	C	A	G
U	**UUU** Phenylalanine (Phe/F)	**UCU** Serine (Ser/S)	**UAU** Tyrosine (Tyr/Y)	**UGU** Cysteine (Cys/C)
	UUA Phenylalanine (Phe/F)	**UCA** Serine (Ser/S)	**UAA** Tyrosine (Tyr/Y)	**UGA** Cysteine (Cys/C)
	UUC Leucine (Leu/L)	**UCC** Serine (Ser/S)	**UAC** **STOP**	**UGC** **STOP**
	UUG Leucine (Leu/L)	**UCG** Serine (Ser/S)	**UAG** **STOP**	**UGG** Tryptophan (Trp/W)
C	**CUU** Leucine (Leu/L)	**CCU** Proline (Pro/P)	**CAU** Histidine (His/H)	**CGU** Arginine (Arg/R)
	CUA Leucine (Leu/L)	**CCA** Proline (Pro/P)	**CAA** Histidine (His/H)	**CGA** Arginine (Arg/R)
	CUC Leucine (Leu/L)	**CCC** Proline (Pro/P)	**CAC** Glutamine (Gln/Q)	**CGC** Arginine (Arg/R)
	CUG Leucine (Leu/L)	**CCG** Proline (Pro/P)	**CAG** Glutamine (Gln/Q)	**CGG** Arginine (Arg/R)
A	**AUU** Isoleucine (Ile/I)	**ACU** Threonine (Thr/T)	**AAU** Asparagine (Asn/N)	**AGU** Serine (Ser/S)
	AUA Isoleucine (Ile/I)	**ACA** Threonine (Thr/T)	**AAA** Asparagine (Asn/N)	**AGA** Serine(Ser/S)
	AUC Isoleucine (Ile/I)	**ACC** Threonine (Thr/T)	**AAC** Lysine (Lys/K)	**AGC** Arginine (Arg/R)
	AUG START Methionine (Met/M)	**ACG** Threonine (Thr/T)	**AAG** Lysine (Lys/K)	**AGG** Arginine (Arg/R)
G	**GUU** Valine (Val/V)	**GCU** Alanine (Ala/A)	**GAU** Aspartic Acid (Asp/D)	**GGU** Glycine (Gly/G)
	GUA Valine (Val/V)	**GCA** Alanine (Ala/A)	**GAA** Aspartic Acid (Asp/D)	**GGA** Glycine (Gly/G)
	GUC Valine (Val/V)	**GCC** Alanine (Ala/A)	**GAC** Glutamic Acid (Glu/E)	**GGC** Glycine (Gly/G)
	GUG Valine (Val/V)	**GCG** Alanine (Ala/A)	**GAG** Glutamic Acid (Glu/E)	**GGG** Glycine (Gly/G)

with several ribosomal RNA (rRNA) molecules. Prokaryotic ribosomes consist of 30S and 50S subunits. Svedberg (S) units are the sedimentation rate of a particle. In eukaryotes, rRNA molecules associate with proteins in the nucleolus to form 40S and 60S subunits. Recognition of the 5′ cap of the eukaryotic mRNA by a ribosome initiates the process of translation.[33]

Each amino acid is encoded by one or more 3-nucleotide sequences, which are collectively known as the genetic code (Table 1-6). Each set of 3 nucleotides of an mRNA that encodes an amino acid is called a codon. As is seen in Table 1-6, the first and second nucleotide positions largely determine which amino acid is encoded by the mRNA codon, while the third base has less effect on which amino acid will be incorporated. In addition to encoding amino acids, certain mRNA codons are used to initiate (START) or terminate (STOP) translation. The genetic code differs slightly between organisms and between mitochon-drial DNA and eukaryotic DNA (Table 1-7). Thus, while one mRNA encodes only one protein sequence, a protein sequence can be encoded by several different mRNA sequences. This is referred to as the degeneracy of the genetic code.

Synthesis of the encoded protein begins at the initiation codon of the mRNA, the first AUG codon after the promoter and encodes a methionine amino acid. This methionine codon establishes the reading frame of the mRNA. The next step in the translation process uses RNA molecules to bridge the information from the sequential mRNA codons to the encoded amino acid in the growing polypeptide chain of the protein. Another set of RNA molecules, tRNA, contain a sequence complementary to each mRNA codon known as the anticodon. The 3′ end of each type of tRNA binds the specific amino acid corresponding to its anticodon sequence. Base pairing of codons with complementary anticodons permits sequential alignment of new amino acids of the polypeptide chain and occurs in the

Table 1-7. Exceptions to the Universal Code in Mammals

Codon	Nuclear Code	Mitochondrial Code
UGA	Stop	Trp
AUA	Ile	Met
AGA	Arg	Stop
AGG	Arg	Stop

small subunit of the ribosome. The large subunit of the ribosome catalyzes the covalent bonds linking each sequential amino acid to the growing polypeptide chain.

Translation ceases when the ribosome encounters a stop codon (UAA, UAG, or UGA). Release factors bound to the stop codon catalyze the addition of a water molecule rather than an amino acid, thus resulting in a COOH terminus to the completed polypeptide chain.[34] Some factors bound to the 3′ untranslated portion of the gene also affect termination.

Structure of Proteins

Just as nucleic acids form various structures via intra- and intermolecular base pairing, proteins also assume various structures depending on the types and locations of amino acids. The primary structure of a protein is the sequence of amino acids from amino terminus (NH) to carboxy terminus (COOH) of the protein. The secondary structure refers to how amino acid groups interact with neighboring amino acids to form structure called an alpha helix or beta sheet. The tertiary structure of a protein is created by amino acids sequentially distant from one another creating intramolecular interactions. The quaternary structure of a protein defines the three-dimensional and functional conformation of the protein. The shape that is ultimately assumed by the protein depends on the arrangement of the different charged, uncharged, polar, and nonpolar amino acids.

Posttranslational Modifications

After generation of the polypeptide chain of amino acids, additional enzymatic changes may diversify its function. These changes are termed posttranslational modifications and can include proteolytic cleavage, glycosylation, phosphorylation, acylation, sulfation, prenylation, and vitamin C– and vitamin K–mediated modifications. In addition, selenium may be added to form selenocysteine. The selencysteinyl-tRNA recognizes the UGA stop codon and adds this unusual amino acid.

Mutations: Genotype Versus Phenotype

Genetic information exists in the form of nucleic acids known as the genotype. In contrast, the encoded proteins function to create a phenotype, an outwardly observable characteristic. Genotypic alterations may or may not cause phenotypic alterations. For instance, missense mutations refer to genetic changes that result in the incorporation of a different amino acid at a specific codon location. These changes may not dramatically alter the protein if the replacement amino acid is similar to the original amino acid (for example, a hydrophobic amino acid replaces another hydrophobic amino acid). However, replacement of an amino acid with a different type of amino acid may significantly change the conformation of the protein and thus change its function. For example, in sickle cell anemia, a valine replaces a glutamic acid at a single position and permits the polymerization of the beta globin molecules to cause stiffening and sickling of the red blood under low oxygen conditions. Different forms of proteins (known as conformers) provide the mechanism for diseases ranging from Creutzfeldt-Jacob disease to Huntington disease. Nonsense mutations describe base changes that replace an amino-acid-encoding codon with a stop codon, which causes premature termination of translation and results in a truncated protein.[35] Truncation may result from the addition or deletion of one or two nucleotide bases, resulting in a shift in the reading frame. Frameshifts often result in premature termination when stop codons are formed downstream from the mutation. Alterations in splice donor or acceptor sites may either erroneously generate or prevent appropriate splicing of the 1° transcript, resulting in a frameshift mutation.[36] Genetic changes in the untranslated portions of the gene affecting the promoter, enhancer, or polyadenylation signals may affect the expression of the gene product and result in a phenotypic change. Not all genotypic changes affect the phenotype. Genetic changes affecting the third base of the codon rarely alter the gene code and would therefore be less likely to cause incorporation of a different amino acid.

With the sequencing of the human genome, numerous single nucleotide polymorphisms have been identified, demonstrating the individual nature of human beings. Numerous studies currently target correlating genotype variations to disease phenotypes. These efforts, in combination with improved understanding of gene structure and function, hold the promise of improved diagnosis, treatment, and patient outcomes in the future.

Acknowledgment

I would like to acknowledge Laurie Sower, PhD, for her helpful advice and editorial assistance and our numerous discussions that greatly facilitated the preparation of this chapter.

References

1. Passarge E. *Color Atlas of Genetics*. 2nd ed. Stuttgart: Thieme; 2001.
2. Willets N. Plasmids. In Scaife DLJ, Galizzi A, eds. *Genetics of Bacteria*. London: Academic Press; 1985:165–195.
3. Hewitt WL. Penicillin-historical impact on infection control. *Ann N Y Acad Sci.* 1967;145:212–215.
4. Livermore DM. Antibiotic resistance in staphylococci. *Int J Antimicrob Agents.* 2000;16(suppl 1):S3–S10.
5. Crick JWJF. A structure for deoxyribonucleic acid. *Nature.* 1953;171:737.
6. Breslauer KJ, Frank R, Blocker H, Marky LA. Predicting DNA duplex stability from the base sequence. *Proc Natl Acad Sci U S A.* 1986;83:3746–3750.

7. Freier SM, Kierzek R, Jaeger JA, et al. Improved free-energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci U S A.* 1986;83:9373–9377.
8. Lewis ME, Arentzen R, Baldino F Jr. Rapid, high-resolution in situ hybridization histochemistry with radioiodinated synthetic oligonucleotides. *J Neurosci Res.* 1986;16:117–124.
9. Felsenfeld G, Groudine M. Controlling the double helix. *Nature.* 2003;421:448–453.
10. Preston BD, Poiesz BJ, Loeb LA. Fidelity of HIV-1 reverse transcriptase. *Science.* 1988;242:1168–1171.
11. Roberts JD, Bebenek K, Kunkel TA. The accuracy of reverse transcriptase from HIV-1. *Science.* 1988;242:1171–1173.
12. Kochetov AV, Ischenko IV, Vorobiev DG, et al. Eukaryotic mRNAs encoding abundant and scarce proteins are statistically dissimilar in many structural features. *FEBS Lett.* 1998;440:351–355.
13. Gotoh O. Homology-based gene structure prediction: simplified matching algorithm using a translated codon (tron) and improved accuracy by allowing for long gaps. *Bioinformatics.* 2000;16:190–202.
14. Olivier M, Aggarwal A, Allen J, et al. A high-resolution radiation hybrid map of the human genome draft sequence. *Science.* 2001;291:1298–1302.
15. McCaughan KK, Brown CM, Dalphin ME, Berry MJ, Tate WP. Translational termination efficiency in mammals is influenced by the base following the stop codon. *Proc Natl Acad Sci U S A.* 1995;92: 5431–5435.
16. Frischmeyer PA, Dietz HC. Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet.* 1999;8:1893–1900.
17. Wilkinson MF, Shyu AB. Multifunctional regulatory proteins that control gene expression in both the nucleus and the cytoplasm. *Bioessays.* 2001;23:775–787.
18. Bruno C, Lofberg M, Tamburino L, et al. Molecular characterization of McArdle's disease in two large Finnish families. *J Neurol Sci.* 1999;165:121–125.
19. Hou VC, Conboy JG. Regulation of alternative pre-mRNA splicing during erythroid differentiation. *Curr Opin Hematol.* 2001;8:74–79.
20. Lipes J, Skamene E, Newkirk MM. The genotype of mice influences the autoimmune response to spliceosome proteins induced by cytomegalovirus gB immunization. *Clin Exp Immunol.* 2002;129: 19–26.
21. Seidl R, Labudova O, Krapfenbauer K, et al. Deficient brain snRNP70K in patients with Down syndrome. *Electrophoresis.* 2001; 22:43–48.
22. Wehner KA, Ayala L, Kim Y, et al. Survival motor neuron protein in the nucleolus of mammalian neurons. *Brain Res.* 2002;945:160–173.
23. Garzon D, Yu G, Fahnestock M. A new brain-derived neurotrophic factor transcript and decrease in brain-derived neurotrophic factor transcripts 1, 2 and 3 in Alzheimer's disease parietal cortex. *J Neurochem.* 2002;82:1058–1064.
24. Sakata N, Yamazaki K, Kogure T, Mukai T. Alternative splicing of Rh blood group polypeptide mRNA produces a novel transcript containing a short nucleotide insertion on human erythroleukemia K562 cells. *Cell Biol Int.* 2001;25:697–703.
25. Liu X, et al. Partial correction of endogenous DeltaF508 CFTR in human cystic fibrosis airway epithelia by spliceosome-mediated RNA trans-splicing. *Nat Biotechnol.* 2002;20:47–52.
26. Phylactou LA, Darrah C, Wood MJ. Ribozyme-mediated trans-splicing of a trinucleotide repeat. *Nat Genet.* 1998;18:378–381.
27. Phylactou LA, Kilpatrick MW, Wood MJ. Ribozymes as therapeutic tools for genetic disease. *Hum Mol Genet.* 1998;7:1649–1653.
28. Lan N, Howrey RP, Lee SW, Smith CA, Sullenger BA. Ribozyme-mediated repair of sickle beta-globin mRNAs in erythrocyte precursors. *Science.* 1998;280:1593–1596.
29. Mansfield SG, Kole J, Puttaraju M, et al. Repair of CFTR mRNA by spliceosome-mediated RNA trans-splicing. *Gene Ther.* 2000;7:1885–1895.
30. Urano Y, Watanabe K, Sakai M, Tamaoki T. The human albumin gene. Characterization of the 5′ and 3′ flanking regions and the polymorphic gene transcripts. *J Biol Chem.* 1986;261:3244–3251.
31. Lin B, Rommens JM, Graham RK, et al. Differential 3′ polyadenylation of the Huntington disease gene results in two mRNA species with variable tissue expression. *Hum Mol Genet.* 1993;2:1541–1545.
32. Boyd CD, Mariani TJ, Kim Y, Csiszar K. The size heterogeneity of human lysyl oxidase mRNA is due to alternate polyadenylation site and not alternate exon usage. *Mol Biol Rep.* 1995;21:95–103.
33. Gallie DR. Protein-protein interactions required during translation. *Plant Mol Biol.* 2002;50:949–970.
34. Chavatte L, Frolova L, Kisselev L, Favre A. The polypeptide chain release factor eRF1 specifically contacts the s(4)UGA stop codon located in the A site of eukaryotic ribosomes. *Eur J Biochem.* 2001;268:2896–2904.
35. Stratakis CA. Mutations of the gene encoding the protein kinase A type I-alpha regulatory subunit (PRKAR1A) in patients with the "complex of spotty skin pigmentation, myxomas, endocrine overactivity, and schwannomas" (Carney complex). *Ann N Y Acad Sci.* 2002;968:3–21.
36. Valentine CR. The association of nonsense codons with exon skipping. *Mutat Res.* 1998;411:87–117.
37. Cline J, Braman JC, Hogrefe HH. PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res.* 1996; 24:3546–3551.
38. Maga G, Shevelev I, Ramadan K, Spadari S, Hubscher U. DNA polymerase theta purified from human cells is a high-fidelity enzyme. *J Mol Biol.* 2002;319:359–369.
39. Kuchta RD, Cowart M, Allen D, Benkovic SJ. Kinetic and structural investigations of the replicative fidelity of the Klenow fragment. *Biochem Soc Trans.* 1988;16:947–949.

Chapter 2

Molecular Pathology Methods

Megan J. Smith-Zagone, Joseph F. Pulliam, and Daniel H. Farkas

Introduction

Molecular pathology is based on the principles, techniques, and tools of molecular biology as they are applied to diagnostic medicine in the clinical laboratory. These tools were developed in the research setting and perfected throughout the second half of the 20th century, long before the Human Genome Project was conceived. Molecular biology methods were used to elucidate the genetic and molecular basis of many diseases, and these discoveries ultimately led to the field of molecular diagnostics. Eventually the insights these tools provided for laboratory medicine were so valuable to the armamentarium of the pathologist that they were incorporated into pathology practice. Today, molecular diagnostics continues to grow rapidly as in vitro diagnostic companies develop new kits for the marketplace and as the insights into disease gained by the progress of the Human Genome Project develop into laboratory tests.

Molecular pathology is a natural extension of anatomic and clinical pathology. As molecular research identifies the most fundamental causes and markers of disease, clinical testing is moving to the nucleus and its genetic material. Underlying mutations responsible for genetic diseases, including cancers, are being discovered and used in molecular diagnostic tests. Some have become routine tests in molecular pathology and are described in detail elsewhere in this book.

In this chapter, fundamental and more advanced molecular biology techniques, as practiced in the molecular pathology laboratory, are reviewed. The entire field of molecular pathology is relatively new, having begun in the 1980s and matured through the 1990s. The new century has brought important advances in automated nucleic acid preparation,[1–3] polymerase chain reaction (PCR), real-time PCR, and deoxyribonucleic acid (DNA) sequencing. More and more, the underlying biochemistry occurring in laboratory instruments may be invisible to the user. The goal of this chapter is to describe the principles of these methods so that practitioners have adequate information for instrument troubleshooting and test interpretation.

Obviously, gene products, that is, proteins and polypeptides, are molecules and could technically be classified within "molecular pathology." This chapter focuses on nucleic acid diagnostics, including investigation of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). The field will eventually move beyond genomics to proteomics. Proteomics is not addressed in this chapter since clinical proteomics is in its most nascent stages and is in extremely limited practice in molecular pathology today.

Basic Science Discoveries: The Foundation

Molecular pathology techniques are rooted in fundamental molecular biology discoveries of the 1940s to 1980s.[4] The clinical laboratory application of molecular biology techniques would not be possible without the discovery by Griffith and Avery that nucleic acid is the genetic material. The foundation of work by Chargaff and Franklin was capitalized on by Watson and Crick, who elucidated the structure of DNA. Understanding DNA structure is seminal to understanding nucleic acid hybridization, which is central to almost all molecular methods used in the clinical molecular laboratory. Additionally, work by Nirenberg (unraveling the genetic code); Wilcox, Smith, Nathans, and others (use of restriction endonucleases for DNA manipulation); Baltimore and Temin (discovery of RNA-dependent DNA polymerase or reverse transcriptase); Britten and Davis (hybridization kinetics); Kornberg and Okazaki (work on DNA polymerases and DNA replication, respectively); Southern (development of solid-phase nucleic acid hybridization, or the Southern blot); Sanger, Maxam, and Gilbert (development of DNA sequencing); Mullis (discovery of PCR for in vitro nucleic acid amplification); and their scientific collaborators and competitors led to a refined understanding of how DNA may be manipulated in vitro for research and ultimately diagnostic purposes.

General Methods

Nucleic Acid Isolation

The first step of most molecular pathology tests is isolation of DNA or RNA from a patient specimen, by either manual or automated methods. Nucleic acid purification begins with lysis of the cells in the sample. Cell lysis liberates cellular macromolecules including proteins, lipids, and nucleic acids. Cell lysis can be accomplished using a detergent solution to break cell membranes and remove lipids. Proteins are enzymatically degraded with protease, usually proteinase K, or selectively precipitated. Protein digestion is performed at about 56°C and will permanently denature many proteins but does not affect nucleic acids. This process is followed by selective extraction that takes advantage of the physical and chemical differences between nucleic acids and other cellular molecules, forming the basis for their isolation. The nucleic acid is then purified from the soluble contaminants produced in the extraction method by precipitation in an ethanol-salt solution. A variation on this theme that combines extraction and purification is the selective adsorption of nucleic acids to silica columns under chaotropic salt conditions. The isolated nucleic acid is then resuspended in a dilute salt buffer, for example, 10 mM Tris/1 or 0.1 mM EDTA pH 7.6 to 8.0 (TE buffer).

The initial lysis step is modified according to the specimen. If the specimen is fresh or frozen solid tissue, it is first homogenized in an appropriate buffer (often TE buffer). If the specimen is formalin-fixed, paraffin-embedded (FFPE) tissue, the paraffin is removed with an organic solvent such as xylenes, followed by rehydration through an alcohol series to a dilute salt buffer before protease digestion. FFPE tissue is also heated during the prolonged protease digestion step to reverse formalin cross-linking between proteins (primarily histones) and nucleic acids. Although this reduces the degree of nucleic acid shearing in subsequent vortexing or centrifugation steps, DNA longer than that packaged into a nucleosome (about 200 base pairs [bp]) is difficult to recover from FFPE tissue. Whole blood specimens may require a centrifugation step to remove erythrocytes prior to recovery of nucleic acids from leukocytes because of the inhibition of PCR by hemoglobin. This cell fractionation step generally is not performed in automated nucleic acid extraction instruments.

Organic (Phenol) Extraction

Nucleic acids have a strong negative charge because of the phosphate groups in the sugar-phosphate backbone, and thus are highly soluble in an aqueous environment. By contrast, proteins, lipids, and carbohydrates contain varying proportions of charged and uncharged domains producing hydrophobic and hydrophilic regions. This difference makes proteins entirely soluble in organic solutions or selective for the interface between the organic and aqueous phases during an organic extraction. This characteristic forms the basis for phenol: chloroform extraction, in which phenol is added to an aqueous solution containing cellular constituents, mixed, and then centrifuged to separate the aqueous and organic phases. If the pH of the extraction is near neutral, both DNA and RNA stay in the aqueous phase, while proteins are in the phenol: chloroform phase or aqueous-phenol interface. If the pH is acidic, the phosphate groups of DNA are preferentially neutralized, driving DNA into the organic phase (or interface) and allowing RNA to be selectively extracted. This method produces high-quality nucleic acids but is relatively labor-intensive, uses hazardous chemicals, and produces liquid organic waste.

Ethanol-Salt Precipitation

Nucleic acids can be precipitated in an aqueous solution by the addition of concentrated ethanol and salt. Ethanol makes the solution hydrophobic, while salt increases the ionic strength of the solution, thereby reducing the repulsion of the negatively charged sugar-phosphate backbone of the nucleic acid. Centrifugation allows the precipitate to be collected and resuspended in a dilute salt buffer (TE buffer).

Chaotropic Salt–Silica Column Extraction

Chaotropic salts such as sodium iodide (NaI) or guanidinium isothyocyanate (GITC) disrupt the structure of water, promoting the solubility of nonpolar substances, such as proteins, in polar solvents, such as water. Saturated chaotropic salts also promote the adsorption of nucleic acids to glass or silica columns. The nucleic acid is purified by a series of washing steps including reducing agents such as sodium azide to further remove contaminants and inhibit remaining enzymes. The nucleic acid is eluted from the column with a dilute, nonchaotropic salt buffer. Since the method is simple, fast, offered in commercial kits by several manufacturers, and adaptable to high-throughput robotic nucleic acid isolation, this method is widely used by clinical molecular laboratories.

RNA Versus DNA Isolation

DNA is the repository of genetic information, which is then transcribed into RNA. RNA forms ribosomes, transfer RNA (tRNA), and messenger RNA (mRNA) for protein translation, and is used for other special functions in the regulation of gene expression that have only recently been discovered. DNA is a hardy molecule present at stable cellular levels, with such well-known exceptions as the amplification and deletion of genes in tumor cells. By contrast, the level of RNA corresponding to a gene can

fluctuate dramatically within a very short time in response to changes in the cell's microenvironment and functional needs. This fluctuation results from changes in both the rate of transcription and degradation of an RNA.

DNA is relatively easy to isolate and store because deoxyribonucleases (DNases) are easily denatured by heating or inhibited by sequestration of divalent cations. RNA, by contrast, is rapidly degraded by a variety of ribonuclease (RNase) enzymes that are replete within the cell and on the skin surface. While this is necessary for cellular homeostasis, the ubiquity of RNases leads to a problem for the molecular analysis of RNA. RNases are very stable, active in virtually any aqueous environment, and can regain their activity after denaturation. As a result, RNA is subject to rapid degradation by RNases in most laboratory settings, making RNA a notoriously labile molecule. The rate of degradation varies among RNA species, leading to further analytical complexity.

RNA analysis depends on successful RNA isolation and preservation. The overall techniques are similar to those described above for nucleic acid isolation, but with the mandatory addition of steps to inhibit or degrade cellular RNases and prevent their reintroduction into the isolated RNA. RNA isolation must be performed promptly after specimen collection, particularly if quantitation of RNA is desired. If RNA isolation is delayed, the sample should be stored at −80°C, or at an intermediary isolation point in a stable buffer with RNase inhibitors. Scrupulously clean laboratory technique is required, including careful cleansing of laboratory equipment with bleach, autoclaving of glassware, preparation of reagents with nuclease-free water, and wearing gloves with frequent glove changes. Addition of GITC or beta-mercaptoethanol to the RNA isolation reagents will inhibit or denature RNases present in the sample. The isolated RNA is rehydrated in water or TE buffer that is nuclease free, and stored at −80°C to further inhibit the activity of any residual RNases.

Despite these difficulties, RNA is valuable in the clinical molecular laboratory for several reasons. Many clinically significant viruses such as human immunodeficiency virus (HIV) and hepatitis C virus (HCV) have RNA genomes. Quantitation of RNA provides an important measure of gene expression, which can be used in the diagnosis or monitoring of disease. In addition, mRNA does not contain introns, which is an advantage when analyzing neoplastic translocations with variable intronic breakpoints such as *BCR-ABL*.[5]

Nucleic Acid Measurement for Quantity and Quality

Nucleic acid quantitation is optional for many protocols that utilize in vitro nucleic acid amplification. Some methods, however, require use of more accurate quantities of nucleic acid, so assessment of the yield and concentration of purified nucleic acids is useful. This is typically done using ultraviolet (UV) spectrophotometry. The absorbance of a nucleic acid solution is measured at several wavelengths. The maximal absorbance for nucleotides is at 260 nm of UV light (A_{260}), while for proteins the maximal absorbance is at 280 nm (A_{280}). Nucleic acids can therefore be quantified by the A_{260} measurement, while the A_{260}/A_{280} ratio provides an estimate of the purity of the sample. Pure DNA has an A_{260} of 1.0 at a concentration of 50 µg/ml and an A_{260}/A_{280} ratio of 1.8, while pure RNA has an A_{260} of 1.0 at a concentration of 40 µg/ml and an A_{260}/A_{280} ratio of 2.0. Lower A_{260}/A_{280} ratios indicate the presence of protein in the solution. Other contaminants can be detected by their absorbance at other wavelengths, such as phenol at A_{270} and guanidinium at A_{230}.

Ethidium bromide (EtBr) intercalates into DNA strands, causing DNA to fluoresce upon illumination with UV light. The fluorescence of EtBr correlates with the number of base pairs of DNA in which the EtBr is intercalated, which is a result of both the size and quantity of the DNA fragment. Therefore, by staining sample DNA with EtBr in an electrophoresis gel and comparing the brightness to mass standards in adjacent lanes, the quantity of DNA can be estimated. This provides a convenient system for quantification of post-PCR DNA prior to sequencing, since the UV spectrophotometer is usually kept in the pre-PCR area (see below for PCR, sequencing, and amplicon carryover contamination). More important, the image of the EtBr-stained sample DNA can be used to assess DNA quality. High-quality, substantially intact DNA forms a single band close to the well serving as the origin of electrophoresis. In contrast, DNA degradation is apparent as a smear of EtBr-stained DNA extending downward from the well. Ethidium bromide is mutagenic and produces light background staining and is therefore being replaced by other intercalating dyes such as SYBR Green.

Electrophoresis

Electrophoresis uses an electric field to separate charged molecules by differential mobility in a sieving matrix that can be either liquid or solid (gel). The differential mobility is determined by the size of the molecule and its conformation, the net charge of the molecule (as modified by pH), temperature, and the pore size of the matrix. DNA, being negatively charged, migrates towards the anode (+) when an electric field is applied to an electrolyte solution. The size of DNA can be modified by restriction endonuclease digestion (see below), rendering DNA fragments small enough to be mobile in the matrix. Conformation can be modified with denaturing conditions prior to or during electrophoresis. Nucleic acids are usually electrophoresed at a slightly alkaline pH to ionize all phosphate groups in the backbone of the molecule.

The pore size of the matrix is determined by the composition and concentration of the polymer. For any given pore size, the mobility of a molecule through the matrix is

inversely proportional to the log of its size. Therefore, for a given size difference between two molecules, the difference in the rate of migration will be substantially less if both molecules are large. The limiting mobility is defined as the rate of migration through the gel at which large molecules can no longer be separated for any given pore size. This may be related to the tendency of sections of long DNA fragments to "snake" through different pores in the gel, retarding the mobility of the fragment. The limiting mobility of gels can be overcome by using pulsed field gel electrophoresis (PFGE), in which the voltage gradient is periodically reoriented.

Polyacrylamide Gels

In clinical molecular laboratories, the matrices used most commonly are acrylamide and agarose. Polyacrylamide gels are formed by cross-linking acrylamide monomers with bisacrylamide in the same salt buffer used for electrophoresis and pouring the solution in a thin space between two glass plates. A comb is inserted at one edge between the plates to form wells for sample insertion. After the gel has formed, the plates are mounted in a vertical electrophoresis unit such that the gel forms a bridge between two buffer chambers. Samples, controls, and sizing standards are loaded into the wells, usually in association with a dye to track the progress of electrophoresis, and glycerol to make the samples sink to the bottom of the wells. Electrodes are attached to the buffer chambers and connected to a power supply providing constant voltage. After electrophoresis, the glass plates are separated and the gel is soaked in EtBr solution. DNA is visualized by EtBr staining under UV light. Polyacrylamide forms very small pores and is useful for high resolution of DNA fragments from 100 to 1000 bp. Single base pair resolution can be achieved, allowing polyacrylamide gels to be used for sequencing under denaturing conditions (see below). However, polyacrylamide gels are thin and fragile, the glass plates are cumbersome to work with, and nonpolymerized acrylamide is a lung irritant and neurotoxin; therefore, alternatives to polyacrylamide gel electrophoresis are desirable in the clinical laboratory. Although more costly, precast acrylamide gels are commerically available to circumvent the biohazards of nonpolymerized acrylamide.

Agarose Gels

Agarose gels are formed by boiling an agarose gel powder until the agarose has completely dissolved in the same buffer used for electrophoresis, optionally adding EtBr, then pouring the solution into a horizontal casting tray. Multiple gel combs can be used to form rows of wells. After cooling and polymerization, the gel is loaded in a horizontal electrophoresis apparatus and covered with buffer in a single chamber. Wells are loaded and electrophoresis performed as described above. Agarose gels have a large pore size. Agarose gels with a concentration of 1% are used to separate DNA fragments of 1 to 20 kilobases (kb), while higher-concentration gels are useful to separate smaller DNA fragments. Agarose gels are thicker and more stable than polyacrylamide gels but do not provide the same degree of resolution. Agarose is safer than acrylamide but still must be handled and disposed of with care if the gel contains EtBr. Other modified agarose compounds are available that can be mixed in various ratios with standard agarose to increase the resolution of agarose gels. Like acrylamide gels, precast agarose gels are commerically available.

Capillary Electrophoresis

Capillary electrophoresis (CE) is a widely used separation technology for analysis of proteins, peptides, chemicals, natural products, pharmaceuticals, and DNA. Capillary electrophoresis systems are commercially available and generally provide more consistent and standardized results with less time and effort than gel electrophoresis. Using CE, DNA fragments are rapidly separated with a high-voltage gradient, because the capillary dissipates heat quickly. Therefore, one CE run takes approximately 0.5 hour or less, and if eight or 16 capillaries are run simultaneously, the process reduces the time from standard electrophoresis, which requires 3 to 4 hours. This is a significant time saving in the clinical laboratory for applications such as sequencing. CE enables more standardized results, maximization of workforce efficiency, increased productivity and throughput, and the potential for error reduction. CE also uses smaller sample volumes.

In CE, electrophoretic separation takes place in a capillary tube ranging in length from 25 to 100 cm and approximately 50 to 75 µm in diameter. Most capillary tubes are made of glass (silica) walls that often are covered with an external polyimide coating. Acid silanol groups impart a negative charge on the internal wall of the capillary. A low-viscosity acrylamide-based flowable polymer acts as the electrolyte solution and sieving matrix within the silica capillary and is responsible for the conductivity of current through the capillary. Polymer concentration affects the pore size and movement characteristics of the DNA.

A small section of the capillary coating is removed at one end of the capillary to create a detection window. The detection window is optically aligned with the detection system of the instrument. The detection system often includes either a diode or argon laser combined with a charge-coupled device (CCD) camera or filter wheel and photomultiplier tube. The opposite end of the capillary and electrode is used for the injection of the sample. Sample injection in CE is frequently performed by electrokinetic injection. In electrokinetic injection, the capillary and electrode are moved into the sample well. The sample enters the capillary when a voltage of 2 to 5 kV is applied

for approximately 5 to 15 seconds. The voltage causes sample ions including DNA to migrate electrophoretically into the capillary in a flat flow profile. Electrokinetic injection produces increased resolution compared to hydrodynamic injection, which produces a laminar (curved) flow profile. After the injection, the capillary and electrode are returned to a buffer reservoir for the separation. The DNA fragments separate by size during migration through the capillary and are detected through the window at the far end of the capillary.

In the clinical molecular laboratory, DNA sequencing and DNA fragment sizing or quantitation are the most common applications performed on CE instruments. One negative aspect of CE as opposed to older polyacrylamide gel technology is that CE is more sensitive to contaminants and DNA concentration. DNA, being negatively charged, migrates into the capillary when voltage is applied. If there are any other charged particles in the sample, they also are injected into the capillary. For example, salt is an ionic competitor. If salt is present, the fluorescent signal intensity of the sample will be greatly reduced because of ionic competition during the brief injection. Proper sample preparation is therefore a key to successful CE.

After a postreaction purification step, if needed for the specific CE application, DNA samples are resuspended in a sample loading solution. High-quality deionized formamide often is used as the sample loading solution. If DNA is denatured prior to CE, the formamide maintains the denatured state of DNA and provides a very stable environment for fluorescent dyes. Following the postreaction purification and resuspension of products, the samples are ready for analysis on the CE instrument. The fragments are injected into the capillary and detected by laser-induced florescence, and data are generated for analysis using software supplied by the manufacturer for different CE applications.

Restriction Endonucleases

Restriction endonucleases (REs) cleave DNA at specific nucleotide recognition sequences. Restriction endonucleases are naturally occurring proteins found in and purified from bacteria. Each bacterial species contains one or more REs, each recognizing a unique sequence of base pairs in double-stranded DNA, called recognition sites (most commonly 4 to 8 bp long). Bacteria use REs to digest and inactivate foreign DNA (such as bacteriophage DNA). The frequency of recognition sites in target DNA for any given RE is inversely proportional to the size of the recognition site. Some REs do not cleave DNA when their recognition sites are methylated; this can be useful in certain clinical laboratory applications such as detection of imprinted genes in genetic diseases or promoter hypermethylation in tumors. Some mutations occur at RE recognition sites and can be detected by a change in the RE digestion pattern of a PCR product or genomic DNA. Unique DNA restriction fragment patterns are generated by digestion with different REs, creating a range of DNA restriction fragment sizes. Fractionated using agarose gel electrophoresis. Restriction endonuclease digestion is commonly used as a component of clinical molecular tests.

Specific Methods

DNA Sequencing

Prior to the performance of a sequencing reaction, the DNA template to be sequenced must be properly purified and quantitated. A sequence-specific complementary primer must be synthesized for use in the sequencing reaction. Sequencing primers must also be high-performance liquid chromatography (HPLC) purified or, at the very least, desalted. Samples must be purified following the sequencing reaction to rid the sample of salts, excess dyes, and excess primers that would compete for injection into the capillary. There are numerous protocols and commercial kits available for the postreaction purification. Ethanol precipitation is the least expensive purification method, but numerous gel-filtration and spin-column methods also can be used.

The ability to sequence DNA has been essential to the field of molecular pathology because sequence information is a prerequisite for PCR, PCR alternatives, and hybridization with probes necessary for successful Southern blot analysis. The method for DNA sequencing developed by Sanger, Nicklen, and Coulson[6] is the basis for most DNA sequencing performed both in clinical laboratories and for the Human Genome Project.

The Sanger sequencing reaction uses a single DNA primer and DNA polymerase with linear amplification rather than the exponential amplification of PCR (see below). Components essential to the Sanger sequencing reaction include (1) an electrophoresis technique capable of clearly distinguishing single nucleotide length differences in DNA strands dozens or hundreds of nucleotides in length, (2) sequence-specific complementary primers, with one primer used in the forward reaction and the other used in the reverse reaction for each DNA template strand, and (3) the addition of small proportions of dideoxynucleoside triphosphates (ddNTPs) in addition to the conventional deoxyribonucleoside triphosphates (dNTPs) used in the polymerase reaction. Dideoxynucleotides differ from deoxynucleotides by having a hydrogen atom attached to the 3′ carbon rather than an OH group, which is present on the deoxynucleotide. Because the ddNTPs lack a 3′-hydroxyl group, elongation of the newly polymerized DNA chain cannot occur once a ddNTP has been incorporated (arabinonucleosides can also be used as inhibitors of elongation). The end result is a set of newly synthesized DNA chains that are complementary to the template DNA but that vary in length, with the length

determined by the point at which the ddNTP was incorporated into the 3′ end of the chain.

In the original paper, for each template four reactions were performed with the addition of a single inhibitor to each, either ddGTP, ddATP, ddTTP or araCTP.[6] The DNA chains were separated by polyacrylamide gel electrophoresis under denaturing conditions and visualized using (α-^{32}P)-dATP on a radio-autograph. These reactions were run in consecutive lanes of the gel, and the complementary DNA sequence was determined by manual inspection based on the size of each chain, and the specific ddNTP in the reaction.

Today, sequencing is very similar except that radioactive labeling has largely been replaced by fluorescent labeling. Two major categories of fluorescent labeling exist. In dye-primer labeling, the sequencing primer is labeled, and the sequencing reaction requires 4 tubes differing only in the incorporation of the specific ddNTP. In dye-terminator labeling, the sequencing primer is unlabeled and instead each ddNTP is labeled with a different fluorophore. This enables the entire sequencing reaction to be performed in a single tube. Dye-terminator labeling has therefore taken precedence in sequencing. Dye-primer labeling is frequently used in fragment analysis for detection of microsatellite instability, loss of heterozygosity, forensic identification, or allogeneic bone marrow transplantation monitoring using short tandem repeat polymorphisms. Automated sequencers recognize both the size of the DNA chain and the fluorescent color of the chain to assign the nucleotide sequence, and also function as precise detectors for fragment analysis.[7]

Conventional DNA sequencing with polyacrylamide gel electrophoresis (whether using manual or automated sequence detection) is time-consuming and labor-intensive. Therefore, the introduction of CE was a welcome change that facilitated the incorporation of sequencing and fragment analysis into the clinical laboratory.[8] The sequencing reaction products must be purified before injection into the CE unit to remove excess salts, dyes, and unincorporated primers that would compete for injection into the capillary. Purification is accomplished by ethanol precipitation or a chaotropic salt–silica column technique. After the postreaction purification step, samples are then resuspended in a sample loading solution containing high-quality deionized formamide to denature the DNA. Formamide also provides a very stable environment for fluorescent dyes. The fragments are injected into the capillary, detected by laser-induced florescence, and rendered into sequence by the analysis software. An electropherogram of the DNA sequence is generated by the detection software by correlating the fluorescent intensity of each dye wavelength corresponding to a specific ddNTP as a function of migration time.

Examples of Applications of DNA Sequencing

1. Congenital adrenal hyperplasia mutation analysis
2. Retinoblastoma mutation analysis
3. *BRCA1* mutation analysis
4. HIV genotyping to monitor drug resistance and sensitivity
5. High-resolution human leukocyte antigen (HLA) typing for allogeneic bone marrow transplantation

Southern Blot

The Southern blot was developed by E.M. Southern in 1975 and was the first molecular biology tool to have a major impact on clinical molecular pathology. The Southern blot is still used today, though it is being replaced by amplification methods. The implementation of Southern blot was based on prior knowledge of nucleic acid isolation, gel electrophoresis, RE digestion, and nucleic acid probe labeling for detection of DNA sequences of interest.

The Southern blot is a relatively labor-intensive, time-consuming clinical laboratory method.[9] High-quality DNA is isolated from a patient specimen, subjected to RE digestion, and then fractionated by gel electrophoresis. “Blotting” is the transfer of fractionated DNA from the gel to a solid support such as a nylon membrane. The DNA is then hybridized to a small piece of complementary DNA labeled in a variety of ways and called a probe. This detection step allows the gene of interest to stand out from the vast background of DNA present in the sample. If the pattern of banding visualized on the membrane is different from the normal pattern, this may be indicative of a mutation.

Because no amplification of target DNA occurs, Southern blot analysis requires a large mass of DNA. The DNA must also be intact and of high molecular weight. Therefore, electrophoresis of the isolated DNA prior to analysis is important for assessing the integrity of the DNA, since only a small degree of DNA degradation is tolerable. Degraded DNA may produce false-negative results if a signal from high-molecular-weight DNA is expected, while false-positive results may occur if partially degraded DNA results in unusually sized bands. Fortunately, most tests in the molecular pathology laboratory today are based on PCR, which is less affected by DNA degradation. Polymorphisms within RE recognition sites also change banding patterns, a principle used to advantage in other molecular tests.

The physical movement of the DNA in the gel to the membrane may be accomplished by manual capillary transfer, automated vacuum transfer, or electrotransfer. DNA in the gel must first be “conditioned”: depurination with dilute HCl and subsequent denaturation with NaOH. Dilute and brief acid treatment causes hydrolysis of the DNA phosphodiester backbone to occur spontaneously at the sites of depurination. This acid induced fragmentation facilitates efficient transfer of the highest-molecular-weight DNA species from the gel to the membrane. Alkali treatment denatures double-stranded DNA (dsDNA) to single-stranded DNA (ssDNA), essential for subsequent nucleic acid hybridization with a labeled ssDNA probe. The

DNA is permanently fixed to the membrane by thoroughly drying the blot in an oven or by exposing the blot to a precise amount of UV irradiation.

The blot is immersed in prehybridization buffer to prepare the DNA on the blot for hybridization with a probe. Prehybridization buffers contain blocking agents included to minimize unwanted nonspecific DNA probe binding that would otherwise contribute to high background on the final image of the Southern blot used to view the results and make diagnostic conclusions. The prehybridization step equilibrates the membrane and blocks sites on the nylon membrane without DNA to prevent the probe from binding nonspecifically and increasing background. A large volume of blocking agent is therefore advantageous. Addition of the labeled probe to the blot begins the hybridization phase of the Southern blot process. A small volume of buffer is used to facilitate probe and target specifically finding each other, thereby promoting hybridization. Hybridization takes several hours to overnight at an appropriate temperature determined by multiple variables: concentrations of the two species; time permitted for hybridization; complexities of the nucleic acids involved; length of the probe and its target and their complementarity to each other (degree of mismatch); pH; temperature; and ionic strength of the buffer used.

DNA probes are labeled before use in hybridization assays to permit visualization of probe-target binding (in reverse hybridization assays, described below, unlabeled probes are immobilized and the target is labeled during the amplification step that precedes hybridization). Such labeling may be accomplished isotopically or nonisotopically. High-specific-activity DNA probes may be generated by in vitro biochemical reactions that synthesize new stretches of DNA from dNTPs, using the probe as a template. One of these dNTPs is labeled with a reporter molecule such as ^{32}P, biotin, or digoxigenin. When incorporated into the newly synthesized DNA, the labeled dNTP, even though it is only one of the 4 dNTPs in the DNA probe, is sufficient to label the entire probe for detection. The probe is then used in vast molar excess relative to target DNA in nucleic acid hybridization to drive the hybridization reaction as quickly as possible.

After hybridization, the blot is washed with buffers containing sodium chloride and detergent to remove excess probe and reduce background. Sodium chloride concentration and stringency are inversely related: the lower the sodium chloride concentration, the more stringent the wash. Increasingly stringent washes remove more nonspecifically bound probe. The temperature of the wash buffer and stringency are directly related: high-temperature washes are more stringent than lower-temperature washes and further contribute to hybridization specificity. When appropriately stringent washing of the blot is complete, only the specific hybrids of interest should remain. Visualization of these specific hybrids, which appear as bands, is achieved by autoradiography for radioactive probes or by luminography for chemiluminescent probes. Hybridization with biotinylated probes is followed by chemical reactions, resulting in insoluble colored precipitates at the site of hybridization on the blot itself that serve as the endpoint (this is also the detection scheme used in the line probe assay; see below). Simple visual inspection is then applied for both isotopic and nonisotopic Southern blots to determine the position where the labeled probe hybridized to its target patient DNA. That position, relative to detection of appropriate controls, allows interpretation.

Northern blotting is an extension of Southern blotting that uses RNA instead of DNA as the target of investigation. Northern blotting is as labor-intensive as Southern blotting but even more problematic due to the highly labile nature of RNA. While northern blotting has been very useful in the research setting to demonstrate the selective expression of genes in various organs, tissues, or cells, it has not become a routine tool in the clinical molecular pathology laboratory.

Examples of Applications of Southern Blotting

1. B- and T-cell antigen receptor gene rearrangement for leukemia and lymphoma[10]
2. Fragile X syndrome diagnosis
3. Myotonic dystrophy diagnosis

Polymerase Chain Reaction

In the mid-1980s in California, Mullis and coworkers developed a method, the polymerase chain reaction (PCR), to amplify exponentially target sequences of DNA.[11] As the name suggests, the method is a DNA polymerase–mediated chain reaction of nucleic acid amplification. Arguably, it is the single most important "invention" that has led to development of a new discipline in clinical laboratory medicine, that is, molecular pathology. Both PCR and Southern blotting are techniques used to investigate specific genomic targets. However, PCR is orders of magnitude more sensitive and much faster, permitting turnaround time of 24 hours or less. PCR lends itself to much higher test volumes than Southern blotting, a crucial point in its acceptance in the clinical laboratory setting. Opportunity for high test volumes, excellent specificity and sensitivity, and the rapid turnaround times of PCR are the principal reasons this technology has spread so quickly in clinical molecular laboratories.

In PCR, a unique sequence of the target nucleic acid of interest is chosen for amplification, for example, oncogene, invading pathogen DNA, genetic mutation. The inherent specificity of the ensuing reaction is provided by two short oligonucleotides, called PCR primers (see Figure 2-1). These short oligonucleotides serve as primers for DNA polymerase–mediated DNA synthesis using denatured target DNA as a template. The two primers are complementary to opposite strands and opposite ends of the targeted DNA template region. Usually the primers bracket

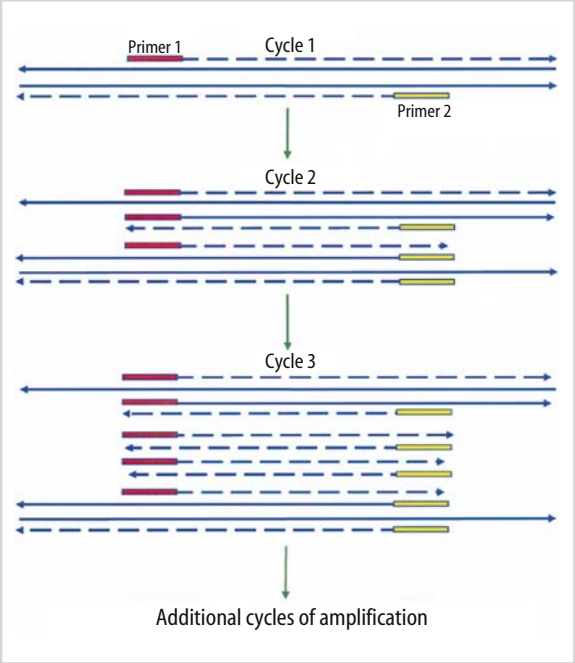


Figure 2-1. The polymerase chain reaction. (Reprinted with permission from Tsongalis GJ, Coleman WB. *Molecular Diagnostics—A Training and Study Guide*. Washington, DC: AACC Press, 2002.)

the area of interest, but one type of PCR (allele-specific PCR; see below) uses primers that overlap the area of interest. Successful PCR depends on temperature cycling, and in the first step of PCR the reaction temperature is raised to 95°C to 98°C to denature the target DNA. After 10 to 60 seconds at this temperature, the temperature is reduced to about 50°C to 70°C, depending on the specific protocol, and held there for usually 10 to 60 seconds. This facilitates hybridization (annealing) between the now-denatured target and the PCR primers, and is thus called the annealing step. This hybridization event is favored over target reannealing because the PCR primers are small and present in vast molar excess, and move more rapidly in solution than larger DNA molecules.

The hybridized PCR primers form local areas of double strandedness with the template DNA, thereby serving as primers for DNA polymerase to bind and synthesize a new strand of DNA, using the target DNA as a template. Subsequent to the initial discovery of PCR, the opportunity for automating the cyclical nature of PCR was realized by using DNA polymerase from hot-spring living bacteria, *Thermophilus aquaticus* (hence the term "*Taq* polymerase"). *T. aquaticus* thrives at very high temperatures, and so its proteins do not denature at the high temperatures needed to denature DNA in the first step of PCR. Catalysis by *Taq* polymerase of a new strand of DNA proceeds at a temperature intermediate to the near-boiling temperature used for denaturation and the relatively lower temperature used for annealing. DNA polymerization occurs during this extension step, typically at 65°C to 75°C. Taken together, these three steps (denaturation, annealing, and extension) define one PCR cycle.

Temperature cycling is automated through the use of an instrument called a thermal cycler. Thermal cyclers hold small capped tubes containing the reagents needed for PCR and cycle between the temperatures needed for the different steps of the PCR.[12] A single PCR tube contains template DNA (<1 ng to 1 µg), *Taq* DNA polymerase, two PCR primers (~15 to 30 nucleotides long), all four dNTPs, Mg^{2+}, and buffer to maintain an elevated pH (~8.4) optimal for *Taq*.

The repetition of the cycles generates exponential amplification of the target DNA because each double-stranded target DNA molecule, theoretically even if there is only one, is replicated after one PCR cycle. Both the original and replicated DNA molecules are then available to function as templates for cycle 2, in true "chain reaction" style, generating another doubling, or four copies of the original target. Cycle 3 ends with eight molecules, and doubling continues with completion of each new cycle. This doubling plateaus in later cycles since reagents, usually dNTPs, become limiting. Additionally, the enzyme may not function at 100% efficiency, and so true exponential amplification is theoretical, although there is a true exponential phase of amplification.

Greater than one billion identical copies of the original target DNA region are generated after 32 cycles of PCR: 2^{32} or more than four billion, the difference owing to the fact that unit-length amplicons are not generated until the end of the second cycle of PCR. Amplicons (PCR products) are defined as replicated target molecules created by PCR. Unit-length amplicons are those whose ends are defined by the primers. During the first cycle, the primers are extended by *Taq* polymerase using template DNA. The termination of this extension is undefined and a function of how far the polymerase moves down the template during the time allotted. The enzyme, therefore, moves beyond the ends of the primer-binding site on the complementary strand. After completion of the first cycle, therefore, the newly synthesized DNA molecules are greater in length than the sequence bracketed on each strand by the primers. In the second cycle, DNA molecules are synthesized from the products of the first cycle whose ends are defined by the two primers. These are so-called unit-length amplicons. While all of the above is true, the practical clinical laboratory difference between one- and four-billion-fold amplification is irrelevant because either number is sufficient for detection of the target, often by electrophoresis with EtBr visualization.

Several factors affect the specificity and sensitivity of PCR. The production of specific PCR amplicons is a function of both the complementarity of the primers to the target DNA and the annealing temperature of the PCR cycle. Heating will denature the primer from its target DNA. The temperature at which the primer melts from the target DNA varies directly with the length of the primer and the guanosine-cytosine (GC) content of the primer, and inversely with the degree of mismatch between the primer and the target DNA. The melting temperature (T_m) of the primer is the temperature at which 50% of the primer is denatured from

the target DNA. If the thermal cycler is programmed to reach an annealing temperature higher than the primer T_m, the efficiency of PCR is compromised and sensitivity decreased. In contrast, if the annealing temperature is substantially less than the primer T_m, the primer can bind to both complementary and noncomplementary DNA, resulting in reduced PCR specificity as nontarget DNA is amplified (and potentially decreased sensitivity as reaction components are used nonspecifically). Therefore, the ideal annealing temperature is slightly less than the T_m of both primers, and the primers should be designed to have a very similar T_m. The annealing temperature can be decreased with subsequent cycles during PCR in a process called "touchdown" PCR. This allows the initial cycles to produce specific products at high annealing temperatures, while later cycles amplify previously generated amplicons more efficiently using lower annealing temperatures, thereby increasing sensitivity (see also the use of touchdown PCR in multiplex PCR, below).

Taq polymerase is very sensitive to mismatches between the primer and the target DNA at the 3′ end of the primer but can withstand considerable noncomplementarity at the 5′ end of the primer. Numerous PCR variations have been designed to take advantage of both these facts. *Taq* polymerase also requires Mg^{2+} as a cofactor for stabilization of primer annealing. Insufficient Mg^{2+} decreases PCR efficiency, while too much Mg^{2+} stabilizes nonspecific primer annealing. Primers with a high GC content may show a narrow range of tolerance for variation from ideal PCR conditions, leading to decreased amplification or nonspecific products. This may be alleviated by using PCR additives such as dimethyl sulfoxide (DMSO) or glycerol, but the success of these additives may need to be determined empirically for different primer pairs. Another strategy to improve specificity is the use of "hot-start" PCR, in which a crucial PCR reactant such as *Taq* is either physically or chemically sequestered from other PCR reagents until denaturation begins. This prevents the generation of nonspecific amplification products by inhibiting the activity of *Taq* until after the initial PCR denaturation step.

PCR is more sensitive than Southern blot hybridization because of the amplification of the target sequence. However, the specificity of the amplified PCR product must be verified. Simple agarose gel electrophoresis coupled with EtBr staining may be used to observe the PCR product(s). When a clinical PCR protocol is established, such gels may be subjected the first time to blot hybridization with a specific probe complementary to the internal, non-primer sequence of the amplicon(s). This exercise proves that the PCR-generated band not only is the correct size and highly likely therefore to be the correct target, but also is a DNA fragment that has high or perfect homology with a known probe. For example, hybridization of a particular 302 bp PCR product band detectable on an agarose gel with a defined cytomegalovirus (CMV) DNA probe confirms that the oligonucleotide primers synthesized based on the CMV sequence and used in the PCR are recognizing CMV-specific DNA and that the PCR is indeed specific for detection of CMV. An alternative method to validate the specificity of the PCR product is to sequence the PCR product. Following this one-time validation analysis, electrophoresis alone, as opposed to blot hybridization or sequencing, may be the assay endpoint.

There have been significant commercial endeavors to automate or semiautomate high-volume PCR-based clinical tests. For example, denatured aliquots of completed PCRs can be added to microtiter plates with wells to which specific DNA probes are bound. In the presence of amplicon, for example, if the patient is infected with the pathogen of interest or a specific mutation is present, the amplicons hybridize to the bound probe and are retained in the well during washing. Subsequent biochemical reactions are used to detect labeled moieties in the amplicons ("built in" to the PCR components), facilitating colorimetric detection of a positive patient reaction by an automated plate reader. Absence of colored product in a well indicates a negative result for that patient specimen, provided that all positive and negative controls are within tolerance limits. This scheme has gained US Food and Drug Administration (FDA) approval for clinical PCR-based detection kits for *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, HCV (qualitative) and HIV.[13] (For a complete list of FDA-approved tests, go to the Resources section at http://www.amp.org/, the home page for the Association for Molecular Pathology). Subsequent generations of automated PCR instrumentation are now available that completely automate the amplification and detection process.[14] The field is moving toward real-time PCR detection (described below).

Another aspect of PCR that is attractive for the clincal molecular laboratory is the ability to use relatively crude extractions of patient specimens rather than highly purified DNA. Cell lysis and subsequent DNA liberation accomplished by boiling or treatment with detergent may be sufficient to process a specimen in preparation for PCR.[15] Conventional PCR-based tests may be completed with turnaround times of as short as 2 to 4 hours, while real-time PCR can be completed in 30 minutes, making this technique attractive for stat testing.

Examples of Applications of PCR

1. Detection of the diagnostic *BCL2*/*IGH* gene rearrangement in follicular lymphoma
2. Detection of *Chlamydia trachomatis* in urine

PCR Variations

PCR-Restriction Fragment Length Polymorphism Analysis

Polymorphisms are inherited differences found among the individuals in a population. The term "polymorphism" is not synonymous with the term "mutation" in that

mutations are variations found less frequently in a population or occur as nongermline changes usually in a tumor cell (somatic mutations). In the case of restriction fragment length polymorphisms (RFLP), DNA sequence differences alter RE recognition sites, manifested either as obliteration or creation of a restriction site. With obliteration of an RE site, the DNA of individuals with an RFLP exhibits a larger restriction fragment of DNA than those without the polymorphism. With creation of a new RE site, RE digestion results in two smaller fragments relative to the unaffected individual. In either case, the polymorphism is detectable by creation of a new restriction fragment pattern, that is, a restriction fragment length polymorphism. In PCR-RFLP, the PCR products are digested by one or a combination of REs and electrophoresed to detect polymorphisms or mutations which are seen as changes in the DNA fragment sizes reflected by changes in the band pattern on the gel.

Examples of Applications of PCR-RFLP Analysis

1. Detection of sickle-cell hemoglobin (HbS)
2. Detection of the *Mnl* I restriction enzyme polymorphism created by the Factor V_{Leiden} mutation[16]

Restriction-Site Generating PCR

Some mutations create or abolish RE recognition sites and can easily be detected by PCR-RFLP. Unfortunately, most polymorphisms or mutations do not alter an RE recognition site. In restriction-site generating PCR (RG-PCR) (and a related research technique called PCR-mediated site-directed mutagenesis [PSDM]), an artificial RE recognition site is generated during PCR using a specially designed PCR primer.[17,18] The primer contains a base mismatch to the template DNA adjacent to the variable base of the mutation that creates an RE recognition site in the PCR product. The mismatched base in the primer is located near or at the 3′ end of the primer, which is near or adjacent to the variable base of the mutation, and together they create a novel restriction site within either the mutant or wild-type amplicon. The presence or absence of the RE recognition site is determined from the pattern of digested fragments by gel electrophoresis. Not all sequences are amenable to the generation of a restriction site, and the amplification efficiency is often decreased due to destabilization of the primer with the mismatch.

Examples of Applications of RG-PCR

1. Identification of *KRAS* codon 12 mutations in colon cancer[19]
2. Identification of mutations in the *CTFR* gene in cystic fibrosis
3. Identification of mutations in the *ATM* gene in ataxia-telangectasis

Multiplex PCR

Multiplex PCR (M-PCR) is a demanding technique used for amplification of several discrete genetic loci with multiple PCR primer pairs in a single reaction. Multiplex PCR simultaneously answers several related questions about a specimen without the need for multiple individual PCR reactions. Multiplex PCR is commonly used for verification that amplifiable nucleic acid is present in the sample, for example, amplification of a housekeeping gene in addition to the gene sequence of interest, and to check for the presence of PCR inhibitors that can prevent amplification of target nucleic acid, for example, coamplification of an exogenously added internal control. Multiplex PCR often requires painstaking optimization of PCR conditions and careful design of the multiple primer pairs to prevent the generation of primer-dimers and other nonspecific PCR products that may interfere with the amplification of specific products. Touchdown PCR can be used with multiplex PCR if the primer pairs have different annealing temperatures. Concentrations of individual primer pairs may need to be optimized to account for different amplification efficiencies and competition between the primer pairs.

Examples of Applications of Multiplex PCR

1. Detection of enterovirus and herpes simplex virus (HSV) nucleic acids in cerebrospinal fluid (CSF)
2. Detection of pathogenic enteric bacteria in stool
3. Analysis of multiple *BRCA1* loci in a breast cancer patient[19]
4. Identification of different bacteria in a respiratory infection[20]
5. Amplification of multide microsatellite loci for bone marrow engraftment analysis

Nested PCR

In nested PCR, two pairs of PCR primers with one set internal to the other (nested) are used to sequentially amplify a single locus. The first pair is used to amplify the locus as in any PCR assay. A dilution of the first PCR reaction then is amplified with the nested primers. Alternatively, semi-nested PCR is performed using one of the original PCR primers and one new internal primer in a second round of amplification. Both nested and semi-nested PCR generate a second PCR product that is shorter than the first one.[21] The logic behind this strategy is that if the wrong locus was amplified incorrectly or nonspecifically, the probability is very low that it would be amplified a second time by a second pair of primers. Thus, nested PCR enhances specificity while also increasing sensitivity. The problem with nested PCR is the high risk of amplicon contamination when the first-round PCR products are used to set up the second round of PCR with the nested primers (see

section below on amplicon contamination control). For this reason, many clinical laboratories do not use nested PCR procedures.

Allele-Specific PCR

Allele-specific PCR (AS-PCR) also is referred to as amplification refractory mutation system (ARMS), PCR amplification of specific alleles (PASA) and PCR amplification with sequence-specific primers (PCR-SSP). AS-PCR is based on the principle that a 3′ mismatch between a PCR primer and the template DNA prevents PCR amplification.[22] AS-PCR is especially useful for detection of single nucleotide polymorphisms (SNPs) or mutations. For AS-PCR, target DNA is amplified in two separate and simultaneous reactions. Each reaction contains an allele-specific primer (either normal or mutant) and a second primer common to both reactions. PCR is performed under stringent conditions, to prevent PCR amplification if a mismatch is present. Genotype is based on amplification in either one of the reactions alone (homozygous normal or mutant) or both reactions (heterozygous). Detection of the amplicon is by either gel electrophoresis or real-time PCR technology (see below). A disadvantage of AS-PCR is that unsuspected nucleotide polymorphisms or mutations located in the DNA template at or adjacent to the 3′ binding site of the primer would prevent amplification, leading to incorrect genotyping.

AS-PCR can detect one mutant allele in the presence of 40 copies of the normal allele. AS-PCR can be combined with M-PCR using multiple allele-specific primers in the same reaction tube. This technique is known as multiplex ARMS, a useful method when a single disease is caused by different mutations in one or more genes. Multiplex PCR-SSP also is commonly used in low-resolution HLA typing, in which multiple primer pairs for HLA loci are used along with control primers that amplify a housekeeping gene to verify that amplifiable DNA is present in each reaction tube.

Examples of Applications of AS-PCR

1. Detection of multiple cystic fibrosis CFTR mutations
2. Detection of alpha-1 antitrypsin deficiency mutations
3. Defection of phenylketonuria mutations

Allele-Specific Oligonucleotide Hybridization

Allele-specific oligonucleotide hybridization (ASOH), also known as dot-blot analysis, is used for genotyping of highly polymorphic regions of DNA. ASOH can be thought of as a variation of the Southern blot, in that patient DNA amplified by PCR is bound to a membrane and hybridized with labeled allele-specific oligonucleotide probes.[23] Reverse dot-blot analysis differs from ASOH in that unlabeled allele-specific oligonucleotide probes are spotted onto different membrane locations and hybridized with labeled PCR amplicons.

For ASOH, the PCR products are denatured and a small amount of denatured (single stranded) amplicon is spotted onto a nylon or nitrocellulose membrane. The amplicon is permanently bound to the membrane by baking under vacuum or UV cross-linking. Amplicons from different specimens can be spotted at different locations to interrogate the genotype of multiple specimens simultaneously. Duplicate membranes are made for each probe type. Each membrane is hybridized with two different labeled oligonucleotide probes (one complementary to the mutant sequence and another to the normal sequence of the same DNA region). The membranes are washed to remove nonspecifically bound probe. Samples that hybridize strongly to only one probe indicate homozygosity for the normal or mutant allele; those that hybridize with both probes are heterozygous. The oligonucleotide probes are labeled and detected by radioactivity (often avoided in clinical molecular laboratories), fluorescence, colorimetry, chemiluminescence, or mass spectrometry. One criticism of ASOH is the potentially ambiguous discrimination of a positive signal. Optimization of the assay and the use of both positive and negative controls help to define and score ASOH results.

Example of Application of ASOH

1. Low-resolution HLA typing

Oligonucleotide Ligation Assay

Oligonucleotide ligation assay (OLA) is a highly specific method for detecting well-defined alleles that differ by a single base.[24,25] The target sequence is initially amplified using PCR and then denatured. A pair of allele-specific oligonucleotide (ASO) probes (one specific for the wild-type allele and the other specific for the mutant allele), a common reporter probe (complementary to a sequence common to both alleles), and DNA ligase are added to the denatured PCR products. The ASO probes are designed to differ from one another only at the terminal 3′ base. The common reporter probe is positioned immediately adjacent to the 3′ terminal end of the ASO probes. If the ASO is complementary to the amplicon, DNA ligase can covalently join the ASO and the reporter probe. If the ASO is not a perfect match to the amplicon, the 3′ base does not anneal with the amplicon, and DNA ligase cannot join the ASO and reporter probes. The ligation products are analyzed by electrophoresis. Alternatively, one of the probes can be biotinylated at the 5′ end and the other probe tagged at the 3′ end with a reporter molecule such as fluorescein or digoxigenin. If ligation occurs, the ligation product is biotinylated at one end, facilitating capture onto a streptavidin-coated microtiter plate. The opposite end contains the reporter label. Washing removes unbound label and the reporter molecule is detected.

Dynamic Allele-Specific Hybridization

Dynamic allele-specific hybridization (DASH) is a temperature-dependent, real-time variation of ASOH.[26] Dynamic allele-specific hybridization begins with amplification of polymorphic DNA regions by PCR in which one primer is biotinylated. The biotinylated product strand is bound to a streptavidin-coated microtiter plate well, and the nonbiotinylated strand is washed away. An oligonucleotide probe complementary to one allele is annealed to a bound PCR product strand, forming a DNA duplex that interacts with a double-strand-specific intercalation dye. On excitation, the dye emits fluorescence proportional to the amount of double-stranded DNA (probe-target duplex) present. The duplex is heated through a temperature range while the fluorescent signal is continually monitored. A rapid fall in fluorescent signal indicates denaturation of the DNA duplex. DNA duplexes with nonhomologous regions denature at a lower temperature than completely homologous DNA. The melting-temperature profile distinguishes homozygosity for either allele alone or a heterozygous mixture of the two.

Reverse Transcription–Polymerase Chain Reaction

Reverse transcription–polymerase chain reaction (RT-PCR) may be thought of as RNA-based PCR. RT-PCR was made possible by the discovery in the early 1970s of retroviral reverse transcriptase (RT), an RNA-dependent DNA polymerase, by David Baltimore and Howard Temin,[27] for which they shared the Nobel Prize in 1975. Reverse transcriptase catalyzes DNA synthesis using RNA as the template, producing a DNA strand complementary to the RNA template, called complementary DNA (cDNA). Because cDNA is not subject to RNase degradation, it is far more stable than the corresponding RNA. Complementary DNA can be treated like any other DNA target in subsequent PCR. Logistically, RT-PCR is trivially more time-consuming than PCR due to the extra enzymatic step of reverse transcription, but there are enzymes that combine reverse transcription and DNA polymerase activities, making RT-PCR more efficient. With the introduction of techniques to successfully isolate and protect RNA from ubiquitous RNases, and to synthesize cDNA by reverse transcription and the discovery of PCR, RNA analysis is now virtually as rapid and sensitive as PCR-based DNA investigation. RT-PCR is a high-volume test method for the clinical molecular laboratory due to its use in the diagnosis and quantitation of RNA virus infections, principally HIV and HCV.

Examples of Applications of RT-PCR

1. HIV and HCV viral load determinations
2. Detection of *BCR/ABL* translocation diagnostic of chronic myelogenous leukemia

Real-Time PCR

Real-time PCR is based on the generation of a fluorescent signal by the PCR process, which is detected during PCR cycling (i.e., in real time) and reflects the amount of PCR product synthesized.[28–30] Different real-time PCR methods use alternative ways to generate a fluorescent signal during PCR. These include an intercalating dye such as SYBR Green that binds the minor groove of DNA, or an oligonucleotide used as a primer or probe and labeled with a fluorogenic dye. Instruments that combine in vitro nucleic acid amplification and real-time detection of the PCR product are dramatically increasing test menus for oncology, infectious diseases, and genetics in clinical molecular laboratories because of the wide range of readily available amplification primers and detection schemes, rapid turnaround time, the potential to eliminate the time and costs associated with electrophoresis, and the reduction in PCR contamination risk.

Real-time PCR is different from conventional PCR in several ways.

- Amplicon generation, temperature profiles, and melt curves are monitored in real time, reducing the time required for post-PCR analysis. In most cases, there is no need for postamplification processing of the PCR products. This eliminates the need for gel electrophoresis and, because the reaction tubes remain closed after PCR starts, there is decreased risk of amplicon carryover contamination within the laboratory.
- Results are more reproducible between runs since quantitation of target is based on amplification cycle threshold in the log-linear phase of amplification rather than traditional endpoint analysis in the PCR plateau phase.
- Real-time PCR methods have a wide dynamic range, up to 10 logs.
- Real-time PCR systems with intercalating dye or fluorogenic probes can be used to perform melting-curve analysis. This adds a check for specificity or potentially the detection of unknown sequence variants (see below).

The simplest real-time PCR method uses intercalating dyes that insert into the stacked bases of DNA PCR products, allowing detection of amplification in real time. These dyes, for example, SYBR Green and EtBr, are nonsequence-specific dyes that increase in fluorescence when bound to double-stranded DNA. Intercalating dyes are often used for melting-curve analysis, qualitative and semiquantitative PCR, product discrimination and purity, and determination of primer and probe melting T_m. Intercalating dyes can be used for quantitative PCR. Results, however, are more specific and accurate with a sequence-specific probe since fluorescence is directly proportional to the amount of specific amplicon produced and reduces the background contributed by primer-dimers or nonspecific PCR prod-

ucts. Intercalating dye fluorescence represents all double-stranded DNA, including primer-dimers and other non-specific product that can be visualized with an endpoint melting-curve analysis.

Most fluorogenic oligonucleotide techniques take advantage of the principle of fluorescent resonance energy transfer (FRET), in which the energy from one dye molecule (the donor) is transferred without the emission of a photon to an acceptor dye molecule when the two are in close proximity. If the acceptor is a fluorophore, a photon is emitted at a characteristic wavelength. However, if the acceptor does not emit a photon, the energy is dissipated and fluorescence from the donor is quenched. The reporter dye can be either the donor (if no FRET takes place) or the acceptor (if FRET does take place) and is defined as the one whose fluorescence correlates with the quantity of desired PCR amplicon. Several fluorogenic techniques are described below.

TaqMan

TaqMan uses a short probe complementary to a non-primer internal sequence of the PCR product. The probe is labeled at the 5′ end with a reporter donor dye and at the 3′ end with an acceptor dye that quenches the reporter when the probe is intact. During the extension phase of PCR, probe bound to an amplicon is cleaved by the 5′ endonuclease activity of *Taq* polymerase, freeing the reporter dye from the quencher and resulting in fluorescence. The fluorescent signal increases proportionally to the number of amplicons generated during the log-linear phase of amplification. To ensure that hydrolysis of the probe occurs, a two-step PCR can be used with annealing and extension taking place at the same temperature (~60°C). Ideally, the TaqMan probe binding site is located near one primer and the size of the amplicon is no longer than 200 to 300 bases. One negative aspect of this format is that once the probe is hydrolyzed, it is unavailable for subsequent reactions or melting-curve analysis, thus requiring an excess amount of probe in the reaction mix with the potential to decrease the PCR efficiency.

Molecular Beacon

A molecular beacon is a longer probe with a 5′ reporter dye and 3′ quencher dye. The probe forms a hairpin loop structure when not bound to target DNA, thereby juxtaposing the dyes with quenching of fluorescence. The loop sequence is complementary to the non-primer amplicon sequence. When the loop of the molecular beacon probe hybridizes to the amplicon during the annealing step of real-time PCR, the reporter dye is separated from the quencher, resulting in fluorescence. For the molecular beacon probe to anneal to the amplicon, the amplicon-probe hybrid must be more stable than the internal base-pairing hairpin so that a fluorescent signal is generated. Generally, DABCYL is the nonfluorescent universal quencher and the other dye is a reporter fluorophore such as FAM, Cy3, TET, TAMRA, Texas Red, ROX, or Cy5.

Hybridization

Hybridization is a two-probe system in which one probe contains a donor dye and the other contains the reporter acceptor dye. The probes are designed to anneal to one strand of the PCR product adjacent to one another and internal to the primers. This juxtaposes the dyes, allowing FRET to occur. This probe format works well with the traditional three-step PCR with annealing at ~55°C (primer specific) and extension at 72°C, the optimal temperature for *Taq* polymerase activity. When DNA polymerase encounters the probes, they are displaced from the target strand rather than hydrolyzed and thus available for the next round of amplification as well as endpoint melting-curve analysis.

Uniprimer (Amplifluor, Sunrise)

Like molecular beacon probes, the uniprimer system uses a hairpin structure in the oligonucleotide to quench fluorescence. The 3′ region of the fluorogenic oligonucleotide is identical to a nonbinding region at the 5′ end of the reverse PCR primer. This allows the fluorogenic oligonucleotide to become a primer for the newly formed amplicon by the third round of PCR. The probe is then opened in the fourth and subsequent rounds of PCR by the polymerase action of *Taq*, allowing fluorescence to occur. The advantage of this system is that the same fluorogenic oligonucleotide can be used in any PCR reaction (universal fluorogenic primer).

Scorpion

Scorpion also uses a hairpin structure in the oligonucleotide to quench fluorescence. The fluorogenic oligonucleotide is part of the reverse primer, and the nucleotides in the hairpin are complementary to the PCR amplicon sequence between the primers. The Scorpion primer unfolds and anneals to the PCR amplicon, allowing fluorescence to take place beginning in the first round of PCR.

Lux

Lux is a variation of real-time PCR that uses a single fluorophore in a primer with a hairpin loop structure. The fluorophore is quenched by the complementary structure of nucleotides in the stem of the hairpin loop. When the primer is incorporated into double-stranded DNA, thus opening the hairpin loop, fluorescence is maximal. The

advantage of this system is lower production costs with the use of only one fluorophore.

Real-Time PCR Method

The following concepts are important for understanding the use of real-time PCR in a clinical diagnostic laboratory. When optimizing real-time PCR, the growth curve of the fluorescent signal versus the number of PCR cycles should be monitored to determine when optimal conditions have been achieved (Figure 2-2). The growth curve should be sigmoidal (S shaped) with three phases: baseline (background signal or lag phase), log-linear (exponential amplification phase), and plateau. For each phase, several characteristics should be assessed. The baseline phase of the curve represents initial cycles of amplification in which accumulation of the specific signal has not yet exceeded the background signal. The fluorescent signal in this phase is from unbound probe or autofluorescing components in the reaction. The log-linear phase of the curve represents exponential amplification of the target (see Figure 2-1) and provides useful information about the reaction. The curve can be described by the following equation: $T_n = T_0(E)^n$, where T_n is the amount of target sequence at cycle n, T_0 is the initial amount of target sequence at cycle 0, and E is the amplification efficiency of the target sequence. The crossing point represents the number of PCR cycles at which the growth curve enters the log-linear phase. There is an inverse linear relationship between the crossing-point cycle number and the number of template copies present in a reaction.

The slope of the log-linear phase is a reflection of amplification efficiency, and the efficiency of the reaction can be determined by identifying the crossing points of known standards and plotting a line of linear regression (see Figure 2-3). The efficiency can then be determined using the following equation: $E = 10^{-1/\text{slope}}$, where E is efficiency and slope is the slope of the standard curve. Using this equation, the slope should be between −3 and −4, with −3.3 indicative of efficiency close to or at 2. The inflection point is the point at which the log-linear amplification curve goes from positive to negative and begins to enter the plateau phase. If there is no inflection point, the curve may represent not amplification of DNA, but rather signal drift. Drift is characterized by gradual increase or decrease in fluorescence without amplification of product.

Plateau is defined as the phase of growth when critical components become rate limiting and amplicon accumulation is minimized or stops. The plateau is also the point at which incremental increase in fluorescent signal stops. As the rate of accumulation slows and enters the plateau phase, the curve levels. Since endpoint measurements are often made in conventional PCR when reaction components are limited, minor sample variations can have a relatively major effect on endpoint product. Real-time PCR focuses on the rate of amplicon accumulation (log-linear phase and cycle threshold or crossing point), not on endpoint signal. The plateau phase can be shortened by decreasing the number of cycles for a product of greater purity. Several factors contribute to the plateau phase: PCR product reannealing versus primer annealing, enzyme or dNTPs becoming limiting, and amplicon buildup with resultant reaction inhibition.

Real-time PCR utilizes the log-linear phase of the amplification curve for data analysis. This method provides a more accurate measurement than endpoint analysis. The cycle at which the curve crosses a specified threshold is called the cycle threshold (Ct), or crossing point (Cp). The Ct value can be used for qualitative or quantitative analysis. A qualitative analysis uses the defined Ct as a pass/fail measurement. A quantitative assay uses the Ct of defined standards of known template concentration to generate a standard curve. Then the Ct values for unknown samples are used to extrapolate the concentration(s) in the unknown samples from the standard curve. Some real-time instrumentation software allows determination of the Ct by a mathematical analysis of the amplification curve, rather than crossing at a set fluorescent signal threshold. Plotting the second derivative of the growth curve generates a peak that corresponds to a point near the baseline of

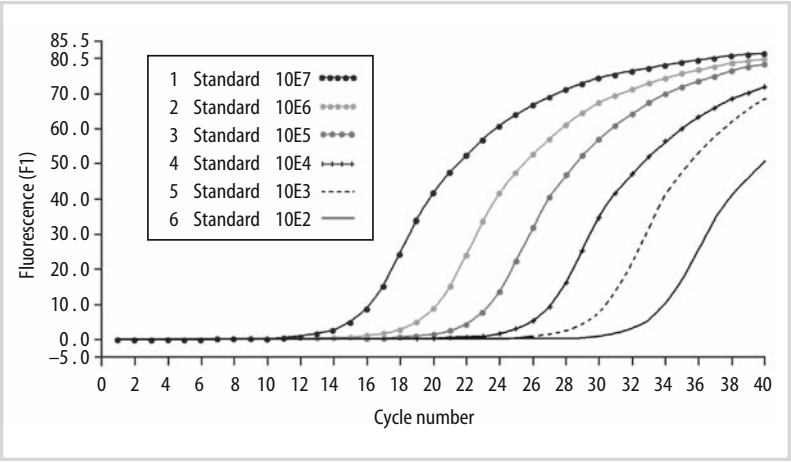


Figure 2-2. Real-time PCR curves for a 10-fold dilution series of a known standard. Data provided by Roche Molecular Diagnostics. (Used by permission. Data © 2003 Roche Diagnostics Corporation, all rights reserved.)

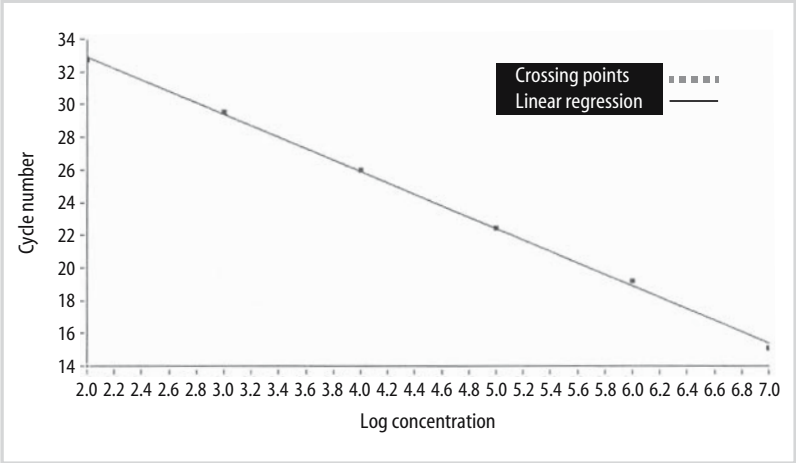


Figure 2-3. Standard curve generated by real-time PCR of a 10-fold dilution series of a known standard.

the growth curve (see Figure 2-4). The cycle at which this peak occurs is designated as the Ct or Cp. This analysis method can provide better run-to-run reproducibility than manually setting the Ct using the primary signal.

DNA Methylation and Methylation-Specific PCR

DNA methylation is a mechanism by which the cell regulates gene expression. Methylation is an enzyme-mediated modification that adds a methyl (-CH_3) group at a selected site on DNA or RNA. In humans, methylation occurs only at cytosine (C) bases followed by a guanosine (G), known as CpG dinucleotides. The CpG dinucleotides are prone to spontaneous mutations and have been selectively depleted from the mammalian genome. However, some regions of DNA have retained CpG dinucleotides and are referred to as CpG islands. The CpG islands are found primarily in the 5′ region of expressed genes, often in association with promoters. When the promoter CpG island is methylated, the corresponding gene is silenced and transcription does not occur. This is one method of silencing imprinted genes, as the transcription repression is passed on through cell division. Aberrant CpG island methylation of tumor-suppressor genes is frequent in cancer and appears to be an important mechanism of neoplastic transformation.

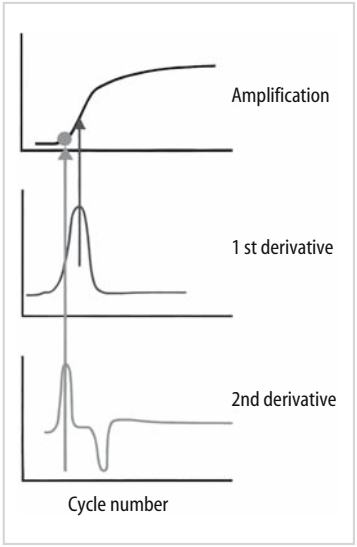


Figure 2-4. Graphical depiction of the second derivative maximum method used to identify the crossing point (Y axis is fluorescence value).

Methylated DNA can be distinguished from unmethylated DNA using sodium bisulfite treatment of DNA, which converts unmethylated C to uracil (U) but leaves methylated C intact.[31] This in vitro treatment can then be followed by one of several methods to distinguish C from U, including restriction endonuclease digestion with methylation-sensitive enzymes, sequencing, or methylation-specific PCR (MSP).[32] MSP of bisulfite-treated DNA uses primer pairs that specifically identify either methylated or unmethylated DNA. The primers are designed to hybridize to regions containing one to three CpG sites concentrated in the 3′ region of the primer to increase the specificity of amplification, and enough non-CpG cytosines to ensure that unmodified DNA is not amplified. Gel electrophoresis is used to detect the presence or absence of the amplicon in each of the two reactions, indicating the presence of unmethylated or methylated alleles or both. A novel modification is the use of quantitative MSP, which combines MSP with real-time PCR to distinguish the high-level CpG methylation in neoplasia from low-level methylation that can occur with aging or in nonneoplastic conditions such as metaplasia.[33]

Examples of Applications of Methylation-Specific PCR

1. Analysis of imprinted genes
2. Clonality assessment based on X chromosome inactivation
3. Abnormal methylation in neoplasia

PCR Variations for Unknown Sequence Variants

Most of the techniques discussed above are used to screen for sequence variants (both mutations and polymorphisms) based on previous knowledge of the variant; i.e., the sequence of the variant is either known or defined by the experimental conditions. By contrast, there is an evolving interest in both research and clinical molecular pathology to identify sequence variants by scanning without prior knowledge of their existence; i.e., the sequence of the variant is unknown. Sequencing is the ultimate screening technique, but is costly and labor-intensive. The goal of the scanning techniques described below (denaturing gradient gel electrophoresis [DDGE], and temperature gradient gel electrophoresis [TGGE], heteroduplex analysis [HA], single-strand conformation polymorphism [SSCP], denaturing high-performance liquid chromatography [DHPLC], and protein truncation test [PTT]) is to identify specimens with possible variant sequences, thereby reducing costs relative to sequencing. Should an unknown variant be detected, for example by a shift in the mobility of the PCR product on a gel or capillary, the PCR product with altered mobility is isolated and sequenced. Melting-temperature analysis in real-time PCR also can be used to identify unknown sequence variants.

Denaturing Gradient Gel Electrophoresis and Temperature Gradient Gel Electrophoresis

Denaturing gradient gel electrophoresis (DGGE)[34,35] and temperature gradient gel electrophoresis (TGGE)[36,37] are similar methods for separating DNA fragments with similar lengths but with different sequences according to their mobilities under a linear gradient of increasingly denaturing conditions. The gradient is created in DGGE with a mixture of urea and formamide, and in TGGE with a combination of water baths and a cooling plate under the gel. Both DGGE and TGGE take advantage of the markedly decreased mobility of partially melted dsDNA compared to either fully annealed dsDNA or ssDNA. Melting within a dsDNA fragment occurs within stretches of base pairs called melting domains. The point at which a domain begins to denature is referred to as the melting temperature (T_m), whether melting was induced by temperature or denaturing chemicals. In general, GC-rich sequences are more resistant to denaturation because of the three hydrogen bonds holding G and C together, as opposed to the two hydrogen bonds between A and T. During electrophoresis, once a dsDNA fragment reaches the point at which the melting domain with the lowest T_m begins to denature, mobility of the fragment through the gel nearly ceases. Fragments that melt early in the gel can therefore be separated from those that melt later. Complete denaturation of the dsDNA can be prevented by adding a GC-rich region to the 5′ end of one of the primers (GC clamp), increasing the sensitivity for detection of sequence variants.

In DGGE and TGGE, the denaturing conditions and the time of electrophoresis should be optimized such that normal sequences migrate to an intermediate position in the gel by the end of electrophoresis. This allows sequence variants creating either a higher or lower T_m to be identified. The denaturing gradient may be perpendicular or parallel to the electric field. Perpendicular gradient gels covering a broad range of denaturing conditions are loaded with normal sequence in all lanes to find the optimal, narrower denaturing gradient (chemical or temperature) for later use in parallel gradient gels. Parallel gradients are then used to run samples but also to optimize the time of electrophoresis by loading the normal sequence to different lanes at different times. Double-gradient DGGE adds a sieving gradient, for example, 6% to 12% polyacrylamide, colinear with the denaturing gradient in the gel matrix, further improving band resolution.

Both DGGE and TGGE work best with DNA fragments less than 500 bp in length. When GC-clamped fragments are analyzed, the sensitivity of detecting a SNP is close to 99%. Following electrophoresis, specific bands can be isolated from the gel and sequenced. DNA fragments with a high GC content are not easily analyzed by DGGE, since all fragments are harder to melt.

Examples of Applications of DGGE or TGGE

1. *APC* gene mutation analysis in familial adenomatous polyposis[38]
2. *CTFR* gene mutation analysis in cystic fibrosis[39]
3. *TCRG* gene rearrangements in lymphoma[40]

Heteroduplex Analysis

Heteroduplex formation results when wild-type and mutant alleles are coamplified, denatured, and allowed to reanneal in a post-PCR annealing step, usually heating and cooling.[41] Some of the strands pair with the complementary strand from the same allele and form homoduplexes. However, some strands pair with a strand from the other allele and form heteroduplexes. Because the heteroduplexes have mismatched base pairs between strands, they form a partially open dsDNA sequence that migrates more slowly in electrophoresis than the fully annealed homoduplexes.

Two types of heteroduplex structures can be formed.[42] When the mismatch consists of one or more single-base mutations, small open areas of dsDNA called "bubble-type" heteroduplexes are formed. When the mismatch is formed by insertions or deletions between the two alleles, a pronounced bending of the dsDNA is produced and referred to as a "bulge-type" heteroduplex. Bulge-type heteroduplexes markedly affect the mobility of the dsDNA, whereas

bubble-type heteroduplexes may be difficult to detect in polyacrylamide gels. Detection of single base-pair mismatches can be enhanced in two ways. Electrophoresis can be performed with mutation detection enhancement (MDE) gels, an altered form of polyacrylamide that enhances separation of heteroduplexes. The post-PCR introduction of a known sequence with a short deletion to form a bulge-type heteroduplex enhances the separation of sequences with base-pair mismatches in a process known as universal heteroduplex generation (UHG).

Examples of Applications of Heteroduplex Analysis

1. *HIV* subtyping
2. *CFTR* gene mutation analysis in cystic fibrosis
3. *NF1* gene mutation analysis in neurofibromatosis type 1

Single-Strand Conformation Polymorphism

The principle of single-strand conformation polymorphism (SSCP) is the differential gel separation of ssDNA that folds into a specific secondary structure based on its sequence.[43–45] For SSCP, the region of interest is amplified and the resulting amplicons are denatured using heat or a denaturation buffer, or both, prior to gel or capillary electrophoresis. Amplicons with different sequences will assume different folding conformations upon denaturation. Conformational differences reflecting sequence changes are detected as differences in electrophoretic mobility of the ssDNA in a nondenaturing polyacrylamide matrix. In general, a wild-type sample generates two bands, one for each of the two strands of the dsDNA product. Bands of mutant ssDNA migrate to positions different from those of the wild-type ssDNA. A homozygous mutant sample generates two bands, but with different migration patterns from the two wild-type bands. If a heterozygous mutant is present, four bands are generated: two with wild-type mobility and two with mutant mobility. Three also can be observed bands in heterozygous specimens if the mutation changes the conformation of only one strand but not the other. The bands with altered mobility can be isolated from wild-type bands in the gel, allowing even rare somatic mutations in tumors to be sequenced.

Temperature, ionic environment, and pH affect the conformation and therefore must be held constant throughout the SSCP run. Accurate temperature control during SSCP increases reliability and is an easily modifiable parameter in repeatable, nonisotopic experiments that may increase sensitivity. SSCP is adversely affected if unincorporated primers are allowed to bind to the ssDNA during denaturing and cooling prior to electrophoresis, or if nonspecific bands are produced by low-fidelity PCR. In SSCP, electrophoretic mobility patterns of variant alleles can be difficult to distinguish from wild type. Another disadvantage is that multiple experimental conditions are required for 100% sensitivity for detection of all sequence variants.

SSCP is most sensitive when the DNA amplicon is less than 200 bp in length. Sensitivity decreases as fragment length increases. This can be overcome by multiplexing differently sized fragments onto a single gel lane and by restriction enzyme digestion prior to electrophoresis. When restriction endonucleases are used, the procedure is referred to as restriction endonuclease fingerprinting–single-strand conformation polymorphism (REF-SSCP). Additionally, SSCP is relatively less sensitive for detecting G to C mutations. The addition of glycerol enhances mutation detection in this circumstance.

Variations of SSCP include RNA-SSCP (rSSCP), dideoxy fingerprinting (ddF), bidirectional ddF (bi-ddF), and SSCP detection of virtually all mutations (DOVAM-SSCP). RNA is more stable and adopts more conformational structures than does ssDNA, allowing enhanced detection using rSSCP. RNA-SSCP is not widely used because of the relative difficulty in producing RNA for analysis. Dideoxy fingerprinting involves a dideoxy Sanger single-primer termination reaction (cycle-sequencing reaction; for additional information on the Sanger reaction, see the section on sequencing, above) followed by nondenaturing electrophoresis. A fingerprint bandshift is indicative of sequence changes. Bidirectional ddF is an advancement of ddF whereby the dideoxy Sanger termination reaction is performed with two opposing primers in the same tube. SSCP detection of virtually all mutations is a recently described modification in which SSCP is performed under different conditions with different buffers and gel matrices that result in overall increased sensitivity for mutation identification.

The detection of an altered SSCP pattern does not identify the exact sequence variation present in the analyzed DNA. Therefore, positive SSCP results require DNA sequence analysis to confirm and identify sequence variation.

Examples of Applications of SSCP

1. Screening for mutations in the adenomatous polyposis coli (*APC*) gene
2. Mutation analysis of the *ATP7B* gene in Wilson disease
3. Mutation analysis in *BRCA1* (in familial breast cancer)
4. Pathogen identification[46]

Denaturing High-Performance Liquid Chromatography

Denaturing high-performance liquid chromatography (DHPLC) is an ion-paired, reversed-phase, liquid chromatography method used to identify mutations, including SNPs and small insertions or deletions, through its ability to separate heteroduplex DNA from homoduplex DNA[47]. DHPLC is conceptually similar to heteroduplex analysis (HA; see above). Conventional HA makes use of a gel matrix to separate homo- and heteroduplex species in a nondenaturing environment, whereas DHPLC uses partially

denaturing conditions in a liquid chromatography column to exaggerate the separation between the two species.

The gene to be studied is first amplified using PCR. High-fidelity PCR is used to prevent the production of PCR artifacts (pseudoalleles) that could produce false-positive results. Optimal amplicon length is between 100 and 500 bp. Purifying the PCR product is usually not necessary, as unincorporated primers, nucleotides, and genomic DNA do not interfere with the analysis. DHPLC requires heteroduplex formation, accomplished by heating and slow cooling. Therefore, for conditions in which only one variant allele type may be present (such as recessive diseases, X-linked conditions in males, or small tumor samples with loss of heterozygosity in all cells), PCR products from normal and patient samples are mixed in equal proportions before heating and cooling to produce heteroduplex DNA and distinguish from homozygous wild-type alleles. The addition of normal PCR amplicons is not required when using DHPLC to test PCR products from heterozygous individuals, which naturally form heteroduplexes when denatured and slowly cooled.

The duplexes are injected into a DHPLC column, and the DNA binds to the stationary matrix. Binding is aided by triethylammonium acetate (TEAA). Because the stability of the binding depends on the temperature, the column is optimally held at the melting temperature of the PCR fragment. The melting temperature can be calculated using a variety of proprietary or free software programs. The DNA is next eluted using acetonitrile, an organic solution that facilitates the subsequent separation of the DNA from the column matrix, and DNA absorbancy is measured at 260 µm. The linear gradient of acetonitrile established in the column allows separation of DNA fragments based on size or the presence of heteroduplexes, or both. All DNA fragments impart a characteristic profile when the absorbance is plotted against elution time. The peak of maximum absorbance is the retention time of that DNA sample at a given acetonitrile concentration. Heteroduplexes are less stable and thus have a lower affinity for the column. The concentration of acetonitrile required to separate heteroduplexes from the column is therefore lower, so heteroduplexes elute from the column earlier than homoduplexes.

The column temperature and gradient conditions can be optimized for the separation of any heteroduplex-homoduplex mixture. Some DNA fragments have more than one melting domain and the analysis may be carried out at more than one temperature. One advantage of DHPLC is that reinjection of the same sample at different temperatures is possible. Other advantages include high detection rates of mutations, rapid separation times per sample, a high degree of automation, and the ability to collect elution fractions and sequence each eluted fragment. Disadvantages of DHPLC include the need for expensive equipment and columns, high-fidelity PCR, and optimization of each reaction required to achieve the highest sensitivity of mutation detection.

Examples of Applications of DHPLC

1. *RET* and *CFTR* mutation detection[48]
2. *BRCA1* and *BRCA2* mutation analysis[49]

Protein Truncation Test

The protein truncation test (PTT) is used to identify mutations that result in premature termination of protein translation. PTT detects mutations at the protein rather than the genomic DNA or RNA levels. Initially developed for Duchenne muscular dystrophy (DMD) testing, PTT now is used more widely,[50] since protein-truncating mutations are associated with multiple types of hereditary cancer syndromes, including breast, ovarian, and colon.

For PTT, the gene segment of interest is amplified by PCR. The amplicons are used for in vitro transcription and translation in a coupled reaction. The resulting proteins are separated using sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis. The presence of a premature termination codon is indicated by the visualization of a lower-molecular-weight protein band than for the wild-type protein. Relatively large gene fragments (2–4 kb) can be analyzed using PTT.

Examples of Applications of PTT

1. Duchenne muscular dystrophy diagnosis
2. *BRCA1* and *BRCA2* mutation detection[51]
3. *APC* mutations in colorectal cancer[52]

Other Forms of In Vitro Nucleic Acid Amplification

PCR is widely used in the clinical laboratory. The proprietary nature of PCR prompted other in vitro diagnostic companies to develop alternative methods of in vitro nucleic acid amplification. Several are described here: ligase chain reaction (LCR); transcription mediated amplification (TMA), strand displacement amplification (SDA); and nucleic acid sequence–based amplification (NASBA).

Ligase Chain Reaction

Ligase chain reaction (LCR)[53,54] is initiated when a mixture of target DNA, thermostable DNA ligase, four oligonucleotide probes, and NAD^+ or ATP is heated to denature dsDNA (both target and complementary probes) in the reaction mixture. Two pairs of complementary probes are used, and, of necessity, their correct design demands a priori knowledge of the sequence of the DNA target. After denaturation and subsequent reaction cooling, the four probes present in the reaction mixture hybridize to their

complementary sequences on each target DNA sister strand. The two probes that hybridize to one sister strand and the two probes that bind to the other sister strand are designed such that when hybridized, the 3′ hydroxyl end of the upstream probe is immediately adjacent to the 5′ phosphate end of the downstream probe. Thermostable DNA ligase enzymatically ligates the two bound probes, thus achieving a "doubling" of the mass of target DNA in the reaction. As the temperature cycling proceeds, a theoretical exponential amplification of the mass of target DNA in the original reaction occurs because the resultant ligated amplicons also serve as targets for probe hybridizations. In practice, amplification is less than exponential, but sufficient to achieve target DNA identification by various detection methods.

There is a tendency for target-independent blunt-end ligation of the probes in the reaction to occur in LCR, which can cause unacceptably high levels of background signal, limiting the assay's sensitivity and specificity. This problem has been solved by use of gap LCR (G-LCR). In G-LCR, the probes are designed such that they cannot be ligated in a target-independent manner because they are not blunt ended. When G-LCR probes hybridize to target DNA, a gap of one or more bases exists between the probes hybridized to the same target strand. This gap is then biochemically "filled" in vitro, thus providing a suitable substrate for DNA ligase, which then performs target-dependent ligation.

Examples of Applications of LCR

1. *Chlamydia trachomatis* detection
2. *Neisseria gonorrhoeae* detection

Transcription-Mediated Amplification

Transcription-mediated amplification (TMA) uses RNA as the template, two primers and two enzymes: reverse transcriptase and RNA polymerase. One primer contains a promoter sequence that binds RNA polymerase. As the amplification process begins, the promoter-containing primer hybridizes to the target RNA at a complementary site (Figure 2-5). Reverse transcriptase then synthesizes a cDNA copy of the target RNA template by extension of the 3′ end of the promoter-primer. The result is an RNA : DNA duplex. The RNA component is degraded by the activity of the enzyme RNase H inherent in reverse transcriptase. The other primer in the reaction mixture hybridizes to the DNA

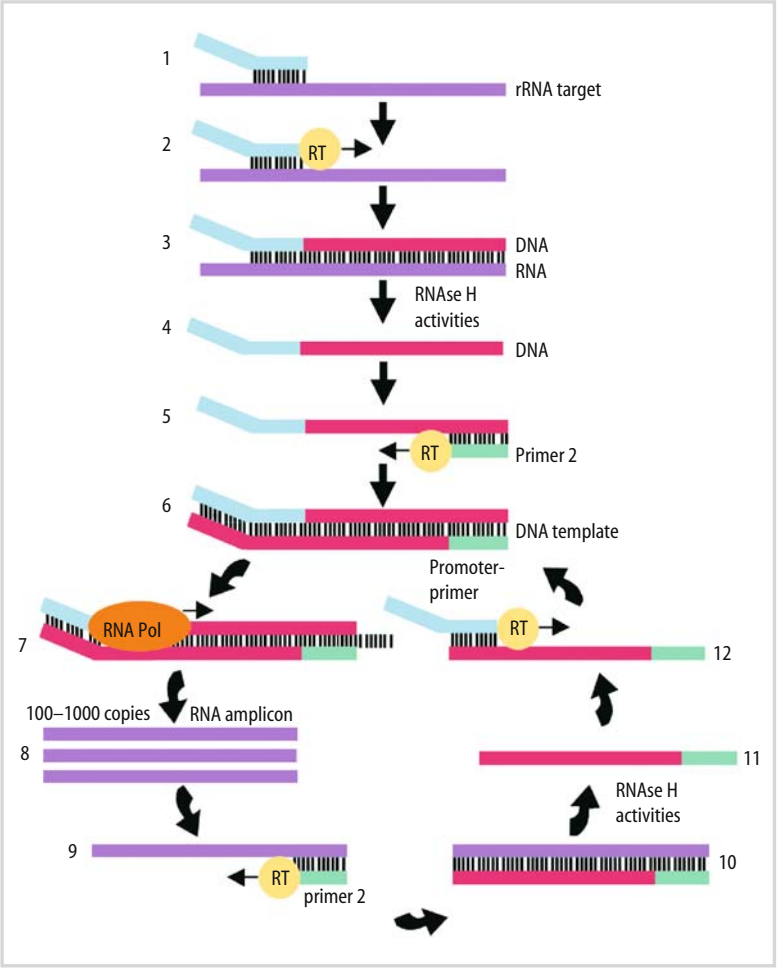


Figure 2-5. Transcription-mediated amplification cycle (TMA):
Step 1. Promoter-primer binds to rRNA target.
Step 2. Reverse transcriptase (RT) creates DNA copy of rRNA target.
Step 3. RNA:DNA duplex.
Step 4. RNAse H activities of RT degrades the rRNA.
Step 5. Primer 2 binds to the DNA and RT creates a new DNA copy.
Step 6. Double-stranded DNA template with a promoter sequence.
Step 7. RNA polymerase (RNA Pol) initiates transcription of RNA from DNA template.
Step 8. 100 to 1000 copies of RNA amplicon are produced.
Step 9. Primer 2 binds to each RNA amplicon and RT creates a DNA copy.
Step 10. RNA:DNA duplex.
Step 11. RNAse H activities of RT degrades the rRNA.
Step 12. Promoter-primer binds to the newly synthesized DNA. RT creates a double-stranded DNA and the autocatalytic cycle repeats, resulting in a billion-fold amplification.
(Reprinted with permission from Gen-Probe.)

copy, and a new DNA strand is synthesized from the end of the primer by RT, generating a dsDNA molecule. The other enzyme in the mixture, RNA polymerase, binds the promoter sequence in the DNA template and initiates transcription. Each of the resulting newly synthesized RNA amplicons reenters the TMA cycle, serving as a template for a new round of replication and exponential expansion of the RNA target. Each DNA template can generate 10^2 to 10^3 copies of RNA amplicon, with the potential for 10^8- to 10^9-fold amplification in less than 1 hour. The process is autocatalytic and isothermal. Acridinium ester–labeled DNA probes are added on completion of the reaction to initiate detection and quantitation based on chemiluminescence.

Examples of Applications of TMA

1. *Chlamydia trachomatis* detection[55]
2. *Neisseria gonorrhoeae* detection
3. HCV detection (qualitative)[56]

Strand Displacement Amplification

Strand displacement amplification (SDA) is an isothermal in vitro nucleic acid amplification technique.[57] Hemimodified DNA is polymerized by using three conventional dNTPs and one containing a 5′-[alpha-thio]triphosphate. The primer(s) is designed with an RE recognition site in the 5′ overhang end. The recognition site is specific for an RE that can nick the unmodified DNA strand at a double-stranded hemiphosphorothioate recognition site, that is, when the newly synthesized strand incorporates the 5′-[alpha-thio]triphosphate nucleotide in the recognition sequence. DNA polymerase lacking 5′ to 3′ exonuclease activity is used to extend the 3′ end at the nick and displace the downstream strand. Nicking and polymerization with re-formation of the hemiphosphorothioate recognition site continuously cycle, generating complementary copies of the DNA target. Linear amplification (called target-generation SDA) occurs when a single primer is used. Exponential amplification (exponential SDA) is achieved by using two primers complementary to opposite DNA strands, with both primers containing RE recognition sites in the 5′ overhang end. Strand displacement amplification has been used in a microarray format,[58] which may become a clinically useful method to combine amplification of low copy number targets with the multiple features of microarray in a format that can be automated.

Examples of Applications of SDA

1. *Chlamydia trachomatis* detection
2. *Neisseria gonorrhoeae* detection

Nucleic Acid Sequence–Based Amplification

Nucleic acid sequence–based amplification (NASBA) is an isothermal method for amplifying nucleic acids using two sequence-specific primers (P1, antisense, and P2, sense; see Figure 2-6), and the coordinated activities of three

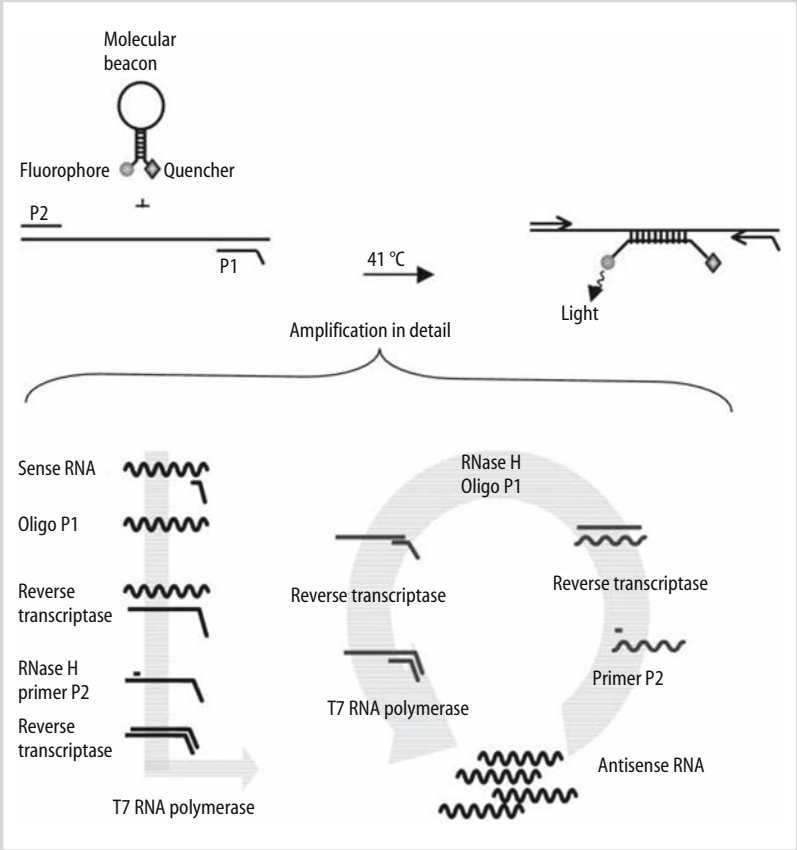


Figure 2-6. Schematic of NASBA reaction. Two primers are used: P1 (antisense) and P2 (sense). The P1 overhang is a promoter sequence for T7 RNA polymerase. A molecular beacon (with fluorophore and quencher, by definition) serving as probe with reporter molecules coupled to NASBA generates a real-time detection system.

enzymes, avian myeloblastosis virus reverse transcriptase (AMV-RT), RNase H, and T7 RNA polymerase.[59] A primer with a T7 RNA polymerase recognition sequence at its 5′ end is used by AMV-RT to transcribe cDNA from the RNA template. The RNA template is destroyed by RNase H. AMV-RT then uses the second primer to synthesize dsDNA. The T7 RNA polymerase uses the dsDNA to synthesize multiple antisense RNA transcripts, and the cycle is repeated. Generally, amplification is approximately 10^{12}-fold in 1 to 2 hours.

In NASBA, nucleic acids serve as amplification templates only if they are single stranded and contain primer-binding regions. NASBA is performed isothermally at 41°C, so RNA is preferentially amplified, because at this temperature genomic DNA remains double stranded and does not bind primers. It is therefore possible to detect RNA in a genomic DNA background without getting false-positive results. Specific DNA amplification using NASBA may be done by introducing a denaturation step before amplification.

Quantitative detection of target nucleic acids is achieved by use of an internal calibrator added at RNA isolation. The calibrator is included at a known concentration, is coamplified during the NASBA process, and is subsequently identified along with the target RNA. Quantitation is based on the analysis of signals generated in real time (one color for calibrator and another for target).

Examples of Applications of NASBA

1. HIV quantitation
2. CMV detection

Amplicon Carryover Contamination

Vast numbers of target DNA copies are generated when PCR and other in vitro nucleic acid amplification techniques are used. By contrast, signal amplification methods (see below) do not generate vast quantities of amplicon and so do not create the potential for amplicon carryover contamination of the laboratory workspace. Amplicons from previous reactions inadvertently introduced into new amplification reactions for the same amplicon are suitable substrates for amplification. Clinical molecular laboratories must therefore take precautions to prevent generation of false-positive results from amplicon carryover contamination.

Amplicon contamination and false-positive results are prevented by using physical barriers and chemical and ultraviolet techniques to destroy amplicons or make them unsuitable for amplification. The physical barriers include large-scale separation of nucleic acid isolation, PCR setup, thermal cycling, and post-PCR analysis in separate areas of the laboratory (different rooms). Air flow is controlled such that air pressure is positive, that is, flows out of the room, in the isolation and PCR setup rooms and is negative in the thermal cycling and post-PCR analysis rooms. Hoods are another way of providing physical separation of the different PCR steps. Small-scale physical separation techniques include the use of barrier pipette tips, frequent glove changes, designated lab coats that do not leave the pre- or post-PCR areas of the laboratory, and PCR tube openers or careful, slow opening of tubes to prevent aerosolization of contents. Real-time PCR reduces the chances of amplicon contamination since the PCR product can be detected and quantified without opening the real-time PCR reaction vessel after PCR.

Chemical techniques include thorough cleansing with bleach of work areas and instruments before and after use. Ultraviolet lights are frequently placed in hoods and work areas. Ultraviolet light creates thymine dimers within amplicons, rendering the amplicons unsuitable as substrates for further amplification. The introduction of isopsoralens in PCR reactions allows DNA cross-linking of amplicons by UV light, also rendering them unsuitable for further amplification. Deoxyuridine may be used in lieu of thymidine in the reaction mixture. Use of deoxyuridine has minimal effect on amplification or product detection, but amplicons with uracil are substrates for uracil-N-glycosylase (UNG). UNG has no effect on DNA that contains only thymidine residues (new patient DNA in subsequent reactions) but digests the uracil-containing amplicons, allowing removal of amplicons before PCR proceeds.[62] So-called UNG sterilization may therefore be performed prior to PCR to rid the reaction of any amplicon contaminants that may be present.

Signal Amplification Methods

Branched DNA Method

The branched DNA method (bDNA)[63,64] is carried out in microwells and begins with the addition of a lysis buffer to a small volume of serum, plasma, or culture supernatant containing cells or virus. The lysis reagent contains detergent to release target nucleic acid, inhibitors to prevent target degradation, and multiple capture extenders (oligonucleotides) that hybridize to specific areas of the target RNA or DNA. In the case of the HIV bDNA assay, the capture extenders hybridize to multiple sequences in the *pol* gene. A common sequence on the capture extenders interacts with capture probes immobilized on the surface of 96-microwell plates, thereby anchoring the target nucleic acid to the plate.

Multiple target probes are added that hybridize to different, conserved sequences on the target RNA or DNA. In the HIV bDNA assay, more than 80 target probes covering a large portion of the 3000 bp of the *pol* gene are used. The target probes contain key sequences that form the foundation for signal amplification, accomplished via the sequential addition of preamplifier (complementary to a region of the target probes), amplifier (complementary to a region of the preamplifier molecule), and alkaline-phosphatase-modified label probes (complementary to portions of the amplifier molecule).

Preamplifier, amplifier, and label probes, as well as the preamplifier region of the binding probes, contain the non-natural nucleotides 5-methyl-2′-deoxyisocytidine (iso^{Me}C) and 2′-deoxyisoguanosine (isoG). These isomers of natural bases can participate in Watson-Crick base pairing with each other but not with cytosine or guanine residues in probes or in DNA or RNA sequences. Incorporation of non-natural bases into the synthetic probe molecules increases the specificity of hybridization by decreasing nonspecific probe interactions, and increases the sensitivity of the assay since higher concentrations of probes can be used.

The series of probes results in formation of large hybridization complexes on the target RNA or DNA. For example, if each hybridization step was 100% efficient in the HIV bDNA assay, each target molecule would be labeled with more than 10,000 alkaline phosphatase molecules. Addition of dioxetane substrate for the alkaline phosphatase results in steady-state chemiluminescence. The luminescent signal is proportional to the amount of target RNA or DNA present in the sample. The amount of target RNA or DNA in a specimen may be calculated by interpolation from a standard curve generated by signals produced from calibrators that contain known concentrations of the specific viral, bacterial, or cellular RNA or DNA. A schematic of this technology is shown in Figure 2-7.

Examples of Applications of bDNA method

1. HIV quantitation[65,66]
2. Hepatitis B virus (HBV) quantitation[67–70]
3. HCV quantitation[71,72]

Hybrid Capture

Hybrid capture (HC) is a signal amplification system based on antibody binding of RNA:DNA hybrids. The first step of HC is the hybridization of unlabeled RNA probes with denatured target DNA. The resulting RNA:DNA hybrids are captured to the surface of microplate wells by an immobilized antibody that recognizes RNA:DNA hybrids. A second anti-RNA:DNA monoclonal antibody conjugated to alkaline phosphatase then is added. Many secondary antibodies bind to each RNA:DNA hybrid. Dioxetane-based substrate is added, which is cleaved by the alkaline phosphatase, producing a chemiluminescent signal. The signal is measured using a microplate luminometer and is proportional to the amount of target DNA in the specimen.[73] The sensitivity of HC can be increased by using a precipitation buffer and high-speed centrifugation.

Examples of Applications of HC

1. Human papilloma virus detection
2. CMV quantitation[74]
3. *Chlamydia trachomatis* and *Neisseria gonorrhoeae* detection[75]
4. HBV quantitation[76]
5. HSV detection[77]

Fluorescence In Situ Hybridization

Fluorescence in situ hybridization (FISH) uses fluorescently tagged DNA or RNA probes to identify genomic

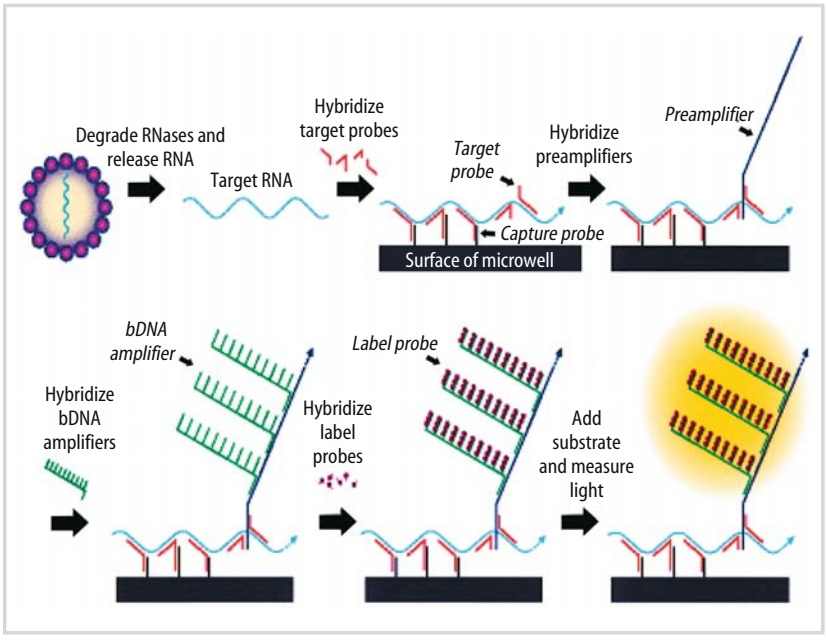


Figure 2-7. Diagram of branched DNA (bDNA) method. (Reprinted with permission of Bayer Healthcare.)

sequences of interest.[78,79] The major advantages of FISH are the ability to utilize FFPE tissue sections, allowing correlation of probe hybridization with tissue morphology, and the increased resolution provided by FISH for identification of specific abnormalities when partnered with conventional cytogenetics. The number and location of the fluorescent signal(s) can identify chromosomal abnormalities including gene amplification, gene deletion, or structural rearrangements such as translocations.

FISH is similar to Southern blot in that sequential steps of denaturation, hybridization, and washing are involved. Slides are prepared in the cytogenetic or histology laboratory. Probe(s) are then applied to the slide along with a nuclear counterstain and reagents to enhance denaturation and reduce background. The slides are sealed and incubated (usually overnight) in a humid environment at high temperature. These conditions denature the probe and patient DNA, allowing hybridization to occur between the probe and its complementary DNA sequence without binding to nonspecific sites. Excess nonspecifically bound probe is washed away, and the pattern of fluorescence is read by fluorescence microscopy. The fluorescent signal(s) can be enhanced by the use of a digital imaging system and computer software.

The specificity of FISH is largely based on the selection of the probe. Ideally, the probe is complementary to the gene of interest; however, if the disease gene is unknown, satellite probes that identify a chromosomal region linked to the disease may be used. Labeled bacterial or yeast artificial chromosomes (BAC or YAC, respectively) are typically used as FISH probes, but short oligonucleotides also can be used with signal amplification techniques. Probes that identify individual whole chromosomes or chromosomal arms are often called "painting probes" due to the colorful patterns they generate. Other probes that hybridize to a specific gene can be used for the detection of deletions or duplications and are called single-copy gene probes. Probes that hybridize to the alpha-satellite regions near centromeres are used in clinical cytogenetics to identify and count individual chromosomes. Probes that hybridize to the subtelomeric portions of chromosomes are used to identify cryptic telomeric abnormalities such as translocations.

Dual-color FISH (dFISH) employs two probes with different fluorescence wavelengths to identify structural chromosomal rearrangements. Each probe generates a characteristic color by itself (split signal) and a third color when the two probes are juxtaposed (fusion signal). A fusion signal indicating chromosomal rearrangement is used to identify disease-causing mutations that predictably involve only two partner genes, for example, *BCR* and *ABL*. However, when a particular gene with multiple potential translocation partners is tested, it is more efficient to have both probes bind to the 5′ and 3′ ends of the particular gene such that the normal allele shows the fusion signal and the rearranged allele shows two split signals. This is the technique used for the *MLL* gene, which is rearranged to over 30 different partner genes in various types of leukemia.

Spectral karyotyping (SKY) and multiplex FISH (M-FISH) are relatively new advancements of conventional FISH that utilize multiple fluorochromes, specialized optics, and image analysis that can simultaneously identify all chromosomes.[80] Comparative genomic hybridization (CGH), also called "copy number karyotyping," is a variation of FISH that detects relative gains or losses of the genome.[81] This method compares the ratios of patient specimen DNA, labeled with one fluorochrome, to that of normal DNA, labeled with a different fluorochrome when hybridized to control chromosomes. CGH is used predominantly in research to identify possible pathways involved in tumor progression, recurrence, or metastasis. Colorimetric probes are used in chromogenic in situ hybridization (CISH), which has the advantage that the signal does not fade with time or require fluorescence microscopy for analysis.

Examples of Applications of FISH

1. Detection of *BCR/ABL* in chronic myelogenous leukemia
2. Detection of *HER2* gene amplification for breast cancer diagnosis and prognosis
3. Detection of *MYCN* amplification in neuroblastoma[82]

DNA Arrays and Chips

Arrays are a relatively recent phenomenon in the field of molecular pathology.[83] The term "array" is jargon for an orderly collection of molecules on solid supports ranging from nylon membranes to printed circuit board to glass slides to silicon surfaces. There are macroscopic arrays, for example, reverse line blots on nylon membranes, and microarrays, for example, DNA chips. Synonyms for microarrays include gene chip, DNA chip, genome chip, biochip, gene array, DNA array, and DNA microarray.

Though most do not think of line probe assays as belonging in a section describing DNA chips, they are indeed arrays. Line probe assays are being used to determine mutations in the HIV genome, to type HCV, and to detect *CFTR* mutations for cystic fibrosis.[84] The method is based on classic nucleic acid hybridization using nylon or nitrocellulose strips as a solid support matrix for hybridization and detection. Generally, reverse hybridization is employed. Oligonucleotide probes specific for mutations or polymorphisms of interest are bound to the membrane strip in a parallel line format. Patient specimen DNA is amplified with biotinylated primers, thereby labeling the amplicon, which is hybridized to probes on the membrane. After hybridization, alkaline phosphatase-labeled streptavidin is added and binds to the biotin of the bound amplicons. Detection is colorimetric following incubation with a specific chromogen, generating a purple-

brown precipitate. The resulting colored lines are read as positive results, hence the name of the assay.

DNA chips, such as those produced by Affymetrix, Nanogen, Clinical Micro Sensors, and other companies, may be thought of as miniaturized ASOH platforms with multiple oligonucleotide probe assays. In one model, the oligonucleotides are synthesized directly onto a solid support, for example, glass or silicon, that is, the "chip." Some companies' chips can hold thousands of unique sequences. The sample DNA is labeled with fluorescent dye, denatured, and hybridized with the oligonucleotides on the chip. The chip is then scanned and the fluorescence patterns are measured with the aid of computer analysis to determine the presence of mutant or wild-type sequences. Gene expression chips use probes that bind to many different cDNA species (up to 33,000 in one array at the time of this writing) to quantitate the level of each cDNA isolated in a sample. Gene expression microarrays with large densities are appropriate for analysis of differential patterns of gene expression between normal and diseased tissues.

Clinical molecular laboratory diagnostic applications for DNA chips currently are limited to those with clinically appropriate density, in the range of 10 to 200. Utility for these platforms is found in SNP and mutation detection[85] as well as pathogen identification. Chips with densities in the thousands have limited, if any, use in clinical laboratories at this time. So much data are generated that more analysis needs to be done to winnow down the numbers to a relevant few genes or transcripts that can be used in clinical assays. This work is progressing well. Furthermore, the informatics software currently being used to analyze these data requires simplification and modification so that useful algorithms for data analysis can be employed for disease diagnosis or prognosis. While these are certainly considerable challenges, it should be noted that DNA array chip technology benefits enormously from robotic manufacturing techniques and rapidly advancing computer software. This has greatly decreased the costs per array feature of DNA chips, giving DNA microarrays enormous potential in the clinical molecular laboratory of the near future.

Examples of Applications of Arrays

1. Line probe assay for HCV genotyping
2. DNA chips for detection of polymorphisms in cytochrome P450 genes for pharmacogenetics

Acknowledgments

The authors thank Zed Penn, Cepheid; Tiffany Murphy, Beckman Coulter; H. James Hnatyszyn, Bayer; Paul van de Wiel and Mike Cronin, bioMerieux; Steve Geiner and Irma Alfaro, Digene; and Andrew Olson and Carl Hilliker, Roche, for their help with different sections of this chapter.

References

1. Kessler HH, Muhlbauer G, Stelzl E, Daghofer E, Santner BI, Marth E. Fully automated nucleic acid extraction: MagNA Pure LC. *Clin Chem.* 2001;47:1124–1126.
2. Fiebelkorn KR, Lee BG, Hill CE, Caliendo AM, Nolte FS. Clinical evaluation of an automated nucleic acid isolation system. *Clin Chem.* 2002;48:1613–1615.
3. Williams SM, Meadows CA, Lyon E. Automated DNA extraction for real-time PCR. *Clin Chem.* 2002;48:1629–1630.
4. Davies J, Reznikoff WS, eds. *Milestones in Biotechnology: Classic Papers on Genetic Engineering*. Boston: Butterworth-Heinemann; 1992.
5. Wasserman LM. A nested reverse-transcriptase-polymerase chain reaction assay to detect *BCR/ABL*. In: Killeen AA, ed. *Methods in Molecular Medicine Molecular Pathology Protocols.* Totowa, NJ: Humana Press; 2001:105–114.
6. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977;74:5463–5467.
7. Esch RK. Basic nucleic acid procedures. In: Coleman WB, Tsongalis GJ, eds. *Molecular Diagnostics for the Clinical Laboratorian.* Totowa, NJ: Humana Press; 1997:55–58.
8. Schmaizing D, Koutny L, Salas-Solano O, Adourian A, Matsudaira P, Ehrlich D. Recent developments in DNA sequencing by capillary and microdevice electrophoresis [review]. *Electrophoresis.* 1999;20:3066–3077.
9. Farkas DH. Specimen procurement, processing, tracking, and testing by the Southern blot. In: Farkas DH, ed. *Molecular Biology and Pathology: A Guidebook for Quality Control.* San Diego: Academic Press; 1993:51–75.
10. Farkas DH. Quality control of the B/T cell gene rearrangement test. In: Farkas DH, ed. *Molecular Biology and Pathology: A Guidebook for Quality Control.* San Diego: Academic Press; 1993:77–101.
11. Mullis K, Faloona F, Schart S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* 1986;51:263–273.
12. Farkas DH. Thermal cyclers. In: *Laboratory Instrument Evaluation Verification and Maintenance Manual.* Northfield, IL: College of American Pathologists; 1998:130–133.
13. Loeffelholz MJ, Lewinski CA, Silver SR, et al. Detection of *Chlamydia trachomatis* in endocervical specimens by polymerase chain reaction. *J Clin Microbiol.* 1992;30:2847–2851.
14. DiDomenico N, Link H, Knobel R, et al. COBAS AMPLICOR™: fully automated RNA and DNA amplification and detection system for routine diagnostic PCR. *Clin Chem.* 1996;42:1915–1923.
15. Greenfield L, White TJ. Sample preparation methods. In: Persing DH, Smith TF, Tenover FC, White TJ, eds. *Diagnostics Molecular Microbiology Principles and Applications.* Washington, DC: American Society for Microbiology; 1993:126–127.
16. Liu XY, Nelson D, Grant C, Morthland V, Goodnight SH, Press RD. Molecular detection of a common mutation in coagulation factor V causing thrombosis via hereditary resistance to activated protein C. *Diagn Mol Pathol.* 1995;3:191–197.
17. Linderman R, Hu SP, Volpato F, Trent RJ. Polymerase chain reaction mutagenesis enabling rapid nonradioactive detection of common β-thalassaemia mutations in Mediterraneans. *Br J Haematol.* 1991;78:100.
18. Sorscher EJ, Huang Z. Diagnosis of genetic disease by primer-specified restriction map modification, with application to cystic fibrosis and retinitis pigmentosa. *Lancet.* 1991;11:1115–1118.
19. Khanna M, Park P, Zirvi M, et al. Multiplex PCR/LDR for detection of K-*ras* mutations in primary colon tumors. *Oncogene.* 1999;18:27–38.
20. McMillin DE, Muldrow LL, Laggette SJ. Simultaneous detection of toxin A and toxin B genetic determinants of *Clostridium difficile* using the multiplex polymerase chain reaction. *Can J Microbiol.* 1992;38:81–83.

21. Sugita T, Nakajima M, Ikeda R, Niki Y, Matsushima T, Shinoda T. A nested PCR assay to detect DNA in sera for the diagnosis of deep-seated trichosporonosis. *Microbiol Immunol.* 2001;45:143–148.
22. Ferrie RM, Schwarz MJ, Robertson NH, et al. Development, multiplexing, and application of ARMS tests for common mutations in the CFTR gene. *Am J Hum Genet.* 1992;51:251–262.
23. Bugawan TL, Begovich AB, Erlich HA. Rapid HLA-DPB typing using enzymatically amplified DNA and nonradioactive sequence-specific oligonucleotide probes. *Immunogenetics.* 1990;32:231–241.
24. Nickerson DA, Kaiser R, Lappin S, Stewart J, Hood L, Landegren U. Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. *Proc Natl Acad Sci U S A.* 1990;87:8923–8927.
25. Jarvius J, Nilsson M, Landegren U. Oligonucleotide ligation assay. *Methods Mol Biol.* 2003;212:215–228.
26. Howell WM, Jobs M, Gyllensten U, Brookes AJ. Dynamic allele-specific hybridization. *Nat Biotechnol.* 1999;17:87–88.
27. Baltimore D. Viral RNA-dependent DNA polymerase. *Nature.* 1970;226:1209–1211.
28. Bustin SA. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol.* 2000;25:169–193.
29. Lay MJ, Wittwer CT. Real-time fluorescence genotyping of factor V Leiden during rapid-cycle PCR. *Clin Chem.* 1997;43:2262–2267.
30. Bernard PS, Wittwer CT. Real-time PCR technology for cancer diagnostics. *Clin Chem.* 2002;48:1178–1185.
31. Olek A, Oswald J, Walter JAA. A modified and improved method of bisulfite based cytosine methylation analysis. *Nucleic Acids Res.* 1996;24:5064–5066.
32. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB. Methylation specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A.* 1996;93:9821–9826.
33. Lo Y, Wong I, Zhang J, Tein M, Ng M, Hjelm N. Quantitative analysis of aberrant p16 methylation using real-time quantitative methylation-specific polymerase chain reaction. *Cancer Res.* 1999;59:3899–3903.
34. Fodde R, Losekoot M. Mutation detection by denaturing gradient gel electrophoresis (DGGE). *Hum Mutat.* 1994;3:83–94.
35. Lerman LS, Beldjord C. Comprehensive mutation detection with denaturing gradient gel electrophoresis. In: Cotton RGH, Edkins E, Forrest S, eds. *Mutation Detection: A Practical Approach.* Oxford: Oxford University Press; 1998:35–59.
36. Rosenbaum V, Riesner D. Temperature-gradient gel electrophoresis: thermodynamic analysis of nucleic acids and proteins in purified form and in cellular extracts. *Biophys Chem.* 1987;26:235–246.
37. Riesner D, Henco K, Steger G. Temperature-gradient gel electrophoresis: a method for the analysis of conformational transitions and mutations in nucleic acids and proteins. *Adv Electrophoresis.* 1991;4:169–250.
38. van der Luijt RB, Khan PM, Vasen HF, et al. Molecular analysis of the APC gene in 105 Dutch kindreds with familial adenomatous polyposis: 67 germline mutations identified by DGGE, PTT, and Southern analysis. *Hum Mutat.* 1997;9:7–16.
39. De Braekeleer M, Mari C, Verlingue C, et al. Complete identification of cystic fibrosis transmembrane conductance regulator mutations in the CF population of Saguenay Lac-Saint-Jean (Quebec, Canada). *Clin Genet.* 1998;53:44–46.
40. Alkan S, Cosar E, Ergin M, Hsi E. Detection of T-cell receptor-gamma gene rearrangement in lymphoproliferative disorders by temperature gradient gel electrophoresis. *Arch Pathol Lab Med.* 2001;125:202–207.
41. Nagamine CM, Chan K, Lau YF. A PCR artifact: generation of heteroduplexes. *Am J Hum Genet.* 1989;45:337–339.
42. Bhattacharyya A, Lilley DM. The contrasting structures of mismatched DNA sequences containing looped-out bases (bulges) and multiple mismatches (bubbles). *Nucleic Acids Res.* 1989;17:6821–6840.
43. Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci U S A.* 1989;86:2766–2770.
44. Hayashi, K. PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods Appl.* 1991;1:34–38.
45. Liu Q, Feng J, Buzin C, et al. Detection of virtually all mutations-single stranded conformational polymorphisms (DOVAM-S): a rapid method for mutation scanning with virtually 100% sensitivity. *Biotechniques.* 1999;26:932–942.
46. Widjojoatmodjo MC, Fluit AC, Verhoef J. Molecular identification of bacteria by fluorescence-based PCR-single-strand-conformation polymorphism of the 16S rRNA gene. *J Clin Microbiol.* 1995;33:2601–2606.
47. Liu W, Smith DI, Rechtzigel KJ, Thibodeau SN, James CD. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. *Nucleic Acids Res.* 1998;26:1396–1400.
48. O'Donovan MC, Oefner PJ, Roberts SC, et al. Blind analysis of denaturing high-performance liquid chromatography as a tool for mutation detection. *Genomics.* 1998;52:44–49.
49. Wagner T, Stoppa-Lyonnet D, Fleischmann E, et al. Denaturing high-performance liquid chromatography detects reliably BRCA1 and BRCA2 mutations. *Genomics.* 1999;62:369–376.
50. Roest PA, Roberts RG, Sugino S, van Ommen GJ, den Dunnen JT. Protein truncation test (PTT) for rapid detection of translation-terminating mutations. *Hum Mol Genet.* 1993;2:1719–1721.
51. Pohlreich P, Stribrna J, Kleibl Z, et al. Mutations of the *BRCA1* gene in hereditary breast and ovarian cancer in the Czech Republic. *Med Princ Pract.* 2003;12:23–29.
52. Zajac V, Kovac M, Kirchhoff T, Stevurkova V, Tomka M. The most frequent APC mutations among Slovak familial adenomatous polyposis patients. *Neoplasma.* 2002;49:356–361.
53. Laffler TG, Carrino JJ, Marshall RL. The ligase chain reaction in DNA-based diagnosis. *Ann Biol Clin* (*Paris*). 1993;51:821–826.
54. Burczak JD, Ching S, Hu H-Y, Lee HH. Ligase chain reaction for the detection of infectious agents. In: Wiedbrauk DL, Farkas DH, eds. *Molecular Methods for Virus Detection.* San Diego: Academic Press; 1995:315–328.
55. Mahony J, Chong S, Jang D, et al. Urine specimens from pregnant and nonpregnant women inhibitory to amplification of *Chlamydia trachomatis* nucleic acid by PCR, ligase chain reaction, and transcription-mediated amplification: identification of urinary substances associated with inhibition and removal of inhibitory activity. *J Clin Microbiol.* 1998;36:3122–3126.
56. Gorrin G, Friesenhahn M, Lin P, et al. Performance evaluation of the VERSANT HCV RNA qualitative assay by using transcription-mediated amplification. *J Clin Microbiol.* 2003;41:310–317.
57. Walker GT, Little MC, Nadeau JG, Shank DD. Isothermal *in vitro* amplification of DNA by a restriction enzyme/DNA polymerase system. *Proc Natl Acad Sci U S A.* 1992;89:392–396.
58. Westin L, Xu X, Miller C, Wang L, Edman CF, Nerenberg M. Anchored multiplex amplification on a microelectronic chip array. *Nat Biotechnol.* 2000;18:199–204.
59. Sooknanan R, van Gemen B, Malek LT. Nucleic acid sequence-based amplification. In: Wiedbrauk DL, Farkas DH, eds. *Molecular Methods for Virus Detection.* San Diego: Academic Press; 1995:261–285.
60. Polstra AM, Goudsmit J, Cornelissen M. Development of real-time NASBA assays with molecular beacon detection to quantify mRNA coding for HHV-8 lytic and latent genes. *BMC Infect Dis.* 2002;2:18.
61. Baeumner AJ, Cohen RN, Miksic V, Min J. RNA biosensor for the rapid detection of viable *Escherichia coli* in drinking water. *Biosens Bioelectron.* 2003;18:405–413.
62. Spadoro JP, Dragon EA. Quality control of the polymerase chain reaction. In: Farkas DH, ed. *Molecular Biology and Pathology: A*

Guidebook for Quality Control. San Diego: Academic Press; 1993: 149–158.
63. Wilber JC. Branched DNA for quantification of viral load. *Immunol Invest.* 1997;26:9–13.
64. Nolte FS. Branched DNA signal amplification for direct quantitation of nucleic acid sequences in clinical specimens. *Adv Clin Chem.* 1998;33:201–235.
65. Elbeik T, Alvord WG, Trichavaroj R, et al. Comparative analysis of HIV-1 viral load assays on subtype quantification: Bayer Versant HIV-1 RNA 3.0 versus Roche Amplicor HIV-1 Monitor version 1.5. *J Acquir Immune Defic Syndr.* 2002;29:330–339.
66. Gleaves CA, Welle J, Campbell M, et al. Multicenter evaluation of the Bayer VERSANT HIV-1 RNA 3.0 assay: analytical and clinical performance. *J Clin Virol.* 2002;25:205–216.
67. Hann HW, Fontana RJ, Wright T, et al. A United States compassionate use study of lamivudine treatment in nontransplantation candidates with decompensated hepatitis B virus-related cirrhosis. *Liver Transpl.* 2003;9:49–56.
68. Hendricks DA, Stowe BJ, Hoo BS, et al. Quantitation of HBV DNA in human serum using a branched DNA (bDNA) signal amplification assay. *Am J Clin Pathol.* 1995;104:537–546.
69. Martinot-Peignoux M, Boyer N, Colombat M, et al. Serum hepatitis B virus DNA levels and liver histology in inactive HBsAg carriers. *J Hepatol.* 2002;36:543–546.
70. Pawlotsky JM, Bastie A, Hezode C, et al. Routine detection and quantification of hepatitis B virus DNA in clinical laboratories: performance of three commercial assays. *J Virol Methods.* 2000;85: 11–21.
71. Beld M, Sentjens R, Rebers S, et al. Performance of the new Bayer VERSANT HCV RNA 3.0 assay for quantitation of hepatitis C virus RNA in plasma and serum: conversion to international units and comparison with the Roche COBAS Amplicor HCV Monitor, Version 2.0, assay. *J Clin Microbiol.* 2002;40:788–793.
72. Trimoulet P, Halfon P, Pohier E, Khiri H, Chene G, Fleury H. Evaluation of the VERSANT HCV RNA 3.0 assay for quantification of hepatitis C virus RNA in serum. *J Clin Microbiol.* 2002;40:2031–2036.
73. Lorincz A, Anthony J. Hybrid capture: a system for nucleic acid detection by signal amplification technology. In: Van Dyke C, Woodfork K, eds. *Luminescence Biotechnology: Instruments and Applications.* Boca Raton, FL: CRC Press; 2002:149–158.
74. Mazzulli T, Drew LW, Yen-Lieberman B, et al. Multicenter comparison of the Digene hybrid capture CMV DNA assay (version 2.0), the pp65 antigenemia assay, and cell culture for detection of cytomegalovirus viremia. *J Clin Microbiol.* 1999;37:958–963.
75. Schachter J, Hook EW III, McCormack WM, et al. Ability of the Digene hybrid capture II test to identify *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in cervical specimens. *J Clin Microbiol.* 1999;37:3668–3671.
76. Kessler HH, Pierer K, Dragon E, et al. Evaluation of a new assay for HBV DNA quantitation in patients with chronic hepatitis B. *Clin Diagn Virol.* 1998;9:37–43.
77. Lorincz A. Hybrid capture method for detection of human papilloma virus DNA in clinical specimens. *Pap Rep.* 1996;7:1–5.
78. Guan X-Y, Zhang H, Bittner M, Jiang Y, Meltzer P, Trent J. Chromosome arm painting probes. *Nature Genet.* 1996;12:10–11.
79. Dyanov HM, Dzitoeva SG. Method for attachment of microscopic preparations on glass for *in situ* hybridization, PRINS, and *in situ* PCR studies. *Biotechniques.* 1995;18:823–826.
80. Schrock E, du Manoir S, Veldman T, et al. Multicolor spectral karyotyping of human chromosomes. *Science.* 1996;273:494–497.
81. du Manoir S, Speicher MR, Joos S, et al. Detection of complete and partial chromosome gains and losses by comparative genomic *in situ* hybridization. *Hum Genet.* 1993;90:590–610.
82. Hachitanda Y, Toyoshima S, Akazawa K, Tsuneyoshi M. N-myc gene amplification in rhabdomyosarcoma detected by fluorescence in situ hybridization: its correlation with histologic features. *Mod Pathol.* 1998;11:1222–1227.
83. Ekins R, Chu FW. Microarrays: their origins and applications. *Trends Biotechnol.* 1999;17:217–218.
84. Descamps D, Calvez V, Collin G, et al. Line probe assay for detection of human immunodeficiency virus type 1 mutations conferring resistance to nucleoside inhibitors of reverse transcriptase: comparison with sequence analysis. *J Clin Microbiol.* 1998;36:2143–2145.
85. Evans JG, Lee-Tataseo C. Determination of the factor V Leiden single-nucleotide polymorphism in a commercial clinical laboratory by use of NanoChip microelectronic array technology. *Clin Chem.* 2002;48:1406–1411.

Section I
Genetics

Section Editor: Debra G.B. Leonard

Chapter 3

Genetic Counseling

Jennifer M. Farmer

Genetic Counseling: The Discipline and the Provider

Genetic counseling is a relatively new and rapidly evolving healthcare service that is increasing in demand as we enter the new era of genomic medicine. Genetic counseling evolved as a combination of disciplines, medical genetics and counseling theory. The original definition of genetic counseling focused on communicating medical information and inheritance and recurrence risk to patients, along with presenting options for responding to the risk in a nondirective manner and helping patients adjust to their conditions.[1] As the practice of genetic counseling has grown over the past 30 years, so have the goals and scope of practice.

As our knowledge of genetic disorders and complex inheritance patterns has expanded, so have the options for molecular-based genetic testing. With this growth, complex ethical and social issues have come to the forefront, such as genetic discrimination. As medical research has advanced, we have come to appreciate the strong influence of genetics in common disorders, such as cancer, diabetes, Alzheimer disease, asthma, and hypercholesterolemia. The burden of passing on an abnormal gene or trait is not limited to individuals and families faced with rare disorders of Mendelian inheritance. It is a reality for everyone. Increasing anxiety about genetic risk for disease and concern about passing on abnormal genes to future generations for common conditions has expanded the need for genetic counseling. There are now subspecialties of genetic counseling, such as prenatal, pediatric, cancer, and neurogenetics. Genetic counselors also are working in clinical molecular diagnostic laboratories providing information and education to the patients and physicians who request tests and try to interpret test results. It has been suggested that by the year 2010 predictive genetic tests will be available for about a dozen common conditions, with interventions and treatments for those who are identified as being at risk.[2]

Genetic counseling is typically provided by specialty-trained healthcare professionals, usually medical geneticists and genetic counselors. Medical geneticists are physicians from various disciplines such as obstetrics, pediatrics, and internal medicine who have obtained specialty training in medical genetics and are board certified by the American Board of Medical Genetics. Genetic counselors generally have a master's degree in genetic counseling, human genetics, or a closely related discipline and are board certified by the American Board of Genetic Counseling.

As genomic medicine evolves into mainstream healthcare, there will be increasing demand for genetic testing, especially related to preventive medicine and pharmacogenetics. Most if not all healthcare providers will need some genetic training to be able to consider appropriate testing for patients, interpret the clinical significance of test results, and in complicated situations provide appropriate referrals to genetic health professionals. Genetic counselors are playing a key role in the genetic education of healthcare providers and are facing a growing group of patients concerned about genetic risk and contemplating complex medical decisions for almost every type of disorder through various stages of the life cycle.

Clinical Genetic Testing and Counseling Setting

Preperinatal Setting

The roots of genetic counseling are often traced to prenatal diagnosis and reproductive decision making. Genetic counselors have been assisting women and couples with difficult reproductive choices for more than three decades. Typically, women referred for genetic counseling and prenatal diagnosis during pregnancy are those who are at increased risk of having a child affected by a genetic disorder. The most common indications are advanced

maternal age (women >35 years of age), abnormal triple marker screening, abnormal ultrasound finding, or family history of an inherited condition.

All women are at risk to have a child with a chromosome abnormality or birth defect with every pregnancy; however, the chromosomal abnormality risk increases with age. At about the age of 35, a woman's risk to have a child with a chromosomal abnormality is roughly equivalent to the risk of a miscarriage from an amniocentesis (1 in 200 to 300). An amniocentesis is a standard procedure performed between 14 to 20 weeks of pregnancy to obtain a sample of amniotic fluid, which contains fetal cells that can be used for cytogenetic and molecular analysis. Another procedure called chorionic villus sampling (CVS), can be done earlier in pregnancy (10 to 12 weeks) to obtain fetal cells for testing.[3] The miscarriage risk associated with CVS is slightly greater than that of amniocentesis at most centers. CVS is offered to women who know that they are at risk to have a child with a genetic condition so that diagnostic information can be obtained earlier in the pregnancy.

Genetic counselors in the prenatal setting explain the risk for genetic conditions due to maternal or paternal age, results from maternal serum screening tests, ultrasound findings, or family history. They present patients with prenatal diagnostic testing choices and discuss management and outcome options. Genetic counselors help patients make informed and autonomous choices through risk and procedure education, while also helping the patient to explore personal, spiritual, and cultural beliefs that affect decision making.[4]

The implementation of cystic fibrosis (CF) screening has caused a change in referral patterns for prenatal genetic counseling. A panel of experts convened for a National Institutes of Health consensus conference[5] recommended that CF genetic testing be available to all adults with a positive family history of CF, a spouse or partner with CF, and all couples planning a pregnancy or seeking prenatal care. These recommendations were adopted by the American College of Obstetricians and Gynecologists and the American College of Medical Genetics with the development of laboratory standards and guidelines.[6] In the majority of cases, patients are getting CF genetic testing without pretest counseling. When a patient has a positive result by the CF screening test, they are referred for posttest counseling to discuss the implications of the results, recurrence risk, and options for prenatal diagnosis. Some posttest counseling scenarios are complicated. For example, some asymptomatic individuals have compound heterozygous mutations associated with a milder CF phenotype. Chapter 10 discusses CF testing and genotype/phenotype associations in detail.

Other pregnancy-related indications for genetic counseling include advanced paternal age, history of infertility, recurrent pregnancy loss, and preconception counseling. Some couples, regardless of family history of a genetic condition, have taken a more proactive approach to family planning and request consultation with a genetic professional to discuss their risk of passing on a genetic condition to their children. Also, genetic carrier testing for disorders more common among individuals of Ashkenazi Jewish ancestry are in demand among couples presenting for preconception counseling and are now offered as a panel ranging from two to eight diseases, depending on the laboratory.

Preimplantation genetic diagnosis (PGD) offers couples who are at risk of having a child with a genetic condition an alternative to prenatal testing. Through in vitro fertilization (IVF), a single cell is removed from an embryo at the 8- to 16-cell stage and tested for the specific genetic condition or familial mutation; then the unaffected embryos are transferred back to the mother. PGD was first successful about a decade ago and is presently being offered for monogenic disorders and chromosome abnormalities.[7] However, it is not widely available as there are still many obstacles in the process related to the highly technical nature of the procedures and the difficulties in performing cytogenetic or molecular analysis on a single cell. The rate of pregnancy among patients undergoing IVF and PGD varies but rarely exceeds about one third.[8] Genetic counseling for couples considering PGD is imperative. Couples need to understand the risks and benefits of these complex procedures and the likelihood of a successful pregnancy. Couples need to weigh these factors against those of standard prenatal diagnosis procedures. PGD also is very expensive and not covered by most healthcare coverage providers.

Diagnostic Setting

Genetic conditions can occur with unique symptoms at all stages of life, from birth defects in a newborn to cognitive changes in an older adult. When a patient has a symptom or medical complaint, a diagnosis is needed, especially for treatable disorders. Initially, genetic disorders were considered untreatable and often investigated in the final stages of evaluation. However, biomedical research advances and recognition of the genetic contribution to more common disorders is changing the evaluation paradigm. Also, the increasing access and availability of genetic testing have improved the diagnostic capabilities for many disorders. When providing genetic counseling in a pretesting scenario for diagnostic purposes, genetic counselors may make testing recommendations, especially when medical intervention is available for the condition.[9] Diagnosis of a genetic condition brings emotional, social, and financial burden for the patient and the family.[10] Unlike many other areas of medicine, genetics does go beyond the patient and includes the impact to the entire family. Whether a diagnosis is made during the neonatal, pediatric, or adult years, the importance of genetic counseling remains. The goals of genetic counseling for the patient and the family following the diagnosis of a genetic condition include:

- Education about the natural history of the condition and medical implications;
- Explanation of the genetic contribution and inheritance and recurrence risks;
- Identification of social and emotional resources;
- Attentiveness to patient's and family's reaction to diagnosis and coping strategies;
- Promotion of the best possible emotional adjustment for the patient and family; and
- Facilitation of access to necessary medical and social services.[11]

Genetic testing for diagnostic purposes is occurring more frequently without pretest counseling in the primary physician's or specialist's office; however, once a positive test result is disclosed, patients are referred for genetic counseling to help them understand the meaning and implications of the result.

Predictive Setting

Predictive genetic testing holds the power to inform individuals, prior to the onset of symptoms, that they are destined to develop or are at increased risk to develop a hereditary disorder. Experience with predictive genetic testing for adult-onset conditions such as Huntington disease and hereditary cancer syndromes has led to the development and strong endorsement of a multidisciplinary approach to predictive genetic testing that includes pre- and posttest genetic counseling protocols.[12,13] This approach allows the patient to explore his or her motives for testing, expectations, the risks and benefits of testing, and coping strategies prior to testing. This is especially important when there is no treatment or medical intervention for the disorder. Most predictive testing protocols require at least two pretest counseling visits (see Table 3-1) to allow the patient time to consider the benefits and risks of testing, develop a support network during the testing process, and ensure voluntary participation in testing. Experiences with Huntington disease testing have shown that patients are at risk for adverse outcomes after disclosure of predictive genetic test results.[15] Result disclosure should always be done in person by a genetic counselor or health professional knowledgeable about the disorder and the implications of the test result.[14] Also, a support person for the patient should be present at the pre- and posttest counseling visits.

Table 3-1. Elements of Pretest Predictive Genetic Counseling[14]

Obtain family history and confirm diagnoses
Review natural history and inheritance of condition, as well as a priori risk
Discuss the benefits, limitations, and risks of testing, and the confidentiality of test results
Discuss motives for testing, anticipated result, psychosocial preparedness, and support system
Present alternatives to testing and assure that testing is voluntary and informed consent is provided

Cancer Setting

During the past 10 years, the supply and demand for diagnostic and predictive cancer genetic testing has rapidly increased, giving rise to the specialty of cancer genetic counseling. Genetic testing for inherited cancer syndromes can be useful for diagnosis and medical management among individuals presenting with a tumor or symptoms. In many cases, testing holds the possibility for treatment and prevention in at-risk individuals. Chapters 17 to 22 highlight the various issues related to molecular testing for this group of conditions, which include variable clinical utility, complex medical management options, difficult dilemmas with approaches to molecular testing, and testing in the research setting or during early transition to the clinical laboratory. The potential risks and benefits associated with testing vary based on the specific hereditary cancer syndrome, as well as the patient and family history. The American Society of Clinical Oncology recommends pre- and posttest counseling for individuals referred for cancer genetic testing.[13] The genetic counselor discusses the details of the genetic testing (detection rate, clinical utility, recurrence risk, etc.) as well as early detection and prevention options for individuals with a positive test result. Often, hereditary cancer syndromes increase an individual's risk for cancer in multiple organ systems, which makes medical management and early detection more complex. For example, some individuals at risk for von Hippel-Lindau syndrome need at least yearly screening for brain and spine hemangioblastomas, retinal angiomas, pheochromocytoma, renal cell carcinoma, and other tumors.

Informed Consent

In the healthcare setting, the process of informed consent is a protection for patients. Prior to diagnostic testing or therapeutic intervention, the provider explains the procedure to the patient along with the risks, benefits, and alternatives so that the patient can voluntarily make informed decisions about diagnostic and treatment options.[16] Depending on state law and laboratory standards, there is great variation in informed consent requirements for genetic testing. Position statements and guidelines for informed consent for genetic testing are available only for certain conditions or groups of conditions. The majority of guidelines exist for predisposition genetic testing. The National Society of Genetic Counselors recommends obtaining informed consent prior to predisposition genetic testing for adult-onset conditions.[14] Guidelines for informed consent prior to genetic testing stress that this is

Table 3-2. Key Elements of Informed Consent for Genetic Testing

Discussion of purpose of the test; procedures involved in testing (blood drawing)
Clinical utility of the test and interpretation of test results
Discussion of risks, benefits, and limitations to testing (including psychosocial, cultural, and financial risks and benefits)
Presentation of alternatives to genetic testing
Procedure for communication of results
Confidentiality of test results
Voluntary nature of informed consent

more than having a patient read and sign a piece of paper. Informed consent should be a communication process that fosters autonomous and informed decision making for the patient.[17,18] Presentation of the key elements of informed consent (Table 3-2) needs to be tailored to the individual patient's learning style, educational and cultural background, and family situation to optimize the utility of informed consent. This is a time-intensive process that cannot routinely be done by primary care physicians during a routine office visit. The informed consent process also applies to genetic testing or collection of tissue or body fluids for research purposes.[19]

Genetic Testing for Children and Adolescents

The benefits and harms of genetic testing need to be carefully evaluated before proceeding with testing in children who may not be able to appreciate the implications of such results. When genetic testing directly impacts medical management or treatment for a child with symptoms or clinical features of a condition, the benefits of testing are clear and the well-being of the child is being promoted. However, when genetic testing does not impact medical management, or the condition in question will occur in adulthood, the implications of testing become more complex and the benefits become less clear. The American College of Medical Genetics and the American Society of Human Genetics wrote "Points to Consider: Ethical, Legal and Psychosocial Implications of Genetic Testing in Children and Adolescents."[20] The recommendations are:

1. Timely medical benefit to the child should be the primary justification for genetic testing in children and adolescents.
2. Substantial psychological benefits to the competent adolescent also may be a justification for genetic testing.
3. If the medical or psychological benefits of a genetic test will not accrue until adulthood, as in the case of carrier status or adult-onset diseases, genetic testing generally should be deferred.
4. If the balance of benefits and harms is uncertain, the provider should respect the decision of the competent adolescent and his or her family.
5. Testing should be discouraged when the provider determines that the potential harms of genetic testing in children and adolescents outweigh the potential benefits.

Education and counseling for the parents and the child, at an appropriate level, should be provided. The benefits and harms related to medical issues, psychosocial issues, and reproductive issues need to be presented and discussed. Children and certainly adolescents have decision-making capacity. The child's competence and wishes should be assessed prior to genetic testing and carefully balanced with parental authority. This is especially true for adolescents who can articulate a specific opinion that differs from that of his or her parents.[20] Assent from the child or adolescent should be obtained in addition to informed consent from the parents.

Transition from Research to Clinical Testing: Role of the Genetic Counselor

Chapter 49 explores the differences between research genetic testing and clinical genetic testing and the transition from the former to the latter. Genetic counseling can be valuable to individuals who are participating in research genetic testing. Patients may have difficulty understanding the purpose, risks, and benefits of the study. In addition, many patients have expectations of receiving research testing results and do not appreciate the limitations of results reported from the research setting. It is strongly recommended that when genetic research results are disclosed to research participants, it be done by a professional able to provide genetic counseling.[19] For many rare conditions, clinical testing is not available; therefore, research testing is the only option for families. When genetic testing is transitioning from the research to the clinical setting, a genetic counselor can be a liaison between the patient and the laboratory and be responsible for informed consent and disclosure of results.

Genetic Counselors Working with the Molecular Pathologist

Genetic testing can be uniquely different from other laboratory tests. One difference is that often it is useful for the laboratory to have information about the patient or family history for interpretation of test results. Gathering the appropriate clinical information and a complete family history is not always straightforward and needs to be tailored to the specific test being offered. A genetic counselor is uniquely trained to understand and interpret the appropriate information needed by the molecular pathology laboratory in order to perform a specific genetic test. Oftentimes it can be challenging to work within a family system to get the necessary information to document a

Table 3-3. Genetic Counselors Working with the Molecular Pathologist

Obtain clinical information and family history data to ensure that the most appropriate test is performed
Facilitate the informed consent process to assure that the patient's autonomy is protected
Interpret clinical validity and test results for referring physicians and patients

genetic diagnosis prior to testing at-risk individuals. For example, if a patient presents requesting testing for spinocerebellar ataxia (SCA) because this condition has been diagnosed in a sibling, it is important to know the specific type of SCA, as there are currently 17 types and commercial testing is available for only about 7 of the 17. If the individual requesting the test does not know the type, one could do testing for all of the commercially available types; however, a negative result would not absolve the individual at risk because the proband may have one of the SCAs not included in the testing. A genetic counselor can work with an at-risk individual to obtain the necessary family history information and documentation so that the most accurate and efficient approach to testing can be used and the interpretation of the results will be more informative. An accurate and comprehensive family history is a valuable tool in a diagnostic evaluation as it can be used as a medical screening tool, establish a pattern of inheritance, identify individuals at-risk, and determine strategies for genetic testing.

Understanding the clinical validity of a genetic test result can be difficult for both healthcare providers and patients. However, information about the sensitivity of the test and the penetrance of mutations is paramount to the interpretation of test results.[9] Each genetic test has its own sensitivity and specificity based on the methodologies and technologies employed by the laboratory performing the test. Many laboratories have genetic counselors on staff. The genetic counselor can be a useful resource for the molecular pathology laboratory to interface with other healthcare providers and the public to provide the necessary education and information to determine the appropriate approach to genetic testing, facilitate the details of ordering a test, and help interpret test results (Table 3-3).

Summary

The goals of genetic counseling are to address the informational and emotional needs of patients and their families.[10] For example, the explanation of risks and benefits associated with a genetic test that is tailored to a patient's educational needs as well as family, social, and cultural background facilitates informed decision making through promoting patient autonomy and informed consent. The key goals of genetic counseling for most patient encounters include:

- Obtain and interpret family medical history information;
- Educate patients so they understand the medical and genetic information (inheritance and recurrence risks) needed to make health-management decisions and "master" their condition;
- Promote informed decision making and informed consent;
- Be aware of nontechnical factors (social, cultural, financial, and emotional factors) that influence patients in the decision-making process;
- Foster genetic competence in patients and families; and
- Identify social and professional resources for patients.

As genetic testing expands and is incorporated into mainstream healthcare, especially for disease prevention and approaches to treatment (pharmacogenetics) for common disorders, pre- and posttest genetic counseling for every test for all patients will not be feasible. However, genetic counseling resources need to be available to healthcare providers and patients to provide education and support and to promote safety in genetic testing.

References

1. Fraser FC. Current issues in medical genetics: genetic counseling. *Am J Hum Genet.* 1974;26:636–659.
2. Collins FS, McKusick VA. Implications of the Human Genome Project for medical science. *JAMA.* 2001;285:540–544.
3. Shulman LP, Elias S. Amniocentesis and chorionic villus sampling in fetal medicine. *West J Med.* 1993;159(special issue):260–268.
4. Biesecker BB, Marteau TM. The future of genetic counseling: an international perspective. *Nat Genet.* 1999;22:133–137.
5. National Institutes of Health Consensus Development Conference Statement on Genetic Testing for Cystic Fibrosis. Genetic testing for cystic fibrosis. *Arch Intern Med.* 1999;159:1529–1539.
6. Grody WW, Cutting GR, Klinger KW, et al. Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. *Genet Med.* 2001;3:149–154.
7. Black SH. Preimplantation genetic diagnosis. *Curr Opin Pediatr.* 1994;6:712–716.
8. Kanavakis E, Traeger-Synodinos J. Preimplantation genetic diagnosis in clinical practice. *J Med Genet.* 2002;39:6–11.
9. Burke W. Genomic medicine: genetic testing. *N Engl J Med.* 2002;347:1867–1875.
10. Weil J. Genetic counseling in the era of genomic medicine: as we move towards personalized medicine, it becomes more important to help patients understand genetic tests and make complex decisions about their health. *EMBO Rep.* 2002;3:590–593.
11. Walker AP. The practice of genetic counseling. In: Baker DL, Schulette JL, Uhlmann WR, eds. *A Guide to Genetic Counseling.* New York: Wiley-Liss; 1998:1–20.
12. International Huntington Association and World Federation of Neurology. Guidelines for the molecular genetic predictive test in Huntington's disease. *J Med Genet.* 1994;31:555–559.
13. American Society of Clinical Oncology (ASCO). Policy statement update: genetic testing for cancer susceptibility. *J Clin Oncol.* 2003;21:2397–2406.
14. McKinnon WC, Baty BJ, Bennett RL, et al. Predisposition genetic testing for late-onset disorders in adults: a position paper of the National Society of Genetic Counselors. *JAMA.* 1997;278:1217–1220.

15. Almqvist EW, Bloch M, Brinkman R, Craufurd D, Hayden MR. A worldwide assessment of the frequency of suicide, suicide attempts, or psychiatric hospitalization after predictive testing for Huntington disease. *Am J Hum Genet.* 1999;64:1289–1292.
16. Andrews LB. Compromised consent: deficiencies in the consent process for genetic testing. *JAMA.* 1997;52:39–44.
17. Geller G, Botkin JR, Green MJ, et al. Genetic testing for susceptibility to adult-onset cancer: the process and content of informed consent. *JAMA.* 1997;277:1467–1474.
18. Sharpe NF. Informed consent and Huntington disease: a model for communication. *Am J Med Genet.* 1994;50:239–246.
19. American Society of Human Genetics (ASHG) Report. Statement on informed consent for genetic research. *Am J Hum Genet.* 1996;59: 471–474.
20. American College of Medical Genetics/ASHG. Points to consider: ethical, legal and psychosocial implications of genetic testing in children and adolescents. *Am J Hum Genet.* 1995;57:1233–1241.

Chapter 4

Pedigree Analysis and Risk Assessment

Catherine Walsh Vockley

The Genetic Family History

The personal and family medical pedigree has evolved from its earliest ancestors in the 15th century to its current form and has become an essential tool in many aspects of the clinical genetics evaluation. Originally used primarily to display relationship information, the pedigree was used for the first time to demonstrate inheritance of traits in the mid-19th century when Pliney Earl published on inheritance of color blindness and Francis Galton described inheritance of artistic ability and genius.[1]

Symbols used to document pedigree information have varied, often depending on personal, professional, or national preferences. The key to functionality for pedigrees, however, is the degree to which they are able to communicate information uniformly to all users. In 1993, a task force of the National Society of Genetic Counselors surveyed genetic counselors regarding interpretation of pedigree symbols and conformity of usage.[2] As many as 17 different symbols were used to depict pregnancy, with 16 different symbols being used to denote miscarriage; in both cases, symbols sometimes had several meanings to different users. It became evident that standardization of symbol usage was needed. The group established a recommended nomenclature for pedigrees, which was published in the *American Journal of Human Genetics* in 1995.[3]

The currently recommended methods for documenting pedigree information including symbols, spatial relationships, and clinical/investigative status are detailed in Figures 4-1 through 4-5. These standards allow recording of traditional relationships as well as those developing as new technologies are applied, particularly in reproductive medicine. They also serve as a uniform baseline for future additions or modifications as the field continues to evolve.

These pedigrees now form the cornerstone for determination of diagnosis, pattern of inheritance, and recurrence risk.[4] Use of pedigree information can impact overall risk assessment, medical management decisions, and feasibility of various testing strategies. In addition, collection of family medical information has aided in the understanding of many unique features of hereditary disorders, including natural history, variability, and gene-gene or gene-environment interactions.

Collection of a family pedigree represents an opportunity to build a relationship with the patient and family and to learn about how the family functions.[4] As the genetic counselor or other healthcare provider explains the purpose of the family history, an atmosphere of open communication and respect can be established. This process provides a window to the social relationships and psychosocial and educational needs of patients and families. In the social sciences, genograms are used to graphically depict family dynamics that influence individual behaviors.[5] This information is also essential for successful counseling of patients in the clinical genetics setting, and while not always recorded in the same fashion, it is a vital part of the process of pedigree gathering. Observations about coping mechanisms, assumptions about disease causation, family hierarchy, key life experiences, stress levels, body language, and religious and ethnic influences all are integrated into consideration about the most effective ways to communicate information about a diagnosis, prognosis, or management plan to patients and families.

Ideally, the pedigree is collected in a face-to-face session. This is usually done prior to or at the beginning of the clinical genetics evaluation, but may be done later, particularly when evaluating a pregnancy or a newborn with an unanticipated, newly diagnosed condition. It is helpful to provide patients with advance notice about the nature of information to be collected, as this facilitates accuracy and completeness. At a minimum, a three-generation pedigree should be collected, including all first-degree relatives (parents, children, full siblings), second-degree relatives (grandparents, aunts, uncles, nieces and nephews, half-siblings, grandchildren), and as pertinent, many third-degree relatives (cousins, great-aunts, great-uncles, great-grandparents). This group can be expanded or condensed, depending on the nature of the referral and patient responses to preliminary questioning about features

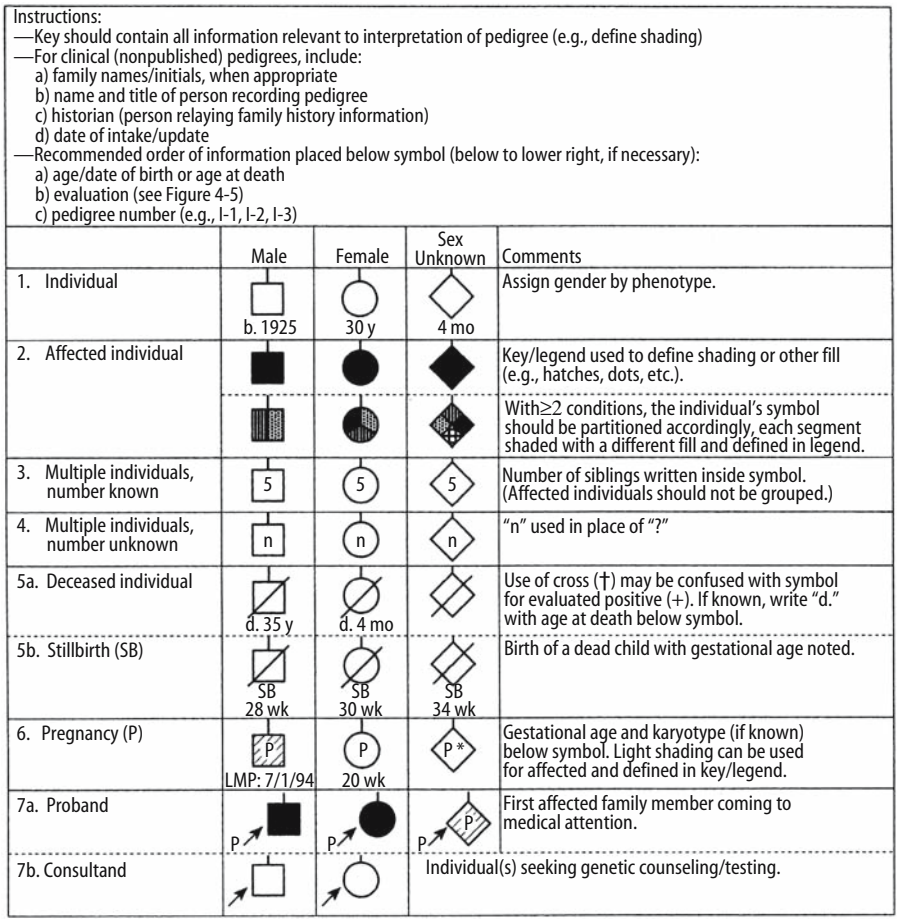


Figure 4-1. Common pedigree symbols, definitions, and abbreviations. (Figures 4-1 to 4-5 reprinted from Bennett RL, Steinhaus KA, Uhrish SB, et al. "Recommendations for standardized human pedigree nomenclature." *American Journal of Human Genetics* 1995;56:745–752, with permission from the University of Chicago Press.)

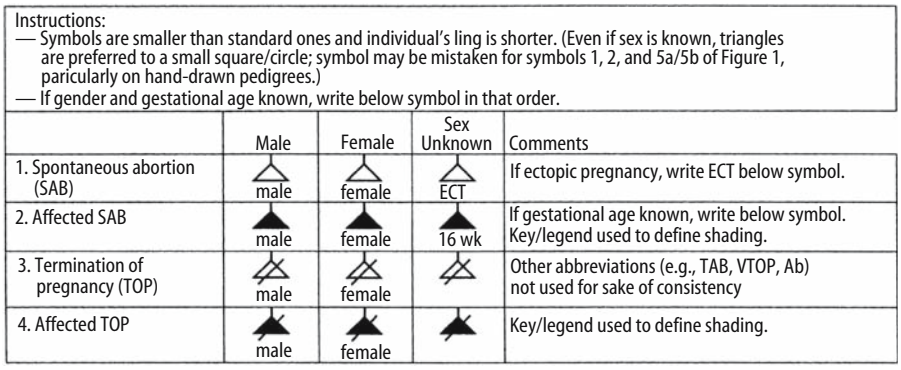


Figure 4-2. Pedigree symbols and abbreviations for pregnancies not carried to term.

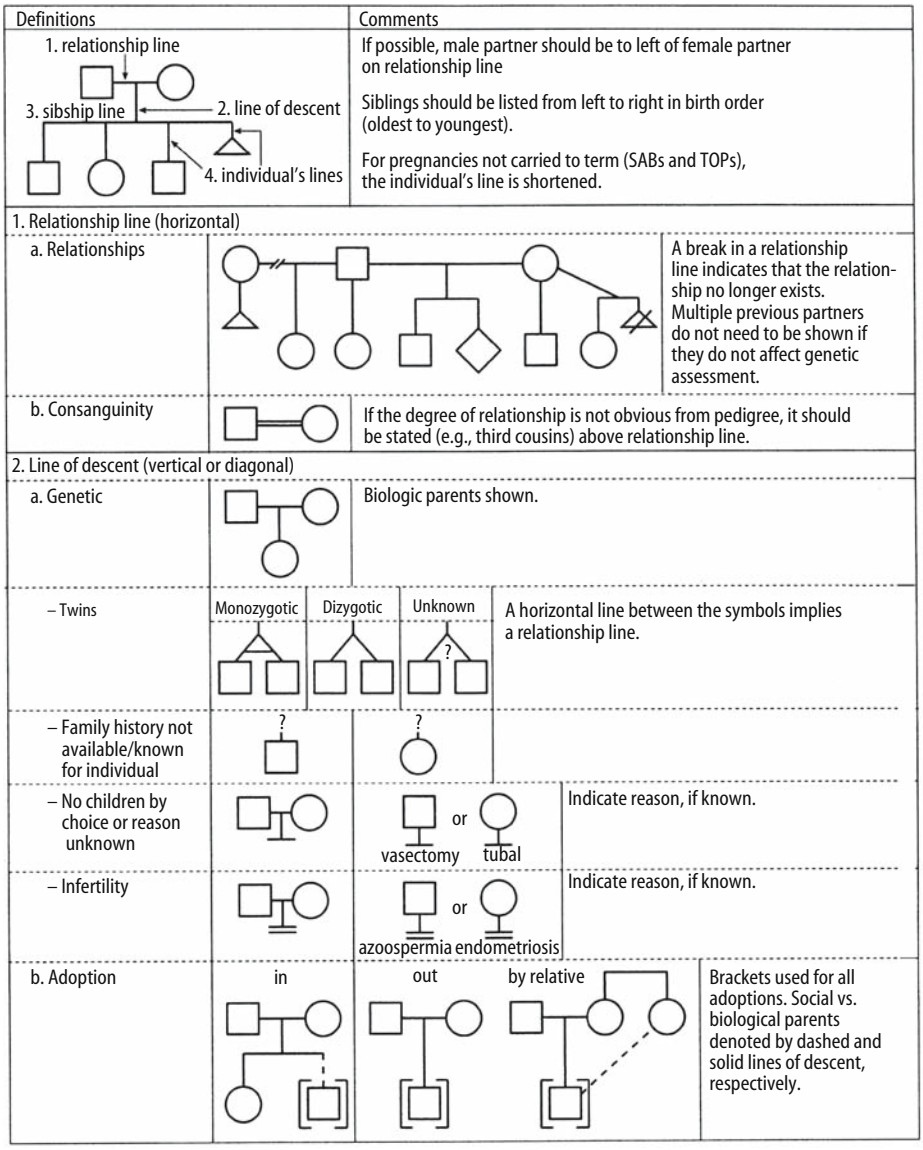


Figure 4-3. Pedigree line definitions.

Definitions:
— Egg or sperm donor (D)
— Surrogate (S)
— If the woman is both the ovum donor and a surrogate, in the interest of genetic assessment, she will be referred to only as a donor (e.g., 4 and 5)
— The pregnancy symbol and its line of descent are positioned below the woman who is carrying the pregnancy.
— Family history can be taken on individuals, including donors, where history is known.

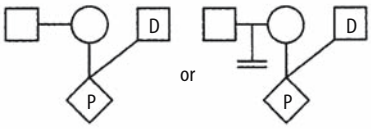
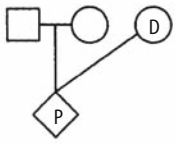
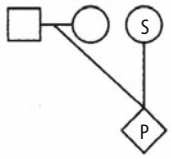
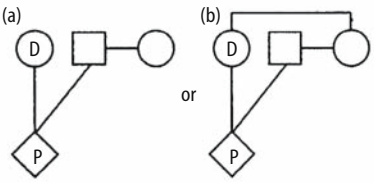
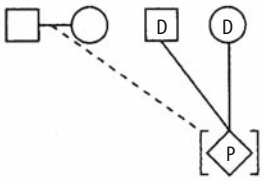
Possible reproductive scenarios		Comments
1. Sperm donor		Couple in which the woman is carrying pregnancy using donor ssperm. No relationship line is shown between the woman carrying the pregnancy and the sperm donor. For a lesbian relationship, the male partner can be substituted with a female partner.
2. Ovum donor		Couple in which the woman is carrying pregnancy using donor egg(s) and partner's sperm.
3. Surrogate only		Couple whose gametes are used to impregnate another woman (surrogate) who carries the pregnancy.
4. Surrogate ovum donor		Couple in which the male partner's sperm is used to inseminate (a) an unrelated woman or (b) a sister who is carrying the pregnancy for the couple.
5. Planned adoption		Couple contracts with a woman to carry a pregnancy using the ovum of the woman carrying the pregnancy and dono sperm.

Figure 4-4. Assisted reproductive technologies symbols and definitions.

Instructions:
— Evaluation (E) is used to represent clinical and/or test information on the pedigree.
  a. E is to be defined in key/legend.
  b. If more than one evaluation, use subscript (E_1, E_2, E_3) and define in key. May be written side by side or below each other depending on available space.
  c. Test results should be put in parentheses or defined in key/legend.
  d. If results of exam/family study/testing not documented or unavailable, may use a question mark (e.g.,"E?").
— Documented evaluation (*)
  a. Asterisk is placed next to lower right edge of symbol
  b. Use only if examined/evaluated by you or your research/clinical team or if the outside evaluation has been personally reviewed and verified
— A symbol is shaded only when an individual is clinically symptomatic.
— For linkage studies, haplotype in formation is written below the individual. The haplotype of interest should be on the left and appropriately highlighted.
— Repetitive sequences, trinucleotides, and expansion numbers are written with affected allele first and placed in parentheses.
— If mutation known, identify and place in parentheses.
— Recommended order of information:
  a. Age/date of birth or age at death
  b. Evaluation information
  c. Pedigree number (e.g., I-1, I-2, I-3)

Definition	Symbol	Scenario	Example
1. Documented evaluation (*)	*	Woman with normal physical exam and negative fragile X chromosome study (normal phenotype and negative test result).	E –*
2. Obligate carrier (will not manifest disease)		Woman with normal physical exam and premutation for fragile X (normal phenotype and positive test result).	E+(100n/35n)*
3. Asymptomatic/ presymptomatic carrier (clinically unaffected at this time but could later exhibit symptoms)		Man age 25 with normal physical exam and positive DNA test for Huntington disease (symbol filled in if/when symptoms develop).	25 y* E+(45n/18n)
4. Uninformative study (u)	Eu	Man age 25 with normal physical exam and uninformative DNA test for Huntington disease (E_1) and negative brain MRI study (E_2).	25 y* E_1u(36n/18n) E_2-
5. Affected individual with positive evaluation (E+)	E+	Individual with cystic fibrosis and positive mutation study, although only one mutation has currently been identified.	E+(ΔF508) Eu E+(ΔF508/u)*
		18- week male fetus with abnormalities on ultrasound and a trisom 18 karyotype.	18 wk* E+(tri 18n)

Figure 4-5. Pedigree symbolization of genetic evaluation/testing information.

relevant to the reason for referral. For example, cancer genetic evaluations may necessitate a more extended family pedigree, while a brief, focused pedigree may suffice when discussing cystic fibrosis carrier testing.

Information that should be collected about each individual in the pedigree is listed in Table 4-1. This, too, may be modified to reflect the nature of the diagnosis under investigation. Ethnicity, consanguinity, and unique biological relationships should be recorded using standard notation. All reported diagnoses or conditions ideally should be confirmed through authorized request and review of medical records. Key records to obtain include pathology reports, test results (particularly for any genetic testing that has been performed), imaging reports, and autopsy reports. In the absence of these documents, family genealogies or death certificates may provide some degree of verification of reported information.

An emerging issue in the use of pedigrees for clinical evaluations and research is the issue of individual confidentiality.[6,7] Each member of the family has a right to expect that medical information will remain confidential. This becomes complicated when one considers the pedigree that may contain both reported ("hearsay") and confirmed information for numerous individuals. Those people may have willingly shared information with the patient but may not want it shared with other family members. If subsequent to an evaluation a patient requests release of his or her pedigree to another family member, a provider should carefully consider the question of ownership of the pedigree information and be attuned to the potential consequences of releasing the (identifiable) information about other family members. Current interpretation of regulations outlined in the Health Insurance Portability and Accountability Act (HIPAA) and other medical records privacy legislation may influence how such information is shared.[8] Professional organizations including the American Society of Human Genetics also have developed position statements on this issue.[9]

Table 4-1. Family History Collection: What to Ask?

For All Family Members
Current age; complete date of birth
Exact relationship to proband
General health status
History of major acute or chronic illness
History of learning problems, diagnosed disabilities, or mental retardation
Highest grade level completed (when relevant)
Employment (when relevant)
Reproductive history, including pregnancies, miscarriages, elective terminations, infertility, and choice not to have children
Gestational age and last menstrual period for ongoing pregnancies
Consanguinity
Targeted questions relevant to the reason for evaluation, for example, key symptoms or features of the condition in question, pertinent evaluations, etc.
Age at death; year of death; cause of death
For Family Member Known to Be Affected by the Condition in Question
Diagnosis
Age at diagnosis
Method of diagnosis
Evaluations and testing completed
Symptoms
Information about ongoing treatment or management plan
Availability of medical records for review

Patterns of Inheritance

One key use of the carefully collected and verified pedigree is determination of the most likely mode of inheritance of a condition in a family. This will have relevance to assessing recurrence risks, approaches to testing, and in some cases, even prognosis. The concept of patterns of inheritance extends from the work of Gregor Mendel, who in the 17th century described transmission of traits associated with single genetic loci.[10] Transmission of human genetic conditions and traits has proven to be more complex, involving not only the single gene patterns first described by Mendel but also chromosomal inheritance, mitochondrial inheritance, and numerous atypical patterns of inheritance, including contiguous gene disorders, imprinting, uniparental disomy, trinucleotide repeat expansion, multifactorial inheritance, mosaicism, epigenetic influences, and synergistic heterozygosity. Undoubtedly, more atypical patterns of transmission will be elucidated as our understanding of the human genome expands. As of June 14, 2006, Victor McKusick's classic reference *Mendelian Inheritance in Man* (12th ed., 1998; http://www.ncbi.nlm.nih.gov/OMIM)[11] lists 16,850 defined gene loci, 2290 of which have been associated with specific clinical entities. There are, however, more than 7500 human traits and/or conditions that have defined classic patterns of inheritance.[10] These primarily fall into three categories, autosomal, X-linked, and Y-linked; however, a number of mitochondrial conditions also have been confirmed.

Autosomal Dominant Inheritance

In classic autosomal dominant inheritance, an affected individual has one non-functional or mutant allele at a particular locus. Each affected individual in a pedigree has a 50% chance of passing the disease-associated mutation to each of his or her offspring. Many factors, however, influence the occurrence of these conditions in families. These will be described as a group following review of the classic modes of inheritance. A key feature of autosomal dominant inheritance is male-to-male transmission of the condition or trait, a pattern not seen in X-linked dominant inheritance, which can be confused with autosomal dominant inheritance on first analysis. Table 4-2 lists additional features of autosomal dominant inheritance, and an example pedigree is shown in Figure 4-6. Codominant inheritance describes equal expression of both alleles of a pair, that is, with equal, coexisting phenotypic effect. An example of this is the ABO blood group.

Autosomal Recessive Inheritance

In autosomal recessive inheritance, an affected individual has two nonfunctional or mutant alleles at a particular locus. One of these is inherited from each of the parents, who are called carriers and who are unaffected by the condition. There is a 1 in 4 (25%) chance of having an affected offspring with each pregnancy of a known carrier couple, and a 2 in 4 (50%) chance that an offspring will be a carrier like the parents. After birth, if a child of a carrier couple is not affected by the condition in question, he or she has a 2 in 3 chance of being a carrier. Risk to future offspring of a known carrier depends on the likelihood that his or her partner is also a carrier. This is influenced by the frequency of the disease gene in the population, which may vary among different populations. Features of autosomal recessive inheritance are listed in Table 4-2, and a pedigree is shown in Figure 4-6.

X-linked Dominant Inheritance

In X-linked dominant inheritance, an affected individual has one non-functional or mutant allele at a locus on an X-chromosome. X-linked dominant conditions can occur in either males or females. Risk for offspring of an affected female is 50%, regardless of the gender of the offspring. Risk to offspring of affected males is gender dependent, with all daughters but no sons inheriting the gene. Many of these conditions, however, are lethal in males, so pedigrees may show overrepresentation of females or increased frequency of miscarriages, presumably of affected male fetuses (see Table 4-2 and Figure 4-6).

Table 4-2. Features of Mendelian Patterns of Inheritance

Autosomal Dominant Inheritance
Male-to-male transmission occurs; both genders can transmit to offspring
Condition occurs in multiple generations
Males and females affected, typically to comparable extent
Variability of clinical findings
Later/adult onset in some disorders
Vertical transmission; affected descendants of affected individuals, unaffected descendants of unaffected individuals (in general)
Homozygotes may be more severely affected than heterozygotes
Homozygosity may be lethal
Occurrence of new mutations
Nonpenetrance; apparent "skipping" of generations
Gender-limited occurrence of conditions (transmission through the unaffected gender)
Germline mosaicism reported
Autosomal Recessive Inheritance
Affected family members are usually in one generation; "horizontal" inheritance
Parental consanguinity or small mating pool may influence disease occurrence
Male and female are affected
Usually consistent in degree of severity among affected family members
Early onset of symptoms more typical
New mutations rare
May see higher frequency of disease in certain ethnic groups
X-linked Dominant Inheritance
No male-to-male transmission
Affected females usually have milder symptoms than affected males
Affected males have no affected sons, but all daughters will be affected
May mimic autosomal dominant inheritance
May be lethal in affected males; paucity of males or overrepresentation of females in the pedigree
Increased occurrence of miscarriage
X-linked Recessive Inheritance
No male-to-male transmission
Males more frequently affected
Carrier females usually unaffected but may have mild symptoms
Affected males in a family are related through females
Occurrence of new mutations, often from maternal grandfather
Y-linked Inheritance
Male-to-male transmission only
Association with increased infertility rates in families
Discrepancy between chromosomal and phenotypic gender

X-linked Recessive Inheritance

Traditional X-linked recessive inheritance is characterized by occurrence of the condition in males with a non-functional or mutant allele on the X-chromosome who are related through females. (See the pedigree in Figure 4-6, and Table 4-2 for additional features.) Typically, carrier females are unaffected; however, due to lyonization (random inactivation of one X chromosome in each cell in a female), carrier females may have mild symptoms. This occurs when, by chance, more of the X chromosomes with the nonfunctional allele remain active in the cells. The likelihood of symptoms in carrier females varies considerably among disorders. Risk to offspring of carrier females is 25% overall, or 50% for affected status if the fetus/offspring is male. Offspring of affected males will not be classically affected, but all daughters will be carriers.

Y-linked Inheritance

In rare cases, one of a limited number of genes on the Y chromosome can be mutated. This can result in disparity between chromosomal and phenotypic gender if the SRY region is involved, or can be associated with genetic/hereditary forms of infertility. This may be identified more frequently as reproductive technologies such as intracytoplasmic sperm injection (ICSI) are used to aid in achieving pregnancies for previously infertile males, due to Y-chromosome deletions, for example (see Table 4-2 and Figure 4-6).

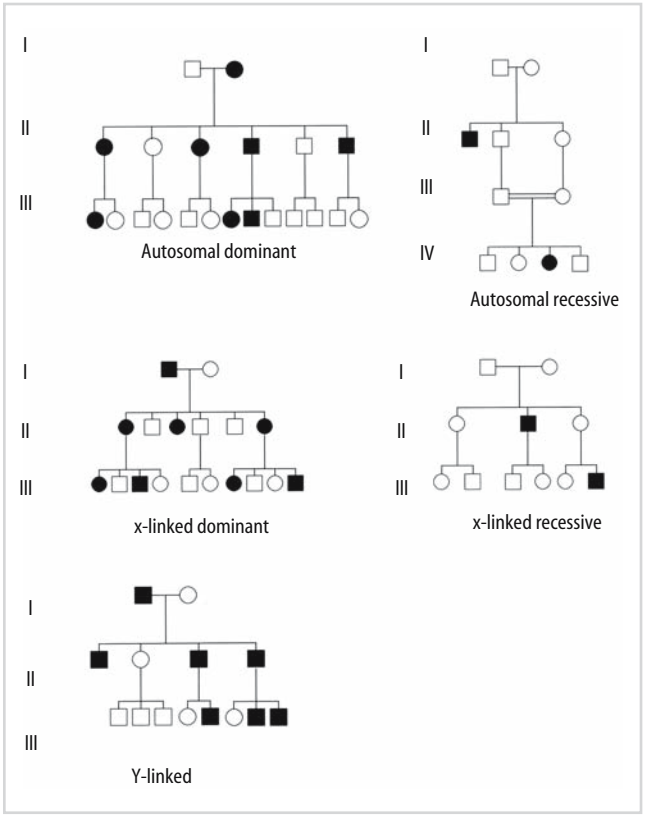


Figure 4-6. Example pedigrees for Mendelian patterns of inheritance.

Non-Mendelian Inheritance Patterns

For a summary of non-Mendelian inheritance patterns, see Table 4-3.

Chromosomal

Chromosome abnormalities can occur sporadically or can be caused by familial transmission of duplications, deletions, or rearrangements that can result in imbalance of genetic material in the offspring.[12] Due to the presence of many genes along the segment of chromosome involved, multiple phenotypic effects usually are seen. Risks to offspring of familial cases depend on parent of origin and size and location of the involved chromosomal segment, and vary depending on loss or gain of material in a particular region. In apparently sporadic cases, parental status with respect to the chromosomal abnormality should be assessed for all cases of offspring with chromosomal rearrangements. Absence of a parental chromosomal abnormality significantly reduces the risk to future offspring.

Contiguous Gene Disorders/ Microdeletion Syndromes

Contiguous gene disorders are the result of loss of several adjacent genes along a segment of chromosome and may consist of symptoms of one known hereditary disorder, more than one closely linked group of hereditary disorders, or either of these in conjunction with mental retardation, dysmorphic features, or both.[13] The condition results from loss of one copy of a group of closely linked genes (haploinsufficiency) that may be detectable by high-resolution chromosome analysis or fluorescence in situ hybridization (FISH) using region-specific probes. Approximately 5% to 10% of monogenic diseases are associated with gene deletions that cannot be detected through routine cytogenetic analysis.[10] The microdeletions occur in regions of repeated genomic sequences that lead to rearrangements (recombination), resulting in loss or gain of genetic material during transmission, both of which have been documented.

Mitochondrial Inheritance

Individuals inherit essentially all their mitochondrial DNA from their mothers; thus, any disease associated with a mitochondrial DNA mutation is transmitted from the mother to the offspring. In each cell, including egg cell progenitors, there may be up to 1000 mitochondria. If a mutation occurs in one of these mitochondria, as the mitochondrion divides over time, the mutation becomes present in a percentage of the overall mitochondrial population in the cell. When the cell divides, the mitochondria are distributed stochastically to the daughter cells. The

Table 4-3. Features of Non-Mendelian Patterns of Inheritance

Chromosomal Disorders

Increased frequency in individuals with 2 or more major birth defects, 3 or more minor birth defects, or 1 major and 2 minor birth defects
Occurrence of multiple pregnancy losses or infertility
Occurrence of mental retardation with dysmorphism
Occurrence of mental retardation with multiple congenital anomalies
Many occur as sporadic conditions with negative family history

Contiguous Gene Disorders/Microdeletion Syndromes

Involvement of multiple organ systems
Negative family history; frequent/isolated or sporadic cases
May appear as recognized single-gene disorder
May involve occurrence of mental retardation with an otherwise recognized hereditary or medical disorder typically lacking mental retardation
May involve occurrence of dysmorphism with an otherwise recognized hereditary or medical disorder typically lacking dysmorphism

Mitochondrial Inheritance

Maternal transmission (fathers do not transmit disease)
Males and females affected
Extreme variability of clinical symptoms; multiple organ systems involved
Multiple generations affected (matrilineal)
Degenerative/neuromuscular disorders predominate
Gender can influence variability of symptoms
Environmental factors may influence symptoms (pseudomultifactorial)

Imprinting

Gender of transmitting parent modifies gene/disease expression (parent-of-origin effects)
May appear to skip generations

Uniparental Disomy

Documentation of only one carrier parent
Single/isolated case in a family

Trinucleotide Repeat Disorders

Anticipation
Increasing severity with subsequent generations
Gender of transmitting parent may influence disease severity
Disorders may have variable age at onset, degree of severity
May see skipping of generations (transmission of premutation)

Synergistic Heterozygosity

Described in inborn errors of metabolism
Variability in severity of symptoms among affected family members
Complex phenotypes, multisystem involvement
Multiple partial enzyme deficiencies in affected individuals
Environmental factors may influence severity of disease

Multifactorial Inheritance

Males and females affected
Gender of affected individual influences recurrence risk
Classically, few affected family members, but now also implicated in common adult-onset disorders
Degree of relationship to affected individual influences recurrence risk
Recurrence risk correlates with number of affected family members

daughter cells may inherit only mutant mitochondrial DNA (homoplasmy), a percentage of mutant mitochondrial DNA (heteroplasmy), or no mutant mitochondrial DNA. The degree of heteroplasmy affects the overall function of the cell or population of cells and thus correlates with disease severity. It is not possible to predict for any given cell what the degree of heteroplasmy will be; thus, it is extremely difficult to predict recurrence risk or severity of disease. Furthermore, different cell populations in different organs can have different degrees of heteroplasmy, yielding a variable multisystem disease (pleiotropy).

Imprinting

Imprinting refers to differential expression of genes depending on the parent of origin. The process is reversible, as it affects the action of the gene but not the gene structure; genes that are passed from a male (imprinted as male) to a female and then passed by the female are reimprinted as female, and so on. This is thought to occur early in development, most likely in the germ cells.[14] A number of disorders have been described that are caused by imprinting. Depending on the underlying mechanism and assuming transmission from the critical parent of origin, recurrence risks could be as high as 50%, particularly if a mutation exists in an imprinting control center that regulates methylation status and, thus, gene expression.

Uniparental Disomy

Uniparental disomy is defined as both copies of all or part of a chromosome in a cell or individual being derived from only one parent. This can appear as heterodisomy (the presence of copies of both of one parent's chromosomes) or homodisomy (a single chromosome or chromosome segment present in two identical copies). This becomes clinically relevant when males and females differentially imprint the chromosomal segment in question, or when the parent who transmits the disomic region carries a mutation in that region.[14] This process has been seen in cystic fibrosis, Prader-Willi and Angelman syndromes and other disorders, and may need to be considered for any autosomal recessive disorder when only one parent is a confirmed carrier, and for X-linked recessive disorders occurring in 46, XX females. The frequency of this phenomenon is unknown.

Trinucleotide Repeat Disorders

Most classic hereditary disorders are caused by static or stable mutations in one or a few genes. For trinucleotide repeat disorders, alterations in the causative gene are unstable, called dynamic mutations, and characterized by a variable number of copies of a tandemly repeated three-nucleotide sequence within the gene.[14] These trinucleotide repeats are normal, do not generally cause disease, and can be inherited stably within certain, usually small, tandem repeat size ranges that are gene specific.

Due to the structure of the repeated gene sequence, however, miscopying during DNA replication can occur, leading to expansion (creation of additional tandem copies of the trinucleotide sequence) or, rarely, contraction (loss of one to five copies of the trinucleotide sequence) of the gene segment. With expansion, the gene segment becomes less stable and thus more likely to expand further. Intermediate lengths of expanded gene segment are called premutations, which are extremely unstable and highly likely to undergo further expansion. Individuals who carry premutations typically do not have symptoms of the associated disorder but may show mild signs or develop associated problems at later ages.

Once the gene segment has expanded into the disease-associated repeat size range, disease symptoms occur in the individual. Degree of disease severity typically correlates with the size of the repeated segment, with earlier age of onset and more severe symptoms with increasing repeat size. The clinical phenomenon of anticipation (earlier onset of disease in subsequent generations) is explained mechanistically by the progressive expansion of the trinucleotide repeat region from one generation to the next, with earlier and more severe disease for each generation. Gender of transmitting parent also influences likelihood and degree of expansion, and is gene specific (the significant parent of origin varies by disease).

Synergistic Heterozygosity

A phenomenon described primarily to date in inborn errors of metabolism, synergistic heterozygosity results from relative decreases in function in several components of a complex biological pathway.[15] Effects of mutations in a single copy of each of multiple genes encoding components of a pathway accumulate and lead to an overall decrease in function of the pathway. This is much more akin to multifactorial or at least polygenic inheritance than classical Mendelian inheritance typical of the majority of inborn errors of metabolism. Recurrence risks depend on the degree of decreased function of each of the components, which components are involved, genetic linkage of the components, or the potential for environmental influences on the pathway, or some combination of these factors.

Multifactorial Inheritance

Multifactorial disorders are the result of interactions among multiple genetic and environmental factors. A threshold effect defines the likelihood of disease based on the relative contributions of each of the factors involved. With a relatively low concentration of contributing factors, no effect will be seen. However, above a critical cutoff of accumulated factors, the condition occurs. Risk to relatives of affected individuals increases as more family members are affected, presumably reflecting the presence of a higher "dose" of critical factors in the family or shared environmental factors. The threshold for affected status

may, however, be different in males and females. In classic conditions, such as pyloric stenosis or neural tube defects, a higher dose of risk factors is needed to push the less-frequently affected gender above the critical threshold; close relatives are therefore more likely to have a similar clustering of risk factors and be above the threshold, particularly if they are of the more commonly affected gender and thus are presumed to have a lower threshold.

Other Factors Affecting Risk and Risk Assessment

Classic and atypical modes of inheritance provide a framework for assessment of risk to close relatives of individuals affected by hereditary disorders. However, many factors influence the ability to clearly define patterns of inheritance in families. From a logistical perspective, family members may not know details about medical conditions in more distant relatives, or relatives may not wish to share those details by medical record request. For some, there may be stigma or guilt attached to discussion of hereditary conditions in themselves or their children. Mechanistically, there are a number of processes that may confound pedigree interpretation (Table 4-4). Variable expressivity and pleiotropy relate, respectively, to the presence of different degrees of severity of symptoms and the presence of varying phenotypic features in affected individuals. These could lead to misclassification of affected status, or failure to recognize the presence of a single clinical entity in affected family members. Further, variability in age of onset, particularly with adult-onset disease, may leave gaps in an otherwise classic pedigree, as can penetrance, or the likelihood that an individual who carries the gene(s) for a condition will show signs or symptoms of that condition. Some conditions show genetic heterogeneity, that is, can be caused by mutations in a number of different genes. While mutations in these genes may be rare, theoretically more than one type of gene mutation could lead to symptoms within a family. Phenocopies, similar conditions with different genetic or nongenetic etiologies or both, may also occur within a family and lead to misinterpretations of patterns of inheritance and, thus, of risk to family members. Small family size or relatively low frequency of the at-risk gender in gender-influenced disorders (sex-limited vs. sex-influenced expression) may result in failure to recognize a hereditary disorder and underestimation of risk.

Table 4-4. Factors Affecting Risk and Risk Assessment

Variable expressivity/pleiotropy
Age of onset
Penetrance
Heterogeneity
Phenocopies
Gender-influenced expression (sex-limited vs. sex-influenced)
Family size/paucity of at-risk gender
Nonpaternity
Consanguinity/inbreeding
Lyonization
New mutation
Mosaicism (somatic vs. germline)
Modifying genes
Environmental effects

Accurate reporting of relationships within a pedigree is critical. Nonpaternity, estimated at 10% in the United States,[4] and consanguinity, or inbreeding (shared common ancestors), are particularly important when considering possible autosomal recessive traits. A recent review confirmed only modest contribution of consanguinity to overall risk,[16] but it can be of critical importance when an autosomal recessive disorder is under consideration in a symptomatic individual due to potential presence of shared nonfunctional genes in related parents. Expression and risk assessment of X-linked disorders are influenced by lyonization, or the random inactivation of one X chromosome in each cell in a female. The percentages of the active and inactive nonfunctional X-chromosome gene could lead to full expression, intermediate symptomatology, or lack of symptoms altogether for an X-linked recessive condition in a female. The occurrence of spontaneous new mutations could lead to failure to recognize risk due to autosomal dominant or X-linked conditions, in particular. Similarly, mosaicism, or the presence of a mixture of at least two populations of cells with some containing a functional and others a nonfunctional gene, could lead to partial expression of a condition in an individual (somatic mosaicism). Mosaicism also could lead to unrecognized or indefinable risk to future offspring if only the germ cells (egg or sperm) are affected or only a percentage of germ cells are affected (germline mosaicism).

Finally, factors outside of the critical gene can influence the expression of that gene and thus the assessment of risk. Expression of some genes is influenced by variant forms of other, so-called modifying genes. Polymorphisms or mutations in these modifying genes can change gene-gene or protein-protein interactions to affect expression of the condition in question. Similarly, environmental factors such as shared environment, dietary practices, and specific exposures (medications, smoke, etc.) may positively or negatively affect gene function or expression of clinical symptoms.

Each of these factors must be carefully considered in the overall diagnostic and risk assessment, initially based on collection of a family pedigree and continued through clinical evaluation, including physical examination and indicated diagnostic testing.

Direct Mutation Analysis and Linkage Analysis

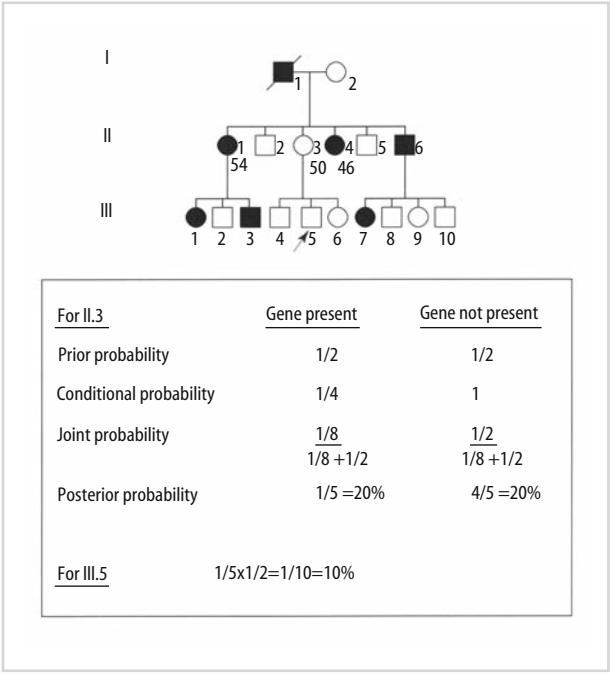
The ability to define mutations or gene regions associated with disease removes much of the art of risk assessment from evaluation of the pedigree and provides a more

definitive answer in many cases. Currently, there are clinical or research tests being done for 1,269 different diseases (http://www.genetests.org, accessed on June 14, 2006), which continues to increase as definitive mutations are identified in newly described disease-associated genes. Methods of gene analysis vary among different laboratories (see chapter 2 and Reference 17). For large deletions and gene rearrangements, Southern blot analysis is used. Dosage analysis (determination of gene copy number utilizing densitometry, multiplex ligation-dependent probe amplification [MLPA] or similar techniques) may be used in cases where an affected individual is not available for study and deletion is a common form of mutation, as is the case in Duchenne muscular dystrophy. Southern blot analysis also may be needed for sizing of large trinucleotide repeats, while smaller repeats can be identified by targeted polymerase chain reaction (PCR) analysis. PCR is also used in conjunction with allele-specific oligonucleotides (ASOs) for analysis of conditions with a single or few common mutations. In disorders where many unique, private mutations have been found, mutation-screening techniques may be utilized, including conformation sensitive gel electrophoresis (CSGE), denaturing gradient gel electrophoresis (DGGE), denaturing high-performance liquid chromatography (DHPLC), two-dimensional gel scanning (TDGS), and single-strand conformation polymorphism (SSCP). Once gene segments with probable variants have been identified by these techniques, DNA sequencing is utilized to verify the presence of a mutation, polymorphism, or variant of unknown significance.

These direct methods of identifying mutations are invaluable when the disease-associated gene is known; however, historically and even today, for many conditions the causative gene has not been identified or is not characterized adequately to allow for mutation-specific testing. In these situations, it is possible to offer an indirect testing method, called linkage analysis, to clarify the risk status of family members if the responsible gene has been localized to a specific genomic map location. For some families, the most significant issue in linkage analysis is the need for specimens from a number of family members, both affected and unaffected, to ensure useful interpretation of results. Linkage analysis requires that the clinical status of the relatives and their relationships be accurately reported for accurate interpretation. Linkage analysis involves determination of "markers" for the disease gene, often variant forms of highly polymorphic short tandem repeats, within or near the genomic map location. It requires that a marker or markers near the genetic locus be informative; that is, key individuals in the family must be heterozygous, or have two different forms of the marker (alleles) at the locus in question. These markers are then tracked as they are passed from one individual to the next. It is thus essential to know which allele(s) is associated with the disease gene and which alleles track with the normal gene copy (setting phase). This typically involves analysis of DNA from a number of affected family members or a carefully selected group of affected and unaffected relatives. In addition, if the marker(s) is closely associated (linked) with the disease gene or is within the gene, presence of the disease-associated marker allele will correlate with presence of the disease gene. If the marker is genetically distant from the disease gene, it may become separated from the disease gene through recombination, and predictions about gene transmission may be inaccurate. Caution is required when doing linkage analysis of very large genes because a marker at one end of the gene may, through recombination, become unlinked from the (unknown) mutation if it resides at the other end of the gene. Accuracy of linkage analysis can be further increased by assessing more than one linked marker, preferably within or flanking opposite ends of the gene.

Bayesian Analysis Used in Risk Modification

When collected pedigrees are used to provide risk assessment, a variety of data may be relevant to the overall assessment. Numerous factors, some listed above, influence the likelihood that a given individual in the family may be affected by the condition in question or may be a carrier of the gene in question. When it is not possible to do direct diagnostic testing for the condition (for example, if the causative gene is unknown), when the affected relative is not available for testing, or for complex traits, it is possible to combine incremental contributors to risk by utilizing Bayesian analysis. Bayesian analysis is a statistical construct that uses information about the likelihood of occurrence of past events or conditions, and the current status of those events or conditions for the individual, to predict the likelihood of a future event or condition, in this case, the presence or absence of a particular gene or genetic condition.[18] Some factors that may be considered in genetic risk assessment using Bayesian analysis include number and pattern of affected and unaffected family members, laboratory data, and natural history of the condition. The probability assigned based on past events is called the prior probability; that based on current information or observations is called the conditional probability. The calculated probability for each possible outcome of an event or condition is the joint probability, and the final probability of one outcome as a percentage of all possible outcomes is the posterior probability. Calculations often utilize data from multiple generations and are usually done in tabular form. In the example pedigree in Figure 4-7 for an autosomal dominant cancer predisposition syndrome affecting males and females equally, based on Mendelian inheritance alone, the risk that individual III.5 is a gene carrier is 25%. However, knowing that 75% of gene carriers have been diagnosed with cancer by age 50, risk can be recalculated as demonstrated. (See chapter 5 for a complete discussion of Bayesian analysis.)



For II.3	Gene present	Gene not present
Prior probability	1/2	1/2
Conditional probability	1/4	1
Joint probability	1/8 / 1/8 +1/2	1/2 / 1/8 +1/2
Posterior probability	1/5 =20%	4/5 =20%

For III.5 1/5x1/2=1/10=10%

Figure 4-7. Bayesian analysis for risk assessment in an autosomal dominant, adult-onset hereditary cancer disorder. Ages of selected individuals in generation II are shown below the pedigree symbols.

References

1. Resta R. The crane's foot: the rise of the pedigree in human genetics. *J Genet Couns.* 1993;2:235–260.
2. Bennett R, Steinhaus K, Uhrich S, O'Sullivan C. The need for developing standardized family pedigree nomenclature. *J Genet Couns.* 1993;2:261–273.
3. Bennett R, Steinhaus K, Uhrich S, et al. Recommendations for standardized human pedigree nomenclature. Pedigree Standardization Task Force of the National Society of Genetic Counselors. *Am J Hum Genet.* 1995;56:745–752.
4. Bennett R. *The Language of the Pedigree: The Practical Guide to the Genetic Family History*. New York: Wiley-Liss; 1999:1–12.
5. Daly M, Farmer J, Harrop-Stein C, et al. Exploring family relationships in cancer risk counseling using the genogram. *Cancer Epidemiol Biomarkers Prev* 1999;8:393–398.
6. Botkin J. Protecting the privacy of family members in survey and pedigree research. *JAMA.* 2001;285:207–211.
7. American Society of Human Genetics Policy Papers and Reports. *Should Family Members About Whom You Collect Only Medical History Information for Your Research Be Considered "Human Subjects"?* Amercian Society of Human Genetics, Bethesda, MD, 2000. http://www.ashg.org/genetics/ashg/pubs/policy/pol-38.htm
8. US Department of Health and Human Services. *Medical Privacy—National Standards to Protect the Privacy of Personal Health Information*. US Department of Health and Human Services; 2003.
9. ASHG statement. Professional disclosure of familial genetic information. The American Society of Human Genetics Social Issues Subcommittee on Familial Disclosure. *Am J Hum Genet.* 1998;63:898–900.
10. Vogel F, Motulsky A. *Human Genetics: Problems and Approaches.* Berlin: Springer-Verlag; 1997.
11. McKusick V. Online Mendelian Inheritance in Man, OMIM [database online]. Baltimore, MD: McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University; Bethesda, MD: National Center for Biotechnology Information, National Library of Medicine, 1966–2005.
12. Tommerup N. Mendelian cytogenetics: chromosome rearrangements associated with Mendelian disorders. *J Med Genet.* 1993;30:713–727.
13. Schmickel R. Contiguous gene syndromes: a component of recognizable syndromes. *J Pediatr.* 1986;109:231–241.
14. Nussbaum R, McInnes R, Willard H. *Thompson & Thompson Genetics in Medicine.* Philadelphia: W.B. Saunders; 2001.
15. Vockley J, Rinaldo P, Bennett M, Matern D, Vladutiu G. Synergistic heterozygosity: disease resulting from multiple partial defects in one or more metabolic pathway. *Mol Genet Metab.* 2000;71:10–18.
16. Bennett R, Motulsky A, Bittles A, et al. Genetic counseling and screening of consanguineous couples and their offspring: recommendations of the National Society of Genetic Counselors. *J Genet Couns.* 2002;11:97–120.
17. Beaudet A, Scriver C, Sly W, Valle D. Genetics, biochemistry and molecular basis of variant human phenotypes. In: Beaudet A, Scriver C, Sly W, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease.* Vol. 1. New York: McGraw-Hill, 2001:14–42.
18. Young I. *Genetic Counseling and the Laws of Probability. Introduction to Risk Calculation in Genetic Counseling.* Oxford: Oxford University Press; 1999:1–14.

Chapter 5

Bayesian Analysis

Shuji Ogino and Robert B. Wilson

Introduction

The purpose of this chapter is to describe basic and general principles of Bayesian analysis for molecular pathologists. Thomas Bayes first described the theorem named after him in an essay on "the doctrine of chances," published posthumously in 1763, and republished in 1958.[1] Analyses based on Bayes' theorem are routinely applied to calculate probabilities in a wide variety of circumstances, not limited to medicine or genetics. In molecular pathology, Bayesian analysis is commonly used to calculate genetic risk, incorporating population data, pedigree information, and genetic testing results. First, Bayesian analysis will be introduced with two simple, concrete examples. In subsequent sections, the general principles illustrated by these examples are discussed and applied to more complex scenarios. For more in-depth treatments, the reader is referred to *Introduction to Risk Calculation in Genetic Counseling* by Young[2] and *The Calculation of Genetic Risks* by Bridge[3] as well as several articles on genetic risk assessment that include advanced Bayesian analyses, particularly for spinal muscular atrophy (SMA)[4,5] and cystic fibrosis (CF).[6–9]

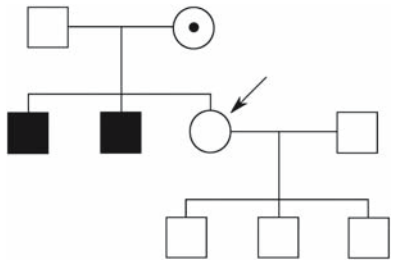
Bayesian Analysis Using Pedigree Information

In the pedigree shown in Figure 5-1a, the two brothers of the consultand (indicated by the arrow) have Kennedy disease (X-linked spinal and bulbar muscular atrophy; Online Mendelian Inheritance in Man [OMIM; database online] #300377), which is caused by a CAG trinucleotide expansion in the androgen receptor (*AR*) gene (OMIM #310200). Because both of the consultand's brothers are affected, we can assume that the consultand's mother is an obligate carrier. Before taking into account her three unaffected sons, the consultand's carrier risk is 1/2, since there is a 1/2 chance that she inherited the mutant X chromosome from her mother. If we take into account that the consultand has three unaffected sons, how does her carrier risk change?

Bayesian analysis starts with mutually exclusive hypotheses. In this example, there are two: that the consultand is a carrier, and that the consultand is a noncarrier. Setting up a table with separate columns for each hypothesis facilitates Bayesian analyses, as shown in Figure 5-1b for this case. The first row of the table comprises the prior probability for each hypothesis. In this example, the prior probabilities are the probability that the consultand is a carrier (1/2), and the probability that she is a noncarrier (also 1/2), *prior* to taking into account the subsequent information that she has three unaffected sons.

The second row of the table comprises the conditional probability for each hypothesis. The conditional probability for each hypothesis is the probability that the subsequent information would occur if we assume that each hypothesis is true. In this example, the subsequent information is that the consultand has three unaffected sons. Thus, the conditional probabilities are the probability that the consultand would have three unaffected sons under the assumption (or condition) that she is a carrier, and the probability that she would have three unaffected sons under the assumption (or condition) that she is a noncarrier. If we assume that she is a carrier, the probability that she would have three unaffected sons is $1/2 \times 1/2 \times 1/2 = 1/8$. This is because she would have to have passed the normal X chromosome three times in succession, each time with a probability of 1/2. If we assume that she is a noncarrier, the probability that she would have three unaffected sons approximates 1, since only in the event of a rare de novo mutation would a noncarrier have an affected son. Thus, the conditional probabilities in this example are 1/8 and 1 (Figure 5-1b).

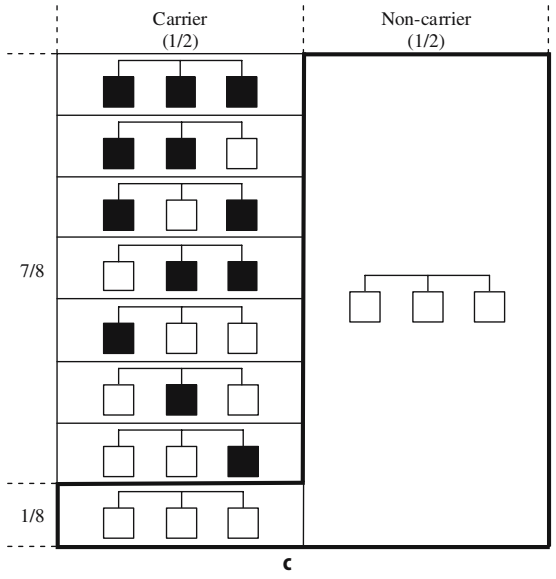
The third row of the table comprises the joint probability for each hypothesis, which is the product of the prior and conditional probabilities for each hypothesis. For the first hypothesis in this example, that the consultand is a carrier, the joint probability is the prior probability that she is a carrier, multiplied by the conditional probability that a



a

	Hypothesis	
	Carrier	Non-carrier
Prior Probability	1/2	1/2
Conditional Probability (of three normal sons)	1/8	~1
Joint Probability	1/16	1/2
Posterior Probability	(1/16) / (1/16 +1/2) = 1/9	(1/2) / (1/16 + 1/2) = 8/9

b



c

Figure 5-1. (a) Pedigree of a family with individuals affected with Kennedy disease (see text). Consultand is indicated by an arrow. (b) Bayesian analysis for the consultand in Figure 5-1a. (c) Schematic representation of the Bayesian analysis of Figure 5-1b. Pedigrees shown in the rectangles represent all possible disease status outcomes for the third generation of the pedigree in Figure 5-1a, given the carrier or noncarrier status of the consultand. Each small rectangle to the left represents one sixteenth of the total area. (See text for full description.)

carrier would have three normal sons, which in this case is $1/2 \times 1/8 = 1/16$ (Figure 5-1b). For the second hypothesis in this example, that the consultand is a noncarrier, the joint probability is the prior probability that she is a noncarrier, multiplied by the conditional probability that a noncarrier would have three normal sons, which in this case is $1/2 \times 1 = 1/2$ (Figure 5-1b).

The fourth row of the table comprises the posterior probability for each hypothesis. The posterior probability for each hypothesis is the probability that each hypothesis is true after (or posterior to) taking into account both prior and subsequent information. The posterior probability for each hypothesis is calculated by dividing the joint probability for that hypothesis by the sum of all the joint probabilities. In this example, the posterior probability that the consultand is a carrier is the joint probability for the first hypothesis (1/16), divided by the sum of the joint probabilities for both hypotheses ($1/16 + 1/2 = 9/16$), or $1/16 \div 9/16 = 1/9$. The posterior probability that the consultand is a noncarrier is the joint probability for the second hypothesis ($1/2 = 8/16$), divided by the sum of the joint probabilities for both hypotheses ($1/16 + 1/2 = 9/16$), or $8/16 \div 9/16 = 8/9$. Thus, taking into account the prior family history, and the subsequent information that the consultand has three unaffected sons, the probability that the consultand is a carrier is 1/9 (Figure 5-1b).

The preceding example is illustrated graphically in Figure 5-1c. The total area represents the total prior probabilities. The left half represents the prior probability that the consultand is a carrier (1/2), and the right half represents the prior probability that she is a noncarrier (also 1/2). Under the hypothesis that the consultand is a carrier, there are eight possibilities, comprising all the permutations of zero, one, two, or three affected sons. The area of the small rectangle that contains three unshaded squares (for three unaffected sons) comprises one eighth of the left half and represents the conditional probability of three normal sons under the hypothesis that the consultand is a carrier. The area of this small rectangle is one sixteenth of the total area and therefore also represents the joint probability that the consultand is a carrier (1/2), and that as a carrier she would have three normal sons (1/8), or $1/2 \times 1/8 = 1/16$.

Under the hypothesis that the consultand is a noncarrier, there is essentially only one possibility, which is that all three sons are unaffected. The area of the larger rectangle that contains the pedigree with three unshaded squares (for three unaffected sons) comprises all of the noncarrier half and represents the conditional probability of three normal sons under the hypothesis that the consultand is a noncarrier. The area of this larger rectangle is one half of the total area and therefore also represents the joint probability that the consultand is a noncarrier (1/2), and that as a noncarrier she would have three normal sons (~1), or $1/2 \times 1 = 1/2$. The reverse-L-shaped box, which is demarcated by a bold line, represents the sum of the joint probabilities, or nine sixteenths of the total area.

Because the consultand has three unaffected sons, the area of the reverse-L-shaped box represents the only component of the prior probabilities needed to determine the posterior probability that the consultand is a carrier. Taking into account that all three of the consultand's sons are unaffected, Bayesian analysis allows us to *exclude* 7/16 of the prior probabilities, those that include one or more affected sons, from consideration. (Note that this explains why the joint probabilities sum to less than 1.) The posterior probability that the consultand is a carrier is therefore the area of the small rectangle with three unshaded squares (for three unaffected sons) divided by the area of the entire reverse-L-shaped box, which represents the only probabil-

ities relevant to the consultand's risk, or 1/16 ÷ 9/16 = 1/9. Likewise, the posterior probability that the consultand is a noncarrier is the area of the larger rectangle with three unshaded squares (for three unaffected sons) divided by the area of the entire reversed-L-shaped box, or 8/16 ÷ 9/16 = 8/9.

Bayesian Analysis Using Genetic Test Results

In the second example, information from a test result modifies the prior risk. In the pedigree shown in Figure 5-2a, the consultand is pregnant with her first child and has a family history of CF (OMIM #219700). CF is caused by mutations in the cystic fibrosis transmembrane conductance regulator gene (*CFTR*; OMIM #602421). The consultand is an unaffected European Caucasian and her brother died years earlier of complications of CF. She undergoes carrier testing for the 23 mutations recommended by the American College of Medical Genetics (ACMG) CF screening guidelines,[10–12] which detects approximately 90% of disease alleles in European Caucasians. The consultand tests negative for all 23 mutations. What is her carrier risk after testing?

As in the first example, the two hypotheses are that the consultand is a carrier and that she is a noncarrier. The prior probability that she is a carrier is 2/3. Because the consultand is unaffected, she could not have inherited disease alleles from both parents. Thus, she either inherited a disease allele from her mother or father, or she inherited only normal alleles; in two of these three scenarios she would be a carrier (shown in Figure 5-2b). The prior probability that the consultand is a noncarrier is 1/3 (Figure 5-2c).

As in the first example, the conditional probability for each hypothesis is the probability that the subsequent information would occur if we assume that each hypothesis is true. In this example, the subsequent information is that the consultand tests negative for all 23 mutations. Thus, the conditional probabilities are the probability that the consultand would test negative under the assumption (or condition) that she is a carrier, and the probability that she would test negative under the assumption (or condition) that she is a noncarrier. If we assume that she is a carrier, the probability that she would test negative is 1/10, since the test detects 90% of European Caucasian disease alleles or carriers. If we assume that she is a noncarrier, the probability that she would test negative approximates 1. Thus, the conditional probabilities in this example are 1/10 and 1 for the carrier and noncarrier hypotheses, respectively (Figure 5-2c).

As in the first case, the joint probability for each hypothesis is the product of the prior and conditional probabilities for that hypothesis. For the first hypothesis in this example, that the consultand is a carrier, the joint probability is the prior probability that she is a carrier (2/3) multiplied by the conditional probability that a carrier of European Caucasian ancestry would test negative (1/10), or 2/3 × 1/10 = 1/15 (Figure 5-2c). For the second hypothesis in this example, that the consultand is a noncarrier, the

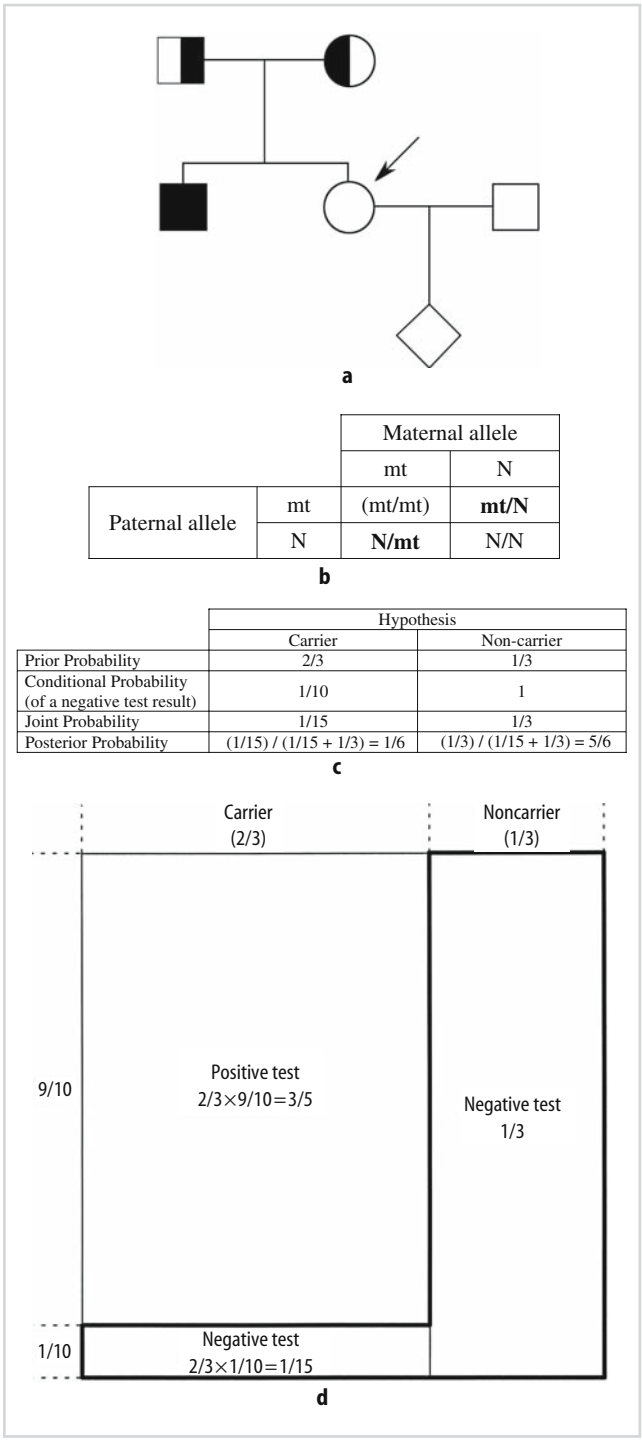


Figure 5-2. (a) Pedigree of a family with an individual affected with CF (see text). Consultand is indicated by an arrow. (b) Possible genotypes of the sibling (consultand in this case) of the affected child prior to genetic testing. The mt/mt genotype (in parentheses) is excluded based on the fact that the consultand is unaffected. mt, mutant; N, normal. (c) Bayesian analysis for the consultand in Figure 5-2a. (d) Schematic representation of the Bayesian analysis of Figure 5-2c (see text).

joint probability is the prior probability that she is a noncarrier (1/3) multiplied by the conditional probability that a noncarrier would test negative (1), or $1/3 \times 1 = 1/3$ (Figure 5-2c).

Finally, the posterior probability is calculated for each hypothesis by dividing the joint probability for that hypothesis by the sum of all the joint probabilities. In this example, the posterior probability that the consultand is a carrier and tests negative for 23 CF mutations is the joint probability for the first hypothesis (1/15) divided by the sum of the joint probabilities for both hypotheses ($1/15 + 1/3 = 2/5$), or $1/15 \div 2/5 = 1/6$ (Figure 5-2c). The posterior probability that the consultand is a noncarrier and tests negative for 23 CF mutations is the joint probability for the second hypothesis (1/3) divided by the sum of the joint probabilities for both hypotheses (2/5), or $1/3 \div 2/5 = 5/6$ (Figure 5-2c).

The preceding example is illustrated graphically in Figure 5-2d. The total area represents the total prior probabilities. The left two thirds represents the prior probability that the consultand is a carrier, and the right third represents the prior probability that the consultand is a noncarrier. Under the hypothesis that the consultand is a carrier, there are two possibilities for the test result: positive or negative. The area of the small rectangle on the lower left comprises one tenth of the 2/3 carrier region of the figure and represents the conditional probability of a normal test result under the hypothesis that the consultand is a carrier. The area of this small rectangle is $1/10 \times 2/3 = 1/15$ of the total probabilities area and therefore also represents the joint probability that the consultand is a carrier (2/3) and that as a European Caucasian carrier she would test negative for all 23 mutations (1/10), or $2/3 \times 1/10 = 1/15$ (Figure 5-2d).

Under the hypothesis that the consultand is a noncarrier, there is essentially only one possibility for the test result, which is negative. The area of the rectangle that comprises all of the 1/3 noncarrier region represents the conditional probability of a negative test result under the hypothesis that the consultand is a noncarrier. The area of this rectangle is one third of the total area and therefore also represents the joint probability that the consultand is a noncarrier (1/3), and that as a noncarrier she would test negative (~1), or $1/3 \times 1 = 1/3$. The reverse-L-shaped box, which is demarcated by a bold line, represents the sum of the joint probabilities, or 2/5 (= 1/3 + 1/15) of the total area.

Because the consultand tested negative, the area of the reverse-L-shaped box represents the only component of the prior probabilities needed to determine the posterior probability that the consultand is a carrier. Taking into account that she tested negative, Bayesian analysis allows us to *exclude* 3/5 of the prior probability, that portion comprising a positive test result, from consideration. (Note, again, that this explains why the joint probabilities sum to less than 1.) The posterior probability that the consultand is a carrier is therefore the area of the small rectangle at the lower left divided by the area of the reverse-L-shaped box, which represents the only probabilities relevant to the consultand's risk, or $1/15 \div 2/5 = 1/6$. Likewise, the posterior probability that the consultand is a noncarrier is the area of the larger rectangle on the right divided by the area of the reverse-L-shaped box, or $1/3 \div 2/5 = 5/6$.

Simple Bayesian Analyses Generalized: Carrier Versus Noncarrier

The preceding Bayesian analyses can be generalized as in Table 5-1. Note that if the correct prior and conditional probabilities can be determined, the rest is simple calculation. Setting up a spreadsheet, as in Table 5-1, facilitates clinical Bayesian analyses.

A very common application of Bayesian analysis in molecular pathology is to calculate carrier risk after a negative test result, as in the second example, above. The need to calculate carrier risk in this scenario stems from the fact that the sensitivity of most carrier tests is, at present, less than 100%; therefore, a negative test result decreases, but does not eliminate, carrier risk. Hypothesis 1 in this scenario is that the consultand is a carrier, and Hypothesis 2 is that the consultand is a noncarrier (Table 5-1). The prior carrier probability ("A" in Table 5-1) depends on whether there is a family history, and if there is, on the relationship of the consultand to the affected family member as shown by the family pedigree. In the absence of a family history, the prior carrier probability is the population carrier risk for that disease. In the case of CF and some other diseases, the appropriate population risk depends on the ethnicity of the consultand. The conditional probabilities ("C" and "D" in Table 5-1) are 1 minus the test sensitivity, and the test specificity, respectively. The remainder of the table is completed through calculation, with the posterior probabilities ("G" and "H" in Table 5-1) representing 1 minus the negative predictive value, and the negative predictive value, respectively. This is shown schematically in Figure 5-3.

For illustration, suppose in the second example above (Figure 5-2) that the consultand's husband is Ashkenazi Jewish, that he has no family history of CF, and that he tests negative for all 23 mutations in the ACMG screening guidelines panel. What is his carrier risk? The carrier risk in Ashkenazi Jewish populations, and therefore the husband's prior carrier risk in the absence of a family history, is approximately 1/25 ("A" in Table 5-1). Thus, his prior probability of being a noncarrier is 24/25 ("B" in Table 5-1). The

Table 5-1. Simple Bayesian Analysis Generalized

	Hypothesis	
	1	2
Prior probability	A	B = 1 – A
Conditional probability	C	D
Joint probability	E = AC	F = BD
Posterior probability	G = E ÷ (E + F)	H = F ÷ (E + F)

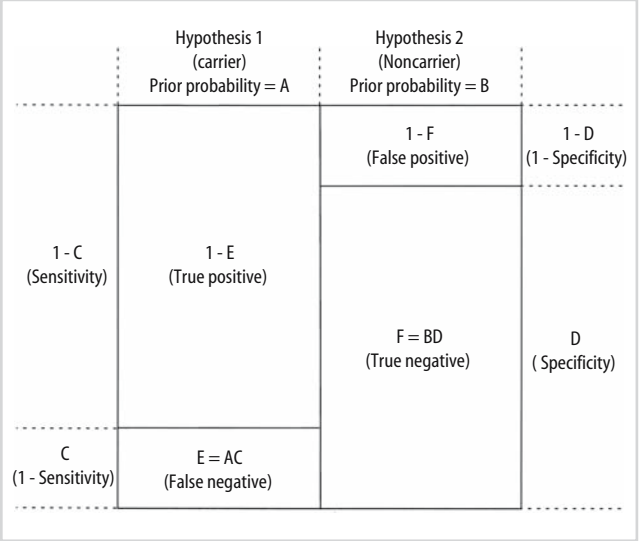


Figure 5-3. Schematic representation of the generalized Bayesian analysis shown in Table 5-1, for the case of a negative carrier test. The small rectangles represent true-positive, false-positive, true-negative, and false-negative rates for a particular consultand; that is, the prior probabilities are influenced by factors such as family history or signs and symptoms, and the sensitivity and specificity of the test are influenced by factors such as ethnicity. For a negative carrier test, the posterior carrier probability (1 minus the negative predictive value) is the false-negative rate divided by the sum of the false- and true-negative rates, or E ÷ (E + F).

ACMG screening guidelines panel of 23 mutations detects 94% of CF mutations in Ashkenazi Jewish populations,[10–12] so the conditional probability of a negative test, under the hypothesis that he is a carrier, is 6% = 3/50 ("C" in Table 5-1). Under the hypothesis that he is a noncarrier, the conditional probability of a negative test approximates 1 ("D" in Table 5-1). (This is generally the case in genetic testing, since noncarriers by definition lack mutations in the relevant disease gene and, hence, unless there are technical problems, essentially always should test negative.) The Bayesian analysis table for this example is shown in Table 5-2. The joint probabilities are the products of the prior and conditional probabilities ("E" and "F" in Table 5-1), and the posterior probabilities ("G" and "H" in Table 5-1) derive from each joint probability divided by the sum of the joint probabilities. The husband's posterior carrier risk after the negative test result is 1/401 (Table 5-2).

What is the risk that the fetus of the mother (consultand) in Figure 5-2 and the father from Table 5-2 is affected with CF? Prior to testing, the risk was the prior probability that the mother was a carrier (2/3), multiplied by the prior probability that the father was a carrier (1/25), multiplied by the probability that the fetus would inherit two disease alleles (1/4), or $2/3 \times 1/25 \times 1/4 = 1/150$. After testing, the risk is the posterior probability that the mother is a carrier (1/6), multiplied by the posterior probability that the father is a carrier (1/401), multiplied by the probability that the fetus would inherit two disease alleles (1/4), or $1/6 \times 1/401 \times 1/4 \cong 1/9600$.

Often, testing is performed on additional family members, and genetic risks need to be modified accordingly. In the example above, testing of both parents of the mother (consultand) would affect her carrier risk calculations. Detection of mutations in both parents using the same mutation test panel would essentially rule out carrier status for the mother, since we would then know that the sensitivity of the test for the mutations she is at risk of carrying is essentially 100%. Alternatively, if the test results for the mother's parents are positive for only one of her parents (for example, her father) and negative for the other parent (her mother), then the sensitivity of the test for the mutations she is at risk of carrying is essentially 50%. The Bayesian analysis for the mother, modified from Figure 5-2c, is shown in Table 5-3a. The conditional probability of a negative test under the hypothesis that she is a carrier has changed from 1/10 to 1/2, which increases the posterior probability that she is a carrier to 1/2. Taken together with her husband's carrier risk of 1/401 (Table 5-2), the risk that

Table 5-2. Bayesian Analysis for an Ashkenazi Jewish Individual Without a Family History of CF Who Tests Negative for the ACMG Screening Guidelines Panel of 23 *CFTR* Mutations

	Hypothesis	
	Carrier	Noncarrier
Prior probability	1/25	24/25
Conditional probability (of negative test result)	3/50	1
Joint probability	3/1250	24/25
Posterior probability	(3/1250) ÷ (3/1250 + 24/25) = 1/401	(24/25) ÷ (3/1250 + 24/25) = 400/401

Table 5-3a. Bayesian Analysis for the Consultand in Figure 5-2a After Testing of the Parents (see text)

	Hypothesis	
	Carrier	Noncarrier
Prior probability	2/3	1/3
Conditional probability (of negative test result)	1/2	1
Joint probability	1/3	1/3
Posterior probability	(1/3) ÷ (1/3 + 1/3) = 1/2	(1/3) ÷ (1/3 + 1/3) = 1/2

Table 5-3b. Alternative Bayesian Analysis for the Consultand in Figure 5-2a After Testing of the Parents (see text)

	Hypothesis		
	Carrier		
	Carrier with Paternal (detectable) Mutation	Carrier with Maternal (undetectable) Mutation	Noncarrier
Prior probability	1/3	1/3	1/3
Conditional probability (of negative test result)	0	1	1
Joint probability	0	1/3	1/3
Posterior probability	0	1/2	1/2

the fetus is affected with CF can be modified to $1/2 \times 1/401 \times 1/4 \cong 1/3200$.

Another way of conceptualizing the Bayesian analysis described above is to separate the carrier hypothesis into two subhypotheses, as shown in Table 5-3b. (See also "Bayesian Analysis with More Than Two Hypotheses," below.) The two subhypotheses are (1) that the consultand is a carrier with a paternal (detectable) mutation and (2) that she is a carrier with a maternal (undetectable) mutation. The prior probability of each hypothesis is 1/3, that is, half of 2/3. The conditional probability of a negative test result, under the subhypothesis that she is a carrier of a detectable paternal mutation, is 0. The conditional probability of a negative test result, under the subhypothesis that she is a carrier of an undetectable maternal mutation, is 1. As in the generalized Bayesian analysis shown in Table 5-1, the joint probability for each hypothesis is the product of the prior and conditional probabilities for that hypothesis, and the posterior probability for each hypothesis is the joint probability for that hypothesis divided by the sum of all the joint probabilities. The posterior probability that the consultand has a detectable paternal mutation is 0, and the posterior probability that she has an undetectable maternal mutation is 1/2 (Table 5-3b).

Simple Bayesian Analyses Generalized: Affected Versus Unaffected

Another common application of Bayesian analysis in molecular pathology is to calculate the risk that a patient is affected with a particular disease after a negative test result. Again, the need to calculate risk in this scenario stems from the fact that the sensitivities of many genetic tests are less than 100%. Hypothesis 1 (in Table 5-1 and Figure 5-3) in this scenario is that the patient is affected, and Hypothesis 2 is that the patient is unaffected. The prior probability ("A" in Table 5-1 and Figure 5-3) usually derives mostly from signs and symptoms but also may depend on aspects of the patient's history, including family history in diseases with a genetic component. As in the CF example, above, the conditional probabilities ("C" and "D" in Table 5-1 and Figure 5-3) are 1 minus the test sensitivity, and the test specificity, respectively. The remainder of the analysis is accomplished by calculation, with the posterior probabilities ("G" and "H" in Table 5-1) representing 1 minus the negative predictive value, and the negative predictive value, respectively.

For example, suppose that a child with clinically typical spinal muscular atrophy type III (type III SMA; Kugelberg-Welander disease; OMIM #253400) tests negative for the homozygous deletion of the *SMN1* gene found in most affected individuals. What is the probability that the child is affected with *SMN1*-linked SMA? The Bayesian analysis for this scenario is shown in Table 5-4a. Wirth et al. found that 17 of 131 individuals with clinically typical type III SMA lacked mutations in both *SMN1* alleles (and therefore were considered to have diseases unrelated to *SMN1*);[13] hence, the prior probability that the child is affected with *SMN1*-linked type III SMA is 114/131, or 0.87. Approximately 6% of individuals with *SMN1*-linked type III SMA have a deletion of one *SMN1* allele and a subtle mutation, undetectable by simple polymerase chain reaction (PCR) testing for a homozygous deletion, in the other *SMN1* allele;[14] hence, the conditional probability of a negative test result under the hypothesis that the child is affected is 6/100 or 0.06. Homozygous deletions of *SMN1*, when present, are highly specific for *SMN1*-linked SMA; hence, the conditional probability of a negative test result under the hypothesis that the child is unaffected with *SMN1*-linked SMA approximates 1. Following the simple calculation rules in Table 5-1, the posterior probability that the child is affected with *SMN1*-linked type III SMA is approximately 0.29 (Table 5-4a).

Suppose that *SMN1* dosage analysis is performed on the child's DNA (i.e., the SMA carrier test), and the result is that the child has one copy of the *SMN1* gene. What is the probability that he or she is affected with *SMN1*-linked SMA? The Bayesian analysis for this scenario is shown in Table 5-4b. Again, the prior probability that the child is affected with *SMN1*-linked type III SMA is 0.87. Because approxi-

Table 5-4a. Bayesian Analysis for a Child with Clinically Typical Type III SMA Who Tests Negative for Homozygous Deletions of the *SMN1* Gene

	Hypothesis	
	Affected	Unaffected
Prior Probability	0.87	0.13
Conditional Probability (of negative test result)	0.06	~1
Joint Probability	0.052	0.13
Posterior Probability	0.29	0.71

Table 5-4b. Bayesian Analysis for a Child with Clinically Typical Type III SMA Who Has One Copy of the *SMN1* Gene by Dosage Analysis

	Hypothesis	
	Affected	Unaffected
Prior probability	0.87	0.13
Conditional probability (of 1-copy test result)	0.06	0.026
Joint probability	0.052	0.0034
Posterior probability	0.94	0.06

Table 5-4c. Bayesian Analysis for a Child with Clinically Typical Type III SMA Who Has 2 Copies of the *SMN1* Gene by Dosage Analysis

	Hypothesis	
	Affected	Unaffected
Prior probability	0.87	0.13
Conditional probability (of 2-copy test result)	0.0009	0.9
Joint probability	0.00078	0.12
Posterior probability	0.006	0.994

mately 6% of individuals with *SMN1*-linked type III SMA have a deletion of one *SMN1* allele and a subtle mutation in the other *SMN1* allele that is detectable as a single copy by dosage analysis,[14] the conditional probability of a single-copy test result under the hypothesis that the child is affected is again 0.06. However, the carrier frequency for SMA in the general population is approximately 1/38;[14] hence, in this scenario the conditional probability of a single-copy test result under the hypothesis that the child is unaffected with *SMN1*-linked SMA is 1/38 or 0.026. Following the simple calculation rules in Table 5-1, the posterior probability that the child is affected with *SMN1*-linked type III SMA is approximately 0.94 (Table 5-4b).

Suppose instead that the result of the *SMN1* dosage analysis is that the child has two copies of the *SMN1* gene. What is the probability that the child is affected with *SMN1*-linked SMA? The Bayesian analysis for this scenario is shown in Table 5-4c. Again, the prior probability that the child is affected with *SMN1*-linked type III SMA is 0.87. Only approximately 9 in 10,000 individuals with *SMN1*-linked type III SMA would be expected to have two subtle, nondeletion mutations, detectable as two gene copies by dosage analysis;[14] hence, the conditional probability of a two-copy test result under the hypothesis that the child is affected is approximately 0.0009. Because more than 7% of unaffected individuals have three copies of the *SMN1* gene, and approximately 2.5% of unaffected individuals have one copy of the *SMN1* gene, for a total of 9.5% of unaffected individuals without two copies of *SMN1*,[14] the conditional probability of a 2-copy test result under the hypothesis that the child is unaffected with *SMN1*-linked SMA is 90.5/100, or approximately 0.9. Following the simple calculation rules in Table 5-1, the posterior probability that the child is affected with *SMN1*-linked type III SMA is only approximately 0.006 (Table 5-4c).

Profiling by proteomics, RNA microarrays, or analysis of single-nucleotide polymorphisms (SNPs), or some combination of these, is likely to play an important role in molecular pathology in the future, and clinical test results will be reported, in many cases, as probabilities or relative risks. For example, suppose that a consultand has a 20% lifetime risk of developing a particular disease (based on family history, physical examination, or clinical laboratory test results, or a combination of these) and that his or her proteomic profile is 16 times more common in those who go on to develop the disease than in those who do not. What is his or her lifetime risk of developing the disease? The Bayesian analysis for this scenario is shown in Table 5-5. Hypothesis 1 (from Table 5-1) is that the consultand will develop the disease, and Hypothesis 2 is that the consultand will not develop the disease. The prior probabilities are 0.2 and 0.8 for Hypotheses 1 and 2, respectively. Because the conditional probability of the proteomic profiling result is 16 times more likely in those who develop the disease than in those who do not, the conditional probabilities ("C" and "D" in Table 5-1) are 16 and 1, respectively. Following the simple calculation rules in Table 5-1, the posterior probability that the consultand will develop the disease is 0.8 (Table 5-5).

Table 5-5. Bayesian Analysis for a Consultand with a 20% Lifetime Risk of Developing a Disease and a Proteomic Profile 16 Times More Common in Those Who Develop the Disease Than in Those Who Do Not

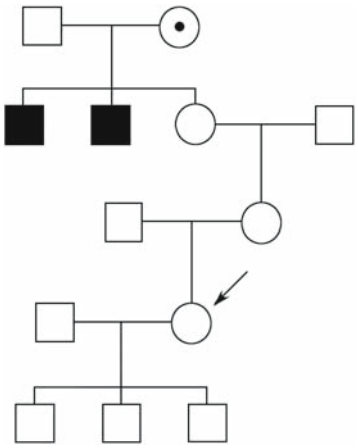
	Hypothesis	
	Affected Eventually	Never Affected
Prior probability	0.2	0.8
Conditional probability (of profiling result)	16	1
Joint probability	3.2	0.8
Posterior probability	0.8	0.2

Note that because posterior probabilities are normalized joint probabilities, the absolute values of the conditional probabilities are unimportant, as long as the ratio (i.e., the odds ratio) between them is correct. This is also true of prior probabilities. For example, in the scenario above, prior probabilities of 1 and 4 can be substituted for 0.2 and 0.8 and the same answer is obtained. Likewise, in the first example of this chapter (Figure 5-1a), prior probabilities of 1 and 1 can be substituted for 1/2 and 1/2, and conditional probabilities of 1 and 8 can be substituted for 1/8 and 1, and the same answer is obtained. Hence, relative risks are easily incorporated into Bayesian analyses.

Bayesian Analyses with More Than One Conditional Probability

Often there is more than one test result, or more than one set of pedigree information, or both, that can be incorporated as conditional probabilities in a single Bayesian analysis. For example, consider the pedigree in Figure 5-4a, in which the two maternal great uncles of the consultand were affected with Duchenne muscular dystrophy (DMD; OMIM #310200), a severe X-linked recessive disease caused by mutations in the *DMD* gene (OMIM #300377). The consultand's maternal grandmother's carrier risk was 1/2, her mother's carrier risk was 1/4, and therefore the consultand's prior carrier risk is 1/8. Suppose that her carrier testing is negative using a highly specific test (an analysis for heterozygous deletions in the *DMD* gene) that detects 2/3 of carriers. Suppose also that her serum creatine phosphokinase (CPK), which is elevated in two-thirds of carriers, is within normal limits. Taking into account her prior probability of 1/8, her normal molecular and CPK test results, and, in addition, her three normal sons, what is the probability that she is a carrier?

The Bayesian analysis for this scenario is shown in Figure 5-4b. Each conditional probability is given its own line. Because the genetic test detects 2/3 of carriers and is highly specific, the conditional probabilities of a negative genetic test result under the hypotheses that she is a carrier and noncarrier are 1/3 and 1, respectively. Because serum CPK is elevated in 2/3 of carriers, the conditional



a

	Hypothesis	
	Carrier	Non-carrier
Prior Probability	1/8	7/8
Conditional probability (of negative genetic test result)	1/3	1
Conditional Probability (of normal CPK result)	1/3	19/20
Conditional Probability (of three normal sons)	1/8	1
Joint Probability	1/576	133/160
Posterior Probability	0.002	0.998

b

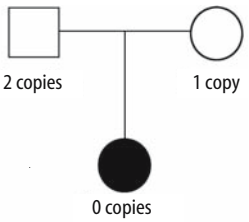
Figure 5-4. (a) Pedigree of a family with individuals affected with DMD (see text). Consultand is indicated by an arrow. (b) Bayesian analysis for the consultand in Figure 5-4a, taking into account her normal carrier test result, her normal CPK test result, and her three normal sons.

probability of a normal serum CPK for the hypothesis that she is a carrier is 1/3. Because 5% of noncarrier women have an abnormal serum CPK (i.e., the normal range is defined as comprising 95% of normal individuals), the conditional probability of a normal serum CPK under the hypothesis that she is a noncarrier is 95% or 19/20. Finally, as in Figure 5-1b, the conditional probabilities of three normal sons under the hypotheses that she is a carrier and noncarrier are 1/8 and 1, respectively. The joint probabilities for each hypothesis are the products of the prior probability, and all conditional probabilities, for each hypothesis (Figure 5-4b). Calculation of posterior probabilities then proceeds exactly as in Table 5-1. In this scenario, taking into account her normal test results and her three normal sons, the consultand's carrier risk is lowered from 1/8 to 0.002, or approximately 1/500.

Bayesian Analyses with More Than Two Hypotheses

In some Bayesian analyses, more than two hypotheses must be considered. For example, consider the pedigree in Figure 5-5a, in which a child with clinically typical type I SMA (type I SMA; Werdnig-Hoffman disease; OMIM #253300) lacks both copies of the *SMN1* gene. By dosage analysis, the child's (unaffected) mother has one copy of the *SMN1* gene and therefore carries one copy of the *SMN1* gene on one chromosome 5, and zero copies of the *SMN1* gene on the other chromosome 5, called the "1 + 0" genotype. However, the child's (unaffected) father has two copies of the *SMN1* gene and therefore could have one of three possible genotypes: (1) two copies of the *SMN1* gene on one chromosome 5 and zero copies of the *SMN1* gene on the other chromosome 5 (the "2 + 0" genotype), (2) one copy of the *SMN1* gene on one chromosome 5 and a subtle mutation in the *SMN1* gene on the other chromosome 5 (the "1 + 1^D" genotype, where "1^D" stands for a "1-copy-disease" allele), or (3) one copy of the *SMN1* gene on each chromosome 5 (the "1 + 1" noncarrier genotype), in which case he passed a de novo deletion of the *SMN1* gene to his affected child. Because the relative frequencies of the various *SMN1* alleles and genotypes in the general population are known,[14] as well as the paternal and maternal de novo deletion rates ($\mu_p = 2.11 \times 10^{-4}$ and $\mu_m = 4.15 \times 10^{-5}$, respectively), the probability that the father is a carrier can be calculated, which obviously has important implications for recurrence risk.

The Bayesian analysis for the father's carrier risk is shown in Figure 5-5b. There are three hypotheses for the father's genotype: 2 + 0, 1 + 1, and 1 + 1^D. The prior probabilities are the relative population frequencies for these genotypes.[14] The conditional probabilities are the probabilities that the father passes a 0-copy allele to his child under each hypothesis. For the 2 + 0 genotype, the conditional probability of passing a 0-copy allele is 0.5, whereas for the 1 + 1 and 1 + 1^D genotypes, the conditional probability of passing a 0-copy allele is the de novo deletion rate of μ_p. As in the generalized Bayesian analysis shown in Table 5-1, the joint probability for each hypothesis is the



a

	Hypothesis (Father's Genotype)		
	2 + 0	1 + 1	1 + 1^D
Prior Probability (relative probability)	1.00×10^{-3}	9.00×10^{-1}	4.58×10^{-4}
Conditional Probability (of passing a 0-copy allele)	0.5	μ_p	μ_p
Joint Probability	5.0×10^{-4}	1.9×10^{-4}	9.7×10^{-8}
Posterior Probability	0.72	0.28	0.00014

b

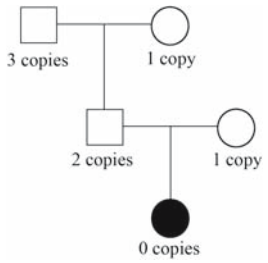
Figure 5-5. (a) Pedigree of a family with an individual affected with type I SMA, with the SMA carrier test results indicated below each individual (see text). (b) Bayesian analysis for carrier risk of the father of the affected child in Figure 5-5a. μ_p, paternal de novo mutation rate.

product of the prior and conditional probabilities for that hypothesis, and the posterior probability for each hypothesis is the joint probability for that hypothesis divided by the sum of all the joint probabilities. The father's carrier risk is the sum of the posterior probabilities of the first (2 + 0) and third (1 + 1^D) columns, or approximately 0.72. The third column contributes little to the carrier risk because the frequency of the 1 + 1^D genotype is low and the conditional probability of a de novo deletion is also low. In contrast, although the frequency of the 2 + 0 genotype is much lower than that of the 1 + 1 genotype, this is counterbalanced by the higher conditional probability of passing a 0-copy allele under the former hypothesis.

Suppose that the father's parents, the paternal grandfather and grandmother of the affected child, are tested and found to have three copies and one copy of the *SMN1* gene, respectively (Figure 5-6a). What is the father's carrier risk? The Bayesian analysis for this scenario is shown in Figure 5-6b. Again, there are three hypotheses for the father's genotype: 2 + 0, 1 + 1, and 1 + 1^D. However, in this scenario, the father's prior probabilities derive from the prior and conditional probabilities of his parents. Because the grandfather has three copies of the *SMN1* gene, his genotype is either 2 + 1 (columns A and C) or 2 + 1^D (columns B and D), and his prior probabilities are the relative population frequencies for these genotypes.[14] Because the (unaffected) grandmother has one copy of the *SMN1* gene, her genotype is 1 + 0, and her prior probability is the relative population frequency of the 1 + 0 genotype for type I SMA in the general population, which is the carrier frequency of 1/38 (2.50×10^{-2}).[14] (Note that because the grandmother must be 1 + 0, simply a prior probability of 1 could be used; as noted above, the absolute values of the conditional probabilities are unimportant, as long as the ratio between them is correct.) The four columns (A through D) show the four possible permutations of grandparental genotypes (prior probabilities) with passage of particular alleles to the father (conditional probabilities) so that he would have a 2-copy SMA carrier test result. Under the hypothesis that the father has a 2 + 0 genotype, he could have inherited a 2-copy "allele" (two copies of *SMN1* on one chromosome 5) from the grandfather (2 + 1) at a probability of 0.5 and a 0-copy allele from the grandmother (1 + 0) at a probability of 0.5 (column A), or he could have inherited a 2-copy allele from the grandfather (2 + 1^D) at a probability of 0.5 and a 0-copy allele from the grandmother (1 + 0) at a probability of 0.5 (column B). Under the hypothesis that the father has a 1 + 1 genotype, he could have inherited a 1-copy allele from the grandfather (2 + 1) at a probability of 0.5 and a 1-copy allele from the grandmother (1 + 0) at a probability of 0.5 (column C). Under the hypothesis that the father has a 1 + 1^D genotype, he could have inherited a 1^D allele from the grandfather (2 + 1^D) at a probability of 0.5 and a 1-copy allele from the grandmother (1 + 0) at a probability of 0.5 (column D).

The father's prior probabilities are the products of the prior and conditional probabilities for the grandparents for each column or permutation. Under the hypothesis that the father has a 2 + 0 genotype, the conditional probability of passing a 0-copy allele to his child is 0.5 (columns A and B), whereas under the hypothesis that the father has a 1 + 1 or 1 + 1^D genotype, the conditional probability of passing a 0-copy allele to his child is the de novo deletion rate of μ_p (columns C and D). As in the generalized Bayesian analysis shown in Table 5-1, the joint probability for each column is the product of the prior and conditional probabilities for that column, and the posterior probability for each column is the joint probability for that column divided by the sum of all the joint probabilities. The father's carrier risk is the sum of the posterior probabilities of columns A (2 + 0), B (2 + 0), and D (1 + 1^D), or approximately 0.999. The father's increased carrier risk in this scenario derives almost entirely from the probability that he has the 2 + 0 genotype; this is unsurprising since the grandfather's 3-copy test result demonstrates the presence of a 2-copy allele in the family. (Note that because the grandmother's prior and conditional probabilities are the same in every column, excluding her data from the analysis will not change the result.)

Suppose instead that the father's parents, the paternal grandfather and grandmother of the affected child, are tested and each is found to have two copies of the *SMN1* gene (Figure 5-7a). What is the father's carrier risk? The Bayesian analysis for this scenario is shown in Figure 5-7b. Again, there are three hypotheses for the father's genotype:

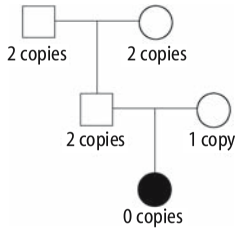


a

	Hypothesis (Father's Genotype)			
	2 + 0		1 + 1	1 + 1^D
	A	B	C	D
Grandfather's Genotype	2 + 1	2 + 1^D	2 + 1	2 + 1^D
Prior probability (relative probability)	7.18×10^{-2}	1.15×10^{-5}	7.18×10^{-2}	1.15×10^{-5}
Conditional probability (of passing a 2-copy allele)	0.5	0.5	--	--
Conditional probability (of passing a 1 or 1^D allele)	--	--	0.5	0.5
Grandmother's Genotype	1 + 0	1 + 0	1 + 0	1 + 0
Prior probability	2.50×10^{-2}	2.50×10^{-2}	2.50×10^{-2}	2.50×10^{-2}
Conditional probability (of passing a 0-copy allele)	0.5	0.5	--	--
Conditional probability (of passing a 1-copy allele)	--	--	0.5	0.5
Father's Genotype	2 + 0		1 + 1	1 + 1^D
Conditional probability (of passing a 0-copy allele)	0.5 0.5		μ_p	μ_p
Joint Probability	2.24×10^{-4}	3.59×10^{-8}	9.47×10^{-8}	1.52×10^{-11}
Posterior Probability	0.999	0.00016	0.00042	0.000000068

b

Figure 5-6. (a) Pedigree of a family with an individual affected with type I SMA with the SMA carrier test results indicated below each individual (see text). (b) Bayesian analysis for carrier risks of the father of the affected child in Figure 5-6a.



a

	Hypothesis (Father's Genotype)													
	2 + 0						1 + 1				1 + 1^{D}			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Grandfather's Genotype	2 + 0	2 + 0	2 + 0	2 + 0	1 + 1	1 + 1^{D}	1 + 1	1 + 1	1 + 1^{D}	1 + 1^{D}	1 + 1	1 + 1^{D}	1 + 1^{D}	1 + 1^{D}
Prior Probability	1.0 x 10^{-3}	1.0 x 10^{-3}	1.0 x 10^{-3}	1.0 x 10^{-3}	0.90	4.6 x 10^{-4}	0.90	0.90	4.6 x 10^{-4}	4.6 x 10^{-4}	0.90	4.6 x 10^{-4}	4.6 x 10^{-4}	4.6 x 10^{-4}
Conditional Probability (of passing a 2-copy allele)	0.5	--	0.5	0.5	--	--	--	--	--	--	--	--	--	--
Conditional Probability (of passing a 0-copy allele)	--	0.5	--	--	2.1 x 10^{-4}	2.1 x 10^{-4}	--	--	--	--	--	--	--	--
Conditional Probability (of passing a 1-copy allele)	--	--	--	--	--	--	1	1	0.5	0.5	1	--	--	0.5
Conditional Probability (of passing a 1^{D} allele)	--	--	--	--	--	--	--	--	--	--	--	0.5	0.5	--
Grandmother's Genotype	2 + 0	2 + 0	1 + 1	1 + 1^{D}	2 + 0	2 + 0	1 + 1	1 + 1^{D}	1 + 1	1 + 1^{D}	1 + 1^{D}	1 + 1	1 + 1^{D}	1 + 1^{D}
Prior Probability	1.0 x 10^{-3}	1.0 x 10^{-3}	0.90	4.6 x 10^{-4}	1.0 x 10^{-3}	1.0 x 10^{-3}	0.90	4.6 x 10^{-4}	0.90	4.6 x 10^{-4}	4.6 x 10^{-4}	0.90	4.6 x 10^{-4}	4.6 x 10^{-4}
Conditional Probability (of passing a 2-copy allele)	--	0.5	--	--	0.5	0.5	--	--	--	--	--	--	--	--
Conditional Probability (of passing a 0-copy allele)	0.5	--	4.2 x 10^{-5}	4.2 x 10^{-5}	--	--	--	--	--	--	--	--	--	--
Conditional Probability (of passing a 1-copy allele)	--	--	--	--	--	--	1	0.5	1	0.5	--	1	0.5	--
Conditional Probability (of passing a 1^{D} allele)	--	--	--	--	--	--	--	--	--	--	0.5	--	--	0.5
Father's Genotype	2 + 0						1 + 1				1 + 1^{D}			
Conditional Probability (of passing a 0-copy allele)	0.5	0.5	0.5	0.5	0.5	0.5	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}	2.1 x 10^{-4}
Joint Probability	1.3 x 10^{-7}	1.3 x 10^{-7}	9.3 x 10^{-9}	4.8 x 10^{-12}	4.8 x 10^{-8}	2.4 x 10^{-11}	1.7 x 10^{-4}	4.3 x 10^{-8}	4.3 x 10^{-8}	1.1 x 10^{-11}	4.3 x 10^{-8}	4.3 x 10^{-8}	1.1 x 10^{-11}	1.1 x 10^{-11}
Posterior Probability	7.3 x 10^{-4}	7.3 x 10^{-4}	5.4 x 10^{-5}	2.8 x 10^{-8}	2.8 x 10^{-4}	1.4 x 10^{-7}	0.997	2.5 x 10^{-4}	2.5 x 10^{-4}	6.5 x 10^{-8}	2.5 x 10^{-4}	2.5 x 10^{-4}	6.5 x 10^{-8}	6.5 x 10^{-8}

b

Grandfather's Genotype (2 copies and asymptomatic)	1+1				2+0			1+1^{D}				
Prior Probability	0.9				0.001			0.00046				
Grandmother's Genotype (2 copies and asymptomatic)	1+1	2+0	1+1^{D}		1+1	2+0	1+1^{D}	1+1		2+0	1+1^{D}	
Prior Probability	0.9	0.001	0.00046		0.9	0.001	0.00046	0.9		0.001	0.00046	
Father's Genotype (2 copies and asymptomatic)	1+1 NC	2+0 C	1+1 NC	1+1^{D} C	2+0 C	2+0 C	2+0 C	1+1 NC	1+1^{D} C	2+0 C	1+1 NC	1+1^{D} C
Conditional Probability (of receiving alleles)	1	0.5 μ_p	0.5	0.5	0.5 μ_m	0.5	0.5 μ_m	0.5	0.5	0.5 μ_p	0.25	0.5
Conditional Probability (of passing a 0-copy allele)	μ_p	0.5	μ_p	μ_p	0.5	0.5	0.5	μ_p	μ_p	0.5	μ_p	μ_p
Joint Probability	1.71x10^{-4}	4.75x10^{-8}	4.35x10^{-8}	4.35x10^{-9}	9.34x10^{-9}	2.5x10^{-7}	4.75x10^{-12}	4.35x10^{-8}	4.35x10^{-8}	2.42x10^{-11}	1.11x10^{-11}	2.22x10^{-11}
Posterior Probability	0.997	2.77x10^{-4}	2.54x10^{-4}	2.54x10^{-4}	5.45x10^{-5}	0.00146	2.77x10^{-7}	2.54x10^{-4}	2.54x10^{-4}	1.41x10^{-7}	6.47x10^{-8}	1.29x10^{-7}
Column	A	B	C	D	E	F	G	H	I	J	K	L

c

Figure 5-7. (a) Pedigree of a family with an individual affected with type I SMA (see text). (b) Bayesian analysis for the father of the affected child in Figure 5-7a. (In the interests of space, only two significant digits are shown.) (c) Alternative Bayesian analysis for the father of the affected child in Figure 5-7a. NC, noncarrier; C, carrier.

2 + 0, 1 + 1, and 1 + 1^D. However, in this scenario, the number of possible permutations of grandparental genotypes (prior probabilities) with passage of particular alleles to the father (conditional probabilities) is dramatically increased. This is because each grandparent could have a 2 + 0, 1 + 1, or 1 + 1^D genotype, and the father could have received a 2-copy allele, a 0-copy allele, a 1-copy allele, or a 1^D allele from either grandparent, in most cases by direct Mendelian inheritance and in some cases from de novo deletions. The organization of the Bayesian analysis in Figure 5-7b is guided by the possible genotypes of the father, which determine the grandparental genotype permutations that need to be considered. Under the hypothesis that the father has the 2 + 0 genotype, he could have received a 2-copy allele from one (2 + 0) grandparent and a 0-copy allele from the other (2 + 0) grandparent, both by direct inheritance (columns A and B), or he could have received a 2-copy allele from one (2 + 0) grandparent by direct inheritance and a de novo deletion allele from the other (1 + 1 or 1 + 1^D) grandparent (columns C, D, E, and F). Under the hypothesis that the father has the 1 + 1 genotype, he must have received a 1-copy allele from each (1 + 1 or 1 + 1^D) grandparent (columns G, H, I, and J). Under the hypothesis that the father has the 1 + 1^D genotype, he must have received a 1-copy allele from one (1 + 1 or 1 + 1^D) grandparent and a 1^D allele from the other (1 + 1 or 1 + 1^D) grandparent (columns K, L, M, and N).

More specifically, under the hypothesis that the father has the 2 + 0 genotype, column A shows the prior probability that the grandfather is 2 + 0 (1.00×10^{-3}), the conditional probability that he passes a 2-copy allele to the father (0.5), the prior probability that the grandmother is 2 + 0 (1.00×10^{-3}), and the conditional probability that she passes a 0-copy allele to the father (0.5). Under the hypothesis that the father has the 1 + 1 genotype, column G shows the prior probability that the grandfather has a 1 + 1 genotype (0.90), the conditional probability that he passes a 1-copy allele to the father (1), the prior probability that the grandmother has a 1 + 1 genotype (0.90), and the conditional probability that she passes a 1-copy allele to the father (1). Under the hypothesis that the father has the 1 + 1^D genotype, column K shows the prior probability that the grandfather has a 1 + 1 genotype (0.90), the conditional probability that he passes a 1-copy allele to the father (1), the prior probability that the grandmother has a 1 + 1^D genotype (4.58×10^{-4}), and the conditional probability that she passes a 1^D allele to the father (0.5).

Again, the father's prior probabilities are the products of the prior and conditional probabilities for the grandparents for each column or permutation. Under the hypothesis that the father has a 2 + 0 genotype, the conditional probability of passing a 0-copy allele to his child is 0.5 (columns A through F), whereas under the hypothesis that the father has a 1 + 1 or 1 + 1^D genotype, the conditional probability of passing a 0-copy allele to his child is the paternal de novo deletion rate of μ_p (columns G through L). As in the generalized Bayesian analysis shown in Table 5-1, the joint probability for each column is the product of the prior and conditional probabilities for that column, and the posterior probability for each column is the joint probability for that column divided by the sum of all the joint probabilities. The father's carrier risk is the sum of the posterior probabilities of columns A through F (2 + 0), and K and L (1 + 1^D), or approximately 1/400. Relative to the previous scenario (Figure 5-6), in which the father also had two copies of *SMN1* but the grandparents had different copy numbers, the father's dramatically decreased carrier risk in this scenario derives from the much lower probability that a 2-copy allele is present in his family, and illustrates the importance of integrating all available genetic testing information into risk assessment calculations.

An alternative organization of the Bayesian analysis shown in Figure 5-7b is shown in Figure 5-7c and is guided by the three hypotheses for the grandparental genotypes: 1 + 1, 2 + 0, and 1 + 1^D. For example, under the hypothesis that both of the grandparents have a 1 + 1 genotype, column A shows the prior probabilities that the grandfather has a 1 + 1 genotype (0.9) and that the grandmother has a 1 + 1 genotype (0.9), the conditional probability that the father received 1-copy alleles from both of the grandparents (1), and the conditional probability that the father passed a 0-copy allele to the affected child (by de novo deletion, μ_p). Under the hypothesis that the grandfather has a 1 + 1 genotype and that the grandmother has a 2 + 0 genotype, column B shows the prior probabilities that the grandfather has a 1 + 1 genotype (0.90) and that the grandmother has a 2 + 0 genotype (0.001), the conditional probability that the father received a 2-copy allele from one of the grandparents (the grandmother in this case) (0.5) and a 0-copy allele from the other grandparent (the grandfather in this case by de novo deletion, μ_p), and the conditional probability that the father passed a 0-copy allele to the affected child (0.5). Under the hypothesis that the grandfather has the 1 + 1 genotype and that the grandmother has the 1 + 1^D genotype, column C shows the prior probabilities that the grandfather has a 1 + 1 genotype (0.90) and that the grandmother has a 1 + 1^D genotype (0.00046), the conditional probability that the father received 1-copy alleles from both grandparents (0.5), and the conditional probability that the father passed a 0-copy allele to the affected child (by de novo deletion, μ_p). Under the hypothesis that the grandfather has the 1 + 1 genotype and that the grandmother has the 1 + 1^D genotype, column D shows the prior probabilities that the grandfather has a 1 + 1 genotype (0.90) and that the grandmother has a 1 + 1^D genotype (0.00046), the conditional probability that the father received a 1^D allele from one of the grandparents (the grandmother in this case) and a 1-copy allele from the other grandparent (the grandfather in this case) (0.5), and the conditional probability that the father passed a 0-copy allele to the affected child (by de novo deletion, μ_p). The father's carrier risk is the sum of the posterior probabilities of columns B, D through G, I, J, and L, or approximately 1/400.

In both approaches (Figures 5-7b and 5-7c), the use of one comprehensive Bayesian analysis table incorporating all necessary information allows simultaneous calculations of the carrier risks of the father, grandfather, and grandmother. Such a comprehensive approach is necessary because the 2-copy test results for the grandparents influence the carrier risk of the father, and the 2-copy test result for the father influences the carrier risks of the grandparents. Using Figure 5-7b, the posterior carrier risk of the grandfather is the sum of the posterior probabilities of columns A through D, F, I, J, and L through N, or approximately 0.0020 (1/500), and the carrier risk of the grandmother is the sum of the posterior probabilities of columns A, B, D through F, H, J, K, M, and N, or approximately 0.0022 (1/450). The posterior probability that all three of them are carriers is the sum of the posterior probabilities of columns A, B, D, F, M, and N, or approximately 0.0015 (1/600). Using Figure 5-7c, the carrier risk of the grandfather is the sum of the posterior probabilities of columns E through L, or approximately 0.0020 (1/500), and the carrier risk of the grandmother is the sum of the posterior probabilities of columns B through D, F, G, and J through L, or approximately 0.0022 (1/450). The posterior probability that all three of them are carriers is the sum of the posterior probabilities of columns F, G, J, and L, or approximately 0.0015 (1/600).

Concluding Remarks

Bayesian analysis plays a central role in genetic risk assessment, and those who offer genetic testing should be proficient. Genetic risk should be assessed as accurately as possible, using all available information at a particular point in time, from the pedigree, from laboratory testing, or from both. Although the technologies for genetic testing will continue to change, Bayesian analysis and genetic risk assessment will remain fundamental aspects of genetic testing and genetic counseling.

References

1. Bayes T. An essay towards solving a problem in the doctrine of chances. *Biometrika.* 1958;45:296–315.
2. Young I. *Introduction to Risk Calculation in Genetic Counseling.* Oxford: Oxford University Press; 1999.
3. Bridge P. *The Calculation of Genetic Risks: Worked Examples in DNA Diagnostics.* Baltimore: Johns Hopkins University Press; 1997.
4. Ogino S, Leonard DGB, Rennert H, Ewens WJ, Wilson RB. Genetic risk assessment in carrier testing for spinal muscular atrophy. *Am J Med Genet.* 2002;110:301–307.
5. Ogino S, Wilson RB. Genetic testing and risk assessment for spinal muscular atrophy (SMA). *Hum Genet.* 2002;111:477–500.
6. Ogino S, Wilson RB. Bayesian analysis and risk assessment in genetic counseling and testing. *J Mol Diagn.* 2004;6:1–9.
7. Ogino S, Wilson RB, Gold B, Hawley P, Grody WW. Bayesian analysis for cystic fibrosis risks in prenatal and carrier screening. *Genet Med.* 2004;6:439–449.
8. Ogino S, Wilson RB, Grody WW. Bayesian risk assessment for autosomal recessive diseases: fetal echogenic bowel with one or no detectable CFTR mutation. *J Med Genet.* 2004;41:e70.
9. Ogino S, Flodman P, Wilson RB, Gold B, Grody WW. Risk calculations for cystic fibrosis in neonatal screening by immunoreactive trypsinogen and CFTR mutation tests. *Genet Med.* 2005;7:317–327.
10. Grody WW, Cutting GR, Klinger KW, Richards CS, Watson MS, Desnick RJ. Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. *Genet Med.* 2001;3:149–154.
11. Richards CS, Bradley LA, Amos J, et al. Standards and guidelines for CFTR mutation testing. *Genet Med.* 2002;4:379–391.
12. Watson MS, Cutting GR, Desnick RJ, et al. Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel. *Genet Med.* 2004;6:387–391.
13. Wirth B, Herz M, Wetter A, et al. Quantitative analysis of survival motor neuron copies: identification of subtle SMN1 mutations in patients with spinal muscular atrophy, genotype-phenotype correlation, and implications for genetic counseling. *Am J Hum Genet.* 1999;64:1340–1356.
14. Ogino S, Wilson RB, Gold B. New insights on the evolution of the SMN1 and SMN2 region: simulation and meta-analysis for allele and haplotype frequency calculations. *Eur J Hum Genet.* 2004;12:1015–1023.

Chapter 6

Developmental Disabilities

Nancy J. Carpenter, Kristin May, Benjamin Roa, and Jack Tarleton

Although the classic childhood phenotypes of many developmental disorders have been established for some time, only in the past decade have the genetic etiologies of some of these disorders been identified. Investigations of the molecular basis of these conditions have resulted in the identification of new genes, leading to insights into the function of new proteins and biochemical pathways. In addition, genetic mechanisms previously unknown in humans, such as genomic imprinting, uniparental disomy, expansion of trinucleotide repeats, and facilitation of deletions and duplications by low-copy repeats, were recognized as the causes of some of these conditions.

This chapter reviews the genetic etiologies of several developmental disorders, including the fragile X, Prader-Willi, Angelman, Rett, and Williams syndromes, and disorders due to cryptic unbalanced chromosome rearrangements. The molecular approaches being applied to the diagnosis of these disorders also are reviewed.

FRAGILE X SYNDROME

Molecular Basis

Named for its association with a chromosomal fragile site observed in many patients (*FRAXA* chromosomal locus Xq27.3), fragile X syndrome (FXS) is the most common cause of inherited mental retardation (MR). FXS results from loss or severe reduction of the protein FMRP, encoded by the *FMR1* (fragile X mental retardation) gene.[1] All patients with FXS have mutations in *FMR1*, as no mutations leading to FXS have been identified in other genes. Both males and females may be affected with FXS, but females are typically less severely affected. Thus, FXS is considered to be X-linked dominant with reduced penetrance in females.[2]

The *FMR1* gene encompasses 38 kilobases (kb) of genomic DNA and has 17 exons.[3] The major *FMR1* messenger RNA (mRNA) produced in most tissues is approximately 4 kb, although several protein isoforms are generated by alternative splicing toward the 3′ end of the mRNA in some tissues. While FMRP can be detected in the nucleus, the majority of the protein associates with translating ribosomes in the cytoplasm, where it acts as a negative translational regulator. FMRP also is known to have a role in neuronal synapse maturation and plasticity. Autopsy samples from FXS patients have shown failure of dendritic spines to assume a normal mature size, shape, and distribution.

The molecular genetics of *FMR1* are complex. A repeated trinucleotide sequence, composed primarily of CGG repeats, is located in the untranslated portion of exon 1, ending 69 base pairs upstream of the translational start. Nearly all mutations (>99%) resulting in FXS occur as instability of the trinucleotide repeat, leading to dramatic expansion of the repeat segment (>200 to a few thousand repeats) accompanied by aberrant hypermethylation of CpG dinucleotides within the gene (full mutations). Relatively rare deletions and point mutations in *FMR1* account for the remaining mutations found in patients with FXS. The mechanism of repeat instability in *FMR1* is believed to be DNA polymerase slippage during DNA replication. AGG repeats, spaced at about 10 repeat intervals within the CGG repeat segment, may mitigate potential repeat instability through disruption of higher-order molecular structures formed during DNA replication. These secondary structures contribute to polymerase slippage, and alleles that lack interrupting AGG repeats are at higher risk for expansion.

The *FMR1* repeat region is naturally polymorphic, with variation of the CGG repeats in normal (i.e., stably inherited) alleles ranging from 5 to 40 repeats, and the vast majority of individuals in the general population have 20 to 40 repeats. Intermediate repeat alleles containing 41 to 59 repeats occasionally have minor variations of a few repeats when transmitted from parent to child, producing no clinical consequences. However, in rare instances, transmission of intermediate alleles with 55 to 59 repeats may expand into pathological alleles.

Interestingly, FXS occurs strictly through maternal inheritance. Individuals with full mutations may inherit a

Table 6-1. Normal and Pathological *FMR1* Allele Types

Allele	Repeat Range	Methylation?
Normal	5–40	No
Intermediate	41–59	No
Premutation	60–~200	No
Full mutation	≥200	Yes
Methylation mosaic	≥200	Variable
Premutation/full mutation repeat size mosaic	Mixed premutation and full mutation	Full mutation may be methylated

similar allele from their mothers or, alternatively, their mothers may have a "premutation" allele. *FMR1* alleles with >60 repeats up to approximately 200 repeats are considered premutations because of potential instability. Individuals with premutations do not have typical characteristics associated with FXS but may transmit an unstable repeat, which undergoes extensive repeat expansion. When transmitted by fathers to their daughters, premutations are not dramatically unstable, and as a result full mutations never arise through paternal inheritance.

Hypermethylation of the *FMR1* promoter region, along with repeat expansion, results in decreased or completely absent transcription and the concomitant loss of FMRP. Patients with partial methylation of a full mutation (methylation mosaics) may have some FMRP expression, resulting in a less severe phenotype. In addition, patients with a mixture of cells having either a premutation or full mutation (premutation/full mutation size mosaics) frequently are identified during molecular testing. These patients usually have MR but may perform at the lower end of normal intellect (IQ > 70). Because methylation is not an all-or-none phenomenon within *FMR1*, the FXS phenotype may encompass a spectrum of possible affectations from mild to severe. Table 6-1 summarizes the classification of *FMR1* alleles.

Clinical Utility of Testing

Due to the presence of unrecognized *FMR1* alterations in unaffected, carrier parents, the first indication of FXS within a family is usually the diagnosis of an affected child. Unfortunately, many families do not learn the FXS diagnosis for long periods after first concerns about their child's development or behavior, and many have subsequent pregnancies before diagnosis for their first child. Such situations highlight the importance of early diagnosis so children and families can receive the benefits of genetic counseling and intervention services. Other than the infrequent deletion or point mutation, which often are spontaneous and not inherited from a parent, mothers of *all* FXS patients are either premutation or full-mutation carriers. In turn, at least one of the mother's parents has an *FMR1* alteration. Consequently, *FMR1* mutations may be present in siblings of an affected individual as well as other extended family members. It is important to remember that the daughters of unaffected males with a premutation (transmitting males) are also unaffected carriers, and that their offspring are at risk for FXS. Many families are known in which an *FMR1* mutation has been transmitted through numerous generations and into family branches unknown to one another.

For developmentally delayed children, *FMR1* molecular testing is diagnostic, as FXS affects development from infancy. However, the nonspecific nature of FXS during early development makes the testing approach one of ruling out FXS in most situations. The hallmark finding in almost all patients with FXS is MR, but the physical and behavioral features of males with FXS are variable prior to puberty. Physical features not readily recognizable in preschool-age boys become more obvious with age: long face, prominent forehead, large ears, prominent jaw, and enlarged testicles (macroorchidism). Motor milestones and speech are frequently delayed, and temperament often is affected (e.g., hyperactivity, hand flapping, hand biting, temper tantrums, and occasionally autism). Females with FXS usually have milder manifestations and as a result are more difficult to diagnose clinically. FXS always should be suspected in males with mild to moderate MR and females with mild MR until shown otherwise by negative *FMR1* analysis.

Women who are full-mutation or premutation carriers have a 50% risk of transmitting their abnormal allele in each pregnancy. While transmission of a full mutation always leads to a child with a full mutation, the risk of a premutation transmission resulting in an affected offspring with a full mutation is proportional to the maternal premutation repeat number. Empirically, the 50% risk of a female carrier producing an affected male child is reduced to 7% if the premutation contains 56 to 59 repeats, 10% for 60 to 69 repeats, 29% for 70 to 79 repeats, 36% for 80 to 89 repeats, and 47% for 90 to 99 repeats; it reaches the maximum, 50%, when a premutation has >100 repeats. Because females have approximately 50% penetrance, the risk for producing an affected female is half that of producing an affected male in any premutation repeat interval category.

Prenatal testing for *FMR1* mutations is available in many clinical molecular laboratories. Genomic DNA isolated from amniocytes obtained during amniocentesis at 16 to 18 weeks gestation or from chorionic villus sampling (CVS) at 10 to 12 weeks gestation can be used for testing. Prenatal molecular analysis proceeds in much the same fashion as that performed on DNA obtained from adult peripheral blood. However, the DNA analysis of CVS may be more complex, as chorionic villi are extraembryonic. Hypermethylation in CVS may be incomplete and not representative of the true *FMR1* methylation status in fetal tissues. Therefore, occasionally a follow-up amniocentesis may be required to resolve ambiguous CVS test results.

General population screening for *FMR1* mutations has been proposed but remains controversial. In comparison to

most disorders already screened in the newborn period, FXS is more prevalent and testing is highly reliable. However, the relatively high costs and the technical complexities of testing must be resolved before population screening is possible. Protein testing of FMRP may be useful for screening populations with MR.

Available Assays

Routine clinical testing for *FMR1* mutations includes molecular assessment of both the trinucleotide repeat number and the *FMR1* methylation status. Standard technical approaches include (1) double-digest Southern blot analysis using a methylation-sensitive restriction enzyme such as *Eag* I, *BssH* II, or *Nru* I along with a methylation-insensitive restriction enzyme such as *EcoR* I or *Hind* III[4] and (2) polymerase chain reaction (PCR) assays specific for the CGG repeat segment of *FMR1* (Figure 6-1).[5] Specialized fragile X chromosome analysis, using special culture techniques to induce fragile sites, is no longer used for diagnosis of FXS due to low sensitivity. While only a very few FXS patients with point mutations in *FMR1* have been identified, clinical molecular testing does not routinely investigate the gene for point mutations, deletions, insertions, or inversions downstream of the repeat segment.

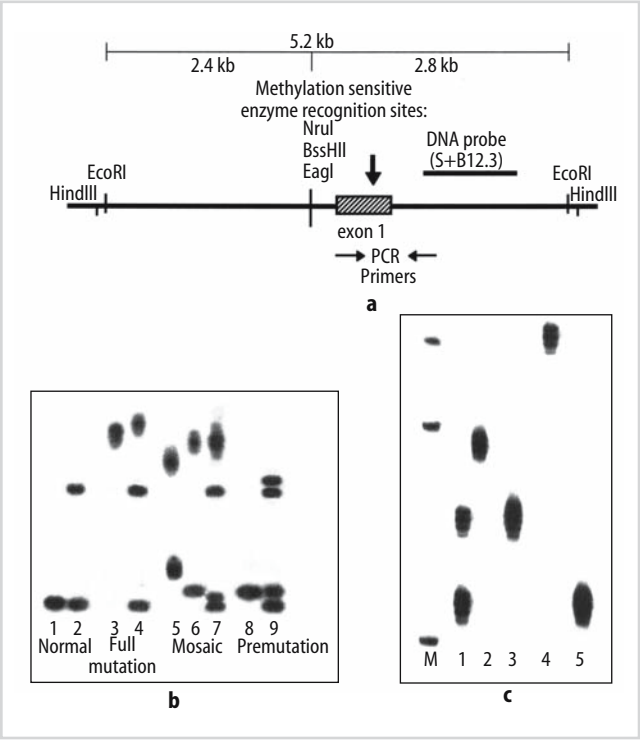


Figure 6-1. Repeat expansion and methylation in *FMR1*. (a) Restriction enzyme map of *FMR1*, with locations of restriction enzyme sites, DNA probe, and PCR primers used in molecular testing. Top line indicates DNA fragments generated using *EcoR* I and *Eag* I as depicted in the Southern blot analysis in Figure 6-1b (2.8 kb and 5.2 kb fragments are detected by the DNA probe StB12.3). The vertical arrow indicates the location of the CGG repeat in exon 1. (b) Southern blot analysis of *FMR1*. Only the 2.8 kb fragment is detected in normal males (lane 1), while both the 2.8 kb and 5.2 kb fragments are detected in normal females (lane 2) due to methylation associated with normal X inactivation. Completely methylated full mutations are depicted in lane 3 (affected male with full mutation) and lane 4 (affected female with full mutation contained on one of her X chromosomes; normal allele on her other X chromosome). Smeary signals occur due to variable repeat expansion within peripheral lymphocytes used for DNA isolation. Mosaic patterns are illustrated in lane 5 (male with partial methylation of full mutation), lane 6 (male with premutation/full mutation mosaicism), and lane 7 (female with premutation/full mutation mosaicism). Lane 8 illustrates a transmitting male with a premutation and lane 9 illustrates a female with a premutation. Both premutations contain approximately 75 repeats. (c) PCR analysis of *FMR1* repeats from five individuals separated on a 6% polyacrylamide gel. Lane 1 contains PCR products from a female with 20 and 30 repeats, respectively, contained within her two normal alleles. Lanes 2, 3, and 5 are males with normal repeat alleles (40, 30, and 20 repeats, respectively), while lane 4 illustrates a male with a 65-repeat premutation allele. Smeary signals result from DNA polymerase stuttering during the PCR amplification.

In most laboratory settings PCR is used to size normal and premutation alleles with typical sensitivity up to 120 to 150 repeats. PCR product yield is inversely proportional to the number of trinucleotide repeats such that little or no product can be obtained when larger repeats are present. Some PCR-based testing protocols may have higher sensitivity regarding detection of larger repeats, yet few laboratories have adopted these practices due to technical difficulties. When used in conjunction with PCR, Southern blot analysis provides a more complete inspection of the gene by detecting multiple possible molecular events, including repeat expansion, DNA methylation, and the relatively rare *FMR1* deletions around the trinucleotide repeat segment. Although it is not routinely performed in most clinical laboratory settings, a few laboratories utilize protein-based testing for FMRP. Since severity of the FXS phenotype appears to inversely correlate with FMRP expression, assessment of FMRP production in patients with methylation mosaicism may be a useful prognostic indicator of disease severity.[6]

Interpretation of Test Results

Because *FMR1* appears to be the only disease-causing gene for FXS, test specificity is 100%. Using both Southern blot analysis and PCR specific for *FMR1*, test sensitivity for repeat expansion can be estimated to be nearly 99%, as only rare point mutations, small deletions/insertions remote from the repeat segment, or gene inversions would be missed. However, these nonrepeat expansion molecular alterations may be underascertained in *FMR1* since gene regions downstream of the repeat segment in exon 1 are rarely investigated, even if repeat expansion is not present. As detected by testing, the presence of cellular mosaicism, in either repeats or methylation, presents potential problems for prediction of FXS severity. Essentially all patients with mutations resulting in reduction of FMRP are impaired, but expression of variable amounts of FMRP may allow some individuals to function at a higher level than expected. These individuals may occasionally have intelligence quotients that are not in the MR range. Prognostication of severity based on testing of a young child should be

predicated with great caution during genetic counseling, because no long-term study exists following the development of individuals with methylation mosaicism.

Laboratory Issues

Due to unusual complexity in *FMR1*, molecular testing should be performed by an experienced molecular pathologist. If the etiology of MR in an individual is unknown, DNA analysis for FXS should be performed as part of a comprehensive genetic evaluation that includes routine cytogenetic analysis. Cytogenetic abnormalities have been identified as frequently as or more frequently than *FMR1* mutations in individuals with MR who are referred for FXS testing. In addition, the use of Southern blotting on DNA isolated from amniocytes for prenatal *FMR1* analysis, with typical 2- to 3-week turnaround times, may lead to stressful situations on occasion regarding the timing of possible pregnancy termination. Utilization of CVS provides additional time for possible pregnancy termination, but equivocal results sometimes occur due to incomplete methylation when a full mutation is present.

PCR-based commercial kits used to estimate repeat copy number are available through some suppliers but not widely utilized. Many laboratories use laboratory-developed methods for both *FMR1* Southern blot analysis and PCR. Patient control cell lines may be purchased from the Coriell Institute (http://coriell.umdnj.edu/). Testing for FXS is routinely included in proficiency tests administered by the College of American Pathologists (CAP).

UNIPARENTAL DISOMY

Molecular Basis

Several developmental disorders arise not just from classical gene mutations but also from the effects brought to bear on gene expression by chromosomal aneuploidy. Aneuploidy occurs in a substantial percentage of all recognized pregnancies, yet most instances result in embryonic lethality with spontaneous abortion during early pregnancy. This high rate of observed aneuploidy suggests the existence of numerous abnormal gametes, either nullisomic or disomic for a particular chromosome, due to meiotic nondisjunction events. Considering the relatively high frequency of aneuploidies, Eric Engel in 1980 hypothesized the potential for rare "gametic complementation" between a gamete nullisomic for a particular chromosome and a gamete disomic for the same chromosome.[7] Although derived from two separate "mistakes," such a union would lead to an apparently normal (2 n or disomic) individual with inheritance of two copies of a chromosome pair (or a chromosomal segment) from one parent and no copy from the other parent, or uniparental disomy (UPD). Maternal UPD occurs when a child has two copies of one of the mother's chromosomes and no copies of that particular chromosome from the father. Paternal UPD occurs when a child inherits two copies of a specific chromosome from the father and no copies of that chromosome from the mother.

Uniparental disomy may or may not cause developmental problems, depending on which chromosome is involved. However, patients identified with UPD indicate that the possible clinical consequences include (1) expression of recessive disorders when only one parent carries a recessive trait, (2) disorders related to parent-of-origin effects (imprinted genes), and (3) residential effects of chromosome aneuploidy (mosaicism). The inheritance of two identical chromosomes, or isodisomy, may occur due to meiosis II nondisjunction, formation of isochromosomes through centromeric misdivision, or mitotic nondisjunction in a monosomic diploid cell. Isodisomy is of particular concern due to the potential expression of recessive disorders when one parent is a carrier of a recessive trait, and for imprinting disorders. Inheritance of two homologous, but nonidentical, chromosomes from one parent is termed heterodisomy, and occurs as a result of a meiosis I nondisjunction. The presence of heterodisomy raises concern related to expression of imprinting disorders.

Gametic complementation is one of several possible mechanisms producing UPD (Table 6-2). The most common mechanism leading to UPD appears to be trisomy rescue. Observations of mosaicism for normal and trisomic karyotypes confined to extraembryonic (placental) tissue obtained by CVS, or confined placental mosaicism (CPM), led to recognition of trisomy rescue. Upon later cytogenetic examination of fetal or neonatal tissue, this mosaicism is not detected and has resolved into an apparently normal disomy. A trisomy may be "rescued" by loss of one trisomic set member through nondisjunction, anaphase lag, or chromosome degradation mediated by centromeric loss. Because the chromosome loss is random, the incidence of UPD in a diploid fetus with known CPM is theoretically 1 in 3. Correction or rescue of a monosomic cell line may occur through early mitotic nondisjunction or endoreduplication of a whole chromosome in a monosomic conception. In addition, chromosomal transloca-

Table 6-2. Mechanisms Leading to UPD

Trisomy rescue
Monosomy rescue
Gametic complementation
Chromosomal translocation
• Centric fusions of acrocentric chromosomes
• Familial heterologous Robertsonian translocation
• Familial homologous Robertsonian translocation
• Heterologous de novo centric fusions
• Homologous de novo centric fusions
• Reciprocal balanced translocations
De novo somatic recombination
Pericentric and paracentric inversions in imprinted domains
Small marker chromosomes containing imprinted genes

Source: Reference 8.

tions, somatic recombination, inversions in imprinted domains, and marker chromosomes also may lead to UPD.

When Engel first conceptualized UPD, he calculated that perhaps 3 in 10,000 individuals have UPD for one of the chromosomes (15, 16, 21, 22, or the sex chromosomes) commonly observed in aneuploidy.[7] Immediately recognized was the potential consequence of isodisomy, resulting in duplication of recessive alleles from a single carrier parent. In 1988, the discovery of cystic fibrosis (CF) in a young girl with maternal UPD for chromosome 7 was the first report of UPD resulting in a recessive condition.[9] The girl's mother was a CF carrier but her father was not. Recessive conditions caused by UPD have been reported for UPD involving chromosomes 1, 2, 4, 5, 6, 7, 8, 9, 11, 13, 14, 15, 16, and X (Table 6-3).

The later discovery of genomic imprinting revealed additional pathological consequences related to UPD.[10] While most genes are expressed from functional alleles derived from both parents, a small minority of genes are normally expressed only from one allele, either the maternal or paternal allele. This differential gene expression depending on the parent of origin results from imprinting, a process initiated in germinal tissue and maintained in somatic tissue by methylation of DNA. Imprinting may be tissue specific and is a normal process for regulating dosage of gene products when normal, biparental inheritance occurs. As a result of DNA methylation, control elements regulate expression of specific individual genes, or whole segments of chromosomes containing several genes, exclusively from either the maternal or paternal alleles. Uniparental disomy for chromosomes containing imprinted genes results in functional loss of gene expression even when no change to the DNA sequence has occurred. Although a small number of genes are affected, several disorders result from imprinting defects or loss of gene expression related to UPD of a chromosome containing imprinted genes (Table 6-4).

Table 6-3. Recessive Disorders Associated with UPD

Disorder	Chromosome	Parent of Origin
Junctional epidermolysis bullosa	1	M, P
Chédiak-Higashi syndrome	1	M
Pycnodysostosis	1	P
Congenital pain insensitivity with anhidrosis	1	P
5-alpha-reductase deficiency	2	P
Abetalipoproteinemia	4	M
Spinal muscular atrophy	5	P
Methylmalonic acidemia	6	P
21-hydroxylase deficiency	6	P
Complement deficiency	6	P
Cystic fibrosis	7	M
Osteogenesis imperfecta	7	M
Congenital chloride diarrhea	7	P
Lipoprotein lipase deficiency	8	P
Leigh syndrome	9	M
Cartilage hair hypoplasia	9	M
β-thalassemia	11	M
Retinoblastoma	13	P
Rod monochromacy	14	M
Bloom syndrome	15	M
α-thalassemia	16	P
Familial Mediterranean fever	16	P
Hemophilia	X	P
Duchenne muscular dystrophy	X	M

Sources: References 8, 10.
M, maternal; P, paternal.

Table 6-4. Disorders Associated with Imprinted Genes

Disorder/Phenotype	Gene(s)	Locus	Maternal or Paternal UPD
Transient neonatal diabetes mellitus	*IGF2R*	6q25–q27	Paternal
Russell-Silver syndrome	*PEG1/MEST*	7q32	Maternal
Beckwith-Wiedemann syndrome	*IGF2*, *H19*, *P57KIP2*, *KVLQT*	11p15	Paternal
Maternal UPD14 syndrome (precocious puberty/ short stature)	unknown	14	Maternal
Paternal UPD14 syndrome (abnormal thorax, short stature)	unknown	14	Paternal
Prader-Willi syndrome	*SNRPN*, *ZNF127*, *FZN127*, *IPW*, *NDN*, *PAR1*, *PAR5*	15q11–q13	Maternal
Angelman syndrome	*UBE3A*	15q11–q13	Paternal
Intrauterine growth retardation	unknown	16	Maternal

Sources: References 8, 10.

Clinical Utility of Testing

Molecular testing is valuable in uncovering UPD as an explanation for imprinting disorders and some recessive diseases. Testing is typically performed for diagnostic purposes on infants, children, or adults when UPD is suspected. Prenatal testing is indicated when confined placental mosaicism is detected. In addition to the clarity a diagnosis brings in most situations, genetic counseling and risk assessments for disorders caused by UPD are greatly affected because recurrence risks are quite different in UPD situations. The recurrence risk for UPD is negligible (<1%) except for that caused by the presence of balanced translocations carried by a parent (perhaps unknown to the parent). Recurrence risks in these situations may be as high as 25%. Uniparental disomy testing after an observation of confined placental mosaicism also

allows informed reproductive choices or may prepare parents for the birth of an affected child.

UPD involving most chromosomes does not cause obvious abnormalities related to imprinting defects, as only a few chromosomes (6, 7, 11, 14, and 15) are known to contain genes that undergo imprinting. Of these, maternally imprinted genes are present on chromosomes 7, 14, 15 and 16, while paternally imprinted genes are contained on chromosomes 6, 11, 14, and 15. Maternally imprinted genes also are suspected, but not confirmed, to be present on chromosomes 2 and 16. No cases are known of either maternal or paternal UPD for chromosomes 3, 12, 18, or 19. In addition, neither paternal UPD for chromosomes 4, 9, or 10, nor maternal UPD for chromosomes 5 or 11, has been reported. Paternal UPD for both the X and Y chromosomes together can only be heterodisomic, of paternal origin, and detected when a paternal X-linked recessive disease is inherited from the father. For many chromosomes, too few examples have been observed to determine a possible phenotype related to UPD.

Available Assays

When UPD is suspected, confirmatory testing is based on detection of polymorphic DNA markers in both parents and the child (Figure 6-2). The markers of choice typically are single nucleotide polymorphisms (SNPs), di-, tri-, or tetranucleotide repeat polymorphisms (short tandem repeats [STRs], microsatellites), or variable number of tandem repeats (VNTRs, minisatellites). These natural variants occur throughout the human genome, and characterized DNA markers are available for all chromosomes.

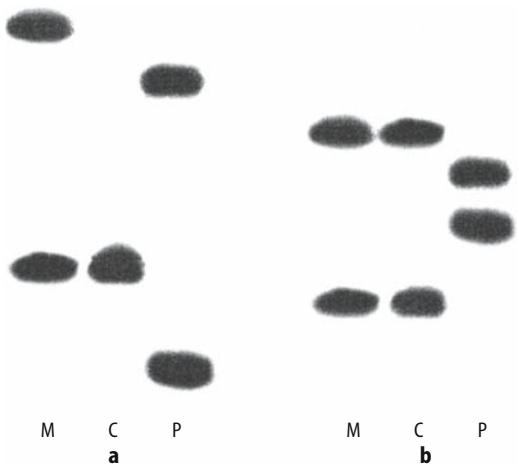


Figure 6-2. DNA marker analysis for isodisomy and heterodisomy. Polyacrylamide gel electrophoresis of a DNA marker indicating typical patterns found when maternal isodisomy (a) or heterodisomy (b) occurs. Parental patterns are opposite if paternal UPD is present (not shown). Autoradiographic signals were produced using PCR incorporating a radioactive label. Maternal DNA is designated as "M," paternal DNA as "P," and DNA from their child with UPD as "C." (a) Maternal isodisomy is indicated by the absence of paternal alleles and the presence of a single maternal allele in the child. (b) Maternal heterodisomy is indicated by the absence of paternal alleles and the presence of both maternal alleles in the child.

While DNA markers can be detected by several laboratory techniques, almost all use PCR amplification to produce sufficient material for analysis. Normal biparental inheritance will result in one of each parental allele for each DNA marker being present in the offspring. Uniparental disomy is present when no alleles are detected from one parent. When no paternal markers are detected, resulting from maternal UPD, paternity can be confirmed by using DNA markers from other chromosomes. Once a marker identifies UPD, additional markers in the chromosomal region can be used to fully explore potential segmental hetero- or isodisomy.

Interpretation of Test Results

DNA marker analysis has a sensitivity approaching 100%, if sufficient markers are informative. However, with the large number of available DNA markers for each chromosome, finding informative markers usually is only a matter of using higher marker density. Segmental disomy may potentially be missed if it involves a chromosomal region that coincidentally is not investigated. The specificity of DNA marker analysis should be 100%, assuming the markers have been correctly mapped to the appropriate chromosome.

Because meiosis I nondisjunction occurs more frequently than meiosis II nondisjunction, heterodisomy will be detected more often than isodisomy. The centromeric regions of heterodisomic homologues nearly always retain heterozygous regions but, crossing over in distal segments may result in homozygosity or heterozygosity. Therefore, informative DNA markers from the centromeric region will be heterozygous, but more distal markers may be heterozygous or homozygous due to recombination events. Isodisomy results in homozygous markers near the centromere, while more distal markers may be heterozygous (partial isodisomy) or homozygous (complete isodisomy). Isodisomy for one chromosomal segment and heterodisomy for another is not uncommon and results from meiotic recombination. In addition, UPD may be detected in only one chromosomal segment in juxtaposition to segments inherited biparentally due to mitotic recombination. Thus, it is important to study markers located near the centromere as well as more distal markers to fully clarify the origins of UPD.

Laboratory Issues

While UPD is readily detected by DNA marker analysis, both parents may not be available for testing. In such cases, analysis of one parent and the child still may be successful for identification of UPD for a particular chromosome. For example, if a mother-child pair is analyzed, the absence of maternal alleles suggests paternal UPD for a specific chromosome, while the converse situation is true with a father-

child pair. Markers from other chromosomes may be used to confirm parentage.

In addition to UPD, absence of DNA marker alleles inherited from a parent may arise as the result of submicroscopic chromosome deletion. Fluorescent in situ hybridization (FISH) analysis is used to confirm microscopic deletions in cases where multiple DNA markers from the same chromosome indicate an absence of marker contribution from one parent in one chromosomal segment yet the presence of normal biparental inheritance in another segment. Routine chromosome analysis (karyotyping) should be used to rule out large deletions as part of a full patient study.

Commercial kits are not available for DNA marker analysis of all chromosomes. However, oligonucleotide primers for amplification of marker loci easily can be synthesized and obtained from numerous companies. Depending on the disorder, cell lines and DNA from patients with specific UPD-associated disorders may be available from the Coriell Cell Repositories (http://coriell.umdnj.edu/). Although proficiency testing for Prader-Willi and Angelman syndromes is available, to date no proficiency test is available specifically to assess proficiency in UPD testing.

PRADER-WILLI AND ANGELMAN SYNDROMES

Molecular Basis

Prader-Willi syndrome (PWS) and Angelman syndrome (AS) are neurodevelopmental disorders caused by a variety of genetic abnormalities involving the proximal part of the long arm of chromosome 15 (q11–q13). Within a 2 megabase (Mb) domain, gene expression depends on the parent of origin resulting from imprinting. Several genes defined above in the region (*SNRPN* [small nuclear ribonuclear protein N], *ZNF127*, *IPW*, *PAR1*, *PAR5*, *PW71*, *NECDIN*) are expressed exclusively from the paternal chromosome in all tissues studied, whereas another gene, *UBE3A* (ubiquitin-protein ligase E3A), is expressed only from the maternal chromosome in brain but has biparental expression in other tissues. Evidence suggests that PWS is caused by loss of expression of one or more of the C/D box small nucleolar RNAs (snoRNAs) encoded within the *SNRPN* locus.[11] On the other hand, AS results from loss of expression of *UBE3A*, which is involved in the ubiquitination pathway targeting certain proteins for degradation.[12] Imprinting of genes in this domain is coordinately controlled by a bipartite imprinting center (IC) overlapping the *SNRPN* promoter and extending 35 kb upstream. Flanking the imprinted region and approximately 4 Mb apart are several low-copy repeats derived from an ancestral *HERC2* gene and other sequences that predispose the region to chromosomal rearrangement by unequal crossing over.

All the genetic mechanisms identified in PWS interfere with the expression of paternally expressed genes in the 15q11–q13 domain.[13] Approximately 70% of PWS patients have a 4 Mb deletion of the paternal chromosome 15q11–13, which occurs sporadically through unequal crossing over. Maternal UPD for chromosome 15 accounts for about 20% of PWS and is usually the consequence of abnormal chromosome segregation during meiosis. In 1% to 5% of cases, microdeletions (6 to 200 kb) that include the IC or epigenetic changes in the IC occur and are inherited in an autosomal dominant manner. Such defects result in both maternal and paternal copies of genes in the imprinted domain having a maternal imprint and therefore lacking expression of the paternally inherited genes. Finally, about 1% of patients have translocations or other structural abnormality involving chromosome 15, which result in deletion or maternal UPD of the 15q11–q13 region.

Four different genetic mechanisms have been identified in 85% to 90% of AS patients, and all cause loss of *UBE3A* expression.[14] A 4 Mb deletion of the maternal chromosome 15q11–q13 region occurs sporadically in about 65% to 70% of cases due to unequal crossing over, as occurs in PWS. Paternal UPD has been detected in about 3% to 5% of patients and is also due to abnormal segregation of chromosome 15 in meiosis. Approximately 7% to 9% of AS patients have an imprinting defect that results from microdeletions of the IC or epigenetic changes in the IC. Such defects cause lack of expression of the maternally inherited *UBE3A* gene in brain because both maternal and paternal copies have a paternal imprint. Point mutations in the *UBE3A* gene (mostly truncating mutations) are found in 4% to 11% of AS patients.[15] Approximately 10% to 15% of patients with a clinical diagnosis of AS have no identifiable chromosomal or molecular abnormality. It is thought that these patients have some undetected abnormality that affects the *UBE3A* gene or have a mutation in another gene in the ubiquitination pathway.

Clinical Utility of Testing

Diagnostic Testing

Prader-Willi syndrome is a disorder with many manifestations related to hypothalamic insufficiency. The major features include infantile hypotonia, hypogonadism, dysmorphic appearance, small hands and feet, hyperphagia and obesity, developmental delay and MR, and characteristic behavior such as temper outbursts, rigidity, and repetitive thoughts and behavior. In infancy, the differential diagnosis includes neuromuscular disorders associated with hypotonia such as congenital myotonic dystrophy. The differential diagnosis in children and adults includes disorders with MR and obesity such as Bardet-Biedl, Cohen, and fragile X syndromes as well as acquired hypothalmic injury.

AS is characterized by microcephaly, gait ataxia, seizures, severe speech impairment, severe development delay or MR, and characteristic behavior, such as inappropriate

laughter and excitability. The differential diagnosis in infancy includes cerebral palsy, inborn errors of metabolism, mitochondrial encephalopathy, and Rett syndrome. Infants with AS sometimes have been clinically misdiagnosed as PWS because of hypotonia, feeding difficulties, and developmental delay, or because the distinctive features of AS are not apparent until later in development.

Prenatal Testing

The risk of recurrence and the type of prenatal testing vary according to the chromosome or molecular defect found in the proband. For this reason, prenatal diagnosis should be undertaken only after the genetic mechanism in the proband has been determined and the parents have received genetic counseling. A low risk of recurrence is associated with PWS and AS due to a deletion or UPD, if the parents' chromosomes are normal; however, prenatal testing may be offered for reassurance. Since mosaicism has been reported in mothers of AS patients with *UBE3A* mutations, prenatal testing should be offered even if the mother is negative for a mutation. Prenatal testing is also appropriate for families without a child with PWS or AS if a deletion of chromosome 15q11–q13 is suspected on chromosome analysis of CVS or amniotic fluid, if trisomy 15 is noted on CVS but a normal karyotype is found on amniotic fluid, and if a de novo chromosome 15 translocation or supernumerary marker chromosome is found by karyotyping.

Testing Parents and Other Family Members of a Proband

Parents of patients with deletions, specifically fathers of PWS patients and mothers of AS patients, should have chromosome and FISH analysis to determine whether they carry balanced subtle chromosome rearrangements or deletions not expressed as an abnormal phenotype in that parent. Chromosome analysis also is appropriate for parents of patients with UPD combined with a Robertsonian translocation to determine whether the translocation is inherited or de novo. Parents also should be tested for mutations that are identified in the proband such as IC deletions or *UBE3A* mutations. If a parent of a PWS or AS patient is a carrier of a mutation or a chromosomal translocation, then the siblings of that parent should be offered testing.

Available Assays

Clinical molecular testing for PWS and AS includes the molecular assessment of the parent-specific imprint within the 15q11–q13 region. Methylation, which is involved in the process of genomic imprinting, has been demonstrated for several of the genes in the imprinted domain, including *SNRPN*. Standard molecular techniques for the methylation analysis at the CpG island in exon α of the *SNRPN* gene include (1) double-digest Southern blot analysis using a methylation-sensitive enzyme such as *Not* I along with a methylation-insensitive enzyme such as *Xba* I,[16] (2) PCR following either *Not* I or McrBC digestion with primers for the *SNRPN* promoter, and (3) methylation-specific PCR based on modifying DNA with bisulfite, which converts all unmethylated cytosines to uracils, followed by amplification using primers specific for the unmethylated and methylated alleles.[17] The assessment of *SNRPN* expression by reverse transcription–PCR (RT-PCR) also may be used for the diagnosis of PWS.

For the purpose of genetic counseling, once the diagnosis of PWS or AS is established by abnormal methylation or *SNRPN* expression testing, further tests should be performed to determine the genetic mechanism responsible for the disorder. FISH of metaphase chromosomes using the probes *SNRPN* for PWS and either *SNRPN* or D15S10 for AS will detect the 4 Mb deletion in the majority of patients. For patients without a deletion, DNA marker analysis should be used to detect UPD. Using specimens from both parents and the affected child. Mutations of the IC account for the remaining PWS patients and some of the AS patients, and referral may be made to a research laboratory for further investigation. Patients with an AS phenotype but normal methylation should be assisted to have testing by a clinical laboratory offering mutation analysis of *UBE3A*.

Interpretation of Test Results

If the methylation pattern or methylation-specific amplification is characteristic of only maternal inheritance, then the diagnosis of PWS is confirmed. If the methylation pattern or methylation-specific amplification is characteristic of only paternal inheritance, then the diagnosis of AS is confirmed (Figure 6-3). Methylation assays detect all cases of PWS and AS caused by deletions, UPD, and IC defects; however, such assays will not detect the rare small deletions not involving the *SNRPN* locus or *UBE3A* mutations. While the methylation assays cannot differentiate among the various mechanisms producing PWS and AS, 99% of PWS patients and approximately 78% of AS patients will be detected. Approximately 50% of the remaining patients with an AS phenotype have mutations in *UBE3A*, which are identified by single strand conformation polymorphism (SSCP) analysis and sequencing.

Two different approaches for the laboratory diagnosis of PWS and AS may be used.[18] The first is to start with an analysis of parent-specific methylation of *SNRPN*. A pattern consistent with biparental inheritance rules out PWS and most cases of AS. If the diagnosis of PWS or AS is confirmed by methylation analysis, FISH and DNA marker analysis should be performed to determine whether the cause is a deletion, a UPD, or an IC defect. The second approach takes into account that deletions are the

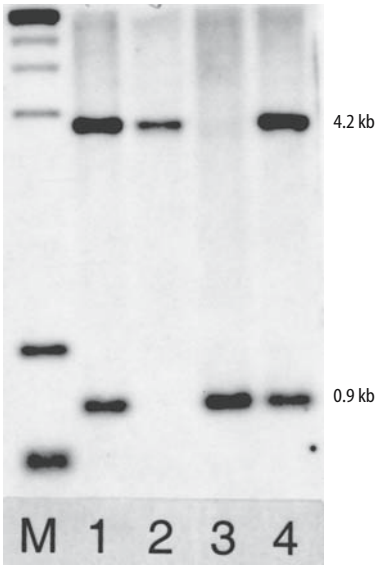


Figure 6-3. Southern blot analysis of PWS and AS. The methylated 4.2 kb (maternal) and the unmethylated 0.9 kb (paternal) fragments generated using *Xba* I and *Not* I are detected by the DNA probe *SNRPN*. Normal individuals exhibit a biparental inheritance pattern (lanes 1 and 4), whereas patients with PWS show a pattern of only maternal inheritance (lane 2) and patients with AS show a pattern of only paternal inheritance (lane 3).

most common cause of PWS and AS. It may be preferable to perform FISH analysis first if the patient is an older child or adult with classic features of PWS or AS. Methylation analysis is appropriate if a deletion is not detected. In the case of abnormal methylation, UPD studies using DNA markers should be performed to determine whether UPD is the cause. For either approach, if UPD is not present, then referral to a clinical laboratory for mutation analysis of *UBE3A* should be considered for patients with an AS phenotype. For patients suspected to have PWS, referral to a research laboratory for further molecular investigation of an IC defect should be considered.

Laboratory Issues

If the etiology of developmental delay or MR in a patient is unknown, DNA analysis should be performed as part of a comprehensive genetic evaluation that includes routine cytogenetic analysis (karyotyping). Cytogenetic abnormalities, including unbalanced translocations and deletions, have been identified in individuals referred for testing for PWS or AS. Although it is not necessary for every laboratory to provide all testing methodologies, a smooth progression through the various testing types may be necessary to determine the genetic mechanism causing PWS or AS, and is facilitated if the laboratory is able to perform at least DNA methylation, UPD, and FISH analyses. If sequencing of the *UBE3A* gene is necessary, testing may be referred to a clinical laboratory offering this test. Mutation analysis of the IC is available only on a research basis.

The *SNRPN* probe for Southern blot analysis is available from American Type Culture Collection (http://www.atcc.org). Many laboratories use laboratory-developed testing methods for methylation-sensitive PCR or RT-PCR analysis. Oligonucleotide primers for amplification of marker loci for UPD analysis can be synthesized or obtained commercially. FISH probes are available from two commercial sources (Cytocell, Inc, and Vysis, Inc). Patient control cell lines may be purchased from the Coriell Cell Repositories (http://coriell.umdnj.edu/). The proficiency tests for molecular genetics provided by the CAP include specimens for PWS and AS testing.

RETT SYNDROME

Molecular Basis

Rett syndrome is an X-linked neurodevelopmental disorder that almost exclusively affects females. In the classic form of the disease, affected girls appear to develop normally until the age of 6 to 18 months followed by a characteristic pattern of regression, which includes deceleration of head growth leading to acquired microcephaly, autistic features, loss of speech and purposeful hand use, irregular breathing patterns, stereotypical hand wringing, and seizures.[19] The frequency of classic Rett syndrome is approximately 1 in 15,000 females. Since the gene responsible for Rett syndrome was identified in 1999, however, a broader range of clinical phenotypes have been associated with mutation of this gene. The gene is *MECP2* on chromosome Xq28, which encodes the methyl-CpG-binding protein 2.[20] MECP2 binds preferentially to methylated DNA via its methyl-CpG-binding domain (MBD), and silences transcription by recruiting corepressor complexes through its transcriptional repression domain (TRD). Instead of serving as a global transcriptional repressor, as initially speculated, protein expression studies suggest that MECP2 may have a specialized role in neuronal maturation. The neuropathological findings in Rett syndrome of individuals with brains indicate that arrested neuronal development is an underlying feature of this disorder.[21]

Clinical Utility of Testing

Independent studies have confirmed *MECP2* as the major causative gene for Rett syndrome by the identification of multiple pathogenic mutations in approximately 95% of classic cases. Approximately 85% of classic Rett patients have point mutations and small rearrangements within the coding region; more recently, large gene rearrangements involving *MECP2* were identified in ~10% of classic cases.[22] Mutations also were found in atypical mild variant cases and in severe early-onset variant cases of Rett syndrome. *MECP2* mutations also have been identified in rare affected males with variable phenotypes ranging from neonatal-lethal encephalopathy and a Rett syndrome–like presentation in mosaic or Klinefelter males, to males with uncharacterized MR. In addition, *MECP2* mutations were

documented in patients with an AS-like presentation and in patients with autistic phenotypes. The variability in phenotypic severity observed in individuals carrying an *MECP2* mutation can result from allelic heterogeneity, as well as the X-inactivation pattern in females.[21] Given the spectrum of neurodevelopmental phenotypes associated with *MECP2* mutations, the clinical utility of molecular testing is significant.

Available Assays

The *MECP2* gene is composed of four exons, which give rise to two distinct MeCP2 isoforms.[23] Given the gene structure and mutation profile for *MECP2*, diagnostic testing for Rett syndrome is recommended to begin with analysis of the entire *MECP2* coding region (exons 1 through 4) by mutation scanning or DNA sequencing. One strategy for mutation scanning is denaturing high performance liquid chromatography (DHPLC), wherein sequence variants that give rise to heteroduplex DNA molecules can be detected with a high degree of sensitivity (95% to 100% under optimized conditions). The specificity of DHPLC is low, however, which requires that positive results be confirmed by DNA sequencing to identify the exact nucleotide change. The use of DHPLC coupled with sequencing has identified multiple recurrent and novel *MECP2* mutations of different types (missense, nonsense, splice-site, frameshifting deletions, and insertions).[24] Figure 6-4 shows representative data for DHPLC and DNA sequence analyses, demonstrating a heterozygous nonsense mutation in the *MECP2* gene (880C→T, R294X). This represents one of the more common truncating mutations seen in Rett syndrome patients. Sequencing is largely considered the gold standard for point mutation detection. Approximately 85% of classic Rett syndrome patients have mutations that are detectable by mutation scanning or sequencing of the *MECP2* gene, which is performed by clinical molecular laboratories (http://www.genetests.org/).

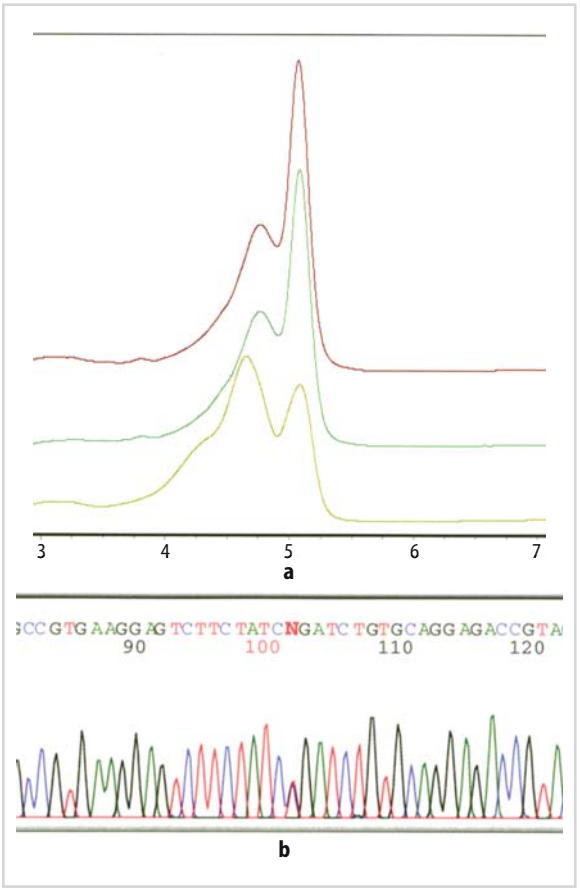


Figure 6-4. (a) Identification of a heterozygous sequence variation in the *MECP2* gene by DHPLC analysis. Elution profiles of *MECP2* exon 4 PCR products corresponding to the wild-type (red and green) and the mutant (yellow) sequences are shown. (b) Identification of the corresponding *MECP2* nonsense mutation by DNA sequencing. A single base change of C to T at nucleotide 880 is identified (N), which predicts an arginine to stop substitution at residue 294 of the MeCP2 protein (880C→T, R294X).

To increase overall mutation detection for the *MECP2* gene, additional testing is available for large *MECP2* gene rearrangements, present in approximately 10% of classic Rett patients. Deletions, insertions, or duplications involving all or part of the *MECP2* gene have been identified by dosage-sensitive DNA testing methods. The classic Southern analysis method has been used to detect copy number differences corresponding to deletions or duplications of specific *MECP2* exons, with detection of associated junction fragments in some cases. Alternative methods for dosage-sensitive analysis of the *MECP2* gene include multiplex ligation-dependent probe amplification (MLPA; MCR Holland, the Netherlands) or quantitative real-time PCR analysis.

Laboratory Issues

One of the issues encountered in sequence-based clinical testing is the interpretation of novel sequence variations, particularly alleles of uncertain pathogenic significance. A valuable resource for laboratories and referring clinicians is provided by the American College of Medical Genetics (ACMG), which has issued recommended standards for interpretation of sequence variations (ACMG Laboratory Practice Committee Working Group 2000; policy statement available at http://www.acmg.net/).[25] Another issue pertaining to Rett syndrome DNA testing at the current time is the value of two-tier testing (e.g., sequencing followed by dosage analysis) to provide comprehensive mutation analysis of the *MECP2* gene. Identification of the mutation in the proband facilitates prenatal testing in subsequent pregnancies. The majority of *MECP2* point mutations are new mutations of paternal origin, with low recurrence risk. Prenatal testing is recommended for fetuses of mothers who are identified to carry a point mutation or large rearrangement in the *MECP2* gene. In cases where the mother is not a carrier, prenatal testing may be sought for parental reassurance due to rare reports of gonadal mosaicism. The current focus on point mutation analysis

within the coding region by sequence analysis does not rule out potential mutations in regulatory elements or other important noncoding regions of the *MECP2* gene. Ideally, barriers to increased uptake of *MECP2* clinical testing may be addressed by improvements in genotype-phenotype correlations, assay design, and clinical services. With the accumulation of additional genetic information and the development of improved DNA technologies, appropriate enhancements can be incorporated into clinical testing of the *MECP2* gene for Rett syndrome.

WILLIAMS SYNDROME

Molecular Basis of Disease

Williams syndrome (WS), also referred to as Williams-Beuren syndrome, is an autosomal dominant disorder that occurs in approximately 1 in 10,000 to 20,000 live births (Online Mendelian Inheritance in Man [OMIM; database online] #194050). Clinical diagnosis is based on the presence of characteristic dysmorphic facial features, growth delay, abnormalities of connective tissue, cardiovascular disease, developmental delay, an outgoing personality type, and a unique cognitive profile of specific mental strengths and weaknesses. A less-frequent finding is hypercalcemia in the first year of life.[26]

Approximately 95% to 99% of individuals with WS have complete deletion of one copy of the elastin gene (*ELN*) located in chromosome region 7q11.23.[26] The most common form of deletion in WS extends well beyond *ELN* in both the 3′ and 5′ directions and results in hemizygosity for genes contained in approximately 1.5 Mb of DNA. The common deletion segment is flanked by repetitive DNA sequences, or "duplicons," that serve as agents for chromosomal rearrangements. This is achieved by creating an opportunity for misalignment of the chromosome 7 homologs during meiotic pairing. Unequal recombination then may yield deletion and duplication products, and in this case, deletion of the 1.5 Mb region between the duplicons results in WS. Deletion also may occur as a result of intrachromosomal rearrangements, and the deletion may involve either the maternal or paternal chromosome 7. These repeats also mediate an inversion of this region, which may result in an increased risk for the WS-causing deletion.[27]

At least 17 genes have been identified in the 1.5 Mb deletion region;[27] however, there is clear evidence for only one of the genes in the pathogenesis of WS. Supravalvular aortic stenosis (SVAS) and peripheral pulmonary artery stenosis are the most common forms of cardiovascular disease in WS. Elastin contributes to the elasticity of tissues such as skin and arteries, and is a major component of the aorta. There is strong evidence that abnormal elastin production is responsible for the arterial disease of WS.[26] Point mutations and intragenic deletions of the elastin gene have been associated with an autosomal dominant form of SVAS. Humans and mouse models hemizygous for *ELN* show that reduced elastin production produces a compensatory increase in lamellar units in response to increased wall stress during development. This results in the obstructive vascular disease of SVAS.[28] Hemizygosity for *ELN* also may contribute to the other connective tissue abnormalities of WS, such as hernias, hoarse voice, joint laxity, and premature aging of the skin.[26]

Evidence for a role of other genes in the clinical aspects of WS is less clear. *LIMKI,* a tyrosine kinase that is expressed in the brain during embryogenesis, may be involved in the cognitive portion of the WS phenotype.[26,29]

Clinical Utility of Testing

Demonstration of deletion of the elastin gene is used to confirm a clinical diagnosis of WS. This deletion is not visible by routine or high-resolution chromosome analysis. Most cases occur sporadically; however, there have been rare cases of multiple affected siblings. Affected individuals have a 50% chance of passing the deletion on to their offspring, and the deletion can be detected prenatally.

Available Assays

FISH is the most commonly used testing method for diagnosis of WS. The analysis is performed with a DNA probe containing a full-length copy of the *ELN* gene. Metaphase chromosome preparations from cultured peripheral blood lymphocytes are hybridized with the probe, and the 7q11.23 region of both chromosomes 7 are analyzed for the presence or absence of a hybridization signal. The assay generally includes a second probe, either for the centromere or a more distal locus on chromosome 7, to serve as an internal hybridization control.

Quantitative PCR can be used to assay the *ELN* copy number; however, there is no significant clinical or cost benefit to PCR over FISH. Whenever available, FISH is the preferred method for assessing genomic copy number because of the value gained by assessing individual cells. FISH testing for WS is usually ordered in conjunction with karyotyping, as other chromosome abnormalities may be part of the differential diagnosis. Addition of FISH testing requires no additional specimen and can be performed on fixed cell pellets after karyotyping is completed. In addition, although it is extremely rare, there have been reports of visible chromosome rearrangements that may result in disruption of the WS region, that would not be detected by PCR.

Interpretation of Test Results

Approximately 95% to 99% of individuals with a WS phenotype have complete deletion of the *ELN* gene, and a FISH assay with the *ELN* gene probe will detect all of these cases.

FISH analysis will not detect the *ELN* mutations that are commonly found in nonsyndromic SVAS; thus, the specificity for the test with respect to WS is very high. Nevertheless, it is possible that a rare *ELN* deletion in an individual with SVAS but without WS may be detected by FISH. As in all genetic disorders, the laboratory result must be interpreted in the context of the individual's complete clinical presentation. Conversely, a negative result should suggest reevaluation of the clinical diagnosis of WS.

Laboratory Issues

The FISH test is performed by cytogenetics laboratories, which routinely process cultured cells. Probes for WS testing are commercially available (sources include Cytocell, Inc, and Vysis, Inc). These probe sets are classified by the US Food and Drug Administration (FDA) as analyte specific reagents. Recommendations for test validation methods have been established by the ACMG (http://www.acmg.net/). Proficiency testing for FISH analysis of microdeletion disorders, such as WS, is available through the CAP.

SUBTELOMERIC DISORDERS

Molecular Basis of Disease

Genetic imbalance caused by a chromosome abnormality frequently results in some form of developmental delay and one or more of the following: abnormal growth, dysmorphic facial features, and congenital abnormalities. Duplications or deletions of very small chromosomal segments are often difficult or impossible to detect by routine or high-resolution chromosome analysis. Though small relative to larger and thus more readily visible rearrangements, these imbalances nonetheless frequently result in an abnormal phenotype due to the resulting imbalance of multiple genes within the affected segment(s).

The subtelomeric region of each chromosome arm is located between the telomere, the functional chromosome cap, and more proximal chromosome-specific sequences.[30] Subtelomeric regions and their adjoining chromosome-specific areas are of particular clinical interest for several reasons. They are particularly gene-dense regions, and shared homology within these segments may facilitate relatively frequent recombination events that may result in loss or gain of genetic material.[31] This suggests that genetic imbalances in these regions may have considerable clinical consequences. Additionally, the majority of all chromosome abnormalities involve imbalance for, or movement of, a segment that contains a subtelomeric locus. Thus, testing panels have been designed to assess the genomic copy number of arm-specific loci within or near the subtelomeric regions of each chromosome for identification of imbalances that are not readily detected by standard chromosome analysis methods.

Clinical Utility of Testing

Because developmental delay is a frequent finding in chromosomal imbalance, the main focus for these assays has been in studies of individuals with idiopathic mental retardation (IMR). An average of 6% of individuals with IMR have an imbalance that involves one or more distal chromosomal segments.[32] Based on these data, subtelomeric testing is the next logical step following a normal routine or (ideally) high-resolution karyotyping study, for individuals whose phenotype suggests an underlying chromosomal etiology. Subtelomeric panel testing is not routinely used for prenatal diagnosis at this time due to constraints of the test methodologies and to potential limitations in test interpretation in the absence of a well-defined phenotype (see Interpretation of Test Results section below).

Available Assays

Telomere FISH tests of metaphase chromosome preparations are currently in use in clinical laboratories. The test is also referred to as a "telomere panel" or "telomere testing," as it is the telomeric or distal end of the chromosome that is assessed. The test utilizes a complete set of chromosome arm-specific probes for loci that are within several hundred kilobases of their respective telomeres.[33] The short arms of the five acrocentric chromosomes are not included since they contain only repetitive sequences, and imbalance of these regions has no clinical consequence. Regions of homology exist for the terminal ends of Xp and Yp, and also for Xq and Yq; therefore, one set of probes hybridizes to both sex chromosomes. This leaves a minimum number of 41 probes needed for complete analysis. The test uses standard FISH methodology in which two or more probes are grouped together into one reaction mixture by using different probe colors, and multiple hybridizations are carried out with each group of probes for one complete test (Figure 6-5).

Research surveys have used PCR analysis of genetic markers, such as VNTRs or STRs, on DNA samples. A disadvantage to this method is that it requires parental DNA samples, which are not always available. STR analysis is more amenable to automation than the FISH assay, and could ultimately be less costly; however, it is more likely to yield false-positive results because of the presence of sequence polymorphisms.[32] The FISH method is more labor-intensive and requires actively dividing cells, but can also determine the position of each signal. This has particular benefit in detecting reciprocal translocations, because a de novo balanced translocation could generate an abnormal phenotype through gene disruption.

A recently introduced testing method uses comparative genomic hybridization (CGH) with subtelomeric FISH probes in a microarray format. The microarray contains bound subtelomeric probes that are hybridized with patient

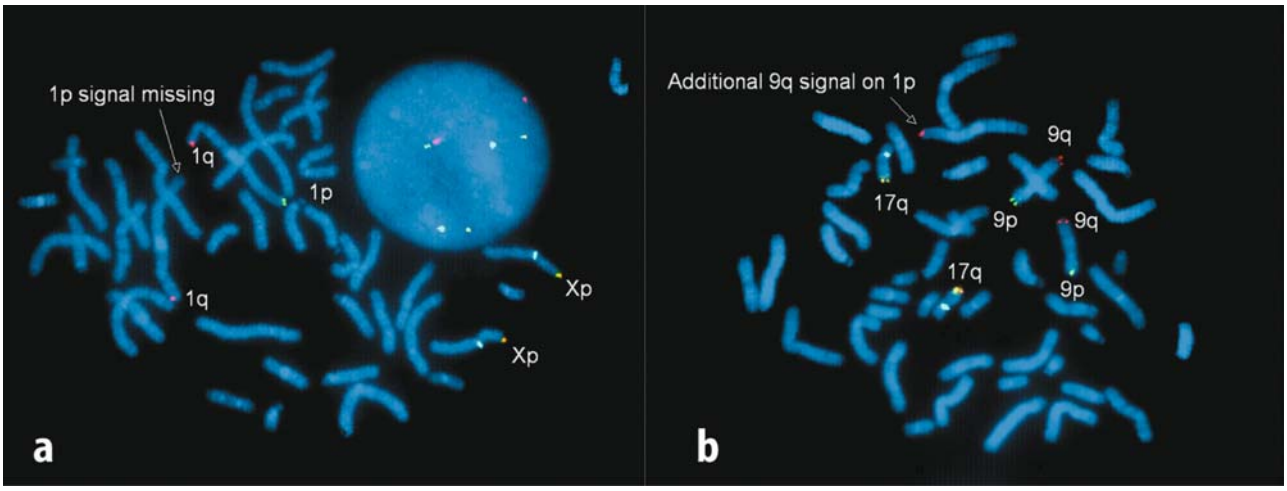


Figure 6-5. Subtelomeric FISH detection of abnormal chromosome 1, which is missing material from the distal short arm (1p) and contains additional material from distal 9q. (a) Hybridization mixture with probes for subtelomeric regions of 1p, 1q, Xp/Yp, and X centromere (control probe; aqua signals). Note absence of green signal on one 1p terminus. (b) Cell from the same individual after hybridization with probes for 9p, 9q, 17q, and 17 centromere (control probe; aqua signals). Note additional 9q signal on the end of 1p. The mother carries a reciprocal translocation between chromosomes 1 and 9. (Courtesy of Emory Genetics Laboratory, Atlanta, GA)

DNA labeled in one color, and a control DNA labeled in a second color. Significant deviation from a 1 : 1 hybridization ratio indicates an area of imbalance in the patient's DNA. All probes can be assessed on a single chip from one DNA specimen, without the need for cultured cells or parental samples, and with the potential for automation.

Interpretation of Test Results

FISH tests have had the widest usage in clinical laboratories. The individual probe sensitivities and specificities are generally very high. One caveat is that the exact origin of an imbalance that involves an additional Xp/Yp or Xq/Yq subtelomeric signal cannot be determined without additional analysis because of the shared homology within these regions. In general, the FISH test is able to detect any imbalance of a segment that contains a probe locus. Because this test has been in clinical use for only a few years, its level of specificity is less certain.

An abnormal result in a proband should ideally be confirmed by FISH analysis with a second probe for the same region. Follow-up performance of a targeted FISH analysis on parental chromosomes to determine whether the imbalance in the proband is due to an inherited rearrangement is very important for genetic counseling. A significant proportion of abnormalities detected by subtelomeric assays are familial in nature, and these cases may have significant risk for recurrence. Reports of families in which identical abnormal signal patterns have been found in a clinically abnormal proband and a clinically normal parent provide a note of caution for interpretation of results. Such findings may indicate locus polymorphisms, imbalances for which there are no obvious clinical significances, or imprinted areas in which the parent of origin determines the clinical consequence of the imbalance.

Some abnormalities detected by this assay fall into categories of well-described deletion or duplication syndromes. Other cases may represent unique situations in which the abnormality consists of a derivative chromosome containing imbalances from two different chromosomal segments. The clinical literature regarding the phenotypes associated with small subtelomeric imbalances is rapidly expanding; however, it must be remembered that this assay does not determine the size of the imbalance. Two individuals with identical abnormal hybridization patterns may have very different phenotypes due to differences in the extent of genetic imbalance.

Laboratory Issues

Two commercial sources for FISH probes are currently available (suppliers are Vysis, Inc, and Cytocell, Inc). A CGH-based microarray test system also is commercially available (Vysis, Inc). The FISH probe sets are classified by the FDA as analyte-specific reagents. Recommendations for FISH test validation methods have been established by the ACMG (http://www.acmg.net/). Proficiency testing is not available for subtelomeric FISH testing; however, the CAP does provide proficiency testing for other types of FISH assays.

References

1. Pieretti M, Zhang FP, Fu YH, et al. Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell.* 1991;66:817–822.
2. Saul RA, Tarleton J, Fragile X syndrome. In: GeneReviews at GeneTests: Medical Genetics Information Resource [database online]. Seattle: University of Washington; 1997–2003.
3. Eichler EE, Richards S, Gibbs RA, et al. Fine structure of the human FMR1 gene. *Hum Mol Genet.* 1993;2:1147–1153.

4. Rousseau F, Heitz D, Biancalana V, et al. Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. *N Engl J Med.* 1991;325:1673–1681.
5. Fu YH, Kuhl DP, Pizzuti A, et al. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell.* 1991;67:1047–1058.
6. Tassone F, Hagerman RJ, Ikle DN, et al. FMRP expression as a potential prognostic indicator in fragile X syndrome. *Am J Med Genet.* 1999;84:250–261.
7. Engel E. A new genetic concept: uniparental disomy and its potential effect. *Am J Med Genet.* 1980;6:137–143.
8. Engel E, Antonarakis SE. *Genomic Imprinting and Uniparental Disomy in Medicine: Clinical and Molecular Aspects.* New York: Wiley-Liss; 2002.
9. Spence JE, Perciaccante RG, Greig GM, et al. Uniparental disomy as a mechanism for human genetic disease. *Am J Hum Genet.* 1988;42: 217–226.
10. Sapienza C, Hall JG, Scriver CR, Beaudet AR, Sly WS, et al. Genome imprinting in human disease. In: Scriver CR, et al., eds. *The Metabolic and Molecular Basis of Inherited Disease,* 8th ed. New York: McGraw-Hill; 2001.
11. Gallagher RC, Pils, B, Albalwi M, et al. Evidence for the role of PWCR1/HBII-85 C/D box small nucleolar RNAs in Prader-Willi syndrome. *Am J Hum Genet.* 2002;71:669–678.
12. Kishino T, Lalande M, Wagstaff J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet.* 1997;15:70–73.
13. Cassidy SB, Schwartz S. Prader-Willi syndrome. In: GeneReviews at GeneTests: Medical Genetics Information Resource [database online]. Seattle: University of Washington; 1997–2003.
14. Williams CA, Lossie AC, Driscoll DJ. Angelman syndrome. In: GeneReviews at GeneTests: Medical Genetics Information Resource [database online]. Seattle: University of Washington; 1997–2003.
15. Malzac P, Webber H, Moncla A, et al. Mutation analysis of UBE3A in Angelman syndrome patients. *Am J Hum Genet.* 1998;62:1353–1360.
16. Beuten J, Sutcliff JS, Casey BM, et al. Detection of imprinting mutations in Angelman syndrome using a probe for exon α of SNRPN [letter]. *Am J Med Genet.* 1996;63:414–415.
17. Kubota T, Das S, Christian SL, et al. Methylation-specific PCR simplifies imprinting analysis [letter]. *Nat Genet* 1997;16:16–17.
18. American Society of Human Genetics/American College of Medical Genetics Test and Technology Transfer Committee. Diagnostic testing for Prader-Willi and Angelman syndromes: report of the ASHG/ACMG Test and Technology Transfer Committee. *Am J Hum Genet.* 1996;58:1085–1088.
19. Hagberg B, Aicardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol.* 1983;14:471–479.
20. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat Genet.* 1999;23: 185–188.
21. Shahbazian MD, Zoghbi HY. Rett syndrome and MeCP2: linking epigenetics and neuronal function. *Am J Hum Genet.* 2002;71:1259–1272.
22. Fang P, Jin W, Glaze DG, Percy A, Zoghbi HY, Roa BB. MECP2 gene deletions account for ~10% of Rett syndrome cases. The American Society of Human Genetics. 2004;476:2652.
23. Mnatzakanian GN, Lohi H, Munteanu I, et al. A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat Genet.* 2004;36:339–341.
24. Buyse IM, Roa BB. Denaturing high-performance liquid chromatography and sequence analysis for *MECP2* mutations in Rett syndrome. In: Potter NT, ed. *Methods in Molecular Biology: Neurogenetics Methods and Protocols.* Vol. 217. Totowa, NJ: Humana Press; 2002:119–130.
25. ACMG Laboratory Practice Committee Working Group. ACMG recommendations for standards for interpretation of sequence variations. *Genet Med.* 2000;2:302–330.
26. Morris CA. Williams Syndrome. In: Cassidy SB, Allanson JE, eds. *Management of Genetic Syndromes,* 2nd ed. New York: Wiley-Liss; 2005:655–665.
27. Osborne LR, Li M, Pober B, et al. A 1.5 million-base pair inversion polymorphism in families with Williams-Beuren syndrome. *Nat Genet.* 2001;29:321–325.
28. Li DY, Faury G, Taylor DG, et al. Novel arterial pathology in mice and humans hemizygous for elastin. *J Clin Invest.* 1998;102:1783–1787.
29. Tassabehji M, Metcalfe K, Karmiloff-Smith A, et al. Williams syndrome: use of chromosomal microdeletions as a tool to dissect cognitive and physical phenotypes. *Am J Hum Genet.* 1999;64: 118–125.
30. Knight SJL, Flint J. Perfect endings: a review of subtelomeric probes and their use in clinical diagnosis. *J Med Genet.* 2000;37:401–409.
31. Mefford HC, Trask BJ. The complex structure and dynamic evolution of human subtelomeres. *Nat Rev Genet.* 2002;3:91–102.
32. Biesecker LG. The end of the beginning of chromosome ends. *Am J Med Genet.* 2002;107:263–266.
33. Knight SJL, Lese CM, Precht KS, et al. An optimized set of human telomere clones for studying telomere integrity and architecture. *Am J Hum Genet.* 2000;67:320–332.

Chapter 7

Neuromuscular Diseases

Thomas W. Prior

DUCHENNE AND BECKER MUSCULAR DYSTROPHIES

Molecular Basis of Disease

Duchenne and Becker muscular dystrophies (DMD and BMD) are X-linked, allelic, neuromuscular diseases characterized by progressive muscular weakness and degeneration of skeletal muscle. Duchenne muscular dystrophy is the most common X-linked recessive lethal disease, with an incidence of approximately 1 in 3,500 newborns, and approximately one third of cases are the result of new mutations.[1,2] Affected children are usually wheelchair bound by the age of 12 years. As the disease progresses, contractures increasingly develop, leading to asymmetrical spinal deformities. Most patients die at about 20 years of age due to pneumonia related to chronic respiratory insufficiency. The allelic disorder BMD has a milder clinical course and slower disease progression. Becker muscular dystrophy has been estimated to occur approximately one tenth as frequently as DMD, with an incidence of about 1 in 35,000. The majority of BMD patients initially experience difficulties between 5 and 15 years of age, although an onset in the third or fourth decade or even later can occur. By definition the affected patients remain ambulatory until 16 years of age or later, thus allowing clinical distinction from patients with DMD.

The *DMD* gene is the largest human gene isolated, spanning more than 2000 kilobases (kb) of genomic DNA, and is composed of 79 exons that encode a 14 kb transcript, which is translated into a protein named dystrophin.[3,4] Dystrophin is a 427 kilodalton (kDa) cytoskeletal protein consisting of 4 domains: (1) an amino terminus that associates with actin or an actin-like protein, (2) a rod domain consisting of long, flexible rows of 24 alpha helical repeats, (3) a cysteine-rich region, and (4) a unique carboxy terminus.[5] Dystrophin has been shown to be tightly associated with a large oligomeric complex of sarcolemmal glycoproteins through its cysteine-rich domain and carboxy terminus, while the amino-terminal domain interacts with actin or an actin-like protein.

By immunochemistry, dystrophin localizes to the cytoplasmic face of the muscle cell membrane and at postsynaptic membrane specializations in neurons. Dystrophin makes up only 0.002% of total muscle protein but up to 5% of the membrane skeleton. Dystrophin is found in skeletal muscle, smooth muscle, cardiac muscle, and brain. There are slightly different forms of dystrophin messenger RNA (mRNA) in different tissues due to different transcription start sites and alternative splicing. The function of dystrophin is not known for certain, but proposed functions for the protein include important roles for the organization and stabilization of the sarcolemma and in protecting muscle fibers from contraction-induced injury. Patients with DMD have very little or no detectable dystrophin, whereas BMD patients have an altered size and/or quantity of dystrophin.[6] However, the disease etiology may be more complex than a simple loss of dystrophin, because several of the dystrophin-associated proteins that interact with dystrophin also are absent. The dystrophin-associated proteins may be directly involved with the calcium flux in the dystrophic fibers. Thus, the loss of dystrophin may be the first of many steps that ultimately lead to muscular dystrophy.

Utilizing complementary DNA (cDNA) probes derived from the 14 kb mRNA and multiplex polymerase chain reaction (PCR) analysis, approximately 65% of the DMD/BMD cases are due to deletions in the dystrophin gene.[7,8] The deletions are nonrandomly distributed and occur primarily in the center (~80%) and less frequently near the 5′ end (~20%) of the gene. The 200 kb region covering intron 44, exon 45, and intron 45 is the major deletion breakpoint region of the gene. The majority of the larger deletions initiate at the 5′ end of the gene.

There is no apparent correlation between the size or location of the deletion and the severity and progression of the disease. One of the largest deletions (35 exons) identified is in a mild BMD patient. Furthermore, sequences deleted in DMD patients often overlap with

deletions in BMD patients. However, it was proposed that if a deletion disrupts the translational reading frame of the dystrophin mRNA triplet codons, then little or no dystrophin will be synthesized, resulting in the more severe disease, DMD.[9] In the milder disease, BMD, the deletion maintains the translational reading frame, and a partially functional protein is produced. The reading frame hypothesis explains the phenotypic differences observed in about 92% of the DMD/BMD cases. One major exception to the reading frame hypothesis has been the identification of BMD patients with an out-of-frame deletion of exon 3 through exon 7. An alternate splicing mechanism or new cryptic translational start site may account for the production of an altered dystrophin protein and the milder phenotype in these patients. A small number of DMD patients with in-frame deletions have also been identified. The more severe phenotype in these patients may be due to the overall effect of the deletion on the protein conformation or may be the result of mRNA instability. Phenotypic variability has even been observed in several patients who share identical gene deletions. Deletion of exon 45, the most commonly observed DMD deletion, has also been associated with the BMD phenotype. Some genetic variability may be due to other molecules involved in destruction of damaged muscle fibers, in muscle regeneration, or in the cellular response to different hormones.

The large gene size, particularly of the introns, which average 35 kb, may account for part of the high deletion rate; however, in addition to size, other factors must be involved. The observed nonrandom deletion pattern may reflect domain-associated variation in chromosomal stability. For instance, complications related to the maintenance of replication, correct transcription, and proper splicing of such a large gene may play an extremely important role.

Partial gene duplications have been identified in 5% to 8% of patients. Unlike the deletion distribution, approximately 80% of the duplications are located at the 5′ end of the gene and only 20% in the central gene region. Out-of-frame duplications occur in DMD patients and in-frame duplications in BMD patients, thus suggesting that the reading-frame genotype-phenotype hypothesis also holds true for duplications.

Small mutations (point mutations and small deletions and duplications) in the dystrophin gene also have been identified in DMD patients.[10] The majority of these mutations have been unique to individual patients and have resulted in a truncated dystrophin protein lacking part or all of the C-terminus. The truncated proteins are presumably unstable, and little or no dystrophin is produced. Therefore, these types of mutations provide little information on structural/functional relationships in the dystrophin protein. The identification of DMD mutations that do not cause protein truncation may provide us with further insight into the function of dystrophin, as well as defining the essential regions and conformations necessary for dystrophin stability. A DMD missense mutation was found in the actin-binding domain.[11] The patient was shown to have correctly localized dystrophin, thus indicating that an intact actin-binding domain is essential for function. The distribution of small mutations is fairly random throughout the gene sequence. However, whereas less than 5% of the gene deletions are found upstream of exon 55, more than 40% of the small mutations are located in this same region of the gene.[12]

Clinical Utility of Testing and Available Testing

The most important impact of DNA testing for DMD/BMD has been on presymptomatic diagnosis and the correct diagnosis of patients with DMD/BMD-like symptoms, as well as for carrier detection and prenatal diagnosis. The western immunoblot is the most sensitive test for the diagnosis of DMD or BMD, but requires a muscle biopsy from the patient. Proteins from the muscle biopsy specimen are separated by electrophoresis, transferred to a membrane, and incubated with antibodies against specific dystrophin domains. Dystrophin from unaffected patients is 427 kDa in size. The complete absence of dystrophin is very specific for the diagnosis of DMD, whereas the presence of an altered molecular weight form or reduced amount of dystrophin, or both, is consistent with BMD. It must be emphasized, however, that the clinical diagnosis and phenotype predictions should never be made exclusively on the basis of dystrophin western blot analysis. Patients with markedly reduced dystrophin levels may have the milder BMD phenotype. Furthermore, western blot analysis is generally not sensitive enough to detect female carriers.

Southern blot analysis using a full-length dystrophin cDNA clone as a probe detects 65% of deletions in affected DMD/BMD patients, as well as gene duplications. The cDNA probes detect the site of the mutation, so meiotic recombination events are irrelevant. Therefore, the chance of diagnostic error is greatly reduced. Diagnostic strategies are based on initially identifying deletions in the affected patient.[13] The deletions are simply detected by examination of Southern blots for the presence or absence of each exon containing genomic restriction fragments, which hybridize to the cDNA probe. However, Southern blot analysis requires the use of a radioisotope, is labor intensive, and is time-consuming.

Before performing Southern blot analysis with the *DMD* cDNA probes, most laboratories perform an initial deletion screen using multiplex PCR[14] to amplify specific deletion-prone exons within the *DMD* gene. Deletion of specific exons from a patient's sample results in absence of the corresponding ethidium bromide–stained amplification products by gel electrophoresis analysis (Figure 7-1). Mul-

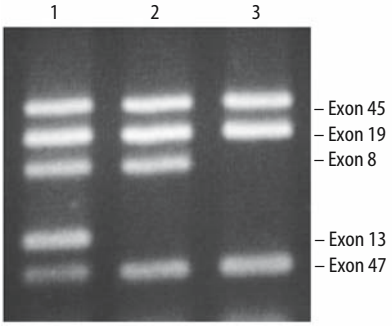


Figure 7-1. Multiplex DNA amplification of DMD exons 8, 13, 19, 45, and 47. Lane 1: normal control; lane 2: DMD patient with exon 13 deletion; lane 3: DMD patient with exon 8 and 13 deletions.

tiplex PCR, using primer sets for about 20 different exons, detects approximately 98% of the deletions in the *DMD* gene. In contrast to Southern blot analysis, which may require several cDNA hybridizations and take several weeks to complete, multiplex PCR testing can be completed in one day. This makes the technique ideal for prenatal diagnosis, when time is critical.

The identification of a deletion in a DMD patient not only confirms the diagnosis but also allows accurate carrier testing for other family members. Carrier status is determined by gene dosage analysis, which is used to assess whether a female at risk of being a carrier exhibits no reduction or 50% reduction in band intensity in those bands that are deleted for the affected male relative. A 50% reduction (single-copy intensity) for the deleted band or bands on the autoradiograph indicates a deletion on one of her X chromosomes and confirms her carrier status. Either Southern blot or quantitative PCR can be used for gene dosage analysis. Dosage determinations permit direct carrier analysis and eliminate the inherent problems of Southern blot analysis without gene dosage (recombinations, noninformative meioses, unavailability of family members, and spontaneous mutations). This is important since unlike affected males, carrier females are generally asymptomatic, and creatine phosphokinase (CPK) is elevated in only approximately two thirds of known carriers.

In 35% of families with undefined mutations, linkage analysis using microsatellite markers must be used for carrier detection and prenatal diagnosis. Microsatellite markers, which correspond to short tandem repeats (di-, tri-, or tetranucleotides) and tend to be highly polymorphic in repeat number, have been found in several locations in the *DMD* gene and have significantly improved linkage analysis for DMD/BMD.[15] The method relies on the coinheritance of the disease gene with those DNA microsatellite sequence variations known to be located within the disease gene. Thus, even when the responsible gene mutation remains unknown, the linkage technique allows the mutation to be traced through an affected family for prediction of carrier or disease risk.

Interpretation of Test Results

The analysis of gene mutations and protein determinations has greatly improved diagnosis, carrier detection, and prenatal counseling. The multiplex PCR deletion test will confirm the clinical diagnosis in approximately 65% of affected DMD and BMD patients. A major advantage of the Southern blot analysis is the detection of an additional 5% to 8% of patients with duplications. The standard multiplex PCR test detects deletions but not duplications. However, if there is any question of the diagnosis after negative results by deletion and duplication molecular testing, western blot analysis of the dystrophin protein should be performed on a muscle biopsy specimen.

When gene dosage testing indicates that the mother does not have the deletion present in the affected child, she still has an uncertain risk of carrier status, due to the possibility of germline mosaicism.[16] Cases of germline mosaicism in DMD have been reported, in which a deletion is transmitted to more than one offspring by a mother who shows no evidence of the mutation in her somatic cells. Cases of germline mosaicism have important counseling implications. First and most obvious is the need to perform carrier studies on all female siblings of affected males, regardless of the outcome of testing for the mother. Furthermore, a negative deletion result in a mother does not rule out a recurrence risk for future pregnancies, and prenatal diagnosis still should be offered. The exact recurrence risk in germline female carriers is unknown because the risk is related to the size of the mutant clone in the mosaic mother. However, in these cases the recurrence risk for subsequent pregnancies is significantly increased relative to what had initially been perceived as a new mutation with a low recurrence risk. It has been estimated that mothers of apparently sporadic DMD cases have an approximate 15% recurrence risk in future pregnancies.

Linkage analysis can provide valuable information but is limited by the possibility of recombination between the polymorphic marker and the unknown mutation, the presence of sporadic mutations, and unavailability of family members. The intragenic recombination rate over the entire length of the DMD gene is estimated to be as high as 12%. The high recombinational error rate can be partially overcome by using microsatellite markers throughout the gene. Linkage analysis results often are extremely limited for extended family members of isolated cases of DMD/BMD, due to the possibility of the occurrence of a new mutation. Linkage analysis indicates only whether the female at risk inherited the same X chromosome as the affected male, not whether she is a carrier of a defective gene. Furthermore, since the gene mutation remains unidentified, a correct clinical diagnosis is essential. This is extremely important with patients presenting with the milder BMD, since this phenotype can overlap with other neuromuscular disorders.

Laboratory Issues

Today, there are many laboratories offering dystrophin-deletion testing. Most laboratories use multiplex PCR, and amplify 16 to 23 deletion-prone *DMD* exons. Since the deletion hotspot is from exon 44 to exon 53, more exons are amplified in this region. Although the 5′ deletions are generally larger, it is important for laboratories to amplify about every third or fourth exon at the 5′ end of the gene; otherwise, smaller deletions are missed. The multiplex strategies are a deletion test and will not detect duplications. Although Southern blot analysis is both costly and labor-intensive, duplications will be detected. Unlike the deletion distribution, duplications are found in about 5% to 8% of the patients and are located primarily at the 5′ end of the gene. The identification of duplications is important both for the confirmation of the diagnosis and for accurate carrier studies for other family members. When laboratories encounter challenging nondeletion types of cases, Southern blot analysis should be performed. Proficiency testing for both *DMD* deletion testing and carrier testing is offered through the College of American Pathologists (CAP), and proficiency specimens are sent to participants twice per year.

MYOTONIC DYSTROPHY

Molecular Basis of Disease

Myotonic dystrophy (DM) is the most common inherited form of muscular dystrophy affecting adults, having an incidence of approximately 1 in 8,000 individuals. DM is an autosomal dominant, multisystem disorder characterized by progressive muscle weakness and myotonia. The diagnosis can be problematic because of the wide range and severity of symptoms. Often, affected individuals have children before they are diagnosed. A severe congenital form of DM results in mental retardation, respiratory distress, hypotonia, and in many cases death due to respiratory complications shortly after birth. The congenital form is seen in the offspring of women who are themselves mildly affected.[17] The clinical genetic phenomenon of anticipation occurs in DM. Anticipation denotes progressively earlier appearance of a disease in successive generations, generally with increasing severity.

The DM mutation has been characterized as an unstable trinucleotide repeat present on chromosome 19q13.3.[18–20] A polymorphic CTG repeat ranges in size from 5 to approximately 30 repeats in the normal population. Mildly affected patients have 50 to 80 repeats, whereas more severely affected individuals have more than 1,000 repeats. The number of repeats varies between affected sibs and increases through generations in parallel with an increasing severity of the disease. Expansion of the CTG repeat between generations accounts for the clinical genetic anticipation phenomenon.

The CTG repeat is located within the 3′ untranslated region of the *DMPK* gene that encodes a protein kinase, named myotonin protein kinase. Since protein kinases are involved in signal transduction pathways in all cells in the body, a defective protein kinase may explain how a single gene defect could result in the diverse symptoms characteristic of DM. However, since the repeat is not in the protein coding region of *DMPK*, the molecular mechanism by which the mutation exerts its dominant expression is difficult to explain. A study using quantitative reverse transcription–PCR (RT-PCR) and a radioimmunoassay demonstrated that decreased levels of mRNA and protein expression are associated with the adult form of DM,[21] suggesting that the dominant nature of DM is likely the result of a dosage-dependent mechanism. However, in another report, expression studies on samples from congenital DM cases demonstrated marked increases in the steady-state levels of the *DMPK* mRNA.[22] In contrast to a reduction of myotonin kinase, the authors proposed that the effect of the mutation would be a nonregulated hyperphosphorylation of kinase substrate by high levels of myotonin kinase.

Another possible disease mechanism is that the repeat expansion affects the expression of other genes in the region. Since the CTG expansion is located in the 3′-untranslated region of the gene, downstream gene expression, like that of DM locus-associated homeodomain protein (DMAHP) may be downregulated. Finally, other studies have suggested a novel type of pathogenic mechanism in which the *DMPK* mRNA with long CUG repeats, and not the protein, results in the pathology. Novel RNA-binding proteins that specifically bind to CUG repeats may be depleted by excessive CUG repeats, or abnormalities in these proteins may disrupt the metabolism of *DMPK* mRNA and other transcripts. Further studies are necessary to define precisely the mechanism by which the repeat expansions cause the DM disease symptoms.

Clinical Utility of Testing and Available Testing

The majority of clinically significant mutations can be identified by Southern blot analysis; however, PCR is essential for small CTG expansions (<100 repeats). Although the PCR test is less expensive and faster than Southern blot analysis, longer repeats often are not reliably amplified. Southern blot analysis can be performed using *Hind* III restriction digestion and the probe pMDY1,[19] which spans the repeat area (Figure 7-2). The probe pMDYI detects a *Hind* III polymorphism with normal alleles of 9.5 and 8.5 kb, the frequencies of which are approximately 0.60 and 0.40, respectively. The polymorphism is due to a 1 kb insertion telomeric to the CTG repeat, and is almost in complete linkage disequilibrium with the CTG repeat mutation in most populations.[20] The mutation is found on the larger 9.5 kb allele, suggesting that there were a limited number

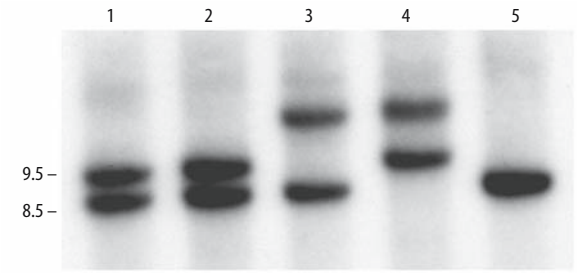


Figure 7-2. Autoradiogram of a Southern blot for myotonic dystrophy. Normal alleles are 8.5 and 9.5 kb. Lanes 1, 2, and 5: unaffected individuals; lanes 3 and 4: expanded and normal alleles in DM patients with repeat expansions.

of ancestral mutations that occurred on a chromosome having the 1 kb insertion. Alternatively, the larger allele may be predisposed to DM mutations. Typical increases in the range of 1 kb to 4 kb are observed on Southern blots in the DM population. Many of the larger expansions are detected as smears, indicating somatic cell heterogeneity of the expanded alleles. Molecular testing is extremely helpful in identifying individuals who are asymptomatic or exhibit equivocal symptoms, such as cataracts. These smaller expansions are often detected using alternate digestions, which reduce the normal size fragment, such *BamH* I, or by PCR amplification across the repeat region. De novo mutations have not been described in DM, which is consistent with the linkage disequilibrium data. To account for the maintenance of the mutation in the population, it was proposed that in DM families there is a high incidence of minimally expanded alleles, which produce few symptoms and are stably transmitted over several generations.[23] Therefore, for counseling purposes, it becomes important to identify in which side of the family the mutation is segregating.

Several cases of reverse mutations have been reported in DM, whereby there is a spontaneous correction of a deleterious expansion mutation upon transmission to an unaffected offspring. The mechanism for the DM reverse mutations remains unknown. A gene conversion mechanism, whereby the normal parental allele replaces the expanded allele, may best account for the reversion events in DM. The reversions may provide an explanation of the nonpenetrance observed in some DM families.

Interpretation of Test Results

The discovery of an expanded repeat sequence in DM has greatly improved our ability to confirm the diagnosis in symptomatic patients, to detect DM carriers who are asymptomatic or who show few of the classical signs of the disease, and for prenatal testing when a parent has been diagnosed with the disorder. Although closely linked restriction fragment length polymorphism (RFLP) markers were available for several years, linkage analysis is not as accurate as direct DNA testing for the mutation, especially when there is an uncertain diagnosis or when key family members are unavailable for testing. Also, linkage test results provide no information regarding the severity of the disease. Molecular testing has largely replaced muscle biopsy, muscle enzyme study, and electromyography as the first diagnostic procedure.

The triplet repeat size does correlate with muscular disability and is inversely related with the age of onset of the disease. However, there is a significant overlap of repeat size in patients with differing severity. When unrelated affected individuals with small to moderate differences in repeat sizes are compared, accurate prediction of the severity of the disease in each case is generally difficult. However, when a child has a significant increase in allele size compared to the parent, the child will almost certainly become symptomatic at an earlier age of onset and will have more severe disease. Lastly, as a result of the somatic heterogeneity observed in DM, genotype/phenotype associations derived from leukocytes may not be as accurate as the measurement of the repeat size in the affected tissue (muscle, heart, others). CTG expansions may be 2- to 13-fold greater in the DNA isolated from skeletal muscle than in the DNA from leukocytes.[24]

Laboratory Issues

There are many molecular pathology laboratories offering DNA testing for DM. The majority of laboratories are using a combination of PCR and Southern blot testing. The PCR allows accurate quantitation of the number of CTG repeats for normal alleles and those with small expansions. For the large CTG expansions, Southern blot analysis is required. All patients with DM have a CTG expansion in the 3′ region of *DMPK*, and no other types of mutations have been identified in the gene. However, the gene for the rarer dominantly inherited myotonic dystrophy type 2 (DM2), which usually results in a more proximal myotonic myopathy and some of the multisystem manifestations of myotonic dystrophy, has been identified.[25] Myotonic dystrophy type 2 is caused by a very large CCTG expansion in intron 1 of the zinc finger protein 9 gene (*ZNF9*). Gene testing for DM2 should be considered in patients with a DM-like phenotype who have a negative test for the *DMPK* CTG expansion. Proficiency testing for DM is available through the CAP, and proficiency specimens are sent to participants twice per year.

SPINAL MUSCULAR ATROPHY

Molecular Basis of Disease

The autosomal recessive disorder proximal spinal muscular atrophy (SMA) is a severe neuromuscular disease characterized by degeneration of alpha motor neurons in the spinal cord, which results in progressive proximal muscle

denervation and atrophy resulting in the symptoms of weakness and paralysis. Spinal muscular atrophy is the second-most common fatal autosomal recessive disorder after cystic fibrosis, with an estimated prevalence of 1 in 10,000 live births.[26] Childhood SMA is subdivided into three clinical groups on the basis of age of onset and clinical course. Type I SMA (Werdnig-Hoffmann disease) is characterized by severe, generalized muscle weakness and hypotonia at birth or within the first three months after birth. Death from respiratory failure usually occurs within the first two years of life. Children affected with Type II SMA are able to sit, although they cannot stand or walk unaided, and survive beyond four years of age. Type III SMA (Kugelberg-Welander syndrome) is a milder form, with onset during infancy or youth, and patients may walk unaided.

The survival motor neuron (*SMN1*) gene has 9 exons and is the primary SMA-causing gene.[27] Two almost identical *SMN* genes are present on 5q13: the telomeric, or *SMN1*, gene that is the SMA-determining gene, and the centromeric, or *SMN2*, gene. The *SMN1* gene exon 7 is deleted in approximately 94% of affected patients, while small, more subtle mutations have been identified in the majority of the remaining affected patients. Deletions of other genes in the SMA region most likely mark the extent of the deletion and may modify the severity of the disease. Although mutations of the *SMN1* gene are observed in the majority of patients, no phenotype-genotype correlation was observed, because *SMN1* exon 7 is absent in the majority of patients independent of the type of SMA. This is because routine diagnostic methods do not distinguish between a deletion of *SMN1* and a conversion event whereby *SMN1* is replaced by a copy of *SMN2*. Several studies have shown that the *SMN2* copy number influences the severity of the disease.[28–30] The number of *SMN2* gene copies varies from 0 to 3 copies in the normal population, with approximately 10% of unaffected individuals having no gene copies of *SMN2*. However, milder patients with type II or III SMA on average have more gene copies of *SMN2* than do type I SMA patients. It has been proposed that the extra *SMN2* gene copies in the more mildly affected patients arise through gene conversions, whereby the *SMN2* gene is copied either partially or totally into the telomeric *SMN1* locus.

Five base pair differences exist between *SMN1* and *SMN2* transcripts, and none of these differences change amino acids. Since virtually all individuals affected with SMA have at least one *SMN2* gene copy, the obvious question that arises is, Why do individuals with *SMN1* mutations have an SMA phenotype? The *SMN1* gene produces predominately a full-length transcript, whereas the *SMN2* copy produces predominately an alternate, exon-7-deleted product. The inclusion of exon 7 in *SMN1* transcripts and exclusion of this exon in *SMN2* transcripts is caused by a single nucleotide difference at +6 in *SMN* exon 7. Although the C-to-T change in *SMN2* exon 7 does not change an amino acid, it does disrupt an exonic splicing enhancer (ESE), which results in the majority of transcripts lacking exon 7. Furthermore, the importance of the exon 7 region was suggested by Talbot et al.[31] by demonstration that a highly conserved tyrosine-glycine (Y-G) dodecapeptide motif is encoded by this exon region and is crucial for the oligomerization and function of the SMN protein. Therefore, SMA arises because the *SMN2* gene cannot completely compensate for the lack of SMN1 protein function when the *SMN1* gene is mutated. However, the small amounts of full-length transcript generated by *SMN2* are able to prevent in utero lethality due to a complete lack of SMN1 protein, and produce a milder type II or III phenotype when the copy number of *SMN2* genes and transcripts is increased.

Recent evidence supports a role for SMN in small nuclear ribonuclearprotein (snRNP) biogenesis and function.[32] The SMN protein is required for pre-mRNA splicing. Immunofluorescence studies using a monoclonal antibody to the SMN protein have revealed that the SMN protein is localized to novel nuclear structures called "gems," which display similarity to and possibly interact with coiled bodies, which are thought to play a role in the processing and metabolism of small nuclear RNAs. SnRNPs and possibly other splicing components require regeneration from inactivated to activated functional forms. The function of SMN is in the reassembly and regeneration of these splicing components. Mutant SMN, such as that present in SMA patients, lacks the splicing-regeneration activity of wild-type SMN. SMA may be the result of a genetic defect in spliceosomal snRNP biogenesis in motor neurons. Consequently, the motor neurons of SMA patients are impaired in their capacity to produce specific mRNAs and as a result become deficient in proteins that are necessary for the growth and function of these cells.

Clinical Utility of Testing and Available Testing

The molecular diagnosis of SMA consists of the detection of the absence of exon 7 of the *SMN1* gene (Figure 7-3). Although this is a highly repetitive region and there is the almost identical centromeric *SMN2* copy of the *SMN1* gene, there is an exonic base pair difference that allows distinction of PCR products of *SMN1* from those of *SMN2* using restriction-site generating PCR (RG-PCR) followed by restriction enzyme digestion. The absence of detectable *SMN1* exon 7 in SMA patients is being utilized as a powerful diagnostic test for SMA, with a sensitivity of approximately 94%. Limitations of this diagnostic test are the inability to detect nondeletion mutations of *SMN1* and the inability to determine carrier status for SMA.

Since SMA is one of the most common lethal genetic disorders, with a carrier frequency of 1 in 40 to 60, carrier

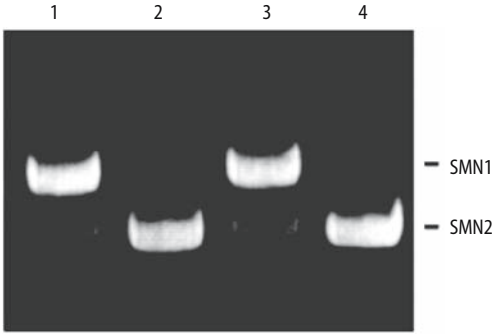


Figure 7-3. Restriction enzyme digestion of RG-PCR products distinguishes *SMN1* from *SMN2* exon 7 PCR products. Lanes 1 and 3: normal controls with *SMN1* present and *SMN2* deletion; lanes 2 and 4: SMA patients with *SMN1* deletion and *SMN2* present.

testing is useful to many families. Carrier detection for the heterozygous state is technically challenging because the SMA region is characterized by the presence of many repeated elements. The *SMN* gene is present in two almost identical copies, *SMN1* and *SMN2*. However, the *SMN2* gene copy number fluctuates: approximately 5% of normal individuals lack the *SMN2* copy, whereas many of the more mildly affected SMA patients have three or more copies of *SMN2*. Thus, a straightforward dosage assay using the *SMN2* gene as the internal control would not be reliable, and the copy number of *SMN2* affects the efficiency of amplification of the *SMN1* exon 7 region. Quantitative PCR assays, using alternative internal controls, are used for the identification of SMA carriers and determine the number of copies of *SMN1* and *SMN2* genes.[28,33]

Interpretation of Test Results

The presence of *SMN1* exon 7 does not entirely exclude a diagnosis of SMA. Although the absence of both copies of the *SMN1* gene is a very reliable and sensitive assay for the majority of SMA patients, about 6% of affected patients have other types of mutations in the *SMN1* gene that will not be detected by PCR deletion testing.[34] Most of these patients will be compound heterozygotes, with one *SMN1* allele deleted and the other allele with a point mutation or other small insertion or deletion. If the clinical suspicion remains high after a negative deletion test, then dosage carrier testing to determine whether there is a single copy of *SMN1* should be considered. A dosage testing result of two copies of the *SMN1* gene greatly reduces the likelihood of SMA.

The carrier test has two limitations. The first is the presence of de novo mutational events in the *SMN1* gene. The de novo mutation rate for this gene has been observed to be approximately 2%, which is high when compared to most autosomal recessive disorders.[35] The second limitation of the carrier test is the finding of two *SMN1* genes on a single chromosome. The allele frequency of the 2-copy *SMN1* chromosome is approximately 2% in the general population. The finding of two *SMN1* genes on a single chromosome has serious genetic counseling implications, because a carrier individual with two *SMN1* genes on one chromosome would have the same dosage result as a noncarrier with two *SMN1* gene copies on each chromosome 5. Approximately 5% of parents of a single affected SMA child have two *SMN1* gene copies by dosage analysis. Thus, although, the finding of normal dosage significantly reduces the risk of being a carrier, there is still a recurrence risk of future affected offspring for individuals with two *SMN1* gene copies. Thus, risk assessment calculations using Bayesian analysis are essential for the proper genetic counseling of SMA families (see chapter 5).

Laboratory Issues

There are many molecular pathology laboratories offering testing for the homozygous deletion of *SMN1*. The majority of laboratories utilize RG-PCR of exon 7 and PCR of exon 8 in two separate reactions, with restriction enzyme digestion to differentiate *SMN1* from *SMN2* PCR products, and report results for both *SMN1* exons 7 and 8. Some laboratories only test for exon 7 deletion. Carrier testing is being offered by a fewer number of laboratories. Proficiency testing for the homozygous *SMN1* deletion is offered by the CAP. These proficiency specimens are sent to participants twice per year. External proficiency testing is not available for dosage carrier testing.

MITOCHONDRIAL ENCEPHALOMYOPATHIES

Molecular Basis of Disease

Mitochondria are semiautonomous replicating cellular organelles with their own genetic material. Each mitochondrion contains multiple copies of the mitochondrial DNA genome (mtDNA), with replication, transcription, and translation machineries separate from the cellular machineries for these functions. Human mtDNA encodes 13 polypeptides of the respiratory chain subunits, 28 ribosomal RNAs (rRNAs), and 22 transfer RNAs (tRNAs) in a circular double-stranded genome of approximately 16.5 kb.

Unlike nuclear DNA, in which each cell contains pairs of chromosomes, one of maternal and the other of paternal origin, mtDNA is inherited exclusively from the mother. This type of transmission is called maternal inheritance. Although both sexes are equally affected by mitochondrial diseases, inheritance of the disorder is from the mother. Mitochondrial mutations are often present in only some of the mtDNA molecules of a cell (heteroplasmy). Heteroplasmy occurs because mitochondria segregate randomly into daughter cells during mitosis, which results in cells containing both mutant and wild-type mtDNA. Thus, the

proportion of heteroplasmic mutation may vary widely between different tissues or even between different cells of the same tissue. The proportion of mutant to wild-type mtDNA plays a role in determining the clinical variability and severity often observed in the mitochondrial disorders. However, the phenoytpe-genotype correlation in the mitochondrial disorders is complex and is influenced by age, the type and extent of respiratory chain disruption caused by the mutation, and the tissue-specific threshold for the pathogenic effect.

The term "mitochondrial encephalomyopathies" is used to describe mitochondrial disorders in which both muscle and the central nervous system (CNS) are affected. These disorders are multisystemic, with diverse clinical features due to defects in the mitochondrial function. This chapter is restricted to those disorders in which the mutation event involves mtDNA, in contrast to the many nuclear genetic disorders that result in mitochondrial pathology. This occurs when the nuclear encoded protein functions in the mitochondria. This chapter discusses Kearns-Sayre syndrome (KSS), mitochondrial encephalomyopathy with lactic acidosis and strokelike episodes (MELAS), and myoclonic epilepsy with ragged-red fibers (MERRF).

Kearns-Sayre Syndrome

The common features of KSS include progressive external ophthalmoplegia (PEO), pigmentary degeneration of the retina, and defects of cardiac conduction.[36] The typical affected patient presents before the age of 20 years with PEO and pitosis. This is followed by the pigmentary retinal degeneration and cardiac conduction block. Other features of the disorder may include ataxia, deafness, dementia, and diabetes mellitus. The most common type of mutation found in KSS is a deletion of mtDNA (Figure 7-4), and almost of all these deletions occur sporadically.[37] Approximately one third of KSS cases are due to a common 4977 bp deletion, which is associated with direct repeats at the deletion junction. The severity of KSS depends on the extent of heteroplasmy and the tissue distribution of structurally altered mtDNA. An extreme KSS phenotype occurs when the frequency of deleted mtDNA in muscle cells is greater than 85%. In contrast, when lower levels of heteroplasmy for the deletion are present, then PEO may be the only symptom.

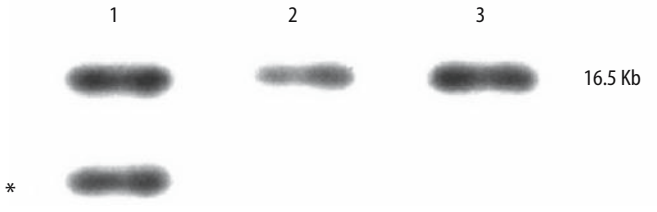


Figure 7-4. Autoradiogram of a Southern blot of muscle mitochondrial DNA. Lane 1: Kearns-Sayre patient with heteroplasmic deletion; lanes 2 and 3: unaffected individuals. The normal allele is 16.5 kb and the deleted allele fragment is indicated by the asterisk.

Mitochondrial Encephalomyopathy with Lactic Acidosis and Strokelike Episodes

MELAS patients are usually normal at birth but develop stunted growth, intermittent vomiting, seizures, and recurrent cerebral insults resembling strokes during the first years of life.[38] An episodic course follows, with recurrent strokelike episodes, only partial recovery, and eventual deterioration, with death from respiratory failure often before 20 years of age. Milder adult-onset cases have been reported. Approximately 80% of all MELAS cases are the result of an A3243G point mutation in the mtDNA gene encoding $tRNA^{Leu}$.[39] The point mutation alters the normal structural conformation of the tRNA, thereby impairing protein synthesis. The A3243G mutation occurs in the heteroplasmic state, with variation among different tissue types. When the mutation is present in greater than 90% of the mtDNA of the muscle tissue, there is increased likelihood of recurrent strokes and classic MELAS manifestations. However, when the heteroplasmy of the mutation is less than 90%, later onset and more moderate symptoms may occur.

Myoclonic Epilepsy with Ragged-Red Fibers

MERRF is a rare maternally inherited disorder in which the full expression includes muscle weakness, myoclonus, generalized seizures, ataxia, and deafness.[40] The hallmark morphologic change seen in the muscle biopsy is the ragged-red fibers. The term "ragged-red fibers" refers to large clumps of abnormal mitochondria that accumulate beneath the sarcolemma and are stained red with the Gomori trichome stain. The majority of MERRF cases are the result of a point mutation (A8344G) in the $tRNA^{Lys}$ gene. The MERRF mutation, like the MELAS $tRNA^{Leu}$ mutation, diminishes overall mitochondrial protein synthesis. Similar to the other mitochondrial disorders, a more classic MERRF phenotype is observed when the mutation is present at higher levels in the muscle and nerve.

Clinical Utility of Testing and Available Testing

The symptoms of the mitochondrial disorders are often nonspecific and are common to many other neuromuscular diseases. As a result, the mtDNA diseases are often considered only after many other diagnoses have been excluded. However, genetic testing for mitochondrial diseases is becoming increasingly available. Southern blot analysis or long-range PCR is used to detect the deletions observed in KSS. Total DNA used for this testing should be obtained from a muscle biopsy. Use of a tissue with rapidly dividing cells, such as blood, will often lead to false-negative results since the heteroplasmy of deleted mtDNA

is shifted toward wild-type mtDNA. PCR-RFLP testing is used for the common tRNA point mutations found in MERRF and MELAS. Patients with MERRF have high levels of the mutant mtDNA in the blood, and therefore blood is an appropriate specimen for testing. In contrast, patients with the MELAS mutation often have low levels of the mutation in the blood, leading to false-negative results when blood is used for testing.

Interpretation of Test Results

If clinical suspicion is strong and the blood results are negative, then mtDNA from a muscle biopsy should be tested. Since the disorders are highly clinically heterogeneous and there is considerable phenotypic overlap, the classic MERRF and MELAS point mutations and the KSS deletions frequently are ordered together and tested as a mitochondrial test panel. The A8344G MERRF mutation is associated with other phenotypes, including Leigh syndrome, myoclonus or myopathy with truncal lipomas, and proximal myopathy. The A3243G MELAS mutation can result in milder phenotypes, including sensorineural deafness with diabetes. Furthermore, these mutations sometimes are found in asymptomatic relatives of the index case.

Laboratory Issues

If the common mtDNA mutations are not detected, more extensive testing for rarer mutations may be useful for diagnosis. Sequencing of the mtDNA is available from several specialty laboratories, which may identify a new mutation. However, since the mtDNA is very polymorphic, new mutations must be carefully verified.

KENNEDY DISEASE

Molecular Basis of Disease

Spinal and bulbar muscular atrophy (SBMA), or Kennedy disease, is a rare X-linked, slowly progressive, adult-onset motor neuropathy.[41] The age of onset is usually between 30 and 50 years and is characterized by muscle cramps, proximal and bulbar weakness, and fasciculation. Endocrine abnormalities, including gynecomastia and testicular atrophy, are common. The disease is caused by a CAG trinucleotide repeat expansion in the coding region of the androgen receptor gene (*AR*).[42,43] The CAG repeat found within the first exon is polymorphic in normal populations, and ranges in length from 10 to 36 repeats. Patients with SBMA have a CAG repeat expansion that does not overlap with the normal population and ranges from 40 to 62 repeats. Similarly to other trinucleotide repeat disorders, the CAG repeat length correlates with disease severity and age of onset. However, considerable variability in age of onset is seen among family members with similar CAG repeat lengths, suggesting that factors other than the size of the repeat modulate the onset and severity of the disease. While *AR* repeats in the unaffected range are stably transmitted, expanded repeats in SBMA patients are transmitted less stably and tend to increase in size by paternal transmission.

The pathogenic mechanism of SBMA expansion involves a gain of a toxic function of the protein product. The mutant allele is both transcribed and translated, arguing against a loss of function mechanism. Individuals with mutations producing a loss of AR protein function have testicular feminization and do not have the motor neuropathy seen in SBMA patients. The exact mechanism by which the expanded polyglutamine tract in the AR protein produces the neuropathy of SBMA is uncertain.

Clinical Utility of Testing and Available Testing

Despite distinct clinical features, SBMA is often misdiagnosed due to the heterogeneity of manifestations and lack of full expression in some family members. Many of the cases are sporadic, and some patients have only mild signs of motor neuron disease and mild gynecomastia. Amytrophic lateral sclerosis, type III SMA, hereditary motor and sensory neuropathy, limb-girdle muscular dystrophy, and facioscapulohumeral muscular dystrophy are included in the differential diagnosis for some patients affected with SBMA.[43] The PCR amplification of the repeat sequence within the first exon of the *AR* gene provides an accurate confirmation of the diagnosis of SBMA, distinct from other neuromuscular disorders. In addition, carrier females and young asymptomatic males may be identified by molecular testing for a repeat expansion with implications for genetic counseling and potential early treatment.

Interpretation of Test Results

The definitive diagnosis of SBMA is by the analysis of the CAG repeat in the *AR* gene. All patients with SBMA have a CAG repeat expansion, and no other related mutations in the *AR* gene have been identified. Some studies have described a correlation between repeat size and age of onset, but interfamilial and intrafamilial variation is large for any given repeat length.

Laboratory Issues

Several molecular pathology laboratories offer DNA testing for SBMA. The PCR test is accurate and inexpensive to perform, and requires only a blood sample. PCR with electrophoretic analysis of the PCR products allows for accurate assessment of the repeat size. External proficiency testing is not available for SBMA.

References

1. Emery AEH. Muscle histology and creatine kinase levels in the fetus in Duchenne muscular dystrophy. *Nature.* 1977;266:472–473.
2. Moser H. Duchenne muscular dystrophy: pathogenetic aspects and genetic prevention. *Hum Genet.* 1984;66:17–34.
3. Hoffman EP, Brown RH, Kunkel LM. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell.* 1987;51:919–928.
4. Koenig M, Hoffman EP, Bertelson CJ, et al. Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell.* 1987;50:509–517.
5. Ahn AH, Kunkel LM. The structural and functional diversity of dystrophin. *Nat Genet.* 1993;3:283–287.
6. Hoffman EP, Fischbeck K, Brown RH, et al. Dystrophin characterization in muscle biopsies from Duchenne and Becker muscular dystrophy patients. *N Eng J Med.* 1988;318:1363–1368.
7. Darras BT, Blattner P, Harper JF, et al. Intragenic deletions in 21 Duchenne muscular dystrophy (DMD)/Becker muscular dystrophy (BMD) families studied with the dystrophin cDNA: location of breakpoints on HindIII and BglII exon-containing fragment maps, meiotic and mitotic origin of mutations. *Am J Hum Genet.* 1988;43: 620–629.
8. Forrest SM, Cross GS, Flint T, et al. Further studies of gene deletions that cause Duchenne and Becker muscular dystrophies. *Genomics.* 1988;2:109–117.
9. Monaco AP, Bertelson CJ, Liechti-Gallati S, et al. An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics.* 1988;2:90–95.
10. Roberts RG, Gardner RJ, Bobrow M. Searching for the 1 in 2,400,000: a review of dystrophin gene point mutations. *Hum Mutat.* 1994;4: 1–11.
11. Prior TW, Papp AC, Snyder PJ, et al. A missense mutation in the dystrophin gene in a Duchenne muscular dystrophy patient. *Nat Genet.* 1993;4:357–361.
12. Prior TW, Bartolo C, Pearl DK, et al. Spectrum of small mutations in the dystrophin coding region. *Am J Hum Genet.* 1995;57:22–33.
13. Prior TW, Friedman KJ, Highsmith WE, et al. Molecular probe protocol for determining carrier status in Duchenne and Becker muscular dystrophies. *Clin Chem.* 1990;36:441–446.
14. Chamberlain JS, Gibbs RA, Ranier JE, et al. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res.* 1988;16:11141–11156.
15. Clemens PR, Fenwick RG, Chamberlain JS, et al. Carrier detection and prenatal diagnosis in Duchenne and Becker muscular dystrophy families, using dinucleotide repeat polymorphisms. *Am J Hum Genet.* 1991;49:951–958.
16. Bakker E, Van Broeckhoven C, Bonten EJ, et al. Germline mosaicism and Duchenne muscular dystrophy mutations. *Nature.* 1987;329: 554–557.
17. Harper PS. Congenital myotonic dystrophy in Britain, II: genetic basis. *Arch Dis Child.* 1975;50:514–521.
18. Brook HG, McCurrach ME, Harley HG, et al. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family. *Cell.* 1992;68:799–808.
19. Fu YH, Pizzuti A, Fenwick R, et al. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science.* 1992;255:1256–1258.
20. Mahadevan M, Tsilfidis C, Sabourin L, et al. Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science.* 1992;255:1253–1255.
21. Fu YH, Friedman DL, Richards S, et al. Decreased expression of mytonin-protein kinase messenger RNA and protein in adult form of mytonic dystrophy. *Science.* 1993;260:235–238.
22. Sabouri LA, Mahadevan MS, Narang M, et al. Effect of the mytonic dystrophy mutation on mRNA levels of the DM gene. *Nat Genet.* 1993;4:233–238.
23. Barcelo JM, Mahadevan MS, Tsilfidis C, et al. Intergenerational stability of the myotonic dystrophy protomutation. *Hum Mol Genet.* 1993;2:705–709.
24. Anvret M, Ahlberg G, Grandell U, et al. Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum Mol Genet.* 1993;2:1397–1400.
25. Liquori C, Ricker K, Mosley M, et al. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science.* 2001;293;864–867.
26. Pearn J. Incidence, prevalence, and gene frequency studies of chronic childhood spinal muscular atrophy. *J Med Genet.* 1978;15:409–413.
27. Lefebvre S, Burglen L, Reboullet S, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell.* 1995;80:155–165.
28. McAndrew PE, Parsons DW, Simard LR, et al. Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMNT and SMNC gene copy number. *Am J Hum Genet.* 1997;60: 1411–1422.
29. Mailman MD, Heinz JW, Papp AC, et al. Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. *Genet Med.* 2002;4:20–26.
30. Velasco E, Valero C, Valero A, Moreno F, Hernandez-Chico C. Molecular analysis of the SMN and NAIP genes in Spanish spinal muscular atrophy (SMA) families and correlation between number of copies of cBCD 541 and SMA phenotype. *Hum Mol Genet.* 1996;5: 257–263.
31. Talbot K, Ponting CP, Theodosiou AM, et al. Missense mutation clustering in the survival motor neuron gene: a role for a conserved tyrosine and glycine rich region of the protein RNA metabolism? *Hum Mol Genet.* 1997;6:497–501.
32. Liu Q, Dreyruss G. A novel nuclear structure containing the survival of motor neurons protein. *EMBO J.* 1996;15:3355–3365.
33. Chen K-L, Wang YL, Rennert H, et al. Duplications and de novo deletions of the SMNt gene demonstrated by fluorescence-based carrier testing for spinal muscular atrophy. *Am J Med Genet.* 1999;85:463–469.
34. Parsons DW, McAndrew PE, Iannaccone ST, Mendell JR, Burghes AHM, Prior TW. Intragenic telSMN mutations: frequency, distribution, evidence of a founder effect, and modification of SMA phenotype by cenSMN copy number. *Am J Hum Genet.* 1998;63: 1712–1723.
35. Wirth B, Rudnik-Schoneborn S, Hahnen E, et al. Prenatal prediction in families with autosomal recessive proximal spinal muscular atrophy (5q11.2–q13.3): molecular genetics and clinical experience in 109 cases. *Prenat Diagn.* 1995;15:407–417.
36. Berenberg R, Pellock J, DiMauro S, et al. Lumping or splitting? "Ophthalmopelgia-plus" or Kearns-Sayre syndrome? *Ann Neurol.* 1977;1:37–42.
37. Moraes CT, DiMauro S, Zeviani M, et al. Mitochondrial DNA deletions in progressive external ophthalmoplegia and Kearns-Sayre syndrome. *N Engl J Med.* 1989;320:1293–1299.
38. Hirano M, Ricci E, Koenigsberger MR, et al. MELAS: an original case and clinical criteria for diagnosis. *Neuromuscul Disord.* 1992;2:125–135.
39. Goto Y, Nonaka I, Horai S. A mutation in the tRNALeu (UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature.* 1990;348:651–653.
40. Silvestri G, Ciafaloni E, Santorelli FM, et al. Clinical features associated with the A → G transition at nucleotide 8344 of mtDNA ("MERRF mutation"). *Neurology.* 1993;43:1200–1206.
41. Kennedy WR, Alter M, Sung JH. Progressive proximal spinal and bulbar muscular atrophy of late onset: a sex-linked recessive trait. *Neurology.* 1968;18:671–680.
42. La Spada AR, Wilson AR, Lubahn DB. Androgen receptor mutations in X-linked spinal muscular atrophy. *Nature.* 1991;353:77–79.
43. Amato AA, Prior TW, Barohn R, Snyder P, Papp A, Mendell JR. Kennedy's disease: a clinicopathologic correlation with mutations in the androgen receptor gene. *Neurology.* 1993;43:791–794.

Chapter 8

Molecular Genetic Testing for Metabolic Disorders

Lisa Edelmann, Yaping Yang, and Ruth Kornreich

Introduction

Inborn errors of metabolism represent a highly diverse group of genetic disorders. Individually the disorders are rare. The most prevalent, phenylketonuria (PKU), affects approximately 1 in 10,000 individuals. However, because numerous metabolic disorders exist, collectively they are estimated to affect as many as 1 in 600 individuals. The clinical consequences of such disorders are broad and can be severe, with progressive neurological impairment, mental retardation (MR), organomegaly, and high morbidity. Their mode of inheritance is usually autosomal recessive but also can be X-linked. Metabolic disorders result from defects in the individual enzymes of pathways that govern many different aspects of metabolism in distinct compartments within the cell.

The onset of disease is most often after birth with the appearance of an apparently normal infant, but in some classes of metabolic disorders multiple congenital anomalies also exist. For most metabolic disorders, disease symptoms present in early infancy or childhood, but in less-severe cases, adolescent or adult onset may occur. Therefore, early recognition with prompt therapeutic intervention when possible is critical for reducing damage due to the metabolic defect. For those diseases that are prevalent and for which early detection and intervention would have a beneficial outcome, neonatal screening is performed in the United States and in several countries around the world. In the United States, each state and the District of Columbia determine the diseases for which newborns are screened and the methods used for screening. With respect to metabolic disorders, all states screen for PKU and congenital hypothyroidism, and all but one screen for galactosemia. A number of states screen for maple syrup urine disease, homocystinuria, biotinidase deficiency, tyrosinemia, and congenital adrenal hyperplasia, or some combination of these. Tandem mass spectrometry has been added to newborn screening programs in many states and can detect more than 20 metabolic disorders, including medium chain acyl CoA dehydrogenase (MCAD) deficiency. DNA testing is currently used as a follow-up to an initial screen for certain disorders, such as MCAD deficiency and PKU.

This chapter discusses the molecular mechanisms of disease and the available genetic testing for selected metabolic disorders. The choice of disorders reflects population prevalence and current availability of molecular testing, as the mutations in many of the metabolic diseases are genetically heterogeneous and diagnoses are still widely dependent on biochemical testing. However, DNA testing is often critical for confirmatory studies, genetic counseling, carrier and prenatal testing, genotype-phenotype correlation, and is widely used for carrier screening for metabolic disorders in certain populations that have a high frequency of specific mutations due to founder effects. As molecular technologies advance, molecular methods will increasingly be used to screen for more metabolic diseases.

AMINO ACIDURIAS (PHENYLKETONURIA)

Molecular Basis of the Disease

PKU is an autosomal recessive disorder caused by the inability of the body to convert phenylalanine to tyrosine. PKU is the most common metabolic disease in caucasians, with an incidence of 1 in 10,000 individuals. About 98% of PKU cases are caused by defects in the phenylalanine hydroxylase (*PAH*) gene. The other 2% are caused by defects in the biosynthesis or regeneration of the cofactor of PAH, 6(R)-L-erythro-tetrahydrobiopterin (BH4). Accumulation of phenylalanine can damage the development of the central nervous system and result in MR. PKU has a spectrum of phenotypes ranging from classic PKU, which is the most severe type with the least tolerance to dietary phenylalanine, to moderate PKU, mild PKU, and mild hyperphenylalaninemia (MHP). Patients with MHP have no clinical symptoms and do not require dietary treatment.

PKU is included in newborn screening programs in all 50 states and is a classic example of a genetic disease that

meets the criteria for newborn screening: relatively high occurrence, availability of fast and economical screening methods, and therapeutic options. With early diagnosis and intervention, including a low-phenylalanine diet, the major disease phenotypes of mental and growth retardation can be prevented.

The *PAH* gene is located on 12q23.2 and spans a genomic region of 90 kilobases (kb). The coding region is about 4 kb and is comprised of 13 exons. More than 400 mutations in *PAH* have been reported to date, most of which are private mutations (http://www.pahdb.mcgill.ca/). The most prevalent European mutations, accounting for approximately two thirds of all mutations, are R408W (31%), IVS12 +1G→A (11%), IVS10–11G→A (6%), I65T (5%), Y414C (5%), R261Q (4%), and F39L (2%).[1]

Clinical Utility of Testing

Molecular diagnosis of PKU serves several purposes, including prognosis, confirmation of clinical and newborn screening results, carrier testing, prenatal diagnosis, and information for genetic counseling. The genotype-phenotype correlations can be used to direct the degree of restriction of phenylalanine in the diet (Table 8-1). Moreover, for patients with mild mutations in the BH4 cofactor-binding region (V190A, R241C, A300S, A313T, E390G, A403V, and P407S), overloading with BH4 can increase PAH activity and may be used as an alternative to dietary restriction.[2] Prenatal diagnosis allows for the termination of an affected fetus or can ensure immediate therapeutic intervention after birth. Proper genetic counseling assists parents in making informed decisions.

Available Assays

Several methods are currently used for the molecular detection of mutations in *PAH* associated with PKU. These methods include:

- Testing for a panel of common mutations with a detection rate of approximately 50%, depending on the number of mutations included.
- Mutation scanning of all 13 exons and the intron-exon junction regions. DNA sequencing detects approximately 94% of mutations; however, this method can be expensive.[3] A recently developed system for mutation scanning, denaturing high-performance liquid chromatography (DHPLC), which also has a high detection rate (~96%) and is more cost-effective than DNA sequencing,[4] may be the method of choice for PKU molecular testing.
- Finally, when molecular analysis fails to detect one or both mutant alleles, linkage studies can be performed and are highly accurate if polymorphic markers within or very closely linked to the *PAH* gene are used.

Table 8-1. Genotype-Phenotype Correlations for the Most Common *PAH* Mutations

Mutation	Prevalence	PAH Activity in COS Cells	Phenotype
R408W	31%	<1%	Classic PKU
IVS12+nt1G→A	11%	<1%	Classic PKU
IVS10–11G→A	6%	Not available	Classic PKU
I65T	5%	26%	Classic PKU Variant PKU Non-PKU PAH
Y414C	5%	50%	Variant PKU Non-PKU PAH
R261Q	4%	<30%	Classic PKU Variant PKU

Source: "PAH Activity in COS Cells" and "Phenotype" from the *PAH*db Phenylalanine Hydroxylase Locus Knowledgebase [database online]. Available at: http://www.pahdb.mcgill.ca/.

Interpretation of Test Results

The heterogeneity of the clinical phenotypes results mainly from the great variety of mutations in the *PAH* gene. Null alleles eliminate almost all the enzyme's activity and cause classic PKU, while mutations with residual PAH activity result in milder forms. Like many single-gene disorders, genotype-phenotype correlations exist in most but not all cases. Environmental factors and/or modifier genes can also play a role in the clinical manifestations of the disease. The correlations of the most common mutations and their biochemical and clinical phenotypes are summarized in Table 8-1.

UREA CYCLE DISORDERS (ORNITHINE TRANSCARBAMYLASE DEFICIENCY)

Molecular Basis of the Disease

Defects in the urea cycle constitute a rare group of disorders resulting in the accumulation of urea precursors, mainly ammonium and glutamine. Ornithine transcarbamylase (OTC) deficiency, the most common inborn error of ureagenesis, is an X-linked disorder. Affected hemizygous males typically present in the neonatal period or later in childhood, with symptoms that include vomiting, lethargy, hypothermia, apnea due to hyperammonemia, and leading to coma or death. Recurrent episodes of metabolic crisis can result in MR. The only available treatment after an acute metabolic episode is liver transplantation, which should be performed as early as possible to prevent brain damage. In 15% to 20% of carrier females, symptoms of disease are evident. Symptomatic carrier females typically have later onset but disease also may be fatal, presumably due to an unfavorable pattern of X-inactivation in the liver.[5]

OTC is a homotrimeric mitochondrial matrix enzyme that catalyzes the synthesis of citrulline from ornithine and carbamyl phosphate, and is found almost exclusively in the liver and intestinal mucosa. Loss of OTC activity results in high plasma glutamine and ammonium, low plasma citrulline and an excess of orotic acid in the urine, a combined metabolic profile that is diagnostic for OTC deficiency. However, a direct assay of OTC activity performed on tissue isolated from a liver biopsy specimen is necessary to obtain unequivocal biochemical results.

The *OTC* gene is located on Xp21 and spans a region of 73 kb that contains ten exons and encodes a protein of 354 amino acids. The overall prevalence of the disease is estimated at 1 in 50,000 in the United States, with similar statistics reported in Japan. Mutations have been identified in all ten exons; however, disease-causing mutations are less frequent in exons 1 and 7, the least conserved exons, most likely reflecting their lesser relevance to the function of the enzyme.[6]

Clinical Utility of Testing

Diagnosis of OTC deficiency by molecular testing is preferable to the more invasive liver biopsy that is necessary for the enzymatic test. Molecular screening for *OTC* mutations identifies approximately 80% of mutations, while the remaining undetected mutations are expected to affect promoter function or splicing.[7] The *OTC* gene has an approximately 50:1 male-to-female mutation ratio, and 80% of male probands inherit the mutation from their mothers, while only 23% of manifesting females inherit the mutation. Therefore, any woman who has a son with OTC deficiency has a 20% a priori risk of having another affected son in her next pregnancy. Due to the inheritability and the severity of the disease with the limited treatment available, molecular screening of at-risk couples for the purpose of prenatal testing may be beneficial. Additionally, mutation identification may be of prognostic value in OTC deficiency (see "Interpretation of Test Results," below).

Available Assays

Only a few laboratories in the United States offer clinical molecular genetic testing for OTC deficiency. One method for mutation screening of the *OTC* gene is single-strand conformation polymorphism (SSCP or DHPLC) analysis of polymerase chain reaction (PCR)-amplified exons with direct sequencing of any SSCP-positive exons.[8] Approximately 230 mutations have been reported (http://www.cnmcresearch.org/OTC/). Most mutations (86%) in the *OTC* gene are point mutations, with G→A transitions accounting for 34% and C→T transitions accounting for 21% of the total. Approximately one third of all point mutations are at CpG dinucleotides, and 15% are at splice junctions. Although the CpG sites are recurrent mutation sites, none accounts for more than 4% of the total single-base substitutions.

Interpretation of Test Results

In general, the genotypic spectrum correlates with the severity of the phenotype, and mutations that result in complete loss of function or amino acid changes near the active site of the protein result in neonatal onset of disease. In contrast, amino acid changes that are not close to the active site and result in protein with residual enzymatic activity are associated with later onset and a milder disease course.[8]

LYSOSOMAL STORAGE DISORDERS

Lysosomal storage disorders are a group of diverse inherited metabolic diseases that result from the disruption of the lysosomal system and catabolism of macromolecules (for review, see Reference 9). Mutations in genes encoding hydrolyzing enzymes, activator proteins, lysosomal membrane proteins, or proteins involved in the posttranslational modification or transport of lysosomal proteins can cause such storage disorders. More than 40 lysosomal storage disorders are known, and they have a collective incidence of approximately 1 in 5,000 to 8,000 live births in the United States. Most of the genes responsible for lysosomal storage disorders have been cloned, permitting gene mutation testing once a diagnosis is established by biochemical analyses. This information is valuable for genotype-phenotype correlation, selection of therapy, and genetic counseling. In this section, two lysosomal storage disorders are discussed: Tay-Sachs disease, which serves as a model for population screening, and Gaucher disease, for which much effort has been concentrated on genotype-phenotype correlations.

TAY-SACHS DISEASE

Molecular Basis of the Disease

Tay-Sachs disease (TSD) is a neurodegenerative disorder resulting from deficiency of the lysosomal enzyme hexosaminidase A (HEX A), resulting in accumulation of the cell membrane glycolipid G_{M2} ganglioside within lysosomes (for review, see References 10 and 11). The clinical course of TSD is characterized by normal development for the first few months of life followed by progressive loss of motor skills, macrocephaly, seizures, blindness, and death usually before 4 years of age. Infantile TSD is always fatal, and there is no effective treatment. There are also later-onset forms with slower disease progression. TSD is an autosomal recessive disease and has a carrier frequency of

approximately 1 in 30 Ashkenazi Jewish individuals and 1 in 250 to 300 in most other populations. Genetic isolates such as the French Canadians of Quebec, Cajuns from Louisiana, and the Amish in Pennsylvania also have carrier frequencies similar to that seen in Ashkenazi Jews. The first carrier screening programs began in 1970 and used the measurement of HEX A activity in serum, leukocytes, or tears. When the *HEXA* gene encoding HEX A was cloned in 1987, disease-associated mutations were identified. Current testing for TSD utilizes both biochemical and molecular testing by various methods.

Clinical Utility of Testing

Carrier detection for TSD, which began in the 1970s and was later endorsed by the American College of Obstetricians and Gynecologists (ACOG) and the American College of Medical Genetics (ACMG), has been a paradigm for population screening. As a result, there has been a 90% reduction in the incidence of TSD in the North American Ashkenazi Jewish population, such that the incidence of TSD is now 3- to 4-fold higher in non-Jews by comparison.

Available Assays

Clinical laboratories use several strategies to incorporate mutation studies into their screening programs. Some laboratories initially screen by enzyme analysis and follow with mutation studies for individuals with a result in the carrier or inconclusive ranges, while other laboratories use DNA studies alone for selected populations. DNA studies are performed used a variety of methods, including PCR amplification followed by allele-specific oligonucleotide (ASO) hybridization or restriction enzyme digestion, allele-specific amplification, TaqMan probe technologies, or ligation chain reaction amplification.

In Ashkenazi Jewish individuals, there are two common mutations in *HEXA* associated with infantile TSD and one associated with an adult-onset form of the disease. A 4 base pair insertion (TATC) in exon 11 accounts for approximately 80% of mutant alleles in this population, and a splice defect in intron 12 (IVS12+1G→C) accounts for another 15%. A missense mutation, G269S, leads to an adult-onset form of TSD and accounts for approximately 2% of carriers.

Interpretation of Test Results

A pseudodeficiency allele, R247W, is present in approximately 2% of Ashkenazi Jewish individuals who are carriers by the enzymatic assay. The R247W variant decreases the activity of HEX A for the artificial substrate used in the laboratory but does not cause TSD since it does not affect HEX A activity for its natural substrate, G_{M2} ganglioside. About 36% of non-Jewish individuals who are carriers by enzyme analysis have a pseudodeficiency allele (32% R247W and 4% R249W). In addition, screening for the three common Ashkenazi Jewish mutations and an additional mutation (IVS9+1G→ A) will identify approximately 95% of Ashkenazi Jewish carriers, but only 40% to 50% of disease-causing alleles in non-Jews. Other populations who are at high risk, such as the French Canadians, Cajuns, and Pennsylvania Dutch, have their own common alleles. Therefore, the mutations included for population screening must target the specific mutations of the ethnic background of the individual being tested.

GAUCHER DISEASE

Molecular Basis of the Disease

Gaucher disease (GD) is another prevalent autosomal recessive lysosomal storage disorder that is found with higher incidence in the Ashkenazi Jewish population. The carrier frequency is 1 in 18 in this population and 1 in 100 in other populations[12] (for comprehensive review on GD, see Reference 13). A defect in the enzyme glucocerebrosidase leads to the accumulation of glucocerebrosides in lysosomal compartments in macrophage/monocyte-derived cells, particularly in the liver, bone marrow, spleen, and lung. Several forms of GD exist. Type 1 GD has a wide range of clinical presentations, with some patients being asymptomatic, but can include bone disease, hepatosplenomegaly, anemia, and thrombocytopenia, but without primary central nervous system involvement. Types 2 and 3 have primary central nervous system involvement that varies by age of onset and rate of disease progression. Type 2 GD patients usually have an earlier age of onset than type 3 patients, with acute disease progression and death by approximately 2 years of age. Type 3 patients have onset in early childhood to adolescence and survive into their first to fourth decade of life. A perinatal-lethal form of GD also can occur, as well as a cardiovascular form characterized by aortic and mitral valve calcification, ophthalmologic abnormalities, and hydrocephalus.

The glucocerebrosidase gene (*GBA*) and its transcribed pseudogene (*ΨGBA*) are located on chromosome 1q21. Almost 200 mutations causing GD have been identified. Many of the mutations are most likely due to gene conversion events with the pseudogene.

Clinical Utility of Testing

The demonstration of deficient glucocerebrosidase activity in leukocytes establishes a diagnosis of GD but is unreliable for carrier detection. Therefore, molecular genetic testing is useful for carrier identification, prenatal testing, and genetic counseling.

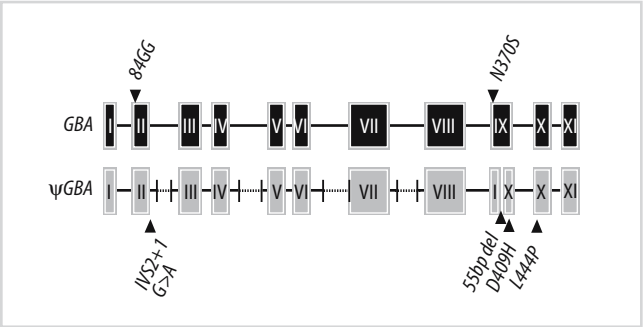


Figure 8-1. Exon/intron structure of the *GBA* gene and the pseudogene copy *ΨGBA*. Exons are indicated by boxes with Roman numeral designations, and introns are indicated by lines. Dotted lines are deletions within the pseudogene relative to the gene. Positions of mutations are indicated by arrow heads. Neither 84GG nor N370S is present in the pseudogene, whereas IVS2+1G→A, 55 bp del, D409H, and L444P are pseudogene specific.

Available Assays

Four mutations (N370S, IVS2+1 G→A, 84GG, L444P) are responsible for approximately 95% of disease-causing alleles in Ashkenazi Jewish individuals and 50% of disease-causing alleles in non-Jewish individuals (Figure 8-1).[12] Most laboratories performing GD mutation studies test for at least these four mutations using standard laboratory techniques. There are several factors to be aware of when designing molecular testing for GD. Primers must be selected that avoid amplification of the pseudogene located 16 kb downstream and approximately 96% identical to the functional gene (Figure 8-1). Recombinant alleles, which are thought to have resulted from unequal crossovers between exons 9 and 10 of the functional gene and pseudogene, contain two or more point mutations, including L444P. If L444P alone is tested, misdesignation of the genotype may occur. This may be important, as the recombinant allele is typically associated with a more severe genotype. Mistyping is also possible when a 55 base pair (bp) deletion in exon 9 is present in combination with the common N370S allele. Homozygosity of N370S would be observed even though the true genotype is N370S/55 bp deletion. Therefore, the 55 bp deletion should be analyzed in patients who are found to be homozygous for N370S.

Interpretation of Test Results

Genotype-phenotype correlations have been widely investigated in GD.[14] While overlaps occur, some generalizations can be made. The presence of an N370S allele is predictive of type 1 disease. Individuals with L444P in the presence of a null allele will usually have type 2 GD, while homozygosity for L444P typically results in type 3 GD. Homozygosity for the D409H allele has been associated with the rarer cardiovascular form of GD. The phenotypic variation observed in GD is due to factors other than genotype alone, presumably involving the interplay of other genes, including modifier genes and environmental factors.

DISORDERS OF CARBOHYDRATE METABOLISM

Galactosemia and the glycogen storage diseases are discussed in this section, which addresses disorders of carbohydrate metabolism.

GALACTOSEMIA

Molecular Basis of the Disease

Galactosemia is an autosomal recessive disorder caused by deficient or absent activity of one of three enzymes involved in the metabolic pathway to convert galactose to glucose: galactokinase (GALK), galactose-1-phosphate uridyl transferase (GALT), and UDP-galactose 4′-epimerase (GALE). The predominant form is classic galactosemia, which is due to a severe reduction or absence of the GALT enzyme, and has an incidence of 1 in 40,000 to 60,000 in European newborns (for review, see Reference 15).

The symptoms of classic galactosemia in neonates include poor feeding, vomiting, failure to thrive, lethargy, jaundice, occasionally diarrhea, and *E. coli* sepsis. The symptoms in an affected newborn can be obviated if a lactose-free diet is initiated within the first two weeks of life. Newborn screening for galactosemia is included in most states in the United States. Newborns with a positive screen are followed up with immediate dietary treatment and confirmatory biochemical analysis.

The *GALT* gene is located at 9q13, is about 4 kb in length, and consists of 11 exons. More than 150 mutations in the *GALT* gene have been reported, most of which are private mutations.[16] Q188R is the most frequent mutation associated with classic galactosemia in many populations, and accounts for 64% of disease alleles in Europeans, 60% to 70% in Americans, and 50% to 58% in Mexican Hispanics.[16,17] Ethnic-specific mutations include K285N, S135L, IVS2–2A→G, and a 5 kb deletion in Caucasian, African American, Hispanic, and Jewish patients, respectively. Due to differing ethnic backgrounds in different regions of the United States, disease allele prevalence may vary in different regions of the United States. For example, Q188R, S135L, K285N, L198P, Y209C, and F171S were reported to be the most prevalent mutations in a study based on individuals from the state of Georgia, while Q188R, K285N, IVS2–2A→G, S135L, and T138M are the most common mutant alleles observed in Texas newborns.[18,19]

Clinical Utility of Testing

Molecular testing is used for confirmation of diagnosis, carrier detection, prenatal diagnosis, prognosis, and genetic counseling. Because the detection rate of molecular testing is less than 100% and biochemical testing is highly accurate, mutation detection is carried out in parallel with biochemical analysis.

Prenatal diagnosis for galactosemia can provide the opportunity for immediate dietary restriction of the newborn. Although galactosemia is considered "treatable," symptoms such as mental and growth delays, speech dyspraxia, abnormal motor function, and premature ovarian failure in women may still occur even with early intervention and lifetime dietary restrictions. These long-term phenotypes are often associated with specific mutations; for example, Q188R can cause premature ovarian failure and speech dyspraxia. Genetic counseling is very important for parents of affected fetuses before a choice regarding pregnancy outcome is made.

Available Assays

Initial molecular testing for the diagnosis of galactosemia focuses on the most prevalent mutations or the N314D mutation associated with Duarte galactosemia (Duarte-2) by quick and cost-effective methods, such as multiplex PCR followed by restriction enzyme digestion.[18] If only one mutation or no mutations are found, screening of all 11 exons and exon-intron boundaries by PCR amplification plus SSCP analysis with confirmation of positives by DNA sequencing or direct DNA sequencing alone is performed.[18] A detection rate of 96% can be achieved by a combination of testing for prevalent mutations and direct DNA sequencing.[19] The remaining 4% of undetected mutations may be due to unknown sequence changes within intronic regions that affect splicing or in the 5′ or 3′ untranslated regions of the *GALT* gene that may affect transcriptional or translational efficiency or both.

Interpretation of Test Results

Genotype-phenotype associations have been established for some mutations.[16] For example, Q188R, K285N, and L196P alleles have undetectable GALT activity and are associated with severe phenotypes, whereas S135L and T138M are less severe and are usually associated with a good prognosis. In addition to the classic form of galactosemia, which has less than 5% of the normal GALT activity, the Duarte-2 variant associated with the N314D variant has 50% of the normal GALT activity. The N314D allele is in linkage disequilibrium with four polymorphisms, IVS4–27G→C, IVS5–24G→A, IVS5+62G→A, and c.-119_-116delGTCA. The symptoms of Duarte-2 patients are mild compared to the classic type, but Duarte-2 still results in long-term phenotypes such as mental and growth delays, speech dyspraxia, abnormal motor function, and premature ovarian failure in women.

GLYCOGEN STORAGE DISEASES

Molecular Basis of the Disease

Glycogen storage diseases (GSD) are a group of heterogeneous genetic disorders characterized by the accumulation of glycogen in tissues. Eight types of GSD that vary significantly in clinical phenotypes, age of onset, and affected organs have been identified, with an overall incidence of 1 in 20,000 to 25,000 live births.[20] They are caused by defects in one of eight genes in glycogen metabolism. Glycogen storage diseases type I to type VII are inherited in an autosomal recessive pattern, and GSD IX is X-linked recessive. A summary of the eight GSD types is presented in Table 8-2. GSD I, II, III, and IV, which are the most common and severe types, are discussed.

GSD I (von Gierke disease) is characterized by hepatomegaly, kidney enlargement, growth retardation, hypoglycemia, hyperuricemia, and hyperlipemia. GSD I has two major subgroups, GSD1a and GSD1b. The subgroup GSD1a is caused by deficiency of glucose-6-phosphatase

Table 8-2. Glycogen Storage Diseases

Disorder(s)	Defective Enzymes	Gene Location	Inheritance Pattern
GSD I (von Gierke disease)	GSD1a: Glucose-6-phosphatase	17q21	AR
	GSD1b: Glucose-6-phosphate translocase	17q23	AR
GSD II (Pompe disease)	Lysosomal acid α-1,4-glucosidase	17q25	AR
GSD III (Cori disease)	Amylo-1-6-glucosidase	1p21	AR
GSD IV (Andersen disease)	Branching enzyme (α-1,4 to α-1,6)	3p14	AR
GSD V (McArdle disease)	Phosphorylase (muscle)	11q13	AR
GSD VI (Hers disease)	Phosphorylase (liver)	14q21–q22	AR
GSD VII	Phosphofructokinase	12q13	AR
GSDIX (GSDVIII)	Phosphorylase kinase	Xp22	XR

(G6Pase), which converts glucose-6-phosphate to glucose and phosphate, the last step in glycogenolysis. The *G6PC* gene encoding G6Pase is located on 17q21. The subgroup GSD1b results from deficiency in glucose-6-phosphate translocase, encoded by *G6PTL* gene located on 11q23. Common mutations vary in different ethnic groups.[20] The prevalent mutations for GSD1a in different ethnic groups are: R83C and Q347X in caucasians; R83C in Ashkenazi Jews; 459insTA and R83C in Hispanics; V166G in Muslim Arabs; R83H and G727T in Chinese; and G727T in Japanese. For GSD1b, two common mutations, G339C and 1211delCT, are present in whites, while W118R is prevalent in Japanese.

GSD II, also known as Pompe disease, is a lysosomal storage disease caused by the inability to degrade glycogen due to defects in acid α-1,4-glucosidase. The phenotypes range from the most severe infantile disorder to juvenile- and late-onset adult myopathy. Patients with the infantile form usually die from cardiomyopathy before they reach 2 years of age. Acid α-1,4-glucosidase is encoded by a gene (*GAA*) located at 17q25, and different forms of the protein are obtained by different proteolytic processing. Common mutations have been identified in different ethnic groups.

Patients affected with GSD III, also known as Cori disease, have symptoms similar to but milder than those associated with GSD I. The gene encoding amylo-1-6-glucosidase has 35 exons with 4596 bp of coding region and a long 3′ UTR of 2371 bp. Molecular testing of GSD III is difficult and impractical due to the large size of the gene and the lack of predominant mutations.

GSD IV, also known as Andersen disease, is caused by glycogen branching enzyme deficiency that results in glycogen that is abnormal and insoluble. Intracellular accumulations occur in the liver, brain, heart, skeletal muscles, and skin fibroblasts. Neonates with GSD IV appear normal at birth but develop hepatomegaly and failure to thrive in the first year of life. Patients develop progressive cirrhosis and usually die of liver failure by 2 to 5 years of age. Mutations in the branching enzyme (α-1,4 to α-1,6) have been identified in a limited number of patients.

Clinical Utility of Testing

Due to the complex nature of molecular testing for GSD (large genes with numerous mutations), enzyme assays are usually used for the diagnosis of GSD. However, for prenatal diagnosis of GSD I, gene-based mutation testing or linkage analysis is the preferred method, since the enzymes are not present in amniocytes or chorionic villi, requiring a liver biopsy to obtain tissue containing the relevant enzyme. For prenatal diagnosis of GSD II, III and IV, DNA testing can be used to complement and confirm biochemical results. Therapy for GSD I, II, III, and IV includes dietary management, and for GSD II, enzyme replacement is available (for complete review on treatment for GSD, see References 20 and 21).

Available Assays

Molecular diagnosis for GSD I by direct DNA testing and GSD II by linkage studies is available clinically.

FATTY ACID OXIDATION DISORDERS (MCAD DEFICIENCY)

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, and skeletal and cardiac muscles. The pathway for fatty acid oxidation occurs in the mitochondria and is complex, involving as many as 20 steps. A number of disorders involving different enzymes in the pathway have been identified. Although the symptoms of the disorders have phenotypic overlap, there are several biochemical measurements that can aid in the diagnosis of these disorders, including plasma carnitine levels which are usually low, plasma acylcarnitines, and urine acylglycines (for review, see Reference 22). The most common of these disorders by far is medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, which is discussed in detail below.

Molecular Basis of the Disease

MCAD is an intramitochondrial enzyme that is encoded by a nuclear gene. The normal function of MCAD is the initial dehydrogenation of acyl-CoAs with chain lengths of 4 to 12 carbons. Defective function leads to the accumulation of metabolites of the medium-chain fatty acids, mainly the dicarboxylic acids, acylglycine in urine, and acylcarnitine in plasma. At present, many states within the United States employ tandem mass spectrometry in their newborn screening program, which allows detection of the abnormal plasma acylcarnitine profile characteristic of MCAD deficiency. These metabolites are at their highest concentration in the blood in the first few days of life, making the newborn period the ideal time for detection. Accordingly, the specificity of this testing is 100%, as no false negatives have been reported. MCAD enzymatic activity also can be assayed in several different tissue types.

Because fatty acid oxidation fuels hepatic ketogenesis, the symptoms of the disorder appear after periods of prolonged fasting or intercurrent infections and include hypoketotic hypoglycemia, lethargy, seizures, coma, and, without treatment, death. Complications of the disease can include hepatomegaly, acute liver disease, and brain damage. The disease typically presents before 2 years of age but after the newborn period. However, individuals have been described who present with symptoms within the first few days of life as well as those who present as adults.

MCAD deficiency is an autosomal recessive disorder that is prevalent in individuals of northwestern European ancestry, with the highest overall frequency of 1 in 4,900 in

Table 8-3. Genotypes of 57 MCAD-Deficient Newborns Detected Using MS/MS to Screen More Than 1.1 Million Newborns (Neo Gen Screening, Pittsburgh, PA, USA)

Mutation Position and Type	Number of Patients Identified
985 A→G/985 A→G	35
985 A→G/199 T→C (exon 3)	8
985 A→G/deletion 343–348	2
985 A→G/other*	5
985 A→G/unidentified	5
799 G→A/254 G→A	1
Unidentified/unidentified	1

Source: Reprinted with permission from Chace DH, Kalas TA, Naylor EW. The application of tandem mass spectrometry to neonatal screening for inherited disorders of intermediary metabolism. *Annual Review of Genomics and Human Genetics* 3:17–45, © 2002 by Annual Reviews, www.annualreviews.org.
*Other mutations:
244 insertion T (exon 4)
362 C→T (exon 5)
489 T→G (exon 7)
IVS 5+1 G→A
IVS 8+6 G→T

northern Germany. The incidence in the United States is somewhat lower and is estimated to be 1 in 15,700. The MCAD gene, *ACADM*, spans a 44 kb region on chromosome 1p31 and contains 12 exons encoding a protein of 421 amino acids. A single founder mutation in exon 11, 985A→G, which results in the substitution of the acidic amino acid, glutamate, for the basic amino acid, lysine (K304E), represents 90% of all alleles in the northern European population. However, recent studies of the US population, attributable to the expansion of newborn screening for MCAD deficiency, indicate that this mutation accounts for 79% of the total mutant alleles in the US population (Table 8-3).[23] The discrepancy between the two results is presumably due to the greater ethnic diversity of the US population.

Clinical Utility of Testing

Molecular genetic testing for MCAD mutations usually is offered as confirmatory testing after the initial diagnosis by biochemical testing. In addition, carrier testing for MCAD deficiency cannot be performed using biochemical metabolite profiles and must be done by molecular genetic testing or by direct assay of MCAD activity in cultured fibroblasts.

MCAD deficiency is a disease that can be treated if promptly diagnosed in the early postnatal period. 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. Although prenatal diagnosis on chorionic villus sampling or cultured amniocytes using biochemical or molecular genetic testing, or both, is possible, with the inherent risks of the procedures, it may offer no advantage to postnatal testing of acylcarnitines and other metabolites characteristic of the disease. Therefore, genetic counseling and discussion of the issues related to this disease are warranted when prenatal testing is being considered.

Available Assays

Because of its high prevalence among individuals with MCAD deficiency, molecular testing for the K304E allele is performed initially by PCR amplification followed by restriction enzyme digestion or other methods that can discriminate between single nucleotide changes, such as ASO hybridization or ligation chain reaction amplification. Clinical testing for this mutation is widely available. When an affected individual is found to be heterozygous for the K304E mutation or in the rare instance when an affected individual is negative for the mutation, gene sequencing is performed on all 12 exons of the *ACADM* gene; however, relatively few laboratories offer screening of the entire *ACADM* gene. Additional mutations have been identified throughout the gene with no obvious mutation hotspot.

Interpretation of Test Results

The majority of mutations identified in *ACADM* are missense mutations located away from the active center of the enzyme, and are thought to affect the overall stability of the protein by affecting proper protein folding (Figure 8-2).[24] Most patients exhibit the classic MCAD phenotype; however, a small subset of patients has been identified that is compound heterozygous for the A304E mutation or for two other mutations, where at least one mutation is present that does not eliminate MCAD activity. These patients are much less likely to experience metabolic decompensation; however, even mildly symptomatic patients should avoid circumstances that could precipitate a metabolic crisis,

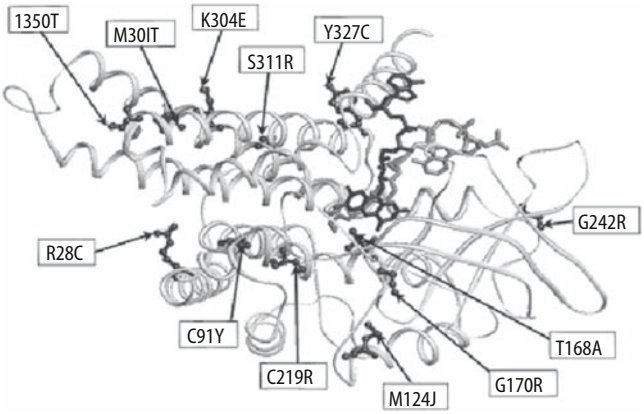


Figure 8-2. Schematic representation of a human MCAD monomer based on the crystal structure with cofactor flavin adenine dinucleotide (FAD) (black) and bound C8-CoA substrate (darker gray). The side chains for residues in which missense mutations have been published are shown in ball-and-stick representation. Only one of these mutations (T168A) is located in close proximity to the active site, forming a hydrogen bond to the flavin N(5) of FAD. (Reprinted from Gregersen N, Andresen BS, Corydon MJ, et al. Mutation analysis in mitochondrial fatty acid oxidation defects: exemplified by acyl-CoA dehydrogenase deficiencies, with special focus on genotype-phenotype relationship. *Human Mutation* 18(3):169–189, by permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc., © 2001.)

since there is variability in age of onset even with classic MCAD deficiency.

PEROXISOMAL DISORDERS (X-LINKED ADRENOLEUKODYSTROPHY)

Genetic disorders of peroxisomal biogenesis and function have severe phenotypic consequences that often result in death in early childhood. A number of important metabolic processes, including β-oxidation of long- and very-long-chain fatty acids and the degradation of H_2O_2, take place in the unique microenvironment of the single-membrane-bound matrix of the peroxisome. Whereas the symptoms of most metabolic diseases manifest after birth, disorders of peroxisome biogenesis, such as Zellweger syndrome, are associated with multiple congenital anomalies (for review, see Reference 25). In this section, X-linked adrenoleukodystrophy (X-ALD) is highlighted, as it is the most common of the peroxisomal disorders and one of the few for which clinical molecular genetic testing is available.

Molecular Basis of the Disease

X-ALD 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. The most common form has an early onset that typically appears at 4 to 8 years of age and results in a progressive irreversible dementia and often death. Less severe presentations of the disorder include adrenomyeloneuropathy (AMN), which has a later age of onset, often adrenal insufficiency, and neurological complications that are limited to the spinal cord and peripheral nerves.[26] Although the disease is inherited in an X-linked recessive manner, up to 20% of carrier females manifest late onset neurological symptoms similar to AMN. More than 93% of X-ALD patients inherit mutations from their mothers, while the remaining 7% carry de novo mutations. The primary biochemical defect is an impaired peroxisomal β-oxidation with the subsequent accumulation of very-long-chain fatty acids (VLCFAs), most notably C26, in the plasma and tissues. Therefore, X-ALD is not a disorder of peroxisomal biogenesis, but rather a specific defect of peroxisomal function. The accumulation of VLCFAs and the accompanying inflammatory response are thought to mediate the severity of the disease phenotype. In the great majority of hemizygous males (99%) and approximately 85% of carrier females, the plasma concentration of VLCFAs is elevated, a measurement that can be used as a diagnostic marker for the disease.[27]

Defects in the peroxisomal membrane protein, ALDP, a member of the ATP-binding cassette family of molecular transporters, cause the severe juvenile form of X-ALD and its milder associated forms. The X-ALD gene, *ABCD1*, is located on Xq28, spans 19 kb, contains 10 exons, and encodes a protein of 745 amino acids. The overall incidence of X-ALD and all variant forms is 1 in 15,000, making it the most common genetic determinant of peroxisomal disease. More than 400 different mutations have been found in the *ABCD1* gene, with the vast majority being point mutations, although deletions and duplications also have been identified (http://www.x-ald.nl/). In addition, mutations in all 10 exons have been reported. No genotype-phenotype correlations are apparent, and wide phenotypic variation has been reported within families.

Clinical Utility of Testing

Molecular testing is most useful for determining the carrier status of at-risk women and for prenatal diagnosis, since 15% of carrier females will not have elevated VLCFA levels and therefore will have a false-negative result by biochemical methods.

Available Assays

Molecular genetic testing of the *ABCD1* gene is available clinically from a few laboratories. Because many of the mutations identified are private mutations specific to a particular family, PCR amplification and SSCP or direct sequencing of all 10 exons has been used successfully to identify mutations in the majority of cases,[28] whereas Southern blotting can be used to assess deletion and duplication status as long as the rearrangement is small enough to be detected with gene-specific probes. Complications can arise during PCR amplification due to the presence of paralogous gene segments of *ABCD1* spanning exons 7 to 10 on chromosomes 2p11, 10p11, 16p11, and 22q11, but can be overcome by choice of primers that avoids amplification from the other chromosomes.[28] Although mutations have been identified throughout the entire *ABCD1* gene, a 2 bp AG deletion in exon 5 was found in 10.3% of families with X-ALD and is therefore the most common mutation identified in the *ABCD1* gene.[29] Interestingly, this mutation, which is associated with all X-ALD phenotypes, does not represent a founder allele and is therefore a mutation hotspot within the *ABCD1* gene.

Interpretation of Test Results

The majority of mutations in *ABCD1* are missense mutations (58.4%), with frameshifts and nonsense mutations accounting for 23.9% and 9.1% of mutation alleles, respectively, and amino acid insertions and deletions and whole exon deletions accounting for 4.6% and 4%, respectively. A recent study of the effects of missense mutations on ALDP stability demonstrated that approximately 70% resulted in absent or reduced ALDP, indicating that most mutations in *ABCD1* result in complete loss of protein function (Figure 8-3).[29] These findings are consistent with the observed lack of genotype-phenotype correlation and lend support to the existence of additional genetic and environmental factors that modify the X-ALD phenotype.

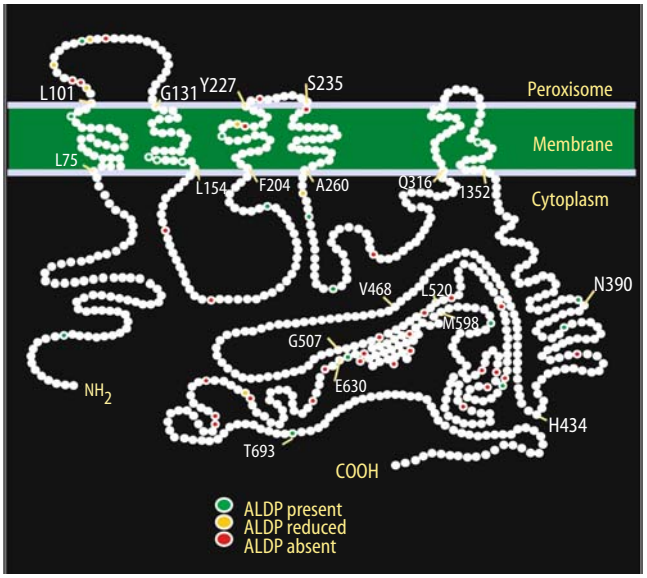


Figure 8-3. A hypothetical model of ALDP is shown. Individual amino acids are represented by circles. Missense mutations may affect the stability of ALDP. The effect of missense mutations on the stability of ALDP has been investigated for 52 independent missense mutations. Of these 52 amino acid substitutions, 31% do not affect ALDP stability and the cells have normal ALDP levels (green circles). Eleven percent of missense mutations result in reduced amounts of ALDP (orange circles). However, at 2 of these 6 positions, missense mutations have been reported that destabilize ALDP and are marked in red. Fifty eight percent of missense mutations result in no detectable ALDP (red circles). Overall, 69% of all missense mutations result in reduced or absent ALDP activity.[29] All other mutations, including in-frame amino acid deletions and truncations near the carboxy terminus, result in the absence of detectable levels of ALDP. (Reproduced from the Web site http://www.x-ald.nl/, with permission © J. Berger and S. Kemp.)

LEUKODYSTROPHIES (CANAVAN DISEASE)

The leukodystrophies are a group of degenerative metabolic diseases that involve the white matter of the brain, resulting predominately from disruption in the synthesis, transport, or catabolism of myelin. Examples include Krabbe disease and metachromatic leukodystrophy (MLD), which affect lysosomal function and degradation of myelin, and Pelizaeus-Merzbacher disease, which is due to the abnormal synthesis of proteolipid protein. Another example, discussed in detail below, is Canavan disease, characterized by loss of axonal myelin sheaths and spongiform degeneration of the brain. Neurological deterioration in most leukodystrophies occurs after a period of normal development, and therapy is usually limited to the alleviation of symptoms.

Molecular Basis of the Disease

Canavan disease (CD) is an autosomal recessive disorder found mainly in Ashkenazi Jewish families and is caused by deficiency in the activity of the enzyme aspartoacylase (for review, see Reference 30). The pathophysiologic relationship between the loss of this enzymatic activity and the development of CD remains to be elucidated. Diagnosis usually is established by the demonstration of increased levels of the substrate N-acetylaspartic acid in urine because enzymatic studies have been shown to be quite variable. Clinical symptoms associated with CD include macrocephaly, hypotonia, severe developmental delay, optic atrophy, poor head control, and death in childhood.

The gene encoding aspartoacylase (*ASPA*), located on the short arm of chromosome 17, is relatively small, with 6 exons spanning 30 kb of genomic sequence. Two point mutations, E285A and Y231X, are responsible for more than 97% of mutant alleles in Ashkenazi Jews. Mutations in non-Jewish individuals are more heterogeneous; however, a panethnic mutation, A305E, accounts for approximately 40% to 48% of non-Jewish European alleles.[31]

Clinical Utility of Testing

The genes responsible for many of the leukodystrophies have been cloned and characterized. However, in most instances the mutations in these genes are diverse. Therefore, biochemical diagnoses are still widely used, although molecular testing may be performed for carrier and subsequent prenatal testing. CD, however, occurs at increased frequency in the Ashkenazi Jewish population and screening for a limited number of mutations is feasible.

CD population screening of Ashkenazi Jewish individuals has demonstrated a carrier frequency of 1 in 40 to 59 (References 32 and 33 and our unpublished data). The carrier frequency for non-Jewish individuals has not been adequately determined, but it is far lower than that seen in the Ashkenazi Jewish population. As the carrier frequency is so high in the Ashkenazi population and the sensitivity of the assay is well above 90%, ACOG and the ACMG recommended in 1998 that carrier screening for CD be performed preconceptually on couples with Ashkenazi Jewish ancestry.

Available Assays

A number of laboratories test for the two common Ashkenazi Jewish mutations, while some also test for A305E. A few laboratories also test for the less-frequent non-Jewish mutation, 433–2A→G. Testing methodologies commonly used for the detection of these mutations include PCR followed by ASO hybridization or restriction enzyme digestion, or allele-specific amplification. Miami Children's Hospital Research Institute holds a patent on the CD gene and testing and requires laboratories to obtain a license, with a royalty fee for each test performed.

Interpretation of Test Results

If carrier screening indicates that both partners are carriers, prenatal testing should be offered. The preferred method of testing is DNA analysis of known mutations.

For couples in which one partner is shown to be a carrier and the other partner is negative by targeted molecular testing, particularly if they are not Jewish, biochemical testing by measurement of the substrate in amniotic fluid is possible when the fetus has been shown to carry the one identified parental mutation.

Laboratory Issues

As discussed in this chapter, molecular genetic testing for metabolic disorders is used primarily as a follow-up to biochemical analyses for confirmation of findings, prognosis, carrier screening, or prenatal testing. Readers interested in learning which laboratories perform individual tests should refer to the GeneClinics Web site (http://www.geneclinics.org), which includes a listing of laboratories certified by the Clinical Laboratory Improvement Amendments (CLIA). As testing for many of these disorders is performed in only a few laboratories, commercial test kits and proficiency testing is limited. Proficiency testing for TSD is available through the California Tay-Sachs Disease Prevention Program. For other disorders, interlaboratory exchange of samples is common practice. Several companies have developed assays for a panel of Ashkenazi Jewish disorders that include TSD, CD, and most likely GD among others. As mutation scanning methods become more routine, molecular testing for metabolic disorders may become more commonplace.

References

1. Nowacki PM, Byck S, Prevost L, Scriver CR. PAH mutation analysis consortium database: 1997. Prototype for relational locus-specific mutation databases. *Nucleic Acids Res.* 1998;26:220–225.
2. Erlandsen H, Stevens RC. The structural basis of phenylketonuria. *Mol Genet Metab.* 1999;68:103–125.
3. Yang Y, Drummond-Borg M, Garcia-Heras J. Molecular analysis of phenylketonuria (PKU) in newborns from Texas. *Hum Mut.* 2001;17:523.
4. Xiao W, Oefner PJ. Denaturing high-performance liquid chromatography: a review. *Hum Mut.* 2001;17:439–474.
5. Brusilow SW, Horwich AL. Urea cycle enzymes. In: Scriver R, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill; 2001:1909–1963.
6. Tuchman M, Jaleel N, Morizono H, Sheehy L, Lynch MG. Mutations and polymorphisms in the human ornithine transcarbamylase gene. *Hum Mut.* 2002;19:93–107.
7. Tuchman M, McCullough BA, Yudkoff M. The molecular basis of ornithine transcarbamylase deficiency. *Eur J Pediatr.* 2000;159(suppl 3):S196–S198.
8. McCullough BA, Yudkoff M, Batshaw ML, Wilson JM, Raper SE, Tuchman M. Genotype spectrum of ornithine transcarbamylase deficiency: correlation with the clinical and biochemical phenotype. *Am J Med Genet.* 2000;93:313–319.
9. Wenger DA, Coppola S, Liu SL. Lysosomal storage disorders: diagnostic dilemmas and prospects for therapy. *Genet Med.* 2002;4: 412–419.
10. Gravel RA, Kaback MM, Proia RL, Sandhoff K, Suzuki K, Suzuki Y. The GM2 gangliosidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill; 2001:3827–3877.
11. Sutton VR. Tay-Sachs disease screening and counseling families at risk for metabolic disease. *Obstet Gynecol Clin North Am.* 2002;29:287–296.
12. Grabowski GA. Gaucher disease: gene frequencies and genotype/phenotype correlations. *Genet Test.* 1997;1:5–12.
13. Beutler E, Grabowski GA. Gaucher disease. In: Scriver R, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill; 2001:3635–3668.
14. Koprivica V, Stone DL, Park JK, et al. Analysis and classification of 304 mutant alleles in patients with type 1 and type 3 Gaucher disease. *Am J Hum Genet.* 2000;66:1777–1786.
15. Holton JB, Walter JH, Tyfield LA. Galactosemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill; 2001:1553–1587.
16. Tyfield L, Reichardt J, Fridovich-Keil J, et al. Classical galactosemia and mutations at the galactose-1-phosphate uridyl transferase (GALT) gene. *Hum Mut.* 1999;13:417–430.
17. Elsas LJ, Langley S, Steele E, et al. Galactosemia: a strategy to identify new biochemical phenotypes and molecular genotypes. *Am J Hum Genet.* 1995;56:630–639.
18. Elsas LJ, Lai K. The molecular biology of galactosemia. *Genet Med.* 1998;1:40–48.
19. Yang Y-P, Corley N, Garcia-Heras J. Molecular analysis in newborns from Texas affected with galactosemia. *Hum Mut.* 2002;19:82–83.
20. Chen Y-T. Glycogen storage diseases. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill; 2001:1521–1551.
21. Hirschhorn R, Reuser AJJ. Glycogen storage disease type II: acid α-glucosidase (acid maltase) deficiency. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill; 2001:3389–3420.
22. Roe CR, Ding J. Mitochondrial fatty acid oxidation disorders. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease.* 8th ed. New York: McGraw-Hill; 2001:2297–2326.
23. 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.
24. Gregersen N, Andresen BS, Corydon MJ, et al. Mutation analysis in mitochondrial fatty acid oxidation defects: Exemplified by acyl-CoA dehydrogenase deficiencies, with special focus on genotype-phenotype relationship. *Hum Mut.* 2001;18:169–189.
25. Sacksteder KA, Gould SJ. The genetics of peroxisome biogenesis. *Annu Rev Genet.* 2000;34:623–652.
26. Moser HW, Smith KD, Watkins PA, Powers J, Moser AB. X-linked adrenoleukodystrophy. In: In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease.* 8th ed. New York: McGraw-Hill; 2001:3257–3302.
27. Moser HW, Loes DJ, Melhem ER, et al. X-Linked adrenoleukodystrophy: overview and prognosis as a function of age and brain magnetic resonance imaging abnormality: a study involving 372 patients. *Neuropediatrics.* 2000;31:227–239.
28. 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;66:128–136.
29. Kemp S, Pujol A, Waterham HR, et al. ABCD1 mutations and the X-linked adrenoleukodystrophy mutation database: role in diagnosis and clinical correlations. *Hum Mut.* 2001;18:499–515.
30. Matalon R, Michals-Matalon K. Recent advances in Canavan disease. *Adv Pediatr.* 1999;46:493–506.
31. Kaul R, Gao GP, Matalon R, et al. Identification and expression of eight novel mutations among non-Jewish patients with Canavan disease. *Am J Hum Genet.* 1996;59:95–102.
32. Kronn D, Oddoux C, Phillips J, Ostrer H. Prevalence of Canavan disease heterozygotes in the New York metropolitan Ashkenazi Jewish population. *Am J Hum Genet.* 1995;57:1250–1252.
33. Sugarman EA, Allitto BA. Carrier testing for seven diseases common in the Ashkenazi Jewish population: implications for counseling and testing. *Obstet Gynecol.* 2001;97:S38–S39.

Chapter 9

Fibroblast Growth Factor Receptor–Related Skeletal Disorders

Peter J. Bridge

INTRODUCTION

Four fibroblast growth receptor genes, *FGFR1* to *FGFR4*, encode receptors for the 18 (or more) fibroblast growth factors. Any of the factors can bind to any of the receptors, although there are preferences. Structurally, the receptors consist of an extracellular region, comprising three immunoglobulinlike domains, a single hydrophobic segment that spans the membrane, and a cytoplasmic tyrosine kinase domain. Binding of the fibroblast growth factor (FGF) ligand to the extracellular domain activates the intracellular tyrosine kinase domain and initiates a signaling pathway that is involved in cell division and differentiation. Mutations that cause hereditary diseases of the skeleton have been discovered in *FGFR1*, *FGFR2*, and *FGFR3* but not so far in *FGFR4*. (The only currently reported variant in *FGFR4*, G388R, is associated with tumor progression and metastasis.) Different gain-of-function mutations in *FGFR1*, *FGFR2*, and *FGFR3* lead to two major categories of disease, the craniosynostoses and the chondrodysplasias. Mutation effects are varied, including ligand-free receptor dimerization that results in constitutive receptor signaling, direct activation of the tyrosine kinase, and modulation of the receptor ligand binding affinity or specificity.[1]

CRANIOSYNOSTOSES AND CHONDRODYSPLASIAS

Craniosynostosis results from the premature fusion of the cranial sutures. The shape of the skull is dependent on which sutures are affected and at what stage of development the premature fusion occurs. Most cases of simple craniosynostosis are sporadic, but up to 10% may be hereditary. Differentiation of sporadic from hereditary craniosynostosis is very important when assessing recurrence risk. The hereditary craniosynostoses are generally characterized by bilateral coronal craniosynostosis (with the exception of Muenke syndrome) in association with specific facial and more variable hand and foot features that delineate the different syndromes. Some types result in cloverleaf skulls due to the premature fusion of all sutures. The craniosynostoses are usually named after the individuals who described them, and all show autosomal dominant transmission with very high or complete penetrance. Craniosynostosis may be a feature of a very large number (>150) of other genetic syndromes. "Splitters" note the variations between the different named syndromes, whereas "lumpers" see them all as fibroblast growth factor receptor (FGFR)-related craniosynostoses. The molecular genetic evidence currently supports both camps: all have *FGFR* mutations, but both types of heterogeneity are known, sometimes with the same syndrome caused by mutations in different genes, and sometimes the same mutation is found in patients with different clinical diagnoses.

The chondrodysplasias are characterized by abnormalities in the growth of long bones, and an autosomal dominant inheritance pattern would apply if the diseases were not lethal. The nomenclature for the chondrodysplasias is descriptive of the associated pathology. A rare FGFR-related disorder, osteoglophonic dysplasia, is associated with features of both craniosynostoses and chondrodysplasias, as well as nonossifying bone lesions that create a hollowed-out appearance of bone on x-ray.

Molecular Testing Methods and Laboratory Issues

For the disorders associated with a limited number of mutations, testing for the point mutations can be performed by a number of methods such as restriction enzyme digestion of polymerase chain reaction (PCR) products, allele-specific PCR, allele-specific hybridization by dot-blot, or in a homogeneous assay format (e.g., TaqMan or Lightcycler). Analysis of most, but not all, of the *FGFR* gene

Table 9-1. A Small Selection of *FGFR3* Mutations That Can Be Detected by Restriction Enzyme Digestion of the Appropriate PCR Products

			Restriction Sites	
Disorder	Mutation	Sequence Change	Create New	Destroy Existing
Thanatophoric dysplasia 1	R248C	CGC → TGC	*Bsp* 1286I	*Hae* II
Thanatophoric dysplasia 1	S371C	AGT → TGT	*Cvi* KI	*Tsp* RI
Achondroplasia	G380R	GGG → AGG	*Sfc* I	*Bsc* GI
Achondroplasia	G380R	GGG → CGG	*Msp* I	*Bsc* GI
Thanatophoric dysplasia 2	K650E	AAG → GAG	*Bsm* AI	*Bbs* I

The use of at least 2 different restriction enzymes is advisable to ensure that all true digestion results can be distinguished from enzyme failure. Restriction enzyme analysis of patient DNA with the K650E mutation is shown in Figure 9-1.

mutations can be accomplished by restriction enzyme digestion of PCR products followed by electrophoretic separation of the digestion products in agarose gels. Whenever possible, two different restriction enzymes are used, one for the normal sequence and one for the mutation sequence. This combination will always give fully interpretable banding patterns, if it can be achieved. If the above plus-minus system cannot be achieved, the next best option is to design the PCR such that the PCR product contains a second constant restriction enzyme recognition site so that the product always is cleaved but yields different patterns in individuals positive or negative for the mutation. Digestion with a second restriction enzyme can be used to confirm a result; for instance, the G375C mutation in *FGFR3* does not destroy a restriction enzyme site but can be detected by the generation of new sites for *Sph* I or *Fat* I. Table 9-1 gives an example of selected mutations found in *FGFR3* and options for testing by restriction enzyme digestion of the DNA (refer to Online Mendelian Inheritance in Man (OMIM),[2] GeneReviews,[3] or Human Gene Mutation Database[4] for current and complete information on mutations in the *FGFR* genes). Figure 9-1 shows the results of restriction enzyme analysis of DNA from normal controls and a patient with the K650E mutation in *FGFR3* (thanatophoric dysplasia, type 2).

The major problem with most restriction enzyme[5] assays for the FGFR-related conditions is that there are often many different mutations (sometimes also in different *FGFR* genes) that can cause the same phenotype, so that many different assays may need to be performed until a positive result is obtained. A better approach, whenever the mutations are clustered, is to sequence the relevant exon(s). This approach will detect any mutation, including mutations not previously identified, in the region sequenced. Figure 9-2 is an automated DNA sequence that shows a C278F mutation in exon 3a of the *FGFR2* gene from a patient with Crouzon syndrome. Note that this sequencing approach would have detected any of the 10 or more mutations in this exon that cause Crouzon syndrome.

A faster screening method has recently become available with the advent of denaturing high-performance liquid chromatography (DHPLC), designed for the analysis of DNA fragments. Figure 9-3 shows superimposed DHPLC traces of exon 3a of the *FGFR2* gene in a normal control and patients with Apert (P253R and S252W) and Crouzon (W290R) syndromes. The P253R, S252W, and W290R mutations have traces that are clearly different from one another as well as from the control. The speed of analysis using DHPLC is a major advantage (4 minutes per sample). Although each patient has a different profile and therefore

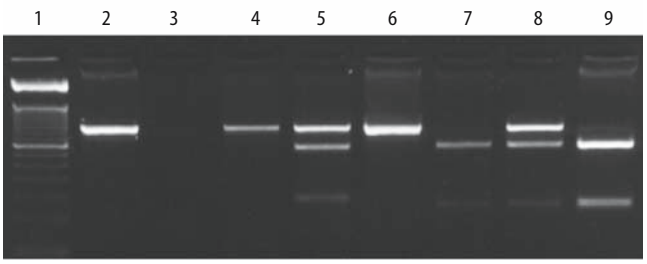


Figure 9-1. Polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP) assay for the K650E mutation in exon 15 of the *FGFR3* gene. Exon 15 of *FGFR3* from patient and control DNA was amplified by PCR, and the resulting PCR products were either not digested or digested with *Bsm* AI or *Bbs* I, and analyzed by agarose gel electrophoresis. The K650E mutation in exon 15 of *FGFR3* creates a new restriction site for *Bsm* AI and destroys a restriction site for *Bbs* I. Lane 1, size standards (50 base pair [bp] ladder); lane 2, undigested PCR product control (480 bp); lane 3, no template PCR control; lanes 4 and 6, normal DNA digested with *Bsm* AI; lane 5, patient DNA digested with *Bsm* AI, which demonstrates an additional cut site compared to the normal DNA in adjacent lanes; lanes 7 and 9, normal DNA digested with *Bbs* I; lane 8, patient DNA digested with *Bbs* I, which demonstrates loss of a cut site compared to the normal DNA in adjacent lanes. These results confirm a diagnosis of thanatophoric dysplasia, type 2.

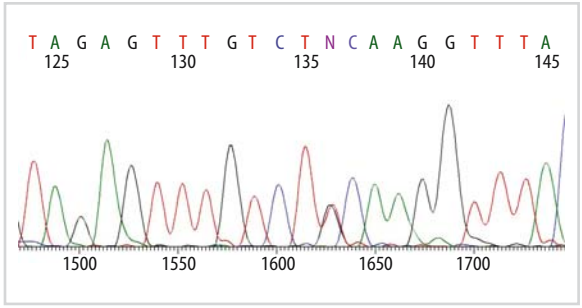


Figure 9-2. Sequence analysis of exon 3a of the *FGFR2* gene. Automated sequence electropherogram (from an ABI377 sequencing instrument; Applied Biosystems, Foster City, CA) of the forward strand of exon 3a of *FGFR2* from a patient with Crouzon syndrome. A TGC to TTC mutation is shown at position 136 in this sequence that would cause a C278F amino acid substitution. This particular mutation creates a new restriction site for *Bbs* I so could also be detected by a PCR-RFLP method for diagnosis or confirmation.

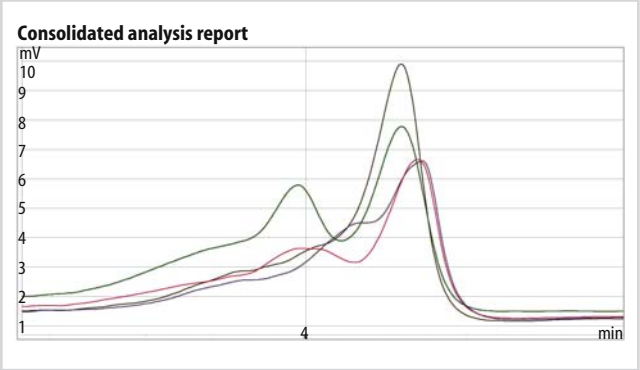


Figure 9-3. DHPLC analysis for patients with Apert or Crouzon syndromes. Superimposed DHPLC traces of exon 3a of the *FGFR2* gene in a normal control (brown trace) and patients with Apert (P253R, pink trace; S252W, green trace) and Crouzon (W290R, blue trace) syndromes. The 3 mutations, each differing from the normal control by only a single nucleotide, have different DHPLC profiles.

a different sequence compared to the normal control, the exact mutation is not identified by DHPLC. Previously run affected controls provide a good indication of which mutation corresponds to which DHPLC profile, particularly where there is a very limited number of mutations associated with the diagnosed disorder. DNA from patients with profiles that do not match any of the reference set must be sequenced to determine the specific mutation that is identified by DHPLC.

The laboratory should ensure that the relevant family members have received adequate genetic counseling and have provided informed consent prior to testing. Regardless of the methods used for detection of *FGFR* mutations, in vitro diagnostic test kits are not commercially available for any of the genes or mutations. Proficiency testing programs are not currently available for *FGFR* testing, so the laboratory must meet proficiency testing requirements by other mechanisms.

Interpretation of Test Results

Novel missense mutations always present a diagnostic dilemma: Are they the cause of the disease or are they benign variants? The task is somewhat easier for the *FGFR* genes because the genes are relatively nonpolymorphic and most changes seem to be pathogenic. Many missense mutations that might seem potentially innocuous on theoretical grounds, because the amino acid changes are relatively conservative (e.g., S252L, I288S, and A314S in *FGFR2*), are actually pathogenic. Nevertheless, interpretation is facilitated if the newly detected mutation is a de novo mutation in an isolated case of the disease or changes an evolutionarily conserved amino acid to one that could not fulfill the same function. For example, a novel 3-base insertion that adds a glycine residue at codon 272 (G272GG) in exon 3a of *FGFR2* was detected in a patient with Crouzon syndrome. The mutation could not be established as a de novo mutation because one parent was unavailable for testing. However, there was precedent for this being a disease-causing mutation in that a very similar type of mutation (also the insertion of 3 bases encoding glycine) had previously been reported at a nearby location in the gene (T268TG). At the time, interpretation was more difficult because more limited experience with testing of the *FGFR* genes had not yet established that almost all changes are pathogenic. Up-to-date listings of known mutations in each *FGFR* gene are available through the Human Gene Mutation Database.[4]

Some disorders are so severe that, although the inheritance pattern theoretically would show dominant transmission, the mutations are not normally transmitted due to their lethality, and all affected individuals have a new mutation. Although the parent who transmitted the mutant allele does not have the syndrome, since they are alive, the possibility of germline mosaicism or more rarely somatic and germline mosaicism in that parent must never be overlooked by either the genetic counselor or the diagnostic laboratory. If a mutation was identified in the affected individual, offering prenatal diagnosis for all future pregnancies is advisable. All *FGFR* test results should be interpreted in the context of the family history by an individual with expertise in clinical molecular genetics.

Comprehensive reviews of the clinical features of each syndrome are available.[6–8] In addition, many valuable online resources are available. Particularly useful resources are OMIM[2] and GeneReviews.[3] A fairly simple description of the various molecular mechanisms of dominance can be found in Strachan and Read,[9] and much more detailed descriptions of the molecular mechanisms as they apply to the skeletal dysplasias can be found in Wilkie.[1,10]

APERT SYNDROME

Molecular Basis of the Disease

All the dominant mutations that cause Apert syndrome are all clustered in exon 3a of *FGFR2*. S252W accounts for approximately 65% of Apert syndrome mutations, and P253R accounts for approximately 35%. These mutations, which are located between immunoglobulinlike domains, increase receptor-ligand binding. There are other very rare mutations in the immediate vicinity of these two mutations.

Clinical Utility of Testing

Apert syndrome is easily diagnosed by clinical criteria. Knowing the specific mutation does not alter treatment but would permit prenatal diagnosis in two scenarios: pregnancies with a risk for Apert syndrome due to germline mosaicism (i.e., there has been a previously affected child and neither parent has the mutation or condition), and pregnancies with a 50% probability of having an affected child since one parent is affected.

Available Assays

Sequence analysis or DHPLC analysis of exon 3a of *FGFR2* will detect mutations.

Interpretation of Test Results

Heterozygosity for S252W or P253R is diagnostic of Apert syndrome.

Laboratory Issues

A polymorphism has been reported that can interfere with PCR amplification of exon 3a of *FGFR2*.[11] If the test is performed using DHPLC, profiles of both of the common mutations as well as normal exon 3a should be established by the laboratory during test validation.

PFEIFFER SYNDROME

Molecular Basis of the Disease

The vast majority of Pfeiffer syndrome mutations are in *FGFR2*. A mutation in *FGFR1* (P252R in exon 5) is found in approximately 5% of Pfeiffer syndrome patients and tends to be associated with a milder phenotype.

Clinical Utility of Testing

Sometimes the parent of a child affected by one of the craniosynostoses is found to have mild symptoms on clinical examination. Testing the parents for the mutation found in the child clarifies recurrence risks. De novo mutations are particularly common for severe forms of the disorder caused by *FGFR2* mutations that occurred in the male germline.

Available Assays

Most mutations are clustered in exons 3a and 3c of *FGFR2*, making sequence analysis the ideal testing method. One strategy for sequencing *FGFR2* for mutations is to first focus on the regions of codons 278 and 342, which are mutation hotspots, and then perform more extensive sequencing if these focused regions do not have mutations.

Interpretation of Test Results

Heterozygous mutations should be compared with tables of known mutations (the online Human Gene Mutation Database) for interpretation. There is significant clinical heterogeneity, and Crouzon, Pfeiffer, or Jackson-Weiss phenotypes sometimes result from the same mutations. Thus, detection of a mutation may indicate a craniosynostosis, but not necessarily which syndrome. In the case of a de novo mutation, recurrence risk due to the possibility of germline mosaicism must be considered.

Laboratory Issues

A polymorphism has been reported that can interfere with PCR amplification of exon 3a of *FGFR2*.[11] Because of the issue of phenotypic heterogeneity, prediction of a specific phenotype based on the mutation can be made only if already clearly delineated by clinical examination.

CROUZON SYNDROME (WITH OR WITHOUT ACANTHOSIS NIGRICANS)

Molecular Basis of the Disease

Crouzon syndrome with acanthosis nigricans is due to the *FGFR3* A391E mutation. All other cases of Crouzon syndrome (without acanthosis nigricans) are due to dominant *FGFR2* mutations.

Clinical Utility of Testing

Sometimes the parent of a child affected by one of the craniosynostoses is found to have mild symptoms by clinical examination. Testing the parents for the mutation found in the child clarifies recurrence risks. Approximately 50% of cases of Crouzon syndrome are due to a de novo mutation that occurred in the male germline.

Available Assays

The mutation A391E in *FGFR3* can be detected by a PCR-RFLP assay. One strategy for sequencing *FGFR2* for mutations is to first focus on the regions of codons 278 and 342, which are mutation hotspots accounting for approximately 60% of mutant alleles in Crouzon syndrome, and then perform more extensive sequencing if these focused regions do not have mutations.

Interpretation of Test Results

Heterozygous mutations should be compared with tables of known mutations (the online Human Gene Mutation Database) for interpretation. There is significant phenotypic heterogeneity, and Crouzon, Pfeiffer, or Jackson-Weiss phenotypes sometimes result from the same mutations. Thus, detection of a mutation may indicate

a craniosynostosis, but not necessarily which syndrome. In the case of a de novo mutation, recurrence risk due to the possibility of germline mosaicism must be considered.

Laboratory Issues

A polymorphism has been reported that can interfere with PCR amplification of exon 3a of *FGFR2*.[11] Because of the issue of phenotypic heterogeneity, prediction of a specific phenotype based on the mutation can be made only if already clearly delineated by clinical examination.

JACKSON-WEISS SYNDROME

Molecular Basis of the Disease

All mutations that cause Jackson-Weiss syndrome are found in *FGFR2*. However, due to overlap of phenotypes, the mutation P252R in *FGFR1* that usually causes Pfeiffer syndrome may occasionally be clinically diagnosed as Jackson-Weiss syndrome.

Clinical Utility of Testing

Sometimes the parent of a child affected by one of the craniosynostoses is found to have mild symptoms on clinical examination. Testing the parents for the mutation found in the child clarifies recurrence risks.

Available Assays

Sequencing of the region encompassing the third immunoglobulinlike domain of *FGFR2* (exon 3c) will detect all known mutations.

Interpretation of Test Results

Heterozygous mutations should be compared with tables of known mutations (the online Human Gene Mutation Database) for interpretation. There is significant phenotypic heterogeneity, and Crouzon, Pfeiffer, or Jackson-Weiss phenotypes sometimes result from the same mutations. Thus, detection of a mutation may indicate a craniosynostosis, but not necessarily which syndrome.

Laboratory Issues

A polymorphism has been reported that can interfere with PCR amplification of exon 3a of *FGFR2*.[11] Because of the issue of phenotypic heterogeneity, prediction of a specific phenotype based on the mutation can be made only if already clearly delineated by clinical examination.

MUENKE SYNDROME

Molecular Basis of the Disease

Muenke syndrome does not always have bilateral craniosynostosis. All cases are due to a single mutation in exon 7 of *FGFR3* (P250R), which is considered part of the diagnostic criteria. This mutation, which is between the second and third immunoglobulinlike domains, is thought to increase affinity of ligand binding.

Clinical Utility of Testing

Mutation testing for P250R is an essential part of the clinical diagnosis, since the phenotype is highly variable. Sometimes the parent of a child affected by one of the craniosynostoses is found to have mild symptoms on clinical examination. Testing the parents for the mutation found in the child clarifies recurrence risks.

Available Assays

The P250R mutation generates a new recognition site for the restriction enzyme *Msp* I in exon 7 of the *FGFR3* gene, so testing can be performed using a PCR-RFLP method.

Interpretation of Test Results

The presence of the P250R mutation in *FGFR3* is diagnostic of Muenke syndrome.

Laboratory Issues

If this test is performed as a stand-alone test, restriction enzyme digestion might be the easiest method; however, if the patient has craniosynostosis and is being evaluated for a broader spectrum of conditions, sequencing or DHPLC analysis of exon 7 may be more useful.

BEARE-STEVENSON SYNDROME

Molecular Basis of the Disease

Beare-Stevenson syndrome has been associated with only two *FGFR2* mutations, S372C and Y375C. In terms of effect on protein, these mutations may be the *FGFR2* equivalents of the S371C and Y373C thanatophoric dysplasia mutations

in *FGFR3*, which result in unpaired cysteine residues that promote ligand-independent dimerization and activation of the receptor.

Clinical Utility of Testing

Beare-Stevenson syndrome is an extremely rare lethal disease, and molecular detection of one of the known mutations in *FGFR2* is very helpful for definitive diagnosis of a condition that extremely few clinicians have seen.

Available Assays

Sequence analysis is the preferred method of testing because of the difficulty in obtaining control specimens for validation of other methodologies. This is a common problem in extremely rare conditions.

Interpretation of Test Results

Heterozygous detection of *FGFR2* S372C and Y375C mutations by sequence analysis is diagnostic of Beare-Stevenson syndrome.

Laboratory Issues

Most laboratories will not have prior experience testing for this condition, nor are they likely to have suitable control specimens. If a laboratory detects one of these mutations, confirmation of the mutation by a laboratory with prior experience testing for Beare-Stevenson syndrome is advisable.

OSTEOGLOPHONIC DYSPLASIA

Molecular Basis of Disease

Dominant mutations in *FGFR1* have recently been identified in patients with osteoglophonic dysplasia.[12] All three of the reported mutations occur at positions that are homologous to mutation sites in other *FGFR* genes associated with other syndromes.

Clinical Utility of Testing

The clinical utility of testing is for confirmation of the diagnosis of osteoglophonic dysplasia and for prenatal testing if a mutation is identified in a previous affected child.

Available Assays

Since the association of osteoglophonic dysplasia with *FGFR1* mutations is a recent finding, clinical testing is unlikely to be available. The reported mutations would be detected by sequencing exons 9 and 10 of *FGFR1*.

Interpretation of Test Results

Results to be used in a clinical setting should be interpreted with caution since there is little data available on mutations in osteoglophonic dysplasia. Further investigation is needed to determine the frequency and locations of *FGFR1* mutations, to determine the frequency of de novo mutations, to elucidate any genotype-phenotype correlations, and to understand the effect of putative mutations on receptor function.

Laboratory Issues

The main laboratory issues are rarity of the disorder and the relative lack of data to assist in the interpretation of results.

ACHONDROPLASIA

Molecular Basis of the Disease

Achondroplasia is characterized by short stature, predominantly with shortening of the upper arms and thighs, more normal length of the forearms, lower legs, and torso, and an average adult height of approximately 4 feet. Other features include prominent forehead, flattened nasal bridge, lordosis, spinal stenosis, tibial bowing, and obstructive apnea. Achondroplasia results from a dominant gain-of-function mutation in the *FGFR3* gene. One very common mutation, a G→A transition at nucleotide 1138 that leads to the substitution of glycine 380 by arginine (G380R), accounts for approximately 97% to 98% of cases. One less common mutation, also G380R (nucleotide 1138 G→C), accounts for a further 1% to 2% of cases. The remainder are very rare mutations, G375C and G346E. When both partners of a couple have achondroplasia, there is a 1 in 4 chance of homozygous achondroplasia for each pregnancy. This is a much more severe disease that, like thanatophoric dysplasia, also affects the ribs and is lethal. Approximately 80% of cases of achondroplasia are the result of a new mutation, making nucleotide 1138 of the *FGFR3* gene the most frequently mutated nucleotide known in humans. Current data support the observation that all cases of de novo mutation occur in the paternally inherited allele of *FGFR3*, and the incidence of achondroplasia correlates with advanced (>35 years) paternal age.

Clinical Utility of Testing

Testing is often performed late in pregnancy following detection of short limbs by ultrasound to differentiate between heterozygous achondroplasia (viable fetus) and thanatophoric dysplasia (lethal). The distinction aids in directing the medical management of the delivery and helps prepare the family for an unfortunate outcome if the fetus has thanatophoric dysplasia. More rarely, testing may be performed in the first trimester because one parent has achondroplasia. If both parents have achondroplasia, prenatal testing can determine whether the fetus has lethal homozygous achondroplasia (25% risk), has heterozygous achondroplasia (50% risk), or does not have any *FGFR3* mutation (25% risk) Note that achondroplastic parents may view heterozygous achondroplasia as the preferred outcome. Testing is performed neonatally to confirm a clinical diagnosis.

Available Assays

Restriction enzyme digestion can detect both of the main mutations. Sequence analysis of exon 10 of *FGFR3* detects very close to 100% of mutations. DHPLC analysis of exon 10 should also detect the same proportion of mutations.[13] A homogeneous assay has been reported that detects G380R mutations (in addition to the hypochondroplasia mutation N540K) using real-time PCR and multiple probes.[14]

Laboratory Issues

Whenever it is necessary to determine whether a fetus has achondroplasia or thanatophoric dysplasia late in pregnancy, the laboratory normally tests for achondroplasia first (if doing so by restriction enzyme assays) because of the relative ease of detecting almost all cases of achondroplasia. If the achondroplasia screen is negative, testing for the wider spectrum of mutations that can cause thanatophoric dysplasia is appropriate. If the original assay involved sequence analysis of exon 10 of *FGFR3*, most of the achondroplasia mutations and about half of the thanatophoric dysplasia mutations would be detected in this single assay.

HYPOCHONDROPLASIA

Molecular Basis of the Disease

Hypochondroplasia is a milder skeletal dysplasia than achondroplasia, although the conditions may overlap (both clinically and molecularly). For example, the same G380R(1138G→A) mutation in *FGFR3* is found in 97% of cases of achondroplasia and 5% of cases of hypochondroplasia. *FGFR3* mutations account for 50% to 75% of cases of hypochondroplasia. The remaining cases do not currently appear to have *FGFR* mutations, and thus the possibility of genetic heterogeneity must be considered. Two different point mutations at codon 540 account for up to 70% of hypochondroplasia alleles.

Clinical Utility of Testing

Testing is often performed late in pregnancy following detection of short limbs by ultrasound. Testing is performed neonatally to confirm a clinical diagnosis.

Available Assays

Restriction enzyme analysis will detect the two N540K mutations in exon 13 of *FGFR3*. The remaining reported mutations are detectable by sequencing exons 10, 13, and 15 of *FGFR3*. A homogeneous assay has been reported that detects the two N540K mutations using real-time PCR and multiple probes.[14]

Interpretation of Test Results

Approximately 5% of individuals who are clinically diagnosed with hypochondroplasia have the common G380R (1138G→A) mutation in *FGFR3* usually associated with achondroplasia.

Laboratory Issues

Because not all cases of hypochondroplasia result from *FGFR* mutations, the laboratory report following a negative finding should include a comment on the possibility of genetic heterogeneity.

THANATOPHORIC DYSPLASIA

Molecular Basis of the Disease

Thanatophoric dysplasia is a lethal form of neonatal dwarfism, with two clinical types. Type 1 is characterized by micromelia with femoral bowing and, uncommonly, a cloverleaf skull. Type 2 is characterized by micromelia without femoral bowing and with uniform presence of moderate to severe cloverleaf skull. Thanatophoric dysplasia type 1 is caused by mutations in *FGFR3* (R248C, S249C, G370C, S371C, and Y373C) that introduce an unpaired cysteine residue into the protein. Inappropriate pairing of

these cysteines is hypothesized to cause activation of the receptor in the absence of its ligand (constitutive receptor activation). The K650E mutation in *FGFR3* causes thanatophoric dysplasia type 2 by a different mechanism. This mutation involves a charge reversal, whereas K650N, which does not have the same charge reversal, causes hypochondroplasia. The X807 read-through mutations are quite rare, but interesting because the natural stop codon is destroyed, resulting in additional amino acids at the carboxyl end of the protein (Hemoglobin Constant Spring is another example). Even more interesting, given the proposed dominant gain-of-function nature of *FGFR3* mutations in thanatophoric dysplasia type 1, is the fact that the presence of 141 additional amino acids at the carboxyl end of the protein should enhance the receptor function.

Clinical Utility of Testing

Thanatophoric dysplasia is the most clinically severe of the FGFR-related skeletal dysplasias. Discovery of a mutation for thanatophoric dysplasia during pregnancy allows the parents to prepare for the birth of a child who usually will not survive and somewhat changes the focus of the medical management of the delivery toward the care of the mother. Several similar lethal forms of dwarfism also are inherited as autosomal recessive traits with a 25% recurrence risk. The finding of a thanatophoric dysplasia mutation reduces the recurrence risk to the risk of germline mosaicism.

Available Assays

Sequence or DHPLC analysis of *FGFR3* exon 10 will detect all of the achondroplasia mutations and about half of the thanatophoric dysplasia mutations. The remaining thanatophoric dysplasia mutations are located in various exons throughout the gene.

Laboratory Issues

As noted above, when this test is being performed late in pregnancy to distinguish between thanatophoric dysplasia and achondroplasia, a screen for achondroplasia mutations is performed first (if the methodology is by restriction enzyme analysis) because 99% of cases are caused by only two mutations.

Acknowledgments

The author thanks Lisa Graham and Leo Dimnik for the results shown in the figures and Jennifer Gerritsen and Karen McElligott for helpful comments on the manuscript.

References

1. Wilkie AOM. Craniosynostosis: genes and mechanisms. *Hum Mol Genet.* 1997;6:1647–1656.
2. OMIM (On-line Mendelian Inheritance in Man) [database online]. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM.
3. GeneReviews. Specific reviews and their authors: Robin NH, Craniosynostosis syndromes (FGFR-related); Bellus GA, Kelly TE, Aylsworth AS, Hypochondroplasia; Francomano CA, Achondroplasia. Available at: http://www.genetests.org.
4. The Human Gene Mutation Database. [database online]. Available at: http://www.hgmd.org [original citation: Krawczak M, Cooper DN. The Human Gene Mutation Database. *Trends Genet.* 1997;13:121–122).
5. Roberts RJ, Vincze T, Posfai J, Macelis, D. Restriction enzyme database. *Nucleic Acids Res.* 2003;31:418–420 [available from Rebase: The Restriction Enzyme Database, database online, at: http://rebase.neb.com/rebase/).
6. Cohen MM Jr, Gorlin RJ, Fraser FC. Craniofacial disorders. In: Rimoin DL, Connor JM, Pyertiz RE, eds. *Emery and Rimoin's Principles and Practice of Medical Genetics.* 3rd ed. New York: Churchill Livingstone; 1997:1121–1147.
7. Rimoin DL, Lachman RS. Chondrodysplasias. In: Rimoin DL, Connor JM, Pyertiz RE, eds. *Emery and Rimoin's Principles and Practice of Medical Genetics.* 3rd ed. New York: Churchill Livingstone; 1997:2779–2815.
8. Cohen MM Jr, MacLean RE. *Craniosynostosis.* 2nd ed. Oxford University Press, United Kingdom, 2000.
9. Strachan T, Read AP, eds. *Human Molecular Genetics.* 2nd ed. Oxford, UK: Bios; 1999.
10. Wilkie AOM. The molecular basis of genetic dominance. *J Med Genet.* 1994;31:89–98.
11. Wong LJ, Chen TJ, Dai P, Bird L, Muenke M. Novel SNP at the common primer site of exon IIIa of *FGFR2* gene causes error in molecular diagnosis of craniosynostosis syndrome [erratum appears in *Am J Med Genet.* 2002;109:166]. *Am J Med Genet.* 2001;102:282–285.
12. White KE, Cabral JM, Davis SI, et al. Mutations that cause osteoglophonic dysplasia define novel roles for FGFR1 in bone elongation. *Am J Hum Genet.* 2005;76:361–367.
13. Su YN, Lee CN, Chien SC, Hung CC, Chien YH, Chen CA. Rapid detection of *FGFR3* gene mutation in achondroplasia by DHPLC system-coupling heteroduplex and fluorescence-enhanced primer-extension analysis. *J Hum Genet.* 2004;49:399–403.
14. Schrijver I, Lay MJ, Zehnder JL. Rapid combined genotyping assay for four achondroplasia and hypochondroplasia mutations by real-time PCR with multiple detection probes. *Genet Test.* 2004;8:185–189.

Chapter 10

Cystic Fibrosis

Jean A. Amos

Molecular Basis of Disease

Cystic fibrosis (CF) is the most common lethal autosomal recessive disease in whites, with an estimated incidence of 1 in 2500 to 3300 live births. Approximately 30,000 children and adults in the United States are affected and approximately 850 individuals are newly diagnosed annually, the majority less than 1 year old. For a current, comprehensive review of clinical CF and molecular diagnostics for this disorder, see http://www.genetests.org/.

The gene mutated in CF, the Cystic Fibrosis Transmembrane Conductance Regulator (*CFTR*), spans approximately 230 kilobases (kb) on chromosome 7q31.2, contains 27 coding exons, produces a 6.5 kb mRNA product, and encodes a CFTR protein of 1480 amino acids with a mass of ~170,000 daltons. The protein CFTR is in the ATP-binding cassette family of transporter proteins, containing five domains: two membrane-spanning domains, a regulatory domain, and two nucleotide-binding domains that interact with adenosine triphosphate (ATP). Mutations that lead to an abnormal CFTR protein cause defective electrolyte transport and defective chloride ion transport in apical membrane epithelial cells, affecting the respiratory tract, pancreas, intestine, male genital tract, hepatobiliary system, and exocrine system, resulting in complex multisystem disease.

CF is characterized by viscous mucus that occurs from faulty transport of sodium and chloride, which leads to salt loss, and affects water content of the extracellular space. CF has a highly variable presentation and clinical course. The clinical features of "classic" CF are dominated by pulmonary disease. The median age of diagnosis is 6 to 8 months, and nearly two thirds of affected individuals are diagnosed by the age of 1 year. Recurrent infections that lead to respiratory failure are the major cause of morbidity. Approximately 90% of CF patients die from pulmonary complications with an overall average survival of 30 years.

Approximately 85% of CF patients have pancreatic insufficiency as a result of obstruction of the pancreatic ducts and subsequent scarring and destruction of exocrine functions. Neonatal meconium ileus occurs in approximately 15% of newborns with CF. Other manifestations include chronic sinusitis, nasal polyps, liver disease, pancreatitis, and congenital bilateral absence of the vas deferens (CBAVD). Males with CBAVD are azoospermic and have an increased frequency of mutations in one or both *CFTR* alleles or an incompletely penetrant mutation (the intron 8 variant 5T allele) in a noncoding region of *CFTR*. These men usually have no pancreatic disease and may have normal, borderline, or elevated sweat electrolytes. A small subset of patients with "atypical" CF have chronic *Pseudomonas* bronchitis, normal pancreatic function, and normal or intermediate sweat electrolytes.

Treatment for CF patients is palliative and includes chest percussion to clear secretions, control of infections by antibiotic therapy, and improvement of nutrition through pancreatic enzyme replacement. The goals of therapy are maintenance of adequate nutritional status, prevention of pulmonary and other complications, encouragement of physical activity, and provision of adequate psychosocial support. Although the development of new therapies for CF based on correction of electrolyte transport is a major research focus, there is no cure for CF.

A diagnosis of CF in a symptomatic or at-risk patient is suggested by clinical presentation and confirmed by a sweat test. In the presence of clinical symptoms (e.g., recurrent respiratory infections), a sweat chloride level above 60 mmol/L is diagnostic for CF. Although the results of this test are valid in a newborn as young as 24 hours, collecting a sufficient sweat sample from a baby younger than 3 or 4 weeks old is difficult. The sweat test also confirms the diagnosis in older children and adults but is not useful for carrier detection.

Since the identification of the *CFTR* gene, more than 1,400 mutations and more than 200 sequence variants have been identified; a complete list can be found at http://www.genet.sickkids.on.ca. Only one mutation is extremely frequent, ΔF508. This mutation is a 3–base pair (bp) deletion in exon 10 that occurs on approximately 70%

of CF chromosomes worldwide; thus, approximately half of CF patients are ΔF508 homozygotes. The majority of mutations are private, occurring only in single families, or rare. G542X is the most frequent mutation in whites after ΔF508, occurring at a frequency of 2% among CF chromosomes. The variety and allelic frequency of *CFTR* mutations vary by ethnic group; for example, the most frequent mutation on Ashkenazi Jewish CF chromosomes is W1282X, and, in this population, ΔF508 is present on only 30% of CF chromosomes.

Clinical Utility of Testing

CF mutation analysis is useful for a variety of clinical indications, including diagnosis, newborn screening, prenatal diagnosis for at-risk pregnancies, and carrier detection. The vast majority of laboratory tests are for carrier detection and risk revision.

Mutation analysis is a useful adjunct to sweat testing for diagnosis of CF, particularly for patients with borderline sweat tests, patients with atypical clinical presentations with normal sweat electrolytes, and at-risk newborns for whom sufficient quantities of sweat cannot be collected. Mutation analysis of a diagnosed proband also is useful for identifying the familial mutation for carrier testing of at-risk relatives. Both genetic and sweat testing are used as follow-up to immunoreactive trypsinogen-level testing in newborn screening programs.

Targeted mutation analysis is used for prenatal testing of pregnancies with a 1-in-4 risk for which both mutations have been identified in the two carrier parents. Prenatal testing often is offered in a lower-risk pregnancy presenting in the second trimester on ultrasound with fetal echogenic bowel (EB). EB is associated with an increased risk for CF in the absence of a positive family history. The risk for CF varies between 2% and 20% depending on the grade of the EB. Mutation analysis of the parents of a fetus with EB may reveal that both are carriers, but most often neither or only one parent is a heterozygote. Prenatal diagnosis can determine that a heterozygous parent has transmitted an identified mutation, but there is no additional testing that can further clarify the fetal CF status. In such cases, Bayesian analysis with incorporation of the grade of EB is used to modify the risk. In practice, scant clinical information, including the grade of the EB, is provided to the clinical laboratory. Furthermore, some clinicians submit a fetal sample without testing the parents, and formal genetic counseling may not be offered to the couple until the laboratory testing has been completed.

Prior to the cloning of the *CFTR* gene, couples learned that they were both CF carriers upon the birth of an affected child. No robust assays were available to assess a subsequent 1-in-4-risk pregnancy. Now, the most common indication for mutation analysis is direct heterozygote detection for carrier risk revision.

In 1997, a National Institutes of Health (NIH) Consensus Development Conference recommended CF carrier testing for all couples planning a pregnancy, regardless of family history. However, at that time, there was no standardization of CF testing among laboratories and no commercially available reagents or kits. The largest challenges to development of standardized mutation panels were the ethnic diversity and admixture of the US population, which complicated the selection of mutations for a standardized screening panel. Using available data from more than 20,000 CF patients, the American College of Medical Genetics (ACMG) and the American College of Obstetrics and Gynecology (ACOG) recommended a panethnic panel of 23 mutations that occur at a frequency of >0.1% in any of the major US ethnic groups, plus reflex testing for four additional sequence variants under specified conditions (see Reference 1 and http://www.acmg.net/resources/policies/pol-005.asp). Simultaneously, these organizations developed and distributed an educational document entitled "Preconceptual and Prenatal Carrier Screening for Cystic Fibrosis" to the memberships of ACMG and ACOG. Thus, CF carrier screening is now the standard of care in the United States.

The ACMG/ACOG minimum core mutation panel is shown in Table 10-1. While intended to be panethnic, it detects primarily mutations that are most frequent in the Ashkenazi Jewish and the non-Hispanic European white populations. Detection frequencies also are provided for the Hispanic (an admittedly mixed ethnic group) and black populations, but mutation detection frequency is unknown for individuals of Asian descent. The use of more extended mutation panels is not recommended by ACOG and ACMG for routine carrier screening. However, a report by Heim et al.,[2] published immediately after the current testing recommendations, suggests that a panel of 64 mutations yields better coverage in all ethnic groups.

Reflex testing for the intron 8 5T/7T/9T sequence variation is specifically recommended for R117H positives, since 5T versus 7T or 9T in cis increases CF severity in a compound heterozygote with a second CF mutation. Since

Table 10-1. Recommended Core Mutation Panel for General Population CF Carrier Screening

Standard Mutation Panel				
G85E		ΔF508	R560T	R1162X
R117H	R334W	1717(–1) G→A	1898(+1) G→A	3659delC
621(+1) G→T	R347P	G542X	2184delA	3849(+10 kb) C→T
711(+1) G→T	A455E	G551D	2789(+5) G→A	W1282X
	ΔI507	R553X	3120+(1) G→A	N1303K
Reflex Tests				
intron 8 5T/7T/9T, I506V, I507V, F508C				

some methodologies do not distinguish between ΔF508 and other alterations in close proximity, a second reflex to three exon 10 sequence variants (F508C, I506V, and I507V) is recommended when unexpected ΔF508 or ΔI507 homozygosity is detected to avoid reporting of false-positive results due to interference by these surrounding polymorphisms.

The ACMG/ACOG panel is not designed for diagnostic testing of CF patients or men with CBAVD. A larger mutation panel may provide increased detection for symptomatic patients. The intron 8 5T allele has been associated with CBAVD, but the intron 8 variation is not in the recommended screening panel because the 5T allele is no associated with classic CF.[3]

The ACMG/ACOG panel is expected to evolve as new data emerge, but the total number of recommended mutations will not substantially increase. As described at http://www.acmg.net/ the American College of Medical Genetics originally recommended a carrier screening panel of 25 mutations that was later decreased to 23. The I148T mutation was deleted based on non-pathogenocity and 1078delT also was dropped from the screening panel based on infrequency.

Available Assays

Prior to 2001, professional recommendations limited CF carrier testing to individuals with a family history and their partners. Testing laboratories offered laboratory-developed assays using standard platforms, including reverse and forward dot blot, polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis, oligonucleotide ligation assays (OLA), sequence specific primer–polymerase chain reaction (SSP-PCR), various methods for exon scanning, and sequencing. Techniques reported by 50 laboratories that participated in a proficiency testing survey from the ACMG/College of American Pathologists (CAP) in 2002 were probe hybridization (38%), OLA (26%), PCR-RFLP (18%), SSP-PCR (16%), sequencing (16%), mutation scanning (8%), and other (26%).

Currently, there are several available analyte-specific reagents (ASRs). Clinical CF tests using these ASRs are considered laboratory-developed tests and must be validated by the laboratory prior to implementation. Technical standards and guidelines for CF mutation testing have been published by the ACMG[5] and are available at http://www.acmg.net. These guidelines address the technology platforms used in commercial ASRs and laboratory-developed tests for CF mutations.

The available ASRs are summarized in Table 10-2 with additional vendors anticipated to develop ASRs in the future. All are robust and include at least the ACMG/ACOG minimum core mutation panel, but vary considerably with respect to criteria that laboratories consider for platform adoption: reagent/royalty costs, footprint of instrumentation, throughput, flexibility, and data analysis. All but the Invader platform are PCR based, requiring a royalty, which varies among academic, hospital, and commercial laboratories; the PCR royalty is built into the Roche ASR. Reagent costs per patient are significant for all of the ASRs. The commercial ASRs are not likely to be submitted for FDA approval in the foreseeable future until the ACMG/ACOG mutation panel has undergone additional review.

The Nanogen and ABI platforms offer both semiautomated detection and data analysis. Both require capital

Table 10-2. Commercially Available Analyte Specific Reagents

Vendor	ABI (Indianapolis, IN)	Innogenetics (Ghent, Belgium)	Nanogen (San Diego, CA)	Orchid Biosciences (Abington, Oxford, UK)	Roche (Indianapolis, IN)	Tm Bioscience (Toronto, Canada)	Third Wave Technologies (Madison, WI)
Platform	PCR target amplification, OLA, and electrophoretic discrimination of alleles	PCR target amplification and reverse line blot	PCR target amplification, chip hybridization, and FRET detection on an electronic microarray	ARMS target amplification and electrophoretic discrimination	PCR target amplification, and reverse line blot	PCR target amplification, allele-specific primer extension, hybridization to universal tags, and sorting on Luminex 100 analyzer	PCR and signal amplification and FRET detection
Product or platform name	Oligo Ligation Assay	Inno-LiPA CFTR	Nanochip	Elucigene CF	CF Gold	TAG-IT CF40	InPlex
Format	16-capillary sequencer	12- and 17-mutation strips	100-site chip	Slab gel	31-mutation strip	96-well plate	96-well plate

FRET, flourescence resonance energy transfer; PCR, polymerase chain reaction, ARMS, amplification refractory mutation system.

purchase of an instrument that has a relatively large footprint. The assay format for both is flexible and easily customized. Potentially, both are amenable to high throughput testing and can be used as a consolidated work platform for other assays.

The Roche and Innogenetics ASRs are reverse line blots that are in widespread use because of early availability and ease of validation in the laboratory. The Inno-Lipa detection system can be semiautomated using a customized instrument. Both are amenable to low and high test volumes but are closed platforms such that additional mutations cannot be added by the laboratory. Current versions require manual analysis of the strip data, but both companies are developing automated data-analysis software.

Both the Roche and Innogenetics probe-hybridization assays include the intron 8 5T/7T/9T variant and the exon 10 polymorphisms in the first-line assay, rather than as a reflex test as recommended by ACMG and ACOG. However, positioning of intron 8 variant probes at the bottom of strips allows strips to be cut to remove the polyT probes from first-line testing. Laboratories obtaining genotypes for the intron 8 polyT variant for all patients are required to report the results, as the technical guidelines require reporting of all patient data. This may be useful for detection of males with CBAVD, since the 5T allele is associated with this condition, but this allele is also present in 5% to 10% of the general population. Several physicians, not recognizing that this allele is benign in the absence of a second mutation, have ordered amniocentesis for couples for which both partners are mutation negative but one member of the couple carries the 5T allele.[5] Because amniocentesis is not without risk, this is considered a misuse of the data.

The Orchid ASR is a manual procedure that features gel-based detection of sequence-specific primer (SSP) amplicons. Although the amplifications are multiplexed, samples from each patient are loaded into several gel lanes and, as such, this closed platform may be best suited for lower-test-volume laboratories.

The Tm Bioscience assay incorporates multiplex PCR and multiplex allele-specific primer extension with a proprietary Universal Tag sorting system on the Luminex 100 xMAP platform. This is a high-throughput assay that incorporates automated software analysis for genotype calls.

The Third Wave Technologies assay requires only pipetting steps and is easily automated using a liquid handler. Thus, this assay is potentially low cost, requiring only a fluorescent plate reader. Detection However, the assay is unique among CF ASRs in that it combines limited cycles of PCR and relies on signal amplification using the Invader FRET technology. Because this assay format is not multiplexible, the detection steps are performed on a custom "InPlex" card, in which reaction products are distributed to wells of dried down reagents, each specific for the detection of a single mutant allele.

Interpretation of Test Results

Mutation analyses have only three possible results, detection of no, one, or two mutations. However, reporting of CF results is complex because: (a) detection in a targeted mutation assay is less than 100% and (b) there are many interpretations for each of the possible results, depending on the indication for the test and the ethnicity and family history of the consultand.

For example, the finding of one mutation in an asymptomatic patient demonstrates the consultand to be a carrier but has a very different interpretation in a diagnostic context. For a symptomatic patient, detection of one mutation could mean that the patient is indeed a CF carrier; however, this patient may also be a compound heterozygote for one identified mutation and a second unidentified mutation that is either private or rare.

Most laboratories offer CF tests using a single orderable test code. Prenatal tests are easily identified based on the fetal sample type (amniocytes or chorionic villus biopsies), but it is not easy to distinguish carrier from diagnostic test orders, since both use whole blood specimens. This issue can be addressed by offering clients three separate test codes for each of the three test indications: carrier, prenatal, or diagnostic testing.

The finding of two mutations in a diagnostic test confirms a diagnosis of CF. However, genotype-phenotype relationships are not well established enough that prognostic statements can be made for the many different possible homozygous and compound heterozygous genotypes. The interpretation of one or no detected mutations for a symptomatic patient must include a recommendation for diagnostic sweat electrolyte analysis.

Since the late 1990s, more than 95% of all CF tests are for carrier screening. Although the finding of one mutation in this context is straightforward, the majority of these analyses are negative, with no mutations detected. In the setting of a negative test result, the physician must be informed that the revised carrier risk, while reduced from the prior risk, is a nonzero number because mutation detection is incomplete. As shown in Table 10-3, prior and revised carrier risks both are dependent on patient ethnicity and family history. Prior carrier risks for patients without a family history of CF are calculated from the ethnic frequency of CF, which is highest among whites (1 in 25), lower in Hispanics (1 in 46), and blacks (1 in 62), and lowest among Asians (1 in 90).

Prior carrier risks for patients with a positive family history are based on pedigree analysis. Meaningful risks vary from a high of 2 in 3 for an individual with an affected sibling to a low of 1 in 16 for a patient with an affected second cousin. The prior risk for a consultand with a more distant affected relative is small relative to the population risk.

Revised risks also are based on the frequency of the tested mutations in the patient's ethnic group. The data in Table 10-3 illustrate that the frequency of mutation detec-

Table 10-3. Prior and Residual CF Risks for Patients with a Negative Carrier Test

A. White, Ashkenazi Jewish: 97% Detection

Affected Relative	Prior Risk	Revised Risk
No affected relative	1 in 25	1 in 801
Sibling	2 in 3	1 in 18
Niece/nephew/half-sibling	1 in 2	1 in 34
Aunt/uncle	1 in 3	1 in 68
First cousin	1 in 4	1 in 101
First cousin, once removed	1 in 8	1 in 234
Second cousin	1 in 16	1 in 494

B. White, European: 90% Detection

Affected Relative	Prior Risk	Revised Risk
No affected relative	1 in 25	1 in 241
Sibling	2 in 3	1 in 6
Niece/nephew/half-sibling	1 in 2	1 in 11
Aunt/uncle	1 in 3	1 in 21
First cousin	1 in 4	1 in 31
First cousin, once removed	1 in 8	1 in 71
Second cousin	1 in 16	1 in 151

C. Hispanic: 57% Detection

Affected Relative	Prior Risk	Revised Risk
No affected relative	1 in 46	1 in 110
Sibling	2 in 3	1 in 2.2
Niece/nephew/half-sibling	1 in 2	1 in 3.3
Aunt/uncle	1 in 3	1 in 5.7
First cousin	1 in 4	1 in 8
First cousin, once removed	1 in 8	1 in 17
Second cousin	1 in 16	1 in 36

D. Black: 69% Detection

Affected Relative	Prior Risk	Revised Risk
No affected relative	1 in 62	1 in 198
Sibling	2 in 3	1 in 2.6
Niece/nephew/half-sibling	1 in 2	1 in 4.2
Aunt/uncle	1 in 3	1 in 7.5
First cousin	1 in 4	1 in 11
First cousin, once removed	1 in 8	1 in 23
Second cousin	1 in 16	1 in 50

tion using the ACMG/ACOG panel ranges from 97% for individuals of Ashkenazi Jewish descent to 57% for Hispanics. Residual risks are not given for Asian patients because these risks are not known. Risk revision is even more complicated for individuals of mixed heritage.

Thus, responsible reporting for CF mutation test interpretation includes risk assessment based on the indication for the test, patient ethnicity, and family history. Often, samples are referred for testing without patient information and, since this information is not an interpretative requirement for routine clinical tests, clients must be educated to provide it. A two-tiered approach for gathering preanalytic information can be used. First, a specialized requisition that solicits patient information can be used for test ordering. When only partial or no information is received on the requisition, a follow-up questionnaire can be faxed to the client. Gathering patient information by telephone is not recommended since such calls interrupt workflow in the client's office and also because the information may not be readily available to the recipient of the call. Using this system, sufficient information is obtained to provide a patient-specific report for more than 85% of all samples referred for carrier screening. Ethnicity and family history are self-reported by patients, and inaccurate information can be transmitted to the laboratory. It is not unusual for a client to state that the patient is white, omit the information that the patient is of Ashkenazi Jewish descent, and simultaneously order several carrier tests for Jewish recessive diseases. In such an instance, follow-up with the client is required to clarify whether the patient is or is not Jewish, and whether the patient's partner is Jewish.

For negative carrier screens, the report can be tailored to the patient's specific ethnicity and family history when this information is available (see Table 10-3), or a comprehensive report covering all possibilities is written when the information is not provided. To emphasize to the ordering physician the impact of patient information on the interpretation, the report also contains the ethnic-family tables shown in Table 10-3.

The published ACMG/ACOG recommendations[1] include well-written model laboratory reports for detection of no or one mutation in a carrier context for individuals with a negative personal and family history. These model reports are also included in the educational materials that have been distributed to the members of ACOG and ACMG. They contain a field for patient ethnicity, but the interpretation is not patient specific. In the model mutation-negative report, the prior and revised carrier risks for all ethnicities are presented in the interpretation field in a tabular format. However, in the author's opinion, it is incumbent upon the laboratory to include specific interpretive information about the patient in negative test reports.

Laboratories can anticipate the most common reporting formats and prepare standard template reports. In practice, as many as 50 templates may be needed, which can then be modified to the patient's specific clinical situation with a minimum number of keystrokes. This approach accommodates high test volumes while maintaining reporting standards. A sample template report for the most common report, a negative test for a patient with a negative family history, is shown in Figure 10-1. Each report contains a patient-specific risk and a statement that the carrier risk of the consultand is reduced but that the risk for having a child affected by CF is also a function of the partner's carrier status. In this way, the nonzero value of the revised risk for a negative test is emphasized as well as the interplay of the genotypes of both parents. Clients have the choice of first testing only one member of the couple, most often a pregnant woman presenting for prenatal care, or testing both members of a couple simultaneously.

Although the ACMG/ACOG mutation panel is designed for carrier screening for individuals without a family history of CF, individuals with an affected relative

CYSTIC FIBROSIS GENOTYPR®
TEST CODE 5356: CARRIER STUDY

SOUCE:	Blood
INDICATION:	Carrier screening, negative history
ETHNICITY:	**ETHNICITY**[a]
PRIOR CARRIER RISK:	**PRISK**[b]
RESULT:	Negative for the mutationns analyzed
REVISED CARRIER RISK:	**RRISK**[c]

INTERPRETATION: Molecular analysis reyealed that this patient is negative for common mutations that comprise **PERCENT1**%[d] of cystic fibrosis (CF) mutations in his/her ethnic group. It is our understanding that this patient has a negative personal and family history of CF.

The revised risk that this patient is a carrir, based on stated ethnicity, negative family and personal history of CF, and thi negative mutation analtsis.is 1 in **RRISK**[b] (0.**PECENT2**%[e]). The risk that this patient will have a child with CF is decreased by this result. This risk, however, is also dependent on the carrier status of his/her partner.

METHOD:
The CF Transmembrane Conductance Regulator Gene was tested for the presence of 25 mutations that are recommended by the American College of Human Genetics for CF camier screening (G85E, R117H, I148T, 621(+1)G>T, 711(+1)G>T, 1078delT, R334W, R347P, A455E, Delta 1507, Delta F508, 1717(-1)G>A, G542X, G551D, R553X. R560T, 1898(+1)G>A, 2184delA, 2789(+5)G>A, 3120+(1)G>A, R1162X, 3659delC, 3849(+10kb)C>T, W1282X and N1303K) using the polymerase chain reactions and allele-specific oligonucleotide hybridization. 6 polymorphisms (intron 8 5/7/9T, I506V, I507V, F508C) were also tested.

GENERAL DISCLAIMER: DNA studies do not constitute a definitive carrier test for CF in all individuals. Thus, interpretation is given as a probability. Risk calculation requires accurate family history and ethnicity information. While DNA analysis is highly accurate, technical sources of error include rare genetic variants that interfere with the analysis.

This test or one or more of its components was developed and its performance characteristics determined by Specialty Laboratories.It has not been cleared of approved by the U.S. Food and Drug Administration (FDA). The FDA has determined that such clearance or approval is not necessary. 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") as qualified to perform high-complexity clinical laboratory testing.

Laboratory results and interpretation reviewed by jean Amos, PhD FACAG

Figure 10-1. Template report for a patient without a family history of CF and a negative test. a, the patient's ethnicity as listed in Table 10-3 (a different template is used when the patient's ethnicity is Asian, not provided, or complex); b, the prior risk as listed in Table 10-3, no affected relative; c, the revised risk as listed in Table 10-3, no affected relative; d, the mutation detection in the specific ethnic group as listed in Table 10-3; e, the revised risk, expressed as a percentage.

often are referred for carrier risk testing, and the same screening mutation panel is used. A footnote can be provided in the negative-test report for a patient without a family history for CF to alert the physician that residual carrier risk after a negative test is modified by this history. The footnote may also include a recommendation that the physician refer patients with a positive family history for genetic counseling and Bayesian calculation of revised carrier risk. Several template reports may be needed for patients with a positive family history and a negative test; the variable portion of such a template is shown in Figure 10-2. All such reports include a patient-specific revised risk calculation and a recommendation for genetic counseling; the prior and residual risks for patients with a positive family history are also shown in Table 10-3.

Often the affected family member has not been tested and the familial mutation(s) is not known, or the mutation has been identified but the information is not available to the at-risk relative. In such cases, the report should state that the familial mutation(s) is not known, provide a patient-specific Bayesian risk, and make a specific recommendation for testing the closest known carrier relative (e.g., "We recommend that this patient's father's sister be referred for mutation analysis to identify the familial mutation."). When the familial mutation(s) is known and analyzed, the residual risk for the consultand is very low, on the order of the false-negative rate of the assay.

Reporting for most of the possible positive results in a carrier test is straightforward and, as recommended by ACMG/ACOG, should always contain a recommendation for genetic counseling and a suggestion that at-risk relatives also may benefit from genetic counseling and mutation analysis for carrier risk revision. As recommended by ACMG/ACOG, interpretation of positive results for the R117H mutation should always be made in the context of the patient genotype for the intron 8 5T/7T/9T sequence, since expression of the R117H mutation in a compound heterozygous child is variable, depending on the specific genotypic combination. The R117H mutation is most severe when on the same allele as the 5T variant. The reflex of an unexpected ΔF508 homozygous result to testing for other exon 10 sequence variations (I506V, I507V, and F508C) is necessary to prevent false-positive results, since some detection platforms (e.g., probe hybridization) do not discriminate the ΔF508 mutation from these polymorphisms.

The laboratory also must be informed when partners are tested, particularly when one or both are found to be carriers. Partners often have different last names, they may be tested at different times or in different laboratories, or the laboratory volume can be high enough that partners are not recognized during CF testing. The nongeneticist physician usually can interpret a 1-in-4-risk for couples who are both carriers, but may not have the experience to provide

CYSTIC FIBROSIS GENOTYPR®
TEST COD 5356: CARRIER STUDY

SOURCE:	Blood
INDICATION:	Positive family history; **DETAILS**[a]
ETHNICITY:	**ETHNICITY**[b]
PRIOR CARRIER RISK:	**PRISK**[c]; familial mutation unknown
RESULT:	Negative for the mutations analyzed
REVISED CARRIER RISK:	**RRISK**[d]

Additional studies and genetic counseling are recommended; see interpretation.

INTERPRETATION: This patient has a family history of cystic fibrosis (CF). We are not aware whether any members of the family have had mutation analysis. Thus, the familial mutation(s) are currently unknown. Based on family history, the prior risk that he/she is a CF carrier is **PRISK**[c]

Molecular analysis revealed that this patient is negative (normal) with respect to common mutations which comprise **PERCENT1**[e]% of CF mutations in his/her ethnic group. Based on positive family history, ethnicity, and this negative mutation analysis, the revised risk that he/she is a CF carrier is **RRISK**[d] (0.**PERCENT2**[f]%).

It would be possible to further revise this patient's camier risk if the familial mutation(s) were identified. To this end, we recommend that this patient,s **RELATIVE**[g], who is an obligate CF carrier, be referred for mutation analysis.

The risk that this patient will have a child with CF is decreased by this result. This risk, however, is also dependent on the carrier status of his/her partner. We recommend that this patient and partner be referred for genetic counseling.

Figure 10-2. Template report for a patient with a family history of CF and a negative test. a, a phrase that describes the closest carrier relative, for example, "patient's mother's sibling is reportedly an obligate carrier"; b, the patient's ethnicity as listed in Table 10-3 (a different template is used when the patient's ethnicity is Asian, not provided, or complex c, the prior risk based on pedigree analysis, as listed in Table 10-3; d, the revised risk, as listed in Table 10-3; e, the mutation detection in the specific ethnic group, as listed in Table 10-3; f, the revised risk, expressed as a percentage; g, the specific relative recommended to be tested, for example, "mother's sibling."

a fetal risk assessment when only one member of the couple is a carrier. Here, the referral for genetic counseling is critical for the single carrier couple. Heathcare providers are very likely to inform the laboratory that one member of the couple has already tested positive when referring the second member for testing, but less likely to alert the laboratory when the couple is tested simultaneously. Again, this information can be solicited on the requisition and follow-up questionnaire, but this is the most difficult preanalytic information to obtain. Healthcare providers need additional education to enhance the probability that they will provide accurate and complete information needed for interpretation of CF test results. Clients must be educated to provide this information so that residual risks to a current or future pregnancy can be stated in the report. Residual fetal risks for couples in a variety of testing circumstances are available in the technical standards and guidelines for CF mutation testing (http://www.acmg.net).

Couples with negative or positive test results may have an affected child in the setting of one parent with a known mutation and the other parent with a rare mutation that is not included in any CF mutation test panel. These births are not avoidable using current technology. Anecdotal reports from the genetic counselors that have seen these patients after the unexpected diagnosis after birth inform us that the couple was not aware of their residual risk after carrier testing. Genetic counseling for couples with one identified CF mutation prior to conception or birth of a child would provide this information to couples.

CF carrier screening is the first large genetic screening program based on DNA analysis in the United States. Little data is available regarding the nongeneticist physician's ability to interpret a complex laboratory report, communicate the results and interpretations to his or her patients, and make appropriate management decisions. The success of this program will depend on many factors, the most important of which is physician education to react appropriately to the complex scenarios that occur in patients and couples.

Laboratory Issues

Several kits for CF testing are commercially available. However, none of the ASRs include controls except the Third Wave Technologies Invader test, which does include some oligonucleotide controls. A complete set of controls for the ACMG/ACOG panel is not commercially available. The Coriell Cell Repository offers a set of DNAs that contains a complete set of the mutations in the ACMG/ACOG panel. The repository also has DNA available for two of the three reflex polymorphisms, F508C and I506V, but not for I507V. Thus, laboratories will have difficulty validating assays using genomic DNA controls, and artificial controls will not be as useful for platforms that rely on electrophoretic separations. Availability of a public source of controls for all the mutations in the ACMG/ACOG mutation panel which is kept current with changes in the recommended mutation panel would be extremely useful for clinical laboratories. A proficiency testing program for CF is available from the CAP and is jointly administered by the ACMG and CAP.

References

1. Grody WW, Cutting GR, Klinger KW, Richards CS, Watson MS, Desnick RJ. Laboratory standards and guidelines for population-based cystic fibrosis carrier screening. *Genet Med.* 2001;3:149–154.
2. Heim RA, Sugarman EA, Allitto BA. Improved detection of cystic fibrosis mutations in the heterogeneous U.S. population using an expanded, pan-ethnic mutation panel. *Genet Med.* 2001;3:168–176.
3. Watson MS, Desnick RJ, Grody WW, Mennuti MT, Popovich BW, Richards CS. Cystic fibrosis carrier screening: Issues in implementation. *Genet Med.* 2002;4:407–409.
4. Rohlfs EM, Zhou Z, Sugarman EA. The I148T CFTR allele occurs on multiple haplotypes: a complex allele is associated with cystic fibrosis. *Genet Med.* 2002;4:319–323.
5. Richards CS, Bradley LA, Amos J, et al. Standards and guidelines for CFTR mutation testing [erratum in: *Genet Med.* 2002;4:471]. *Genet Med.* 2002;4:379–391.

Chapter 11

Deafness

Richard J.H. Smith and Guy Van Camp

Introduction

Recent advances in the molecular biology of hearing and deafness are being transferred from the research laboratory to the clinical arena. This transfer of knowledge is enhancing patient care by facilitating the diagnosis of hereditary deafness. Traditionally, hereditary deafness has been distinguished from nongenetic causes of deafness by otologic, audiologic, and physical examinations, complemented by a family history and ancillary tests such as temporal bone computed tomography, urinalysis, thyroid function studies, ophthalmoscopy, and electrocardiography. Even using this test battery, an unequivocal distinction between genetic and nongenetic causes of deafness often is difficult. If comorbid conditions are identified, the deafness may fall into one of more than 400 recognized types of syndromic hearing loss, but if hearing loss segregates as the only abnormality, diagnosing the deafness as nonsyndromic and inherited is challenging.[1]

The relative contributions of syndromic and nonsyndromic deafness to the total deafness genetic load vary with age of ascertainment. Because syndromes are generally straightforward to recognize, most are noted at birth. In aggregate, syndromes account for about 30% of prelingual deafness. The majority of congenital hereditary deafness, however, is nonsyndromic, and this relative contribution increases with age, reflecting the greater occurrence and diagnosis of postlingual nonsyndromic as compared to postlingual syndromic deafness in late childhood and adulthood.

In many cases, the diagnosis of autosomal recessive nonsyndromic deafness (ARNSD) is a diagnosis of exclusion. Current data suggest that even with a thorough history and physical examination, in simplex families (only one affected child) errors in diagnosis are made one third of the time; that is, a child is diagnosed with ARNSD when the correct diagnosis is congenital acquired deafness.[2] This type of error precludes appropriate genetic counseling and can be a source of concern and anxiety for the family.

Molecular genetic testing offers a potential means to unequivocally diagnose inherited deafness. Allele variants of many genes are known to cause hereditary deafness;[3] however, extreme heterogeneity and the relatively small contribution each gene makes to the total deafness genetic load make it impractical to offer complete mutation screening of all genes known to cause inherited deafness. Currently, clinical mutation screening is offered for three genes that cause nonsyndromic hearing loss and deafness (DFNB), *GJB2*, *SLC26A4*, and *WFS1*, and one gene that causes syndromic deafness, *EYA1*. The clinical utility of screening *GJB2* and *SLC26A4* is based on three facts. First, both genes cause types of deafness that are difficult to diagnose without genetic testing (*GJB2* for DFNB1; *SLC26A4* for DFNB4 and Pendred syndrome); second, the relative contribution of these two genes to the total genetic deafness mutation load is high; and third, both genes are relatively easy to screen by molecular methods. Mutation screening of *WFS1* (nonsyndromic autosomal dominant deafness, DFNA6/14) is offered because the audioprofile of affected persons is unique, and in the case of *EYA1*, the occasional association of branchio-oto-renal syndrome with prenatal lethality has been the impetus to develop molecular genetic testing. In this chapter, genetic testing of these four genes is reviewed.

DFNB1(*GJB2*)

Molecular Basis of Disease

In 1994, Guilford et al. mapped the first locus for ARNSD to chromosome 13q12-13 and named it DFNB1 (Online Mendelian Inheritance in Man [OMIM; database online] #220290).[4] Three years later, the deafness-causing gene at this locus was identified as *GJB2*.[5] *GJB2* encodes a transmembrane protein called Connexin 26 (Cx26) that oligomerizes with five other connexins to form a connexon. Connexons in adjoining cells join to form gap junctions, or conduits, that facilitate the rapid exchange of

electrolytes, second messengers, and metabolites from one cytoplasm to another.[6] Interestingly and unexpectedly, mutations in Cx26 have been found in approximately 50% of persons with severe-to-profound congenital ARNSD in several worldwide populations.[7–11]

Immunohistochemical studies of Cx26 expression in rat cochleae have demonstrated that two groups of cells are interconnected via gap junctions. The first group, nonsensory epithelial cells, includes interdental cells of the spiral limbus, inner and outer sulcus cells, sensory supporting cells, and cells within the root process of the spiral ligament. The second group, the connective tissue cell gap junction system, includes fibrocytes within the spiral ligament and spiral limbus, basal and intermediate cells of the stria vascularis, and mesenchymal cells, which line the scala vestibule and interconnect the two populations of cell types. Expression of Cx26 in the vestibular labyrinth is similar.[12,13] These studies suggest that the Cx26 gap junction system plays a role in potassium recycling, facilitating the rapid transport of K^+ ions through the supporting cell network to the stria vascularis, thereby helping to maintain the unique potassium-sodium endolymph balance.[14]

Although more than 80 different deafness-causing allele variants of *GJB2* have been reported,[15] in populations of European descent a single mutation predominates, 35delG. This mutation reflects the deletion of one deoxyguanosine from a string of 6, resulting in a shift in reading frame and premature protein truncation. Based on an analysis of single nucleotide polymorphisms (SNPs) tightly linked to the 35delG mutation, this mutation segregates on a common haplotype background and arose as a result of a founder effect about 10,000 years ago.[16]

Today, the carrier frequency for the 35delG mutation in the midwestern United States is approximately 2.5%, and in this population roughly two thirds of persons with Cx26-related deafness are 35delG homozygotes.[2,17] Of the remaining persons with Cx26-related deafness, most are 35delG heterozygotes, and carry a second, noncomplementary mutation. Consistent with a founder origination is the observation that in some populations the 35delG mutation is rare. For example, in the Ashkenazi Jewish and Japanese populations, the 167delT and 235delC mutations, respectively, are most common.[18,19]

Clinical Utility of Testing

Establishing a molecular diagnosis of *GJB2*-related deafness is important clinically since children with this type of deafness can avoid further diagnostic tests and are not at increased risk for medical comorbidity. Bony abnormalities of the cochlea are not part of the deafness phenotype, and developmental motor milestones and vestibular function are normal.[11,20,21] The rare exceptions include a child with bony cochlear overgrowth noted at surgery,[22] a child with asymmetry of the right modiolus,[22] a child with vertigo, migraine, and unilateral weakness,[11] and a child with marked prematurity and maturational vestibular weakness.[21]

GJB2-related deafness occasionally cosegregates with a skin abnormality characterized by hyperkeratosis of the palms and soles, often with peeling, known as palmoplantar keratoderma (PPK), but this occurrence is rare and is associated only with specific *GJB2* allele variants.[23] Vohwinkel syndrome (VS), a specific form of PPK and deafness caused by the D66H mutation, has the additional component of autoamputation secondary to bandlike circumferential constrictions of the digits.[24] However, as a general rule, comorbid conditions are uncommon with *GJB2*-related deafness, and vision, intelligence, electrocardiography, and thyroid function are normal.[2,11,25,26]

Genetic screening of *GJB2* to establish a diagnosis of DFNB1 does have important limitations. First among these are the limitations in genetic testing itself.[27] Although *GJB2* is a small gene with only a single coding exon, the more than 80 different allele variants associated with ARNSD are scattered throughout the gene, making mutation screening of the entire coding sequence essential. In spite of this thorough approach, the identification of a single deafness-causing allele variant is not uncommon, implying the presence of a "missed" mutation in a noncoding region or coincidental carrier status in a person with deafness of another etiology.

A second limitation is limited genotype-phenotype correlation for predicting the degree of deafness. Among persons with Cx26 deafness, the degree of deafness can vary from mild to profound (mild [<40 dB], 1.7%; moderate [40–55 dB], 10.3%; moderately severe [56–70 dB], 7.8%; severe [71–90 dB], 30.2%; profound [>90 dB], 50.0%), even among persons with the same mutations.[2,11,19,28,29] Typically, the audiogram has a downsloping or flat pattern.[21] Selective midfrequency loss is rare,[28] and selective low-frequency loss has not been described. Symmetry between ears is normal, although one fourth of individuals have intra-aural differences of up to 20 dB.[11,20] The loss tends to be stable, with neither improvement in hearing nor fluctuation in hearing level over the long term.[21]

Available Assays

A number of mutation detection strategies can be used to screen *GJB2* for nucleotide changes, including restriction enzyme digestion, allele-specific polymerase chain reaction (AS-PCR) analysis, single-strand conformation polymorphism (SSCP) analysis, heteroduplex analysis (HA), denaturing high-performance liquid chromatography (DHPLC), and direct sequencing. Of these methods, SSCP and HA were among the first methods to be used to detect genetic polymorphisms and remain very popular because of their simplicity. However, the gold standard for establishing the identity of unknown nucleotide sequences is direct sequencing.

Interpretation of Test Results

The theoretical sensitivity of *GJB2* mutation screening based solely on AS-PCR detection of the 35delG mutation (defined as the subportion of the population with *GJB2*-related deafness and the 35delG mutation divided by the total population with *GJB2*-related deafness) has been calculated at 96.9% (range, 95.4% to 98.0%).[2] Calculated specificity (defined as the subportion of the population with *GJB2*-unrelated deafness not coincidentally carrying the 35delG mutation divided by the population with *GJB2*-unrelated deafness) is equally high at 97.4% (range, 97.0% to 98.0%). The observed sensitivity and specificity of 94% and 97%, respectively, are comparable to these values.[2]

These calculations assume that the population is randomly mating with respect to *GJB2*. The existence of population substructure, particularly endogamous subpopulations, results in a decreased proportion of heterozygotes (Wahlund's effect), with an overestimation of sensitivity for the population as a whole. Other assumptions made in these calculations include complete penetrance, lack of ascertainment bias (i.e., equal referral rates regardless of genotype), and negligible heterozygote selection advantage, spontaneous mutation rate, and migration effects. Deviation of the actual population from Hardy-Weinberg equilibrium due to these factors is likely to be minimal and does not affect the order of magnitude of the figures obtained, with the possible exception of assortative mating among the deaf.[2]

This mutation screening strategy of 35 delG mutation testing, however, misses *GJB2*-related deafness caused by non-35delG allele variants. More comprehensive mutation screens are based on DHPLC or direct sequencing. We have tested the sensitivity of DHPLC by screening a panel of 55 individuals carrying 52 combinations of 48 distinct *GJB2* sequence variants. Amplicons were analyzed at 62°C, 60°C, and 58°C to increase DHPLC detection efficiency, since all mutations cannot be detected at 62°C (L90P and I230T are not detected at 62°C). DHPLC wave profiles were analyzed for differences in shape and retention time compared to wild-type samples in four separate runs to test detection repeatability. On this basis, we have found DHPLC more sensitive at detecting *GJB2* allele variants than SSCP analysis, the comparative detection rates being 98.1% and 82.3%, respectively. Other authors have reported the sensitivity of DHPLC to range from 95% to 100%.[30–33]

Laboratory Issues

Determination of *GJB2*-related deafness is dependent on the identification of mutations in the DNA of affected individuals. Mutation screening of only exon 2 of *GJB2* by any technique is incomplete because there are two common noncoding, noncomplementary DFNB1-causing mutations that must be considered in persons heterozygous for a known *GJB2* deafness-causing allele variant. These mutations are the intron 1 splice donor mutation (IVS1+1G→A) and the large 5′ 342 kilobase (kb) deletion that includes a portion of *GJB6* and an additional upstream sequence (Δ[*GJB6*-D13S1830]). Based on the relative frequency of *GJB2* allele variants in the general population, the frequency of noncoding *GJB2* mutations associated with deafness at the DFNB1 locus, and phenotype-genotype correlations,[2] the existence of at least one additional mutation associated with the DFNB1 phenotype that is outside the coding region of *GJB2* is predicted.

For mutation screening by DHPLC, we use the acetonitrile gradient and partial denaturing temperature predicted by Wavemaker software, but add analysis 2°C above and below the predicted temperature to detect all possible mutations. To increase column life, cleanup and equilibration durations are extended and hot washing of the column is performed every 200 injections. DHPLC standards are run every 200 injections to confirm column reliability, since the ability of the column to detect standards is directly related to its ability to detect sequence variants in *GJB2* and other genes of interest. Water quality is checked by testing resistivity and total organic content. The purity of PCR products is verified by analyzing a sample at 50°C prior to analysis at partial denaturing conditions on the DHPLC.

PENDRED SYNDROME AND DFNBH (*SLC26A4*)

Molecular Basis of Disease

Mutations in *SLC26A4* (formerly known as *PDS*) cause Pendred syndrome (PS; OMIM #274600), an autosomal recessive disorder characterized by sensorineural deafness and goiter.[34] The deafness is congenital and associated with temporal bone abnormalities that range from isolated enlargement of the vestibular aqueduct (dilated vestibular aqueduct, DVA) to Mondini dysplasia, a more complex malformation that also includes cochlear hypoplasia, an anomaly in which the normal cochlear spiral of 2.5 turns is replaced by a smaller coil of 1.5 turns. Both DVA and Mondini dysplasia are easily recognized by computed tomography or magnetic resonance imaging.[35]

The thyromegaly in PS is the result of multinodular goitrous changes in the thyroid gland that develop in the teenage years,[36] although affected persons typically remain euthyroid with elevated serum thyroglobulin levels. The perchlorate discharge test is often abnormal. In this test, a person is given radiolabeled iodide and its localization to the thyroid is measured. Potassium perchlorate, a competitive inhibitor of iodide transport into the thyroid, then is administered. Normally, the amount of iodide in the thyroid will remain stable, reflecting rapid oxidation of iodide to iodine as it is incorporated into thyroglobulin. However, in a person with PS, iodide transport into the thyrocyte is delayed and so when perchlorate is

administered and blocks the sodium-iodide symporter, cytoplasmic iodide leaks back into the bloodstream. This back leakage is quantifiable as a change in thyroid radioactivity, with a positive result reflecting a drop in radioactivity of greater than 10%.[37]

In addition to PS, mutations in *SLC26A4* cause DFNB4 (OMIM #600791), a type of autosomal recessive nonsyndromic deafness in which, by definition, affected persons do not have thyromegaly.[38] No other physical abnormalities cosegregate with the deafness, although abnormal inner ear development, and in particular DVA, can be documented by temporal bone imaging. Together, DFNB4 and PS are estimated to account for 1% to 8% of congenital deafness.

Functional studies suggest that some of the observed differences between PS and DFNB4 are due to the degree of residual function of the encoded protein pendrin. While the function of pendrin is not fully determined, by homology it is thought to function in the transport of negatively charged particles (particularly chloride, iodide, and bicarbonate) across cell membranes. Mutations that abolish all transport function are more likely to be associated with the PS phenotype, while retained minimal transport ability appears to prevent thyroid dysfunction, although sensorineural deafness and temporal bone anomalies still occur, as in DFNB4.[39]

Many clinical studies have demonstrated intrafamilial variability, at times making the distinction between DFNB4 and PS difficult. The perchlorate discharge test is not a reliable test to resolve phenotypic ambiguities and is not consistently positive.[36] For example, in one family with two affected siblings, one child demonstrated the classic features of PS with severe-to-profound deafness, goiter, and a positive perchlorate discharge test, but the other child had only mild sensorineural deafness and no goiter.[40] In another study in which six individuals had confirmed PS, only three had a positive perchlorate washout of greater than10%.[41]

In a large clinical study of *SLC26A4* mutations in relation to temporal bone abnormalities, deafness-causing mutations were demonstrated in approximately 80% of multiplex families segregating DVA or Mondini dysplasia but in only 30% of simplex families.[42] These data suggest that allele variants of *SLC26A4* are a major genetic cause of these temporal bone abnormalities.

Since thyroid enlargement is an unreliable clinical indicator of disease and the perchlorate discharge test can be ambiguous, several investigators have recommended genetic testing of *SLC26A4* to establish a clinical diagnosis.[36,42] To date, 62 mutations have been reported in a total of 116 families.[43] Most of these mutations have been reported in only single families; however, 15 mutations are more common, and four (L236P, IVS8+1G→A, E384G, and T416P) account for approximately 60% of the total PS genetic load.[42] This broad spectrum of deafness-causing allele variants means that mutation screening of *SLC26A4* must include an analysis of all 20 protein-encoding exons (2–21) in addition to the splice donor site of exon 1.

Clinical Utility of Testing

Clinical data suggest that the major genetic cause of DVA and Mondini dysplasia is mutations in *SLC6A4*.[42] Because simplex cases include both genetic and nongenetic causes of DVA and Mondini dysplasia, a corollary is that most sporadic cases of DVA (~80%) and many sporadic cases of Mondini dysplasia (~40%) are not genetic and therefore are unlikely to recur in a family. This fact can be used to modify recurrence risks.

There is no concordance between specific *SLC26A4* allele variants and audiogram configuration,[42] although some mutations may be associated more frequently with specific temporal bone anomalies. For example, Masmoudi et al. studied two families segregating for L445W and found that while affected persons showed phenotypic variability with respect to thyroid disease, the temporal bone imaging revealed only DVA.[44]

Available Assays

A number of mutation detection strategies can be used to screen *SLC26A4* for nucleotide changes, including SSCP, HA, DHPLC, and direct sequencing. Because of its large size, high-throughput mutation screening of *SLC26A4* is challenging. The application of SSCP and HA is relatively insensitive and unnecessarily laborious, while direct sequencing is expensive. These constraints have made DHPLC an attractive alternative for mutation screening of *SLC26A4*.

Interpretation of Test Results

To compare the sensitivity of SSCP and DHPLC for detection of allele variants of *SLC26A4*, Prasad et al. screened a panel of 55 individuals segregating 41 different sequence-verified coding mutations. All 41 allele variants of *SLC26A4* were identified by DHPLC by their elution profile, for a detection rate of 100%. Nineteen mutations were detected at all three partial denaturing temperatures, ten mutations were detected at two of three temperatures, and 12 mutations were detected at only one temperature. Of the four common mutations, L236P and E384G showed discrete elution profiles at all three temperatures, but T416P and IVS8+1G→A were detected at only two and one temperatures, respectively. Mutations were tested multiple times from different samples to confirm test-retest reliability. Elution profiles of the four most common *SLC26A4* mutant alleles were distinct and easily could be differentiated from one another.

SSCP detected 26 (63%) of these same allele variants. Of the missed mutations, five (V138F, G209V, FS400, G672E, H723R) have been reported in more than one family. Three of the most common mutations (L236P, IVS8+1G→A, T416P) were detected by SSCP, although detection of L236P and IVS8+1G→A was not consistent.[45]

The finding that DHPLC sensitivity for detecting *SLC26A4* allele variants is 100% is similar to results reported by Taliani et al.[33] for DHPLC screen of *PROC*, a gene that encodes protein C, for variant identification. More than 200 different mutations in this gene are associated with an increased susceptibility to venous thromboembolism.[33] Other authors have reported the sensitivity of DHPLC to range from 95% to 100%.[30–32]

Laboratory Issues

To ensure a high level of accuracy and reliability and to optimize cost-effectiveness and turnaround time, a number of testing parameters must be considered. Amplicons should be maintained between 200 and 1,000 base pairs for accurate DHPLC analysis, intronic primers should be selected sufficiently far upstream and downstream of splice sites, and PCR conditions must be optimized for high-stringency exon amplification. Although the optimal acetonitrile gradient and partial denaturing temperature are determined using Wavemaker software, additional testing must be performed 2°C above and below the predicted temperature to detect all possible mutations.

While other investigators have been able to correlate DHPLC chromatogram profile with mutation type,[33] this correlation may be unreliable. The profile for a given *SLC26A4* allele variant differs from column to column and even in the same column, based on column life and buffer constitution. However, all heterozygote mutations are distinguished from wild-type samples. Because DHPLC does not distinguish homozygote allele variants due to the limitations inherent in HA, homozygotes can be detected by analysis following pooling of the unknown samples with sequence-verified wild-type DNA at a ratio of 2 : 1. By coupling additional automated instrumentation with DHPLC, high-throughput, accurate, reliable, and cost-effective mutation screening of persons with a Pendred syndrome/DNFB4 phenotype is possible.

Detection of only a single mutation is more common for testing of simplex families.[42,46] In three multiplex families segregating single mutations, Southern hybridization, and real-time PCR have failed to identify abnormalities in the "normal" allele, although all affected persons within a family had the same parental "normal" allele.

WOLFRAM SYNDROME AND DFNA6/14 (*WFS1*)

Molecular Basis of Disease

Mutations in *WFS1* cause autosomal dominant low-frequency sensorineural hearing impairment (LFSNHI) at the DFNA6/14 loci (OMIM #600965).[47,48] Deafness is bilaterally symmetrical and affects frequencies below 4,000 Hz; hearing is most impaired at the lowest frequencies, giving the DFNA6/14 audiogram an upsloping configuration.[49–52] DFNA1 also is characterized by LFSNHI, but in contrast to DFNA6/14 deafness, DFNA1 deafness is rapidly progressive and ultimately affects all frequencies.[53] Progression of DFNA6/14 deafness is minimal, although with aging the consequences of presbycusis result in flattening of the audiogram.

Mutations in *WFS1* also cause Wolfram syndrome, an autosomal recessive disease characterized by diabetes insipidus, diabetes mellitus, optic atrophy, and deafness, giving rise to the acronym DIDMOAD for this disease.[54,55] Minimal diagnostic criteria are diabetes mellitus and optic atrophy, and in addition to diabetes insipidus and sensorineural deafness, peripheral neuropathy, urinary-tract atony, and psychiatric illness can occur. Remarkably, the hearing loss in DIDMOAD syndrome is in the high frequencies.[56,57]

The observed phenotypic differences between DIDMOAD and DFNA6/14 appear to have a genotypic correlate. Sixty-five percent of persons with DIDMOAD carry inactivating mutations in *WFS1*, suggesting that loss of function of the encoded protein is causally related to Wolfram syndrome; most of these mutations lie in predicted transmembrane domains. In contrast, all disease-causing DFNA6/14 allele variants have missense mutations, and with one exception these amino acid changes are in the fifth intracellular domain of WFS1.[58,59] The protein lacks significant homology to published DNA or protein sequences, but secondary structure predictions suggest that it is has nine helical transmembrane segments. Biochemical studies suggest that wolframin is an endoglycosidase H-sensitive membrane glycoprotein predominantly located in the endoplasmic reticulum (ER).[60] Although the function of wolframin within the inner ear is unknown, it may play a role in the canalicular reticulum, a specialized ER that maintains intracellular ion homeostasis. Functional studies are necessary to test this hypothesis and to determine how different mutations in *WFS1* give rise to different phenotypes.

Clinical Utility of Testing

WFS1 contains eight exons (exon 1 is noncoding), encompasses 33.4 kb of genomic DNA, and transcribes a messenger RNA (mRNA) of 3.6 kb that encodes an 890 amino acid protein with a predicted molecular mass of 100 kilodaltons (kDa).[54,55] Mutations in *WFS1* are a major cause of LFSNHI in families demonstrating an autosomal dominant segregation pattern; however, mutation screening of *WFS1* in simplex cases is unlikely to identify abnormal allele variants.

Available Assays

A number of mutation detection strategies can be used to screen *WFS1* for nucleotide changes, including DHPLC and direct sequencing. A *WFS1* mutation database is available

to track the latest information on *WFS1* mutations in DIDMOAD and LFSNHI. This database lists all known mutations and polymorphisms with corresponding references.[61] Electronic submission of new mutations and polymorphisms is encouraged.

Interpretation of Test Results

Only noninactivating mutations are found in LFSNHI, and nearly all are located in the C-terminal region of the protein, suggesting a dominant-negative effect.[58] Although no clear-cut genotype-phenotype correlation has been drawn for DIDMOAD, homozygosity or compound heterozygosity for missense mutations is rarely found, and when it does occur, affected persons have a relatively mild phenotype that includes optic atrophy, diabetes mellitus, and sometimes hearing impairment, but excludes diabetes insipidus and other clinical findings. Persons carrying two deletions, insertions, nonsense mutations, or splice site mutations rarely have a mild phenotype.[59]

The disease spectrum in DIDMOAD has focused attention on the role of *WFS1* in diabetes mellitus and psychiatric diseases. The R456H, H611R, and I720V allele variants have been significantly correlated with type 1 diabetes mellitus in the Japanese, with a marginally significant association between R456H and type 2 diabetes mellitus.[62] Other case-control studies from the United Kingdom and Spain appear to confirm the involvement of *WFS1* in the pathogenesis of type 2 diabetes mellitus.[63,64] The exact role of *WFS1* in the etiology of diabetes is not known.

Laboratory Issues

Over 70 coding allele variants of *WFS1* have been reported, including synonymous and nonsynonymous changes.[61] Many of these changes have been found in persons with DIDMOAD and DFNA6/14, and cosegregate with known disease-causing mutations. Those that have been detected in controls are likely to be benign polymorphisms, but the effect of other mutations (K193Q, L432V, L499F, G576S, A559T, A671V, A684V, R708C, R818C, D866N, V871M) on disease phenotype is more difficult to judge, as they have been described both as disease-causing mutations and as polymorphisms.[59]

BRANCHIO-OTO-RENAL-SYNDROME (*EYA1*)

Molecular Basis of Disease

Mutations in *EYA1* cause branchio-oto-renal (BOR) syndrome (OMIM #113650).[65–69] Disease prevalence is estimated at 1 in 40,000 in the general population, and the syndrome is reported to occur in about 2% of profoundly deaf children.[70] Clinical expression is highly variable within and among families, but typical manifestations include branchial arch anomalies (preauricular pits, branchial fistulae, and pinnae abnormalities), hearing loss (conductive, sensorineural, or mixed), and renal hypoplasia.[71]

Phenotypic features that occur in more than 20% of affected persons are classified as major. Hearing loss and preauricular pits are most prevalent, affecting approximately 90% and 80% of individuals, respectively. Branchial cleft fistulae (~50%), lop-ear deformity (~35%), and stenotic external auditory canals (~30%) also are common. Temporal bone abnormalities can be identified in most individuals with hearing impairment examined by computed tomography, and renal anomalies are identified in approximately 65% of individuals examined by ultrasound or excretory urography.[71]

Hearing loss is mixed (~50%), conductive (~25%), or sensorineural (~25%), and ranges from mild to profound, but is most commonly severe (~35%), and is progressive in approximately 25% of affected persons. Temporal bone abnormalities include stenosis and atresia of the external auditory canal, malformation, malposition, dislocation or fixation of the ossicles, and reduction in size or malformation of the middle ear cavity. In the inner ear, the most common anomaly is cochlear hypoplasia. Enlargement of the cochlear and vestibular aqueducts and hypoplasia of the lateral semicircular canal also are found. Major renal anomalies include agenesis (most common), hypoplasia, and dysplasia. Calyceal diverticula, ureteral pelvic junction obstruction, hydronephrosis, pelviectasis, calyectasis, and vesico-ureteral reflux also are seen.[71]

In some families, anticipation may appear to be present; however, a study of seven three-generational families assessed for anticipation yielded conflicting results. In four of these families, the degree of hearing loss showed anticipation, but in the remaining three, a generational loss did not occur.[71] With respect to renal disease, generational progression was present in three families, but in one family, the reverse trend was found.[71] There is also no evidence for a parent-of-origin effect. Marked renal defects have been reported in six live-born offspring (including bilateral renal agenesis in three individuals) who had affected fathers and in 4 live-born offspring of affected mothers.[72–77] An excess of unexplained fetal deaths, presumably a consequence of bilateral renal agenesis, occurred in all of these families.

Clinical Utility of Testing

EYA1 contains 16 exons, encompasses 164.8 kb of genomic DNA, and transcribes an mRNA of 1.7 kb that encodes a 559 amino acid protein with a predicted molecular mass of 61.2 kDa. Mutations in *EYA1* are a major cause of hearing loss in families segregating a BOR phenotype; however, mutation screening of *EYA1* in simplex cases is unlikely to identify abnormal allele variants.[43]

Available Assays

Mutation screening of *EYA1* currently is available by SSCP analysis, with sequencing of exons in which band shifts are identified. An *EYA1* mutation database is available to track the latest information on *EYA1* mutations in BOR syndrome. This database lists all known mutations with corresponding references. Electronic submission of new mutations and polymorphisms is encouraged.[43]

Interpretation of Test Results

Initial reports on BOR syndrome described a patient with a complex cytogenetic rearrangement on chromosome 8q[78,79] and another patient with 8q12.2-q21 deletion,[80] indicating that complex genomic rearrangements can cause this phenotype. In the first two mutation analyses of *EYA1*, three large deletions are described.[65,66] Current estimates suggest that about one fifth of *EYA1* mutations resulting in BOR syndrome represent complex genomic rearrangements, perhaps indicating that the *EYA1* region is unstable.[81] These rearrangements cannot be detected by commonly used mutation screening procedures, making it impossible to use a single approach for mutation screening of *EYA1*.

Laboratory Issues

Although numerous reports have confirmed that alterations in *EYA1* cause BOR syndrome,[66–70] in many of these reports only 20% of affected patients have confirmed *EYA1* mutations.[81] To explain this low mutation detection rate, some investigators have hypothesized that mutations in another gene tightly linked to *EYA1* also cause BOR syndrome. However, using Southern analysis, other investigators have been able to detect mutations in families linked to chromosome 8, in which SSCP analysis and direct sequencing fail to detect *EYA1* mutations.[81] This finding underscores the major drawback of many PCR-based mutation detection procedures in the study of an autosomal dominant disorder, the inevitable amplification of the normal allele that can mask deleted or disrupted exons.

Conclusion

With continued advances in our understanding of the molecular biology of hearing and deafness, the clinical management of deafness will become more sophisticated. Molecular genetic testing will be used to unequivocally diagnose many different forms of inherited deafness, providing prognostic information for patients and their families. Novel habilitation options will be developed that will be applicable to select persons with specific types of genetic deafness. At the forefront of these advances will be the ability to offer precise genetic testing that is rapid, robust, and cost-effective.

Acknowledgments

This work was supported in part by grants DC02842 and DC03544 (RJHS).

References

1. Smith RJH, Green GE, Van Camp G. Hereditary hearing loss and deafness [GeneClinics Web site]. 2003. Available at: http://www.geneclinics.org/.
2. Green GE, Scott DA, McDonald JM, et al. Carrier rates in the midwestern United States for GJB2 mutations causing inherited deafness. *JAMA*. 1999;281:2211–2216.
3. Van Camp G, Smith RJH. [Hereditary Hearing Loss home page]. 2003. Available at: http://webhost.ua.ac.be/hhh/.
4. Guilford P, Ben Arab S, Blanchard S, et al. A non-syndromic form of neurosensory, recessive deafness maps to the pericentromeric region of chromosome 13q. *Nat Genet.* 1994;6:24–28.
5. Kelsell DP, Dunlop J, Stevens HP, et al. Connexin 26 mutations in hereditary non-syndromic sensorineural deafness. *Nature.* 1997;387:80–83.
6. Bruzzone R, White TW, Paul DL. Connections with connexins: the molecular basis of direct intercellular signaling. *Eur J Biochem.* 1996;238:1–27.
7. Zelante L, Gasparini P, Estivill X, et al. Connexin 26 mutations associated with the most common form of non-syndromic neurosensory autosomal recessive deafness (DFNB1) in Mediterraneans. *Hum Mol Genet.* 1997;6:1605–1609.
8. Denoyelle F, Weil D, Maw MA, et al. Prelingual deafness: high prevalence of a 30delG mutation in the connexin 26 gene. *Hum Mol Genet.* 1997;6:2173–2177.
9. Estivill X, Fortina P, Surrey S, et al. Connexin-26 mutations in sporadic and inherited sensorineural deafness. *Lancet.* 1998;351:394–398.
10. Scott DA, Kraft ML, Carmi R, et al. Identification of mutations in the connexin 26 gene that cause autosomal recessive nonsyndromic hearing loss. *Hum Mutat.* 1998;11:387–394.
11. Cohn ES, Kelley PM, Fowler TW, et al. Clinical studies of families with hearing loss attributable to mutations in the connexin 26 gene. *Pediatrics.* 1999:103;546–550.
12. Kikuchi T, Adams JC, Paul DL, et al. Gap junction systems in the rat vestibular labyrinth: immunohistochemical and ultrastructural analysis. *Acta Otolaryngol.* 1994;114:520–528.
13. Kikuchi T, Kimura RS, Paul DL, et al. Gap junctions in the rat cochlea: immunohistochemical and ultrastructural analysis. *Anat Embryol (Berl).* 1995;191:101–118.
14. Tekin M, Arnos KS, Pandya A. Advances in hereditary deafness. *Lancet.* 2001;358:1082–1090.
15. Estivill X, Gasparini P [Connexin-deafness homepage]. 2003. Available at: http://davinci.crg.es/deafness/.
16. Van Laer L, Coucke P, Mueller RF, et al. A common founder for the 35delG connexin 26 (*GJB2*) gene mutation in non-syndromic hearing impairment. *J Med Genet.* 2001;38:515–518.
17. Kelley PM, Harris DJ, Comer BC, et al. Novel mutations in the connexin 26 gene (GJB2) that cause autosomal recessive (DFNB1) hearing loss. *Am J Hum Genet.* 1998;62:792–799.
18. Morell RJ, Kim HJ, Hood LJ, et al. Mutations in the connexin 26 gene (GJB2) among Ashkenzi Jews with nonsyndromic recessive deafness. *N Eng J Med.* 1998;339:1500–1505.
19. Kudo T, Ikeda K, Kure S, et al. Novel mutations in the connexin 26 gene (GJB2) responsible for childhood deafness in the Japanese population. *Am J Med Genet.* 2000;90:141–145.

20. Denoyelle F, Marlin S, Weil D, et al. Clinical features of the prevalent form of childhood deafness, DFNB1, due to a connexin-26 gene defect: implications for genetic counseling. *Lancet.* 1999;353:1298–1303.
21. Green GE, Mueller RF, Cohn ES, Avraham KB, Kanaan M, Smith RJH. Audiological manifestations and features of Connexin 26 deafness. *Audiolog Med* 2003;1:5–11.
22. Kenna MA, Wu BL, Cotanche DA, et al. Connexin 26 studies in patients with sensorineural hearing loss. *Arch Otolaryngol Head Neck Surg.* 2001;127:1037–1042.
23. Hohl D. Towards a better classification of erythrokeratodermias. *Br J Dermatol.* 2000;143:1133–1137.
24. Maestrini E, Korge BP, Ocana-Sierra J, et al. A missense mutation in connexin26, D66H, causes mutilating keratoderma with sensorineural deafness (Vohwinkel's syndrome) in three unrelated families. *Hum Mol Genet.* 1999;8:1237–1243.
25. Fukushima K, Sugata K, Kasai N, et al. Better speech performance in cochlear implant patients with GJB2-related deafness. *Int J Pediatr Otorhinolaryngol.* 2002;62:151–157.
26. Green GE, Scott DA, McDonald JM, et al. Performance of cochlear implant recipients with GJB2-related deafness. *Am J Med Genet.* 2002;109:167–170.
27. Smith RJ. Mutation screening for deafness: more than simply another diagnostic test. *Arch Otolaryngol Head Neck Surg.* 2001;127: 941–942.
28. Mueller RF, Nehammer A, Middleton A, et al. Congenital non-syndromal sensorineural hearing impairment due to connexin 26 gene mutations—molecular and audiological findings. *Int J Pediatr Otorhinolaryngol.* 1999;50:3–13.
29. Sobe T, Vreugde S, Shahin H, et al. The prevalence and expression of inherited connexin 26 mutations associated with nonsyndromic hearing loss in the Israeli population. *Hum Genet.* 2000;106:50–57.
30. Ellis LA, Taylor CF, Taylor GR. A comparison of fluorescent SSCP and denaturing HPLC for high throughput mutations screening. *Hum Mutat.* 2000;15:556–564.
31. Liu WG, Smith DI, Rechtzigel KJ, et al. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. *Nucleic Acids Res.* 1998;26:1396–1400.
32. O'Donovan MC, Oefner PJ, Roberts SC, et al. Blind analysis of denaturing high-performance liquid chromatography as tool for mutation detection. *Genomics.* 1998;52:44–49.
33. Taliani MR, Roberts SC, Dukek BA, et al. Sensitivity and specificity of denaturing high-pressure liquid chromatography for unknown protein C gene mutations. *Genet Test.* 2001;5:39–44.
34. Everett LA, Glaser B, Beck JC, et al. Pendred syndrome is caused by mutations in a putative sulphate transporter gene (PDS). *Nat Genet.* 1997;17:411–422.
35. Phelps PD, Coffey RA, Trembath RC, et al. Radiological malformations of the ear in Pendred syndrome. *Clin Radiol.* 1998;53:268–273.
36. Reardon W, Coffey R, Chowdhury T, et al. Prevalence, age of onset, and natural history of thyroid disease in Pendred syndrome. *J Med Genet.* 1999;36:595–598.
37. Morgans ME, Trotter WR. Association of congenital deafness with goitre: the nature of the thyroid defect. *Lancet.* 1958;1:607–609.
38. Li XC, Everett LA, Lalwani AK, et al. A mutation in *PDS* causes non-syndromic recessive deafness. *Nat Genet.* 1998;18:215–217.
39. Scott DA, Wang R, Kreman TM, et al. Functional differences of the PDS gene product are associated with phenotypic variation in patients with Pendred syndrome and nonsyndromic hearing loss. *Hum Mol Genet.* 2000;9:1709–1715.
40. Johnsen T, Sorensen MS, Feldt-Rasmussen U, et al. The variable intrafamiliar expressivity in Pendred's syndrome. *Clin Otolaryngol.* 1989;14:395–399.
41. Yong AML, Goh SS, Zhao Y, et al. Two Chinese families with Pendred's syndrome—radiological imaging of the ear and molecular analysis of the pendrin gene. *J Clin Endocrin Metab.* 2001;86:3907–3911.
42. Campbell C, Cucci RA, Green GE, et al. Pendred syndrome, DFNB4 and *PDS*—Identification of eight novel mutations and phenotype-genotype correlations. *Hum Mut.* 2001;17:403–411.
43. Chang E, Kolln K, Smith RJH. [Pendred syndrome/BOR homepage]. 2003. Available at: http://www.medicine.uiowa.edu/pendredandbor/.
44. Masmoudi S, Charfedine I, Hmani M, et al. Pendred syndrome: phenotypic variability in two families carrying the same PDS missense mutation. *Am J Med Genet.* 2000;90:38–44.
45. Prasad S, Kölln KA, Cucci RA, Trembath RC, Van Camp G, Smith RJH. Pendred syndrome and DFNB4—Mutation screening of SLC26A4 by denaturing high-performance liquid chromatography and the identification of seven novel mutations. Am J Med Genet 124A:1–9, 2004.
46. Coyle B, Reardon W, Herbrick J, et al. Molecular analysis of the PDS gene in Pendred syndrome (sensorineural hearing loss and goitre). *Hum Mol Genet.* 1998;7:1105–1112.
47. Bespalova IN, Van Camp G, Bom SJ, et al. Mutations in the Wolfram syndrome 1 gene (*WFS1*) area a common cause of low frequency sensorineural hearing loss. *Hum Mol Genet.* 2001;10:2501–2508.
48. Young TL, Ives E, Lynch E, et al. Non-syndromic progressive hearing loss DFNA38 is caused by heterozygous missense mutation in the Wolfram syndrome gene *WFS1*. *Hum Mol Genet.* 2001;10:2509–2514.
49. Kunst HPM, Marres HAM, Huygen PLM, et al. Autosomal dominant non-syndromal low-frequency sensorineural hearing impairment linked to chromosome 4p16 (DFNA14): statistical analysis of hearing threshold in relation to age and evaluation of vestibulo-ocular functions. *Audiology.* 1999;38:165–173.
50. Brodwolf S, Böddeker IR, Ziegler A, et al. Further evidence for linkage of low-mid frequency hearing impairment to the candidate region on chromosome 4p16.3. *Clin Genet.* 2001;60:155–160.
51. Huygen PLM, Bom SJ, Van Camp G, et al. The clinical presentation of the DFNA loci where causative genes have not yet been cloned: DFNA4, DFNA6/14, DFNA7, DFNA16, DFNA20 and DFNA21. In: Cremers CWRJ, Smith RJH, eds. *Advances in Otorhinolaryngology*. Basel: Karger; 2002;98-106.
52. Bom SJH, Van Camp G, Cryns K, et al. Autosomal dominant low-frequency hearing impairment (DFNA6/14): a clinical and genetic family study. *Otol Neurotol.* 2002:23:876–884.
53. Lynch ED, Lee MK, Morrow JE, et al. Nonsyndromic deafness DFNA1 associated with mutation of a human homolog of the Drosophila gene diaphanous. *Science.* 1997;278:1223–1224.
54. Inoue H, Tanizawa Y, Wasson J, et al. A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome). *Nat Genet.* 1998;20:143–148.
55. Strom TM, Hörtnagel K, Hofmann S, et al. Diabetes insipidus, diabetes mellitus, optic atrophy and deafness (DIDMOAD) caused by mutations in a novel gene (wolframin) coding for a predicted transmembrane protein. *Hum Mol Genet.* 1998;7:2021–2028.
56. Cremers CWRJ, Wijdeveld PG, Pinckers AJ. Juvenile diabetes mellitus, optic atrophy, hearing loss, diabetes insipidus, atonia of the urinary tract and bladder, and other abnormalities (Wolfram syndrome): a review of 88 cases from the literature with personal observations on 3 new patients. *Acta Paediatr Scand Suppl.* 1977;264:1–16.
57. Higashi K. Otologic findings of DIDMOAD syndrome. *Am J Otol.* 1991;12:57–60.
58. Cryns K, Pfister M, Pennings RJE, et al. Mutations in the WFS1 gene that cause low frequency sensorineural hearing loss are small non-inactivating mutations. *Hum Genet.* 2002;110:389–394.
59. Prasad S, Kölln KA, Cucci RA, Trembath RC, Van Camp G, Smith RJH. Pendred syndrome and DFNB4-Mutation screening of SLC26A4 by denaturing high-performance liquid chromatography and the identification of seven novel mutations. *Am J Med Genet*. 2004;124A:1–9.
60. Takeda K, Inoue H, Tanizawa Y, et al. WFS1 (Wolfram syndrome 1) gene product: predominant subcellular localization to endoplasmic reticulum in cultured cells and neuronal expression in rat brain. *Hum Mol Genet.* 2001;10:477–484.
61. Sivakumaran TA, Lesperance MM. WFS1 Gene Mutation and Polymorphism Database [database online]. 2003. Available at: http://www.khri.med.umich.edu/research/lesperance_lab/lfsnhl.shtml.
62. Awata T, Inoue K, Kurihara S, et al. Missense variations of the gene responsible for Wolfram syndrome (WFS1/wolframin) in Japanese: possible contribution of the Arg456His mutation to type 1 diabetes

as a nonautoimmune genetic basis. *Biochem Biophys Res Commun.* 2000;268:612–616.
63. Minton JAL, Hattersley AT, Owen K, et al. Association studies of genetic variation in the WFS1 gene and type 2 diabetes in U.K. populations. *Diabetes.* 2002;51:1287–1290.
64. Domenech E, Gomez-Zaera M, Nunes V. WFS1 mutations in Spanish patients with diabetes mellitus and deafness. *Eur J Hum Genet.* 2002;10:421–426.
65. Abdelhak S, Kalatzis V, Heilig R, et al. Protein, nucleotide, a human homologue of the Drosophila eyes absent gene underlies branchio-oto-renal (BOR) syndrome and identifies a novel gene family. *Nat Genet.* 1997;15:157–164.
66. Abdelhak S, Kalatzis V, Heilig R, et al. Protein, nucleotide, clustering of mutations responsible for branchio-oto-renal (BOR) syndrome in the eyes absent homologous region (eyaHR) of *EYA1*. *Hum Mol Genet.* 1997;6:2247–2255.
67. Kumar S, Kimberling WJ, Weston MD, et al. Identification of three novel mutations in human *EYA1* protein associated with branchio-oto-renal syndrome. *Hum Mutat.* 1998;11:443–449.
68. Kumar S, Deffenbacher K, Cremers CW, et al. Branchio-oto-renal syndrome: identification of novel mutations, molecular characterization, mutation distribution, and prospects for genetic testing. *Genet Test.* 1998;1:243–251.
69. Usami S, Abe S, Shinkawa H, et al. *EYA1* nonsense mutation in a Japanese branchio-oto-renal syndrome family. *J Hum Genet.* 1999;44:261–265.
70. Fraser FC, Sproule JR, Halal F. Frequency of the branchio-oto-renal (BOR) syndrome in children with profound hearing loss. *Am J Med Genet.* 1980;7:341–349.
71. Chen A, Francis M, Ni L, et al. Phenotypic manifestations of branchio-oto-renal syndrome. *Am J Med Genet.* 1995;58:365–370.
72. Carmi R, Binshtock M, Abeliovich D. The branchio-oto-renal (BOR) syndrome: report of bilateral renal agenesis in three sibs. *Am J Med Genet.* 1983;14:625–627.
73. Cremers CWRJ, Fikkers-Van Noord M. The earpits-deafness syndrome: clinical and genetic aspects. *Int J Pediatr Otorhinolaryngol.* 1980;2:309–322.
74. Greenberg CR, Trevenen CL, Evans JA. The BOR syndrome and renal agenesis. *Prenatal Diagn.* 1988;8:103–108.
75. Van Widdershoven J, Monnens L, Assmann K, et al. Renal disorders in the branchio-oto-renal syndrome. *Helv Paediatr Acta.* 1983;38: 513–522.
76. Chitayat D, Hodgkinson KA, Chen MF, et al. Branchio-oto-renal syndrome: further delineation of an underdiagnosed syndrome. *Am J Med Genet.* 1992;43:970–975.
77. Fitch N, Sorolovitz H. Severe renal dysgenesis produced by a dominant gene. *Am J Dis Child.* 1976;130:1356–1357.
78. Gu JZ, Wagner MJ, Haan EA, et al. Detection of a megabase deletion in a patient with branchio-oto-renal syndrome (BOR) and tricho-rhino-phalangeal syndrome (TRPS): implications for mapping and cloning of the BOR gene. *Genomics.* 1996;31:201–206.
79. Haan EA, Hull YJ, White S, et al. Tricho-rhino-phalangeal and branchio-oto syndromes in a family with an inherited rearrangement of chromosome 8q. *Am J Med Genet.* 1989;32:490–494.
80. Vincent C, Kalatzis V, Compain S, et al. A proposed new contiguous gene syndrome on 8q consists of branchio-oto-renal (BOR) syndrome, Duane syndrome, a dominant form of hydrocephalus and trapeze aplasia: implications for the mapping of the BOR gene. *Hum Mol Genet.* 1994;3:1859–1866.
81. Vervoort V, Smith RJH, O'Brien J, et al. Genomic rearrangements of *EYA1* account for a large fraction of families with BOR syndrome. *Eur J Hum Genet.* 2002;10:757–766.

Chapter 12

Coagulation Disorders

Rajiv K. Pruthi and John A. Heit

Introduction

Reduced to its simplest form, the vascular system is a series of interconnected biologic conduits filled with fluid moving under pressure. This system is subject to injury and must repair itself in order to avoid exsanguination. This latter function is provided by the hemostatic system. Thus, normal hemostasis is a reparative process and consists of three major mechanisms: (1) vasospasm, which is the main mechanism for controlling bleeding after transection/avulsion injuries of large arteries, arterioles, and veins; (2) platelet plug formation; and (3) the procoagulant system. The procoagulant system consists of the mechanisms by which fluid blood is converted into an insoluble hemostatic thrombus. While inherited or acquired disorders of all three mechanisms can cause clinically abnormal bleeding, this chapter reviews the molecular basis of inherited disorders of the procoagulant system (e.g., hemophilia A and B, von Willebrand disease). Inherited vascular and platelet disorders associated with abnormal bleeding are not discussed here.

Consider also that, once activated, the reparative (e.g., hemostatic) response to injury must be constrained such that the thrombus is limited to the site of injury, and must be downregulated such that the conversion of fluid blood to solid thrombus does not propagate throughout the lumen of the vessel. This coagulation-limiting function is provided by the anticoagulant system. Inherited disorders of the anticoagulant system (e.g., familial thrombophilia) are associated with venous thromboembolism and complications of pregnancy. Although biologically plausible, the association between familial thrombophilia and arterial thrombosis is unproven and is not reviewed here.

INHERITED BLEEDING DISORDERS

The initial reparative response to endothelial injury is platelet adhesion to the exposed subendothelial matrix. Platelets circulate preferentially along the vessel luminal surface, as opposed to circulating evenly distributed throughout the blood. Thus, platelets can be viewed as "surveillance" blood cells that are "searching" for sites of endothelial injury. Platelet adhesion at an injury site is mediated by the ligand von Willebrand factor (VWF). VWF tethers platelets to the site of injury by binding to collagen within the subendothelial matrix exposed by the injury, and to the constitutively active platelet receptor glycoprotein (GP)-Ib-IX-V complex. Firm platelet adhesion at the injury site is mediated by subsequent binding of the platelet GP-$\alpha_2\beta_1$ collagen receptor to subendothelial matrix collagen. VWF binding to platelet GP-Ib-IX-V generates transmembrane signals that activate platelets, converting the constitutively inactive platelet fibrinogen receptor GP-$\alpha_{IIb}\beta_3$ (GP-IIb/IIIa) to the fibrinogen-binding conformation. Subsequent platelet-to-platelet cohesion is mediated by fibrinogen or VWF binding to platelet GP-$\alpha_{IIb}\beta_3$. In addition, a scramblase within the platelet membrane phospholipid bilayer flips negatively charged phospholipids from the inside to the outside bilayer leaflet of the membrane. The negatively charged phospholipids support assembly of the plasma procoagulant factors on the external surface of the platelet membrane. Finally, the contents of the platelet dense granules (ADP, serotonin, etc.) and α-granules (fibrinogen, VWF, factors V [F5] and XI [F9], etc.) are released, which feed back to activate additional platelets and support the plasma procoagulant system.

The plasma procoagulant system has been characterized as a "cascade" of amplifying enzymatic reactions leading to activation of the final serine protease enzyme, thrombin. This cascade is initiated by exposure of the coagulation activator, tissue factor (TF), to circulating blood. TF is located within the wall of blood vessels and normally is sequestered from the circulation. TF is exposed to blood after blood vessel injury. Exposed (or expressed) TF binds to circulating factor VII_a (F7a) to form a "Factor X-ase" activation complex, which either cleaves ("activates") factor X (F10) to factor Xa (F10a) directly, or activates factor IX (F9) to factor IXa (F9a) (Figure 12-1). F9a binds factor VIIIa (F8a) to form a second Factor X-ase activation complex. F10a binds F5a

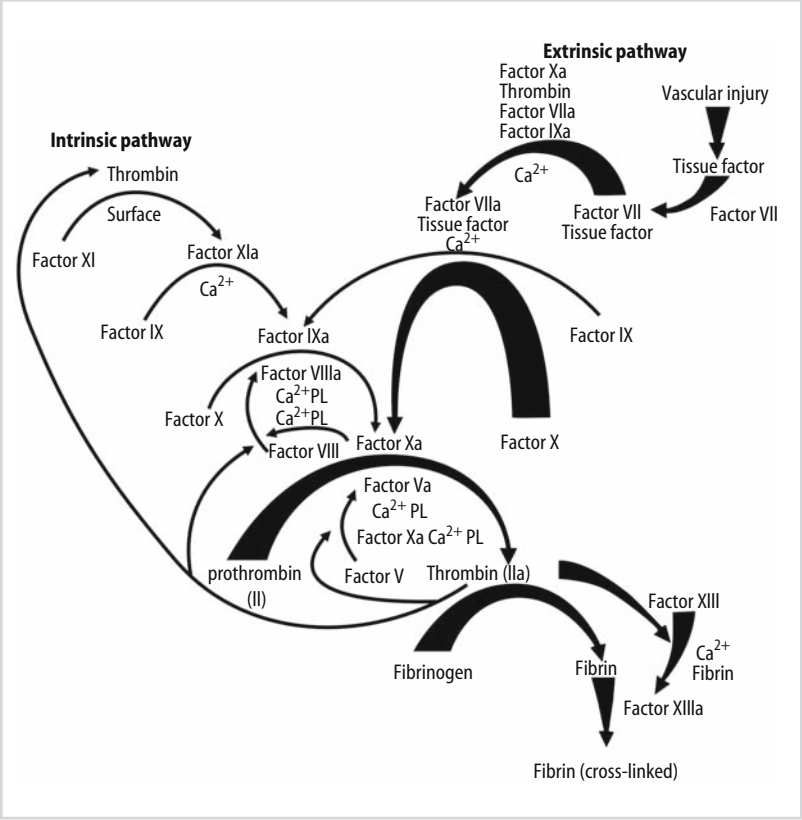


Figure 12-1. The coagulation cascade. a, activated factor; Ca^{2+}, calcium; PL, phospholipid. (Reprinted with permission from Davie EW, Fujikawa K, Kisiel W. The coagulation cascade: initiation, maintenance and regulation. *Biochemistry.* 1991;30:10363. Copyright 1991 American Chemical Society.)

to form the "prothrombinase complex" that activates prothrombin (F2) to thrombin (F2a). The Factor X-ase and the prothrombinase complexes assemble on negatively charged phospholipids within the outer membrane of activated platelets. Thrombin produces a hemostatic thrombus by cleaving fibrinogen (F1) to form fibrin monomers, by activating platelets, and by activating factor XIII (F13) to factor XIIIa (F13a), which cross-links strands of fibrin monomers to form an insoluble hemostatic thrombus. In a feedback amplification loop, thrombin also increases its own production by activating F5, F8, and factor XI (F11).

HEMOPHILIA A

Hemophilia A (HA) is an X-linked recessive bleeding disorder due to a deficiency in F8. HA affects approximately 1 in 10,000 live male births, among all ethnic populations. Approximately 30% of HA cases result from new mutations occurring in families with no apparent family history of HA. The diagnosis of HA is established based on reduced or absent F8 activity and is classified into severe (<1% F8 activity), moderate (1% to 5%), or mild (>5% to 40%) disease; the estimated prevalence of each is 43%, 26%, and 31%, respectively.[1]

Molecular Basis of Disease

The *F8* gene is located on the long arm of the X chromosome (Xq28), is 186 kilobases (kb) in length, and has 26 relatively short exons, ranging from 69 to 262 base pairs (bp), and two long exons, exon 14 (3106 bp) and exon 26 (958 bp) (Figure 12-2). The resulting messenger RNA (mRNA) is approximately 9 kb, of which the coding sequence is 7053 nucleotides. The intron/exon boundaries roughly correlate with the F8 protein domains. The introns are large (14 to 23 kb), with intron 22 being that largest (32 kb). A CpG island in intron 22 acts as a bidirectional promoter for two additional genes. The first, termed *F8*-associated gene A (*F8A* or *INT22H1*), is an intronless gene approximately 2 kb long within intron 22 of *F8* and is transcribed in the opposite direction from *F8*. The second, *F8*-associated gene B (*F8B*), is 2.5 kb long and transcribed in the same direction as *F8*. The *F8A* and *F8B* transcripts originate within 122 bases of each other, and the functions of their potential protein products are unknown. The *F8A* sequence is replicated at least twice, approximately 500 kb telomeric to the *F8* gene and close to the tip of the X chromosome, termed *INT22H2* (proximal) and *INT22H3* (distal). *INT22H2* and *INT22H3* are about 100 kb apart and transcribed in the same direction as the *F8* gene. These three homologous repeats are involved in the intron 22 inversion mutation, which is a frequent rearrangement of the *F8* gene resulting in severe HA. F8 is predominantly expressed in the liver. The putative promoter region is located 300 nucleotides 5′ of the gene, and although a TATA box is not essential for transcription, liver-enriched transcription factors (e.g., HNF1, NFkB, C/EBPa, and C/EBPb) interact with the *F8* promoter region. The *F8* gene encodes a precursor protein of 2351 amino acids consisting of a 19 amino acid leader peptide followed by 2332 amino acids in the mature protein. The mature protein has

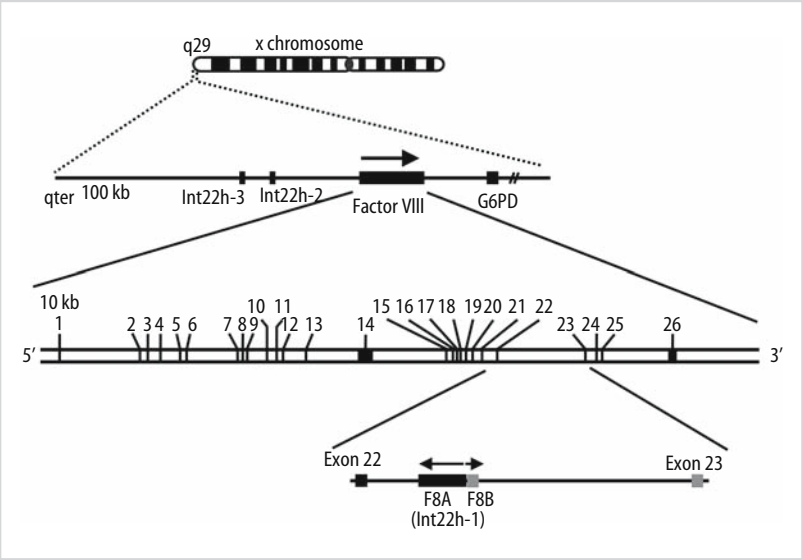


Figure 12-2. The factor VIII gene. G6PD, glucose 6 phosphate dehydrogenase; kb, kilobases. (Reprinted with permission from Kazazian HH, Tuddenham EGD, Stylianos EA. Hemophilia A: deficiency of coagulation factor VIII. In: Scriver CR, Beaudet AL, Valle D, et al., eds. *The Metabolic and Molecular Bases of Inherited Disease*. Copyright 2001 McGraw-Hill.)

several homologous domains termed A1, A2, B, A3, C1, and C2.

Proteolytic cleavage, by F2a or F10a, in the presence of phospholipid surfaces results in activation of F8 to F8a (Figure 12-3). Although cleavages at 740 or 1721 have no effect on coagulant activity, the cleavages at 372 or 1689 are important for F8 procoagulant activity. Cleavage at 1689 releases F8 from VWF, permitting F8 interaction with phospholipids and platelets. Missense mutations affecting these cleavage sites have been found in patients with HA and result in cross-reacting material positive (CRM+)-HA with normal levels of F8 antigen but low activity (1% to 7%). Although missense mutations have been found in the A2 domain in HA, and the importance of the A2 domain to F8 coagulant activity has been confirmed by in vitro studies, the exact role of the A2 subunit remains unknown.

The B domain is cleaved during proteolytic activation. Since B-domain-deleted F8 molecules are expressed at 5- to 10-fold higher levels than non-B-domain-deleted F8, the B domain may have a role in intracellular processing or secretion of F8 or both. Mutations in this region have been reported in HA.

VWF-bound F8 is protected from inactivation by activated protein C (APC). The putative VWF binding region of F8 is thought to be at the N-terminus of the light chain of F8 and in the C2 domain. The binding site for F9a has been localized to the A2 domain and regions of the light chain. In addition, the binding site for F10 is localized to the C-terminus of the A1 domain. Binding to phospholipids, which are important for F10 activation by F9a and F8a, occurs in the C1 and C2 domains of the light chain F8. No deleterious mutations in HA have been identified in the inactivation cleavage sites of F8.

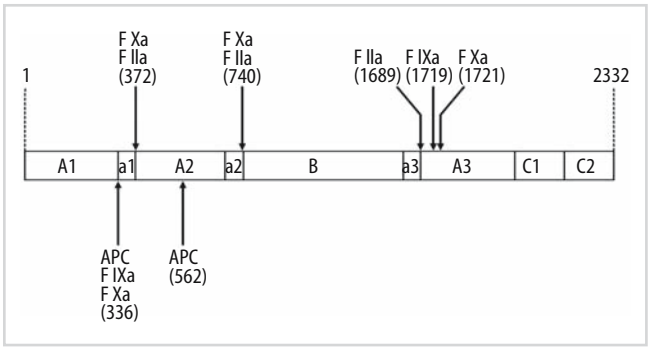


Figure 12-3. F8 peptide showing domains (A1 to C1), cleavage sites (arrows) for thrombin (FIIa), activated factor IX (FIXa), activated factor X (FXa), and activated protein C (APC). Amino acids are numbered in parentheses. The heavy chain cleavage site (A1-a1-A2-a2) is linked to the light chain cleavage site (a3-A3-C1-C2) via the B domain. (Reprinted with permission from Pruthi RK, Nichols WL. Autoimmune factor VIII inhibitors. *Curr Opin Hematol* 6:314. Copyright 1999 Lippincott Williams and Wilkins.)

Mutations in the *F8* Gene

Reported mutations and polymorphisms are cataloged in an international database, the Haemophilia A Mutation, Structure, Test and Resource Site, available on the Internet and updated periodically: http://europium.csc.mrc.ac.uk/WebPages/Main/main.htm. References for the amino acid numbering system used below can be accessed at this Web site.

Currently, no mutations have been found in the *F8* promoter region. Approximately 40% of HA patients with severe disease have an inversion at the tip of the X chromosome that disrupts the *F8* gene.[2] Homologous recombination can occur when the *F8A* gene (*INT22H1*) in intron 22 recombines with one of the two homologous regions (*INT22H2* or *INT22H3*) telomeric to the *F8* gene probably as a result of folding over of the tip of the X chromosome. Upon unfolding, exons 1 to 22 are inverted and placed about 500 kb upstream of exons 23 to 26 and oriented in the opposite direction. Depending on which repeat *F8A* pairs with, the inversion may be termed type I (distal) or type II (proximal). Rarely type III inversion mutations occur in patients with a third extragenic copy of *F8A*. Given that the majority of inversions originate in male meiosis, almost all

mothers of patients with the inversion mutation are carriers. An additional inversion of exon 1 of the *F8* gene also occurs and affects up to 5% of patients with severe HA.[3]

Given the large size of the *F8* gene, deletions are common and account for about 5% of characterized mutations. Typically these result in severe disease with <1% F8 activity. However, specific deletions of 156 bp in exon 22 or 294 bp in exons 23 and 24 are associated with moderate disease, likely due to in-frame splicing of exon 21 to exon 23, or exon 22 to exon 25. Patients with large deletions are susceptible to formation of F8 inhibitor (antibodies) in response to therapy with F8 concentrates. In an analysis of the HA database, up to 40% of patients with deletions develop F8 inhibitors, whereas up to 60% of patients with single base pair changes resulting in nonsense mutations and 15% of patients with single base pair changes resulting in missense mutations develop inhibitors.[4]

The remaining patients typically have single base pair changes (resulting in missense, frameshift, or splice junction mutations), insertions, or duplications. Single base pair changes that result in missense mutations are spread throughout the *F8* gene. Although the structure-function relationships of some of the missense mutations are known or can be deduced (e.g., alteration of the VWF binding site or thrombin cleavage site), the structural consequences of most missense mutations remain undefined.

Although HA predominantly affects males, some female carriers have reduced F8 levels and may have clinically significant bleeding. The molecular basis of symptomatic females includes lyonization of the normal X chromosome and Turner syndrome (XO karyotype), where the dominant mutant *F8* gene is responsible for production of the abnormal F8. Rarely, mating of a carrier female with an affected male, an X-autosome translocation involving a breakpoint within the *F8* gene, and uniparental isodisomy have been implicated.

Polymorphisms in the *F8* Gene

Polymorphisms present within the *F8* gene (intragenic) or outside the *F8* gene (extragenic) have been used to assign haplotypes (combinations of polymorphisms) for linkage analysis. The putative defective *F8* gene can be tracked with polymorphisms that are closely linked to the gene. The carrier frequency of such polymorphisms varies depending on the ethnicity of the study population and needs to be considered when studying patients of diverse ethnic origin. A complete listing of polymorphisms is available in the *F8* mutation database (http://europium.csc.mrc.ac.uk/WebPages/Main/main.htm).

Clinical Utility of Testing

The diagnosis of HA is established by assaying plasma F8 activity. Given that F8 normally relies on binding to VWF for survival, all patients with a low F8 activity should be excluded for a diagnosis of Von Willebrand disease (VWD). In addition, F8 is a labile factor resulting in a loss of up to 15% F8 activity in suboptimally processed plasma specimens. Thus, mild reductions in F8 activity should prompt repeat testing with careful attention to specimen handling. While a reduced F8 activity in at-risk females typically confirms carrier status, a normal F8 activity does not exclude carrier status. In such cases, molecular genetic testing is the only option for diagnosis.

Knowledge of the causative *F8* mutation in probands does not alter clinical management but may be useful in predicting the risk of developing F8 inhibitors.[4] Mutation identification also is useful for focusing on the relevant region of the *F8* gene in carrier testing of family members. Rarely, adopted asymptomatic females with no access to their familial *F8* genotype will need their entire *F8* gene analyzed for carrier testing. For carriers, prenatal diagnosis during pregnancy provides useful information for management of labor and delivery and, occasionally, pregnancy termination. Advances in reproductive technology may permit preimplantation genetic diagnosis and circumvent the need for prenatal testing.[5]

Available Assays

Several laboratories offer genetic testing for HA. A complete list of laboratories that have registered with GeneTests can be found on that Web site.[6] In general, testing can be divided into direct and indirect testing, and selected laboratories offer prenatal testing.

Direct DNA Analysis

Testing for the inversion mutation should be restricted to patients with severe HA (F8 activity <1%). For the proband with severe HA or for an at-risk female with a family history of severe HA, direct DNA analysis for the most common intron 22 inversion mutation is typically performed by restriction digestion and Southern blot analysis. This detects the more common types I and II inversions and the rare type III inversion. A polymerase chain reaction (PCR) assay has also been described;[7] however, this assay may not detect the rare atypical inversions. For the remaining patients with severe HA, testing for the intron 1 inversion mutation should be considered.[3] Detection of deletions in the *F8* gene in carriers by PCR amplification of individual exons is not useful. Thus, either linkage analysis or Southern blot analysis is required.

For the rest of the patients with severe HA, at-risk females, or those patients with mild or moderate HA, advances in molecular techniques have permitted efficient screening or direct sequencing of all the relevant regions of the *F8* gene.[8] The large size of the *F8* gene, with 26 exons, makes sequence analysis of the entire gene a labor-intensive task. Thus, alternate strategies have been utilized

on a research basis, which involve analysis of *F8* mRNA by reverse transcription PCR (RT-PCR)[9] or other single-strand conformation polymorphism (SSCP)-based screening techniques.[10]

Indirect DNA Analysis (Linkage Analysis)

The principle underlying genetic linkage analysis is the tendency for alleles close together on the same chromosome to be transmitted together as an intact unit through meiosis. Polymorphisms, intragenic or extragenic, are typically single base pair changes or simple sequence repeats (e.g., CA_n repeats). Disadvantages of linkage analysis include the need for DNA samples from the proband and both parents, and the requirement that the proband's mother be heterozygous and thus "informative" for the polymorphisms. For a polymorphism to be "informative," the DNA sequence at a locus must differ on the maternally inherited and paternally inherited chromosomes. Most families (up to 90%) are informative at one or more DNA polymorphic loci, if both intragenic and extragenic polymorphisms are analyzed. The distance between the disease-causing mutation and the polymorphism is related to the risk of recombination between the gene and the polymorphism, which may lead to false-positive or false-negative results. The recombination risk is lowest if intragenic markers are used. Rarely, families will be informative only for extragenic polymorphisms, and misdiagnosis can occur in up to 6% of carriers as a result of linkage disequilibrium. When an intragenic polymorphism is used for the diagnosis of carrier or affected status, the chance of an error is less than 1%. A major drawback of linkage analysis is the possibility of unmasking nonpaternity in any given family. A complete listing and frequency of polymorphisms is available on the *F8* mutation database Web site.

Interpretation of Test Results

Reduced F8 activity is virtually diagnostic of HA provided VWD and specimen artifact have been excluded. Rare instances of other genetic disorders that can have low F8 activity include VWD type 2N (see section on VWD) and rare combined deficiencies of F5 and F8. The latter occurs as a result of mutations in *ERGIC53*, which encodes a protein necessary for efficient transport of F5 and F8 from the endoplasmic reticulum to the Golgi apparatus. F5 and F8 levels are typically in the 10% to 15% range. Molecular genetic testing is available only on a research basis.[11]

Detection of the presence of a mutation determines the genotype of the proband and can be used to confirm the carrier status of at-risk females. Certain well-defined mutations (e.g., inversions, deletions, insertions, splice junction mutations, and nonsense mutations) have obviously deleterious effects. Determining the deleterious nature of previously uncharacterized missense mutations poses a challenge. However, such a mutation in hemizygous genes (F8 and F9) and absence of additional mutations provide reasonable evidence of the effect of the mutation on the function of the protein. Other criteria typically considered in predicting the effect of a missense mutation include analysis of the degree of conservation of the respective residue among other species and presence of a similar mutation in other patients. The best evidence of the deleterious nature of a missense mutation is in vitro confirmation of its effect on protein function.

Laboratory Issues

As with any genetic test performed on peripheral blood leukocyte genomic DNA, high-quality DNA extracted from whole blood is essential for the quality of the testing. The specimen needs to be clearly identified and information on allogeneic bone marrow or peripheral blood stem cell transplant in the individual needs to be provided.

HEMOPHILIA B

Hemophilia B (HB) is an X-linked recessive bleeding disorder due to a deficiency of F10 and is clinically indistinguishable from HA. HB affects 1 in 30,000 live male births across all ethnic groups. Up to 30% of HB cases occur in families with no prior family history of HB. The diagnosis is established based on plasma F10 activity, which, like HA, permits classification of HB into severe (<1% F10 activity), moderate (1% to 5%), or mild (>5% to 40%) disease.

Molecular Basis of Disease

The *F9* gene is located on the long arm of the X chromosome (Xq27.1) and is 38 kb long with eight exons of varying lengths (25 to 1935 bp). The resulting mRNA is approximately 3 kb, of which the coding sequence is 1390 nucleotides. The intron/exon boundaries roughly correlate with the F9 domains and bear a high degree of homology to members of the vitamin K–dependent protein family (F7, F10, and protein C) (Figure 12-4). The *F9* gene encodes a precursor protein of approximately 454 amino acids

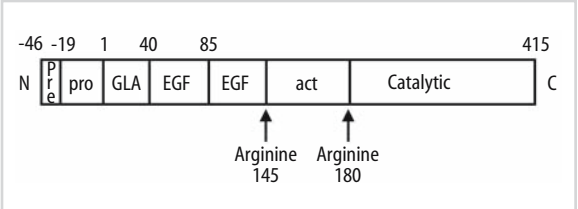


Figure 12-4. Factor IX peptide showing domains. N, N-terminus; C, C- terminus; Pre, pre-propeptide; pro, propeptide; GLA, GLA domain; EGF, EGF domain, act, activation peptide.

consisting of a propeptide followed by a glutamic acid-rich (GLA) domain, two epidermal growth factor (EGF) domains, an activation peptide, and a catalytic domain.

F9 is activated to F9a when cleaved by F7a-TF and activated factor XI (F11a). Cleavage releases the activation peptide, resulting in a circulating light chain and heavy chain connected by a disulfide bond. Numerous posttranslational modifications are necessary for the normal function of F9a, including tyrosine sulfation, serine phosphorylation, and *O*- and *N*-linked glycosylation. The relatively small size of the *F9* gene lends itself to detailed molecular analysis, which has been cataloged in an online database, Haemophilia B Mutation Database, which is updated regularly; the number system used below can be accessed at this Web site (http://www.kcl.ac.uk/ip/petergreen/haemBdatabase.html).

Mutations in the *F9* Gene

The majority of the *F9* mutations are single base pair changes that result in missense, frameshift, or nonsense mutations. Short deletions (<30 nucleotides) account for approximately 7%, larger deletions approximately 3%, and insertions approximately 2% of mutations. Many of the single base pair changes occur at CpG doublets that are hotspots for mutation. However, a subset of repeated mutations is due to a founder effect, which typically results in mild disease. Mutations have been detected in all regions of the *F9* gene including the poly(A) signal.

Missense mutations account for the majority of mutations that typically result in mild disease unless the mutations occur in residues critical for normal F9 function. Selected mutations in the promoter region of the *F9* gene result in a unique phenotype, termed hemophilia B Leyden, which is characterized by severe disease at birth with progressive amelioration of severity through adolescence and puberty. However, some *F9* promoter mutations (e.g., Brandenburg mutation at −26) result in lifelong severe disease. Nonsense mutations in the signal peptide and propeptide regions lead to severe HB; however, missense changes leading to retention of F9 within hepatic cells have been described (e.g., Ile30 and Ile19). Lack of cleavage of the propeptide leads to a dysfunctional F9 molecule (e.g., Arg4).

Mutations in the GLA domain disrupt γ-carboxylation (posttranslational modification) that is important for normal F9a binding to collagen, activated platelets, and endothelial cells. Mutations in the EGF domain results in disruption of F9 binding to calcium that is essential to its procoagulant activity, as well as cofactor F8. Mutations in the catalytic domain typically disrupt the catalytic triad (His 221, Asp 269, and Ser 365) essential for F9a protease function.

An unusual F9 variant, due to mutation at Ala10, is characterized by normal baseline F9 activity. However, warfarin therapy results in a severe and disproportionate reduction in F9 activity (typically to <1%) and causes bleeding in patients being treated with warfarin who have an apparently therapeutic international normalized ratio (INR). Indicative of such a situation is a disproportionate prolongation of the activated partial thromboplastin time (aPTT), which should be assessed by clotting factor assays and reveals the F9 deficiency.

Polymorphisms in the F9 Gene

Eight common polymorphisms have been described in different ethnic populations of European and African descent. These polymorphisms, however, are much less common in Asian and other populations. The most informative polymorphism documented in the Asian population is *Hha* I (allele frequency = 0.17); several recent reports have described additional polymorphic loci in these populations, facilitating molecular diagnosis of nonwhite carriers and patients with HB. A polymorphism within the *F9* coding region, Ala148Thr, occurs within the activation peptide. This does not correlate with F9 activity or antigenic levels. The Thr allele occurs with a frequency of 0.3 in the white population, but is much less frequent in the African American population (0.053 to 0.15) and Asian population (<0.01).

Clinical Utility of Testing

The diagnosis of HB is established by assaying F9 activity, rather than by genetic testing. Given that F9 is a vitamin K–dependent protein, all patients with mild to moderate reductions in F9 activity should be evaluated to exclude vitamin K deficiency. Reduced F9 activity in at-risk females typically confirms HB carrier status. However, normal F9 activity does not exclude the possibility of being a carrier. In this circumstance, molecular genetic testing would be the only option for diagnosis of HB.

Molecular Genetic Diagnostic Testing of the Proband

Knowledge of *F9* genotype does not alter clinical management of HB patients; however, it may predict the risk of developing F9 inhibitors and anaphylaxis in response to F9 concentrate therapy.[12] Although knowledge of the proband's genotype is important for carrier testing of at-risk female family members, the size of the *F9* gene renders it more amenable to analysis for the specific *F9* mutation causing HB.

Molecular Genetic Carrier Testing for HB

For a discussion applicable to molecular genetic carrier testing for HB, see the Clinical Utility of Testing section on HA.

Available Assays

A complete listing of laboratories offering HB genetic testing that have registered with GeneTests can be found on that Web site.[6] In general, testing can be divided into direct and indirect testing, and selected laboratories offer prenatal testing.

Direct DNA Analysis

Generally, most patients with HB have mild disease. Thus, given that up to 25% of white patients with mild HB have one of three founder mutations (Gly60Ser, IIe397Thr, or Thr296Met),[13] a logical first step is to perform limited testing for these founder mutations. For those HB patients in whom a founder mutation is not identified and for severe HB patients and at-risk carriers, the logical next step is screening or sequencing regions of functional significance in the *F9* gene. Although the majority of HB patients have one mutation, about 1% of HB patients are compound heterozygous for two mutations.

For probands with deletion of part of their *F9* gene, carrier testing by PCR amplification of individual exons will likely not be useful. Thus, either linkage analysis or Southern blot analysis is required for deletion diagnosis.

Indirect DNA Analysis (Linkage Analysis)

As discussed above for HA, indirect DNA analysis in the proband does not provide any information on the proband's genotype, but does define a haplotype that identifies the abnormal *F9* gene, which is useful for carrier testing of family members. Given the relatively small size of the *F9* gene, direct sequencing also is feasible for carrier testing when the HB mutation is not a large *F9* gene deletion.

Interpretation of Test Results

For a discussion applicable to interpretation of test results for HB, see the "Interpretation of Test Results" section for HA.

Laboratory Issues

For a discussion applicable to laboratory issues for HB, see the "Laboratory Issues" section for HA.

VON WILLEBRAND DISEASE

Von Willebrand disease (VWD) is an autosomal disorder characterized by a deficiency of von Willebrand Factor (VWF) and is the most commonly recognized congenital bleeding disorder, with a prevalence varying from 0.82% to 2%.[14] Diagnosis is established based on a personal and family history of abnormal clinical bleeding and reduced plasma VWF antigen levels or functional activity (ristocetin cofactor activity) or both, and analysis of size distribution of the VWF multimers.

Molecular Basis of Disease

The *VWF* gene is located near the tip of the short arm of chromosome 12, spans 180 kb, and consists of 52 exons. The intron-exon boundaries roughly correlate with the VWF domains. The presence of an unprocessed partial pseudogene located on chromosome 22q11.2 corresponds to exons 23 through 34 of the *VWF* gene. VWF is synthesized in endothelial cells and megakaryocytes and is secreted from storage sites (platelet alpha granules and endothelial cell Weibel-Palade bodies) into the plasma. VWF is a 2,813 amino acid peptide consisting of a 22 amino acid signal peptide, a 741 amino acid propeptide, and a mature 2,050 amino acid peptide (Figure 12-5). After removal of the signal peptide, the pro-VWF dimerizes by disulfide bonds at the C-terminal ends and further polymerizes at the N-terminal ends, resulting in multimers ranging in size from 0.5 to more than 10 million daltons. Binding sites for various ligands have been localized to different domains of the VWF subunit.

VWD can be divided into two broad categories based on quantitative (types 1 and 3) or qualitative (type 2) abnormalities of plasma VWF. Quantitative abnormalities

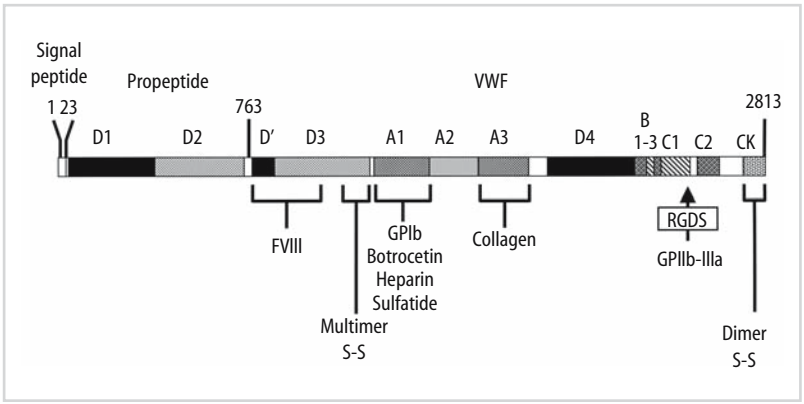


Figure 12-5. Von Willebrand peptide domains. FVIII, factor VIII; S-S, disulphide bridge; GPIb, Glycoprotein Ib; VWF, von Willebrand factor; GPIIb–IIIa, glycoprotein IIb–IIIa; RGDS, arginine (R), glycine (G), aspartic acid (D), serine (S) binding sequence. (Reprinted with permission from Sadler AJE. Von Willebrand disease. In: Scriver CR, Beaudet AL, Valle D, et al., eds. *The Metabolic and Molecular Bases of Inherited Disease*. Copyright 2001 McGraw-Hill.)

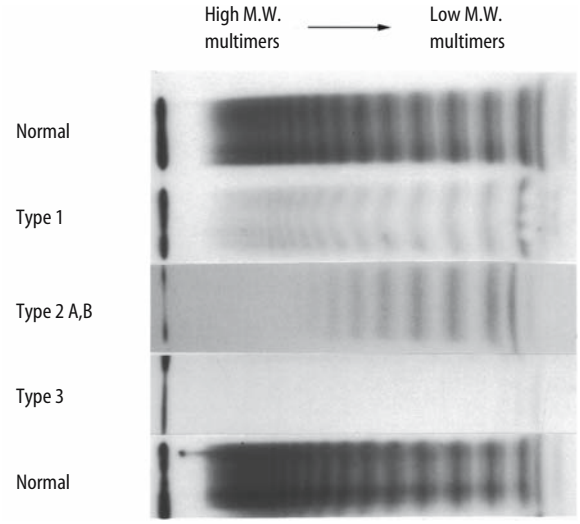


Figure 12-6. Von Willebrand factor multimer distribution. Types 1, 2, and 3 refer to types of von Willebrand disease. M.W., molecular weight. (Reprinted with permission from Bleeding disorders: an overview and clinical practice. In: Tefferi A, ed. *Primary Hematology.* Totowa, NJ: Humana Press; 2001:303. By permission of Mayo Foundation for Medical Education and Research.)

include a mild reduction of qualitatively normal VWF (type 1) or absent VWF (type 3). In type 2 VWD, the plasma VWF exhibits defective structure and function, and typically results in absence or reduction of the larger VWF multimers (Figure 12-6).

Mutations and polymorphisms in the *VWF* gene are currently being cataloged in an international database, which can be accessed via the Internet. References for the amino acid number system used below can be accessed at this website (http://www.vwf.group.shef.ac.uk/index.html).

Type 1 Von Willebrand Disease

Type 1 VWD accounts for approximately 70% of cases of VWD and is typically autosomal dominant. Diagnosis is established by assays of VWF demonstrating proportionate reduction in plasma VWF antigen, ristocetin cofactor activity, and F8 activity, with a normal distribution of VWF multimers. Although type 1 is the most commonly diagnosed variant of VWD, little is known of the molecular pathogenesis. Currently, apart from a few sporadic reports, few mutations in type 1 VWD have been characterized. Although type 1 VWD appears to be linked to the VWF locus, animal data is suggestive of locus heterogeneity as an explanation for the mild quantitative deficiency of VWF associated with type 1 VWD (e.g., defects in glycosylation of the VWF protein).[15] However, selected patients with type 1 VWD are a result of heterozygous inheritance of a type 3 VWD defect, which includes nonsense, frameshift, or deletion mutations.

Type 1 VWD has heterogeneous clinical manifestations with variable penetrance that may be due to concordant or discordant reductions in platelet VWF. Some pedigrees are characterized by significant reductions in VWF levels, which may be due to defects in multimer synthesis or secretion.

Type 3 von Willebrand Disease

Type 3 VWD is autosomal recessive and is characterized by a severe reduction in VWF antigen, ristocetin cofactor activity, and a concordant reduction in F8 activity, resulting in a more severe phenotype. Type 3 VWD mutations include frameshifts, deletions, and nonsense mutations. Although most patients are typically compound heterozygous for such VWF mutations, homozygosity has been demonstrated in a few consanguineous families. Although most patients with type 3 VWD appear to have two defective VWF alleles, many have clinically unaffected parents. This circumstance poses a challenge in providing genetic counseling.

The prevalence of type 3 VWD is 0.5 to 3 per million and, based on the Hardy Weinberg equilibrium, heterozygotes should occur at an expected frequency of at least 1,400 to 3,500 per million population. However, the currently estimated prevalence of type 1 VWD is at least 10-fold lower. Although VWF levels may be lower in parents of patients with type 3 VWD, there is clearly variable expression, with most parents being asymptomatic.

Type 2 Von Willebrand Disease

Type 2 variants of VWD are characterized by qualitatively abnormal VWF with defective stability, function, or multimer distribution and include types 2A, B, M, and N.

Type 2A Von Willebrand Disease

Type 2A VWD is autosomal dominant and accounts for approximately 75% of all type 2 VWD. Type 2A VWD patients have a variable reduction in VWF antigen, with a discordant reduction in ristocetin cofactor activity, indicative of a qualitative VWF abnormality. The higher and intermediate plasma VWF multimers are reduced or absent, and the lower-molecular-weight multimers are relatively increased or have an abnormal infrastructure. Plasma and platelet multimer abnormalities may be concordant or discordant depending on the underlying molecular defect.

Missense mutations resulting in type 2A VWD occur predominantly in the A2 domain of the *VWF* gene and result in abnormal VWF patterns by two distinct mechanisms. The first group includes mutations impairing the assembly and secretion of normal VWF multimers, resulting in decreased higher-molecular-weight VWF multimers in both plasma and platelets. The second group includes mutations that result in normal assembly and secretion of VWF; however, the mutant VWF has an increased sensitivity to proteolytic degradation in plasma, resulting in decreased plasma high-molecular-weight multimers but a

normal platelet VWF multimer pattern. The cleavage site in a subset of patients with type 2A VWD was shown to be the Tyr842-Met843 bond, where mutations may result in a conformational change, resulting in increased sensitivity to proteolysis.

Type 2B Von Willebrand Disease

Type 2B VWD is autosomal dominant and accounts for approximately 20% of all type 2 VWD. Type 2B VWD patients have a variable reduction in the VWF antigen, a discordant reduction in ristocetin cofactor activity, with a loss of the higher and intermediate plasma VWF multimers, but normal distribution of platelet VWF multimers. The type 2B variant is distinguished from type 2A VWD by the presence of mild-to-moderate thrombocytopenia. This occurs as a result of an increased affinity of VWF for platelet GPIb, resulting in spontaneous binding of VWF to platelets and rapid clearance of the platelet-bound larger multimers from plasma. In addition, platelet aggregation in response to ristocetin demonstrates an exaggerated response. The few cases described of normal multimer distribution with hyperresponsiveness to ristocetin appear to represent a mild form of type 2B VWD. Causative missense mutations occur in the A domain of the *VWF* gene and result in a dominant gain-of-function phenotype. Mutations in the A domain likely disrupt a regulatory site that normally inhibits the binding of the A1 domain to platelet GPIb.

Type 2B VWD must be distinguished from a pseudo-VWD or platelet-type VWD, which is similar in presentation. Patients with platelet-type VWD have a primary platelet defect resulting from mutations in the platelet GPIb/IX receptor.

Type 2M Von Willebrand Disease

In type 2M VWD patients, although the VWF antigen and the distribution of VWF multimers are normal, the ristocetin cofactor activity is reduced, reflecting a functional defect of the VWF multimers. There may be uncleaved proVWF or ultrahigh-molecular-weight multimers. Type 2M mutations occur in the A1 domain of the *VWF* gene and result in decreased binding affinity of VWF for platelet GPIb.

Type 2N (Normandy) Von Willebrand Disease

Mutations in the F8 binding domain of VWF result in suboptimal binding of F8 to VWF. This binding defect results in a shorter half life of plasma F8, and thus plasma F8 activity is reduced. Levels of VWF antigen and ristocetin cofactor activity are normal, as is the VWF multimer distribution. The type 2N subtype mimics mild HA but has an autosomal recessive pattern of inheritance rather than the X-linked recessive pattern of HA. In a recent international survey, VWD Normandy was detected in 4.8% (58 of 1198) of patients previously diagnosed as having mild HA. Three *VWF* gene mutations (Thr791Met, Arg816Trp, and Arg854Gln) account for 96% of type 2N patients.[16] Type 2N VWD should be considered in patients with a diagnosis of "mild HA" with a non-X-linked inheritance pattern. Typically, heterozygotes have normal F8 levels and homozygotes have reduced F8 activity. However, apparent heterozygotes with low F8 levels typically have inherited a second allele, resulting in VWD type 1 (compound heterozygotes).

Clinical Utility of Testing

Currently, the most significant impact of molecular testing on clinical management and genetic counseling is the differentiation of type 2N VWD and mild HA. Both have a mild to moderate reduction in F8 activity, with normal levels of VWF antigen and ristocetin cofactor activity. The autosomal inheritance pattern and the need for use of VWF concentrates rather than pure F8 concentrates make this an important distinction. Differentiation of VWD types 2A and 2B also provides useful information that alters clinical management. Although patients with type 2B VWD are characterized by the presence of variable degrees of thrombocytopenia, the differentiation from type 2A by platelet count is not always possible. The distinction between types 2A and 2B is clinically significant because treatment with vasopressin (DDAVP) is contraindicated in patients with type 2B due to the potential for worsening the thrombocytopenia.

Currently, for type 1 VWD patients, given the mild phenotype, the lack of well-characterized mutations, and the implication of locus heterogeneity, genetic testing is not useful. In contrast, type 3 VWD patients have well-characterized mutations and a severe phenotype. Genotyping the index patient would likely not affect clinical management; however, identification of the specific mutations is useful for genetic counseling and prenatal diagnosis.

Available Assays

Direct DNA testing to identify the specific mutation in the *VWF* gene is useful for the differentiation of type 2N from mild HA, and type 2A from type 2B. Methods include restriction fragment length polymorphism (RFLP) or direct sequencing. Indirect testing by linkage analysis is available for VWD; however, it is rarely indicated except for severe type 3 VWD. Laboratories offering testing for VWD are listed in GeneTests.[6] Currently, no laboratory offers screening or sequencing for the *VWF* gene on a clinical basis.

Interpretation of Test Results

Detection of homozygous mutations in the F8 binding domains of *VWF* correlates with the presence of type 2N VWD. However, not all mutations in *VWF* have been characterized at a structural level and must be interpreted cautiously. The presence of a heterozygous mutation in the setting of low F8 activity could reflect the presence of a second unidentified mutation in the *VWF* gene (compound heterozygote). Although most known mutations in types 2A and 2B VWD are clustered within exon 28 of the *VWF* gene, type 2 variants with mutations outside this domain have been described.

Laboratory Issues

The large size of the *VWF* gene precludes efficient, cost-effective screening or sequencing for clinical testing. In addition, (co)amplification of the unprocessed pseudogene, homologous to exons 23 to 34 of the *VWF* gene, complicates mutation analysis of PCR products amplified from genomic DNA. However, availability of primers designed specifically to amplify the *VWF* gene allows for analysis of the *VWF* gene sequence distinct from the pseudogene. Mutation analysis of mRNA has been performed on a research basis but is not practical for clinical testing given the instability of mRNA in transported specimens and the complexity of the testing.

INHERITED THROMBOPHILIAS

The recognized plasma components of the anticoagulant system include antithrombin (AT), protein C (PC), and protein S (PS). PC is a circulating vitamin K–dependent zymogen that is activated to activated protein C (APC), the active enzyme, by the thrombin-thrombomodulin complex. Thrombomodulin is an integral membrane protein on the luminal surface of endothelial cells that binds thrombin and changes the thrombin activity such that thrombin no longer clots fibrinogen or activates platelets. Instead, thrombomodulin-bound thrombin activates PC to APC. APC functions as an anticoagulant by inactivating (via enzymatic cleavage) F5a and F8a in the presence of PS (Figure 12-7). Procoagulants F5a and F8a are cofactors that, together, markedly accelerate the rate of thrombin generation. Thus, F5a and F8a are key hemostasis regulatory points; inactivation of these cofactors markedly downregulates thrombin generation. Antithrombin is a serine protease inhibitor (SERPIN) and acts as a pseudo-substrate for thrombin to irreversibly inhibit thrombin and procoagulants F9a, F10a, F11a, and F12a by covalently binding the enzymatic active sites. Antithrombin anticoagulant activity is markedly enhanced (catalyzed) by glycosaminoglycans (e.g., heparin).

Genetic deficiencies of plasma AT, PC, or PS activity to approximately half of normal plasma activity is strongly associated with venous thromboembolism (VTE). Comparing family members with inherited AT, PC, or PS

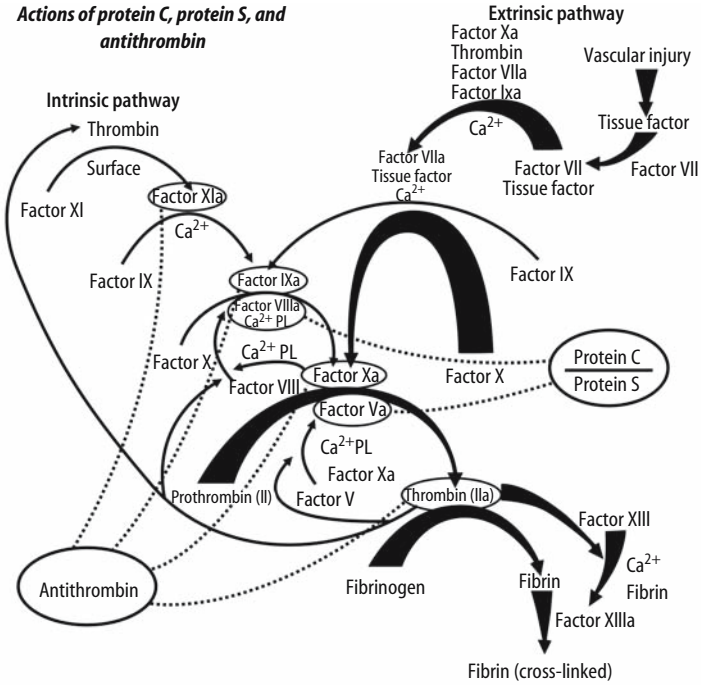


Figure 12-7. Actions of protein C, protein S, and antithrombin. a, activated factor; Ca^{2+}, calcium; PL, phospholipid. (Modified with permission from: Davie EW, Fujikawa K, Kisiel W. The coagulation cascade: initiation, maintenance and regulation. *Biochemistry* 1991;30:10363. Copyright 1991 American Chemical Society.)

deficiency to family members with no defect, the lifetime VTE risk is increased 7- to 8-fold. Individuals with homozygous PC or PS deficiency may develop severe full-thickness dermal and limb thrombotic infarction as neonates (e.g., purpura fulminans neonatalis). Homozygous AT deficiency appears to be incompatible with life except for mutations that impair heparin binding to antithrombin. Altogether, deficiency of AT, PC, or PS is found in approximately 5% of VTE patients. Mutations causing impaired expression or function of AT, PC, or PS are rare and distributed throughout their respective genes. Moreover, there is minimal data correlating specific mutations with a unique clinical thrombosis phenotype. Thus, functional testing of plasma activities of AT, PC, or PS is preferred over genetic testing for diagnosis.

FACTOR V LEIDEN MUTATION

Molecular Basis of Disease

Initial APC cleavage at arginine (R) 506 of F5a is required for optimal exposure and subsequent rapid inactivation of F5a by subsequent APC cleavage at positions R306 (in the presence of phospholipid and protein S) and R679 (Figure 12-8). The discovery of three unrelated patients with idiopathic VTE, whose plasma was resistant to the anticoagulant effect of exogenously added APC, provided exciting new insights into the etiology of VTE.[17] Early epidemiological data suggested that activated protein C resistance (APC-R) was familial with an autosomal dominant inheritance pattern. Procoagulant F5 isolated from APC-R patient plasma was resistant to inactivation by APC, triggering an intensive search for the genetic explanation.[18] Subsequent work identified a single point mutation, called Factor V Leiden, of guanine (G) to adenine (A) at nucleotide 1691 within exon 10 of the *F5* gene on chromosome 1 (1q21–25).[19] Factor V Leiden encodes for substitution of a glutamine (Q) for arginine (R) at amino acid position 506 within the heavy chain of F5, at one of three APC cleavage sites (R306, R506, R679). Factor V Leiden promotes thrombosis by impaired downregulation of the generation of thrombin and by inhibition of fibrinolysis.

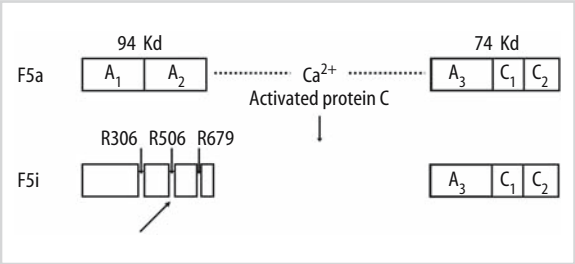


Figure 12-8. Cleavage sites for inactivation of activated factor V. F5a, activated factor V; F5i, inactivated factor V; Kd, kilodalton; Ca^{2+}, calcium; *Arrows,* cleavage sites; R, arginine.

Factor V Leiden is a founder mutation, arising between 21,000 and 34,000 years ago, after the evolutionary divergence of Africans from non-Africans and of Caucasoid from Mongoloid subpopulations. Worldwide, the Factor V Leiden carrier frequency ranges from 2% in southern Europe to 15% in southern Sweden, and generally declines in native populations moving from west to east toward Asia, Africa, Australasia, and ultimately, the Americas. In the United States, about 3% to 7% of asymptomatic white populations of northern European or Scandinavian ancestry are heterozygous carriers. Factor V Leiden is much less common in other U.S. populations, with a carrier frequency of 1.2% in African Americans, 2.2% in Hispanics, 1.2% in Native Americans, and 0.45% in Asian Americans.

Clinical Utility of Testing

Patients with Factor V Leiden (or another familial thrombophilia) may present with one or more clinical phenotypes, including VTE, recurrent fetal loss, complications of pregnancy, and possibly arterial thrombosis.

VTE is a common disease, with an average annual incidence of over 1 per 1,000 person-years.[20] VTE is also a lethal disease, mostly due to pulmonary embolism. Almost one third of all pulmonary embolism patients die within 7 days, and one quarter die suddenly.[20] For patients who die suddenly, available time is insufficient to recognize, diagnose, and begin therapy to alter the course of the disease. To improve survival and prevent complications, the incidence of VTE must be reduced, either by avoidance of high-risk exposures or by prophylaxis during such exposures. Independent clinical risk factors for VTE include confinement to a hospital (with or without surgery) or nursing home, trauma, malignant neoplasm (with or without chemotherapy), central venous catheterization, placement of a transvenous pacemaker, prior superficial vein thrombosis, and serious neurologic disease with extremity paresis. Among women, additional risk factors include pregnancy and the puerperium, oral contraception, hormone (estrogen and/or progesterone) replacement therapy, and selective estrogen receptor modulator (SERM) therapy (e.g., tamoxifen, raloxifene). In addition to these risk factors, the patient age affects VTE risk, with a markedly increased risk with increased age for both men and women.[20]

While clinical risk factors can identify populations at risk, most persons that are exposed to such risk factors do not develop VTE. Moreover, about 25% of VTE patients have no obvious risk exposure.[21] The discovery of APC-R and Factor V Leiden has provided new insights into the etiology of VTE. APC-R is the most commonly recognized familial thrombophilia among patients with VTE. The Factor V Leiden prevalence among incident VTE cases is 12% to 21%. Heterozygous Factor V Leiden carriers have an 8-fold increased risk for VTE, while the risk among

homozygous carriers is increased 80- to 100-fold. VTE risk increases due to interaction between Factor V Leiden and other genetic disorders such as AT, PC, or PS deficiency. Genetic and clinical risk factors also interact to compound the risk of VTE. For example, VTE risk increases 30-fold among heterozygous Factor V Leiden women carriers receiving oral contraceptives, 7- to 16-fold during pregnancy or the puerperium, and 13- to 15-fold for postmenopausal women carriers receiving estrogen replacement therapy.[22] Factor V Leiden carriers also may be at increased risk for VTE after surgery. The incidence of VTE among Factor V Leiden carriers also varies by age, ranging from 2 to 3 per 1,000 person-years for ages 15 to 30 years, to 7 to 11 per 1,000 person-years for age 60 years and older. However, compared to that for patients with AT, PC, or PS deficiency, the lifetime probability of developing VTE is considerably less for Factor V Leiden heterozygotes. Only 2.4% of Factor V Leiden carriers develop thrombosis by age 65 years.[23] Thus, the vast majority of Factor V Leiden heterozygotes do not develop symptomatic VTE.

During counseling of patients or family members, it is most useful to estimate the incidence of VTE by age and exposure. For example, the incidence of VTE among women of childbearing age is about 1 per 10,000 woman-years.[24] The risk for women of childbearing age is increased about 3-fold (e.g., to an incidence of 3 per 10,000 woman-years) among women taking oral contraceptives. The incidence of VTE among women who are heterozygous for Factor V Leiden in this age group is about 6 per 10,000 woman-years (e.g., a relative risk of 6 compared to women noncarriers), and that incidence is increased to about 30 per 10,000 woman-years for heterozygotes taking oral contraceptives. While this represents a 30-fold increased relative risk, the incidence is still only 0.3% per year (i.e., 99.7% of carriers on oral contraceptives will not develop VTE during that year). Only the patient can decide whether this risk is acceptable.

VTE also recurs frequently, with an estimated cumulative recurrence of 30% by 10 year.[20] Independent predictors of recurrence include older age, obesity, malignant neoplasm, and extremity paresis. The risk of recurrent VTE is increased approximately 10-fold for patents with VTE and deficiency of AT, PC, or PS. While about 50% of patients with VTE are Factor V Leiden heterozygotes, there are conflicting data as to whether Factor V Leiden heterozygotes are at a higher risk for recurrent VTE compared to individuals who do not have Factor V Leiden.[22] However, the risk of recurrence is significantly increased for Factor V Leiden patients who also are heterozygous for the prothrombin 20210G→A gene mutation (relative risk, 2.6 to 9.1), particularly for individuals without other clinical risk factors (relative risk, 5.1 or 4.0). The risk of recurrent VTE is 3-fold higher for individuals who are homozygous for Factor V Leiden.

Factor V Leiden also is a mild risk factor for recurrent pregnancy loss and possibly other serious obstetrical complications (e.g., preeclampsia, fetal growth retardation, and placental abruption), possibly due to thrombosis of placental vessels and impaired placental perfusion.[22]

Available Assays

In most circumstances, a functional assay of plasma APC resistance is the preferred initial test. In the original assay, aPTT was measured before and after addition of a standardized amount of APC that was sufficient to prolong the aPTT about 2-fold. The result was usually expressed as the APC-R ratio (defined as the ratio of the aPTT clotting times in the presence and absence of APC). An APC-R ratio of less than 2.0 was indicative of APC-R. About 95% of patients with APC-R were carriers for Factor V Leiden. Although specific, the original APC-R assay was relatively insensitive. Factor V Leiden carriers can have an APC-R ratio as high as 2.9, and the assay did not reliably distinguish heterozygotes from homozygotes. Moreover, the original assay was uninterpretable if the baseline aPTT was prolonged due to warfarin or heparin anticoagulation, factor deficiency, or a lupus anticoagulant or other specific factor inhibitor.

A modified (second-generation) APC-R functional assay that overcomes the limitations of the original assay is now widely available and FDA approved. In this assay, the patient's plasma is first mixed with F5-deficient plasma that contains a heparin neutralizer. The addition of the F5-deficient plasma corrects for deficiencies of other coagulation proteins and may dilute the effect of some lupuslike anticoagulants. The modified assay is essentially 100% sensitive and specific for Factor V Leiden and accurately distinguishes heterozygotes from homozygotes. In addition, the second-generation assay is unaffected by heparin or warfarin anticoagulation. However, the assay may still be uninterpretable if the baseline aPTT (after mixing with F5-deficient plasma) is still prolonged due to a lupus anticoagulant or specific factor inhibitor. Moreover, the assay will miss patients with acquired APC-R. Each laboratory must determine its own normal range for this test.

Direct DNA testing for Factor V Leiden is widely available for diagnosis. Commonly used molecular methods use PCR amplification of the region surrounding the mutation followed by restriction enzyme digestion (PCR-RFLP), allele-specific PCR amplification, or allele-specific hybridization. Semiautomated assay methods include fluorescence detection of PCR products with allele-specific hybridization probes, and non-PCR signal amplification methods based on either enzymatic hybridization mismatch recognition using fluorescent allele-specific probes or linked fluorescent allele-specific pyrophosphorolysis-kinase reactions.[22] Each laboratory must insure that its method can distinguish Factor V Leiden from an uncommon polymorphism at nucleotide 1696 (A→G). In 2003, Roche Diagnostics Corporation (Indianapolis, IN)

obtained FDA approval for in vitro diagnostic test kits for Factor V Leiden and the prothrombin 20210G→A gene mutation.

Recommendations for Clinical Use of Factor V Leiden Testing

Testing for Factor V Leiden is recommended for patients with the following events:[22]

- a history of recurrent VTE;
- a first VTE before 50 years of age;
- a first VTE without clinical risk factors at any age;
- a first VTE at an unusual anatomic site such as the cerebral, mesenteric, portal, or hepatic veins;
- a first VTE at any age in a subject with a first-degree family member with a VTE before 50 years of age;
- a first VTE related to pregnancy, the puerperium, oral contraception, or hormone replacement therapy;
- unexplained pregnancy loss during the second or third trimester.

Testing for Factor V Leiden is controversial for the following individuals:

- young women smokers (age <50 years) with a myocardial infarction;
- older patients (age >50 years) with a first provoked VTE in the absence of cancer or an intravascular device;
- patients with a first VTE related to SERM therapy;
- selected cases of women with unexplained severe preeclampsia, placental abruption, or intrauterine growth retardation.

After appropriate counseling, testing for Factor V Leiden also may be indicated in asymptomatic adult family members of probands with known Factor V Leiden mutations, especially those with a strong family history of VTE at a young age (<50 years), and asymptomatic female family members that are pregnant or are considering oral contraceptives or pregnancy.

Factor V Leiden testing is not recommended for the following purposes:

- general population screening;
- routine initial test during pregnancy;
- routine initial test prior to or during oral contraceptive use, hormone replacement therapy, or SERM therapy;
- prenatal test, newborn initial test, or as a routine test in asymptomatic prepubescent children;
- routine initial test in patients with arterial thrombotic events.

However, Factor V Leiden testing may be appropriate for patients with unexplained arterial thrombosis without atherosclerosis or for young patients who smoke.

Clinical Management

Individuals who are heterozygous or homozygous for Factor V Leiden with a first lifetime deep vein thrombosis or pulmonary embolism should be treated in standard fashion, initially with heparin (either unfractionated or low-molecular-weight heparin), followed by oral anticoagulation with a vitamin K antagonist (target INR 2.5; therapeutic range 2.0–3.0).[22] In general, 3 to 6 months of oral anticoagulation therapy is recommended after a first lifetime VTE for patients with Factor V Leiden, especially if the event was associated with a transient clinical risk factor (e.g., surgery, oral contraceptive use, pregnancy, or the puerperium). The need for lifelong anticoagulation after a first episode of VTE for patients with Factor V Leiden has not been established by appropriate clinical trials. Therefore, indefinite anticoagulation should be recommended only after careful consideration of the risks and benefits. Indefinite anticoagulation may be recommended for patients with Factor V Leiden who have an idiopathic or life-threatening VTE (especially in the presence of reduced cardiopulmonary functional reserve), have more than one hereditary thrombophilia, are homozygous for a hereditary thrombophilia, or have additional persistent clinical risk factors (e.g., malignant neoplasm, serious neurologic disease with extremity paresis, antiphospholipid antibodies). Hereditary thrombophilia patients (or any patient) with recurrent unprovoked VTE should be considered for indefinite anticoagulation therapy.

Women with Factor V Leiden and a history of unprovoked VTE should receive prophylactic anticoagulation with heparin or low-molecular-weight heparin during pregnancy and for at least 6 weeks postpartum. Routine anticoagulation therapy is not recommended for individuals with Factor V Leiden who also have atherosclerotic arterial occlusive disease. However, for patients with myocardial infarction or stroke and Factor V Leiden, anticoagulation therapy for secondary prevention may be appropriate.

Anticoagulation therapy is not recommended for asymptomatic individuals with Factor V Leiden. Factor V Leiden carriers (with or without previous VTE) should receive appropriate prophylaxis when exposed to risk factors. Standard prophylaxis recommendations are sufficient for most types of surgery. A possible exception is an asymptomatic Factor V Leiden carrier undergoing hip-replacement surgery, who might be at increased risk of symptomatic VTE for several weeks after surgery. These patients should receive extended out-of-hospital prophylaxis, especially in association with obesity or prolonged immobilization.

Prophylactic anticoagulation is not routinely recommended in pregnant Factor V Leiden carriers with no history of thrombosis. Decisions about anticoagulation should be individualized based on the genotype (heterozygous or homozygous) and coexisting risk factors. Asymptomatic women who do not receive anticoagulation should

be followed closely throughout pregnancy and given prophylaxis during the puerperium.

PROTHROMBIN 20210G→A MUTATION

Molecular Basis of Disease

The prothrombin (Pt) 20210G→A mutation is a relatively common polymorphism that affects the 3′-terminal nucleotide of the prothrombin gene 3′-untranslated region.[25] Individuals with this polymorphism have increased plasma prothrombin levels. The Pt 20210G→A mutation causes a gain of function due to increased recognition of the polyadenylation cleavage signal, increased 3′ end RNA processing, mRNA accumulation, and increased protein synthesis.[26] A family-based study has shown close linkage between a quantitative trait locus determining plasma prothrombin activity and the Pt 20210G→A mutation.[27] Plasma prothrombin concentration is a major determinant of plasma thrombin generation potential. Thrombin generation is increased in the plasma of heterozygous and homozygous Pt 20210G→A mutation carriers. High plasma prothrombin levels also inhibit APC-mediated inactivation of procoagulant F5a, further augmenting thrombin production.

Based on haplotype analyses, the Pt 20210G→A mutation also appears to be a founder mutation, arising 20,000 to 30,000 years ago, and after the divergence of Africans from non-Africans and of Caucasoid from Mongoloid subpopulations. In Europe, the overall carrier frequency is 2%, ranging from 1.7% in northern Europe to 3.0% in southern Europe. In the United States, the estimated overall carrier frequency is 1% to 2%. The mutation is uncommon among African Americans and rarely seen among Asian Americans and Native Americans.[28]

Clinical Utility of Testing

The Pt 20210G→A mutation has been associated with a 2- to 3-fold increased risk of VTE incidence. The VTE risk among Pt 20210G→A carriers is compounded by deficiency of AT, PC, or PS, or by Factor V Leiden. Women with the Pt 20210G→A mutation who are taking oral contraceptives have a 16-fold increased risk for deep vein thrombosis or pulmonary embolism, and a 150-fold increased risk of cerebral vein thrombosis. While the VTE risk among individuals homozygous for the Pt 20210G→A mutation is likely higher than the risk among heterozygotes, this mutation is a relatively weak risk factor compared to deficiency of PC or PS. Compared to noncarriers with VTE, the risk of recurrent VTE does not appear to be increased for individuals heterozygous for the Pt 20210G→A mutation. However, the risk of recurrence is increased for individuals with both a heterozygous Pt 20210G→A mutation and Factor V Leiden.

Available Assays

Although Pt 20210G→A carriers have statistically higher plasma prothrombin activity, the normal range for plasma prothrombin activity is quite broad. Thus, plasma prothrombin activity cannot accurately distinguish Pt 20210G→A carriers from noncarriers and is not of value for diagnosis. Direct DNA testing for the specific mutation is required for diagnosis. Molecular testing methods available for the Pt 20210G→A mutation are essentially the same as for Factor V Leiden. Each laboratory must ensure that its method can distinguish the Pt 20210G→A from an uncommon polymorphism at nucleotide 20209 (C→T). In 2003, Roche Diagnostics Corporation (Indianapolis, IN) obtained FDA approval for in vitro diagnostic test kits for Factor V Leiden and the prothrombin 20210G→A gene mutation. Recommendations for prothrombin 20210G→A mutation testing and management are essentially the same as for Factor V Leiden.

HYPERHOMOCYST(E)INEMIA

Molecular Basis of Disease

Homocysteine is a non-protein-forming, sulfhydryl amino acid that is an intermediary between methionine and cysteine (Figure 12-9). Homocysteine is formed by intracellular demethylation of dietary methionine. Homocysteine is converted to cysteine in a two-step transsulfuration pathway that requires initial condensation of homocysteine with serine to form cystathionine. The latter reaction is catalyzed by cystathionine-β-synthase (CBS) and requires the essential cofactor pyridoxal 5′-phosphate (vitamin B_6). In the second step, catalyzed by cystathionine γ-lyase, cystathionine is hydrolyzed to α-ketobutyrate and cysteine. Homocysteine can be remethylated to methionine via two

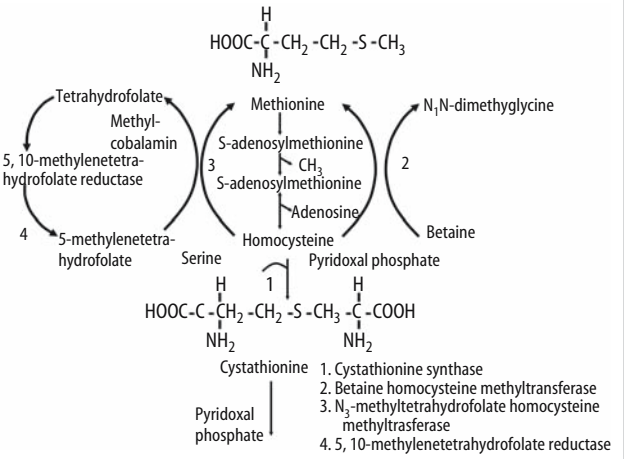


Figure 12-9. Homocysteine metabolism pathway.

pathways. In the first, a methyl group is donated by methyltetrahydrofolate in a reaction catalyzed by methyltetrahydrofolate-homocysteine methyltransferase and requiring cobalamin (vitamin B_{12}). Tetrahydrofolate (folic acid) is remethylated to methyltetrahydrofolate in a reaction that includes the intermediary 5,10-methylenetetrahydrofolate and requires the enzyme methylenetetrahydrofolate reductase (MTHFR). In the second pathway, betaine (trime-thyl-glycine) donates a methyl group to remethylate homocysteine to methionine in a reaction requiring betaine-homocysteine methyltransferase. Remethylation is primarily responsible for regulation of fasting homocysteine levels, while transsulfuration mainly regulates higher homocysteine levels as occur in the postprandial state or after methionine loading.[29]

Homocystinuria is a rare inherited disorder affecting 3 to 5 per million of the general population, and usually is caused by severe deficiency of CBS. Most patients are homozygous or compound heterozygous for one or more of three mutations (833T→C, 919G→A, or 1224A→C) within the *CBS* gene located in the subtelomeric region of chromosome 21 (21q22). Heterozygotes often have normal basal plasma homocysteine levels but develop hyperhomocysteinemia after a methionine load.[29] Although rare, severe MTHFR deficiency also can cause homocystinuria. Homozygotes for the common *MTHFR* 677C→T mutation who become folate deficient may develop mild hyperhomocysteinemia. This mutation encodes for substitution of a valine for alanine at amino acid position 223. Approximately 12% of the U.S. population is homozygous for this mutation.

Clinical Utility of Testing

Patients with homocystinuria are at markedly increased risk for both atherosclerotic arterial occlusive disease and VTE. About 25% of patients with homocystinuria develop a thrombotic vascular occlusive event by age 16 years, and about 50% develop such an event by age 29 years.[30] Of these, about half are arterial occlusive events (e.g., stroke, myocardial infarction, or peripheral artery thrombosis), and the remainder are VTE. The event rate is significantly reduced by vitamin therapy in B_6-responsive patients.[29]

Homocystinuric patients generally have plasma homocyst(e)ine levels of 100 to 300 µmol/L. Multiple case-control studies have found that milder hyperhomocysteinemia (e.g., 15 to 100 µmol/L) is a risk factor for both arterial occlusive disease and VTE, with odds ratios of about 2.2 to 3.0.[29] However, there is concern that these studies may have been confounded because plasma homocysteine levels may be elevated after thrombotic vascular events. Indeed, although fewer in number, prospective studies have provided conflicting results. Thus, hyperhomocysteinemia appears to be a weak risk factor for thrombotic arterial occlusive disease, while the risk for VTE remains uncertain. Heterozygotes for the *MTHFR* 677C→T mutation are not at risk for hyperhomocysteinemia, thrombotic arterial occlusive disease, or VTE. While homozygotes are at increased risk for hyperhomocysteinemia, homozygosity for the *MTHFR* 677C→T mutation in the absence of hyperhomocysteinemia is not an independent risk factor for either arterial or venous thrombosis. There are no studies showing that reduction in plasma homocysteine levels by therapy with vitamins B_6 or B_{12} or folic acid reduces the risk for arterial or venous thrombosis. Testing for hyperhomocysteinemia is recommended for patients with documented atherosclerotic arterial occlusive disease (e.g., coronary artery, cerebrovascular, or peripheral vascular), while testing patients with VTE remains controversial. Genotyping for the *MTHFR* 677C→T mutation may provide insight into the etiology of hyperhomocyst(e)inemia but does not influence therapy or warrant family counseling.

Available Assays

Acceptable assays for plasma homocyst(e)ine include HPLC and immunoassay.[28] Each laboratory should determine its own gender- and local population-specific reference ranges. The basal plasma homocysteine level should be determined first, and if elevated, further investigation regarding the potential cause as well as therapeutic intervention should be considered. It is unclear whether plasma homocysteine levels should be tested after an overnight fast. Because plasma homocysteine levels can be elevated for several months after myocardial infarction or stroke, testing should be delayed accordingly. In addition to vitamin deficiency, impaired renal function and hypothyroidism are other common causes of hyperhomocyst(e)inemia. Vitamin B_{12} deficiency should be excluded prior to beginning therapy since high-dose folic-acid therapy can precipitate acute B_{12} neuropathy. If the basal homocysteine level is normal, methionine loading (0.1 g/kg body weight or 3.8 gm/m^2 body surface area with measurement of plasma homocysteine 4 to 6 hours after the load) should be considered since 25% to 40% of symptomatic patients develop hyperhomocyst(e)inemia only after methionine loading. Therapy includes folic acid (0.5 to 1.0 mg/day), vitamin B_{12} (400 to 1000 µg/day), and/or vitamin B_6 (20 to 50 mg/day).[30]

Acknowledgments

Funded, in part, by grants from the National Institutes of Health (HL66216), the Centers for Disease Control and Prevention (TS306), U.S. Public Health Service; the Doris Duke Charitable Foundation Innovation in Clinical Research; and Mayo Foundation.

References

1. Soucie JM, et al. The Hemophilia Surveillance System Project Investigators. *Am J Hematol.* 1998;59:288.

2. Antonatakis SE, et al. Factor VIII gene inversions in severe hemophilia A: results of an International Consortium Study. *Blood.* 1995;86:2206.
3. Bagnall RD, et al. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A. *Blood.* 2002;99:168.
4. Gilles JGG, et al. Factor VIII inhibitors. *Thromb Haemost.* 1997;78: 641.
5. ESHRE Preimplantation Genetic Diagnosis Consortium: data collection III. *Hum Reprod.* 2002;17:233.
6. GeneTests: Medical Genetics Information Resource [database online]. Seattle, WA: University of Washington and Children's Health System, 1993–2003. Updated weekly. Available at: http://www.genetests.org.
7. Liu Q, et al. Single tube polymerase chain reaction for rapid diagnosis of the inversion hotspot mutation in hemophilia A. *Blood.* 1998;92:1458.
8. Higuchi M, Kazazian HH Jr, Kasch L, et al. Molecular characterization of severe hemophilia A suggests that about half the mutations are not within the coding regions and splice junctions of the factor VIII gene. *Proc Natl Acad Sci U S A.* 1991;88:7405.
9. Naylor JA, et al. Detection of three novel mutations in two haemophilia A patients by rapid screening of whole essential region of factor VIII gene. *Lancet.* 1991;337:635.
10. Buzin CH, et al. Scanning by DOVAM-S detects all unique sequence changes in blinded analyses: evidence that the scanning conditions are generic. *Biotechniques.* 2000;28:746–50.
11. Nichols WC, et al. Mutations in an endoplasmic–Golgi intermediate compartment protein cause combined deficiency of coagulation factor V and VIII. *Cell.* 1998;93:61.
12. Warrier I, et al. Factor IX inhibitors and anaphylaxis in haemophilia B. *J Pediatr Hematol Oncol.* 1997;19:23.
13. Ketterling RP, Bottema CD, Phillips JA 3rd, Sommer SS. Evidence that descendants of three founders constitute about 25% of hemophilia B in the United States. *Genomics.* 1991;10:1093.
14. Sadler JE. Von Willebrand disease. In: Scriver CR, Beaudet AL, Valle D, et al., eds. *The Metabolic and Molecular Bases of Inherited Disease.* 8th ed. New York: McGraw-Hill, 2001.
15. Mohlke KL, Purkayastha AA, Westrick RJ, et al. MVWF, a dominant modifier of murine von Willebrand factor, results from altered lineage-specific expression of a glycosyltransferase. *Cell.* 1999;96: 111.
16. Mazurier C, Meyer D. Factor VIII binding assay of von Willebrand factor and the diagnosis of type 2N von Willebrand disease-results of an international survey. On behalf of the Subcommittee on Von Willebrand Factor of the Scientific and Standardization Committee of the ISTH. *Thromb Haemost.* 1996;76:270.
17. Dahlbäck B, Carlsson M, Svensson PJ. Familial thrombophilia due to a previously unrecognized mechanism characterized by poor anticoagulant response to activated protein C: prediction of a cofactor to activated protein C. *Proc Natl Acad Sci U S A.* 1993;90:1004–1008.
18. Sun X, Evatt B, Griffin JH. Blood coagulation factor Va abnormality associated with resistance to activated protein C in venous thrombophilia. *Blood.* 1994;83:3120–3125.
19. Bertina RM, Koeleman BP, Koster T, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature.* 1994;369:64–67.
20. Heit JA, Silverstein MD, Mohr DN, et al. The epidemiology of venous thromboembolism in the community: state of the art. *Thromb Haemost.* 2001;86:452–463.
21. Heit JA, O'Fallon WM, Petterson TM, et al. Relative impact of risk factors for deep vein thrombosis and pulmonary embolism: a population-based study. *Arch Intern Med.* 2002;162:1245–1248.
22. Press RD, Bauer KA, Kujovich JL, Heit JA. Clinical utility of Factor V Leiden (R506Q) testing for the diagnosis and management of thromboembolic disorders. *Arch Pathol Lab Med.* 2002;126:1304–1318.
23. Rodeghiero F, Tosetto A. Activated protein C resistance and factor V Leiden mutation are independent risk factors for venous thromboembolism. *Ann Intern Med.* 1999;130:643–650.
24. Vandenbroucke JP, Koster T, Briet E, Reitsma PH, Bertina RM, Rosendaal FR. Increased risk of venous thrombosis in oral-contraceptive users who are carriers of factor V Leiden mutation. *Lancet.* 1994;344:1453–1457.
25. Poort SR, Rosendaal FR, Reitsma PH, Bertina RM. A common genetic variation in the 3′-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood.* 19;88:3698–3703.
26. Gehring NH, Frede U, Neu-Yilik G, et al. Increased efficiency of mRNA 3′ end formation: a new genetic mechanism contributing to hereditary thrombophilia. *Nat Genet.* 2001;28:389–392.
27. Soria JM, Almasy L, Souto J, et al. Linkage analysis demonstrates that the prothrombin G20210A mutation jointly influences plasma prothrombin levels and risk of thrombosis. *Blood.* 2000;95:2780–2785.
28. McGlennen RC, Key NS. Clinical and laboratory management of the prothrombin G20210A mutation. *Arch Pathol Lab Med.* 2002;126: 1319–1325.
29. Key NS, McGlennen RC. Hyperhomocyst(e)inemia and thrombophilia. *Arch Pathol Lab Med.* 2002;126:1367–1375.
30. Mudd SH, Skoyby F, Levy HL, et al. The natural history of homocystinuria due to cystathionine beta-synthase deficiency. *Am J Hum Genet.* 1985;37:1–31.

Chapter 13

Hematologic Disorders: Hemochromatosis, Hemoglobinopathies, and Rh Incompatibility

Daniel B. Bellissimo

HEMOCHROMATOSIS

Molecular Basis of Disease

Hereditary hemochromatosis (HHC) is an autosomal recessive disorder of iron metabolism resulting from excess iron storage in the liver, skin, pancreas, heart, joints, testes, and pituitary gland. If left untreated, life-threatening complications such as cirrhosis, diabetes, liver cancer, and cardiomyopathy may result. Iron overload and the resulting clinical complications can be avoided by early diagnosis and periodic phlebotomy to reduce the body's iron stores.

HHC was known to be associated with the human leukocyte antigen (HLA)-A3 and later was linked to HLA-A on the short arm of chromosome 6. In 1996, a gene related to the HLA class I family was identified as a gene for HHC and subsequently called *HFE*.[1] The gene located at 6p21.3 is approximately 10 kilobases (kb) and resides approximately 4.6 megabases telomeric from the HLA-A gene. Extensive reviews on the *HFE* gene and HHC are available.[2–4] The role of *HFE* in HHC and iron homeostasis is supported by the fact that an *HFE* knockout mouse develops iron overload similar to that seen in HHC.

Two allelic variants of the *HFE* gene, C282Y (845G→A) and H63D (187C→G), are significantly correlated with HHC, and most of the clinical studies have focused on these variants. These two variants are usually not found on the same chromosome. Several other *HFE* allelic variants have been described, but little is known regarding their phenotypic effects. However, the S65C mutation has been implicated in HHC.[2] Although the *HFE* gene appears to be involved in the majority of HHC patients, there are populations in which hemochromatosis does not show linkage to *HFE*.[3] Mutations in two other genes, *TFR2* (7q22) and *SLC11A3* (2q32), have been shown to cause HHC.[5,6] In addition, a juvenile form of HHC has been linked to 1q21 and 19q13. For all these mutations, HHC is inherited in an autosomal recessive manner, except for *SLC11A3*, which is autosomal dominant.

The HFE protein is a 343 amino acid transmembrane protein that is homologous to major histocompatibility (MHC) class I molecules and associates with β_2-microglobulin. The HFE protein is found in the crypt cells of the duodenum, the site of dietary iron absorption. At the cell surface, HFE complexes with the transferrin receptor and may decrease the receptor's affinity for iron-loaded transferrin or block tranferrin's binding to the receptor.[3] The C282Y mutation disrupts a conserved disulfide bridge within HFE, which disrupts HFE secondary structure required for binding to β_2-microglobulin and subsequent transport to the cell surface. The C282Y HFE protein remains in the Golgi apparatus rather than being transported to the cell surface. The H63D mutation protein product is expressed on the cell surface, but likely cannot interact normally with the transferrin receptor and therefore does not properly modulate the receptor's binding of transferrin.

The frequency of the C282Y and H63D mutations in different ethnic groups has been reviewed.[2] These mutations are most common in populations of European descent. A pooled analysis of several white population studies found that the frequency of C282Y homozygosity was 0.4% and heterozygosity was 9.2% (Table 13-1). The frequency of H63D homozygosity was 2% and heterozygosity was 23%. Compound heterozygosity (C282Y/H63D) was 2%. The frequency of *HFE* genotypes in HHC cases was 77.5% for C282Y/C282Y, 5.3% for C282Y/H63D, and 1.5% for H63D/H63D. The pooled odds ratio (OR) for each *HFE* genotype was determined in another study.[7] The C282Y/C282Y genotype had the highest risk for iron overload (OR = 4383). For the C282Y/H63D, H63D/H63D, and C282Y heterozygote (C282Y/+), the ORs were 32, 5.7, and 4.1, respectively. The high frequency of HHC patients with the C282Y/C282Y genotype suggests that the penetrance for this genotype is very high. However, additional studies designed to estimate the penetrance suggest that only a minority of C282Y homozygotes develop clinical disease resulting from iron overload. Currently, the best estimates of the penetrance range from 1% to 50%. The penetrance

Table 13-1. *HFE* Genotype Frequencies and Their Contribution to Hereditary Hemochromatosis

Genotype	Population Frequency	Pooled Odds Ratio	95% CI	% HHC Cases
C282Y/C282Y	0.4%	4383	1374–>10,000	77.5
C282Y/H63D	2.0%	32	18.5–55.4	5.3
H63D/H63D	2.0%	5.7	3.2–10.1	1.5
C282Y/+	9.2%	4.1	2.9–5.8	3
H63D/+	23%	1.9	1.5–2.5	3

Source: Data are from References 2 and 7.

of the C282Y/H63D and H63D/H63D genotypes is much lower, ranging from 0.3% to 1.4% for C282Y/H63D and 0.04% to 0.2% for H63D/H63D.

Clinical Utility of Testing

Diagnosis of HHC is based on clinical, biochemical, histological, and molecular studies, specifically tests such as transferrin saturation, ferritin concentration, evaluation of iron stores by liver biopsy, and genotyping for C282Y and H63D mutations in the *HFE* gene. Due to the reduced penetrance of C282Y, the even lower penetrance of H63D, and the presence of other mutations within *HFE* or in other genes related to HHC, the diagnosis or ruling out of HHC cannot be made with a molecular test alone. However, in patients with clinical and biochemical symptoms of iron overload, molecular testing can be useful in confirming the diagnosis of HHC.

After the diagnosis of HHC is made in the proband, other family members may be at higher risk than the general population for developing iron overload. These individuals' iron status can be monitored as appropriate. The *HFE* genotyping test can be used to identify individuals with the same genotype as the proband; these individuals would be at higher risk for developing iron overload than family members who did not inherit the mutations present in the proband. At least at this time, the benefit and value of using genotyping to direct genetic counseling and monitoring of asymptomatic family members must be weighed against the possible stigmatization and genetic discrimination that could occur by identifying individuals who are at increased risk for HHC.

Since simple phlebotomy can prevent iron overload and its severe clinical complications, population screening for HHC may be justified. However, given the low penetrance of the *HFE* genotypes, the majority of individuals identified with *HFE* mutations would be unlikely to develop clinical symptoms associated with HHC and yet might be subject to stigmatization and/or genetic discrimination as a result of the testing. At this time, population screening for HHC is not recommended until further details regarding the benefits and risks of screening are available.[4] Requests for prenatal testing for HHC are highly unusual, and such testing would not be ethical because HHC is a treatable adult onset disorder with low penetrance.

Genetic counseling is recommended for *HFE* mutation results to ensure an accurate assessment of the risk of disease. HHC caused by the C282Y and H63D mutations is inherited in an autosomal recessive manner, so the usual risk to siblings of a proband is 25%. However, due to the high carrier frequency of the C282Y mutation in the population (9% to 11%), the possibility exists that the parents of an affected individual can be heterozygous or homozygous for the mutation or that the spouse of an individual with HHC may also be a mutation carrier. The risk assessment should include the most current estimates of the disease penetrance.

Available Assays

As the C282Y and H63D mutations are single nucleotide polymorphisms, they can be accurately detected using a variety of methods. The first method described for these mutations was an oligonucleotide ligation assay (OLA).[1] The primers described by this group to amplify the regions of the *HFE* gene have been used in a variety of assays. However, polymorphisms in one primer binding site sequence has been described[8,9] and has the potential to result in a diagnostic error. Using primers that do not include the polymorphisms eliminates this potential error. The possible interference of the S65C mutation with the detection of H63D should also be considered in assay design. A number of additional assay methods for these mutations have been described, including–polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP), allele-specific PCR (AS-PCR), allele-specific oligonucleotide (ASO) hybridization, single-nucleotide primer extension, and real-time PCR methods (reviewed in Reference 10).

The most common method for detection of C282Y and H63D is PCR-RFLP. C282Y is detected by digestion of the PCR product with *Rsa* I, and H63D is detected by using *Bcl* I or *Mbo* I. The digested PCR products are analyzed by gel electrophoresis. The advantage of this assay is that optimization is relatively straightforward and expensive equipment is not required. The disadvantage is a longer turnaround time for testing. AS-PCR assays distinguish the normal and mutant alleles directly only by additional gel electrophoresis, so the post-PCR processing is reduced (Figure 13-1). The AS-PCR reactions require careful optimization, and multiple reactions are required to determine

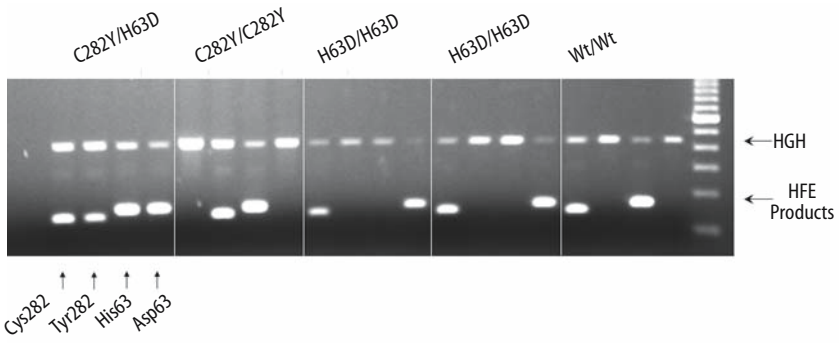


Figure 13-1. Allele-specific PCR assay for Cys282Tyr (C282Y) and His63Asp (H63D) mutations of the *HFE* gene. Each sample is tested with 4 reactions. The first and second reactions detect the Cys282 and Tyr282 mutations, respectively. The third and fourth reactions detect His63 and Asp63 mutations, respectively. Each reaction contains a control PCR product generated from the human growth hormone gene (HGH). Examples of C282Y/H63D compound heterozygous, C282Y homozygous, and two H63D homozygous genotypes are shown.

the genotype. Real-time PCR instruments, such as the Lightcycler (Roche Diagnostics, Indianapolis, IN), couple PCR with fluorescent hybridization probe analysis for mutation detection. A melting curve of the hybridization probe fluorescence can be used to detect changes in thermal stability and therefore discriminate single base mutations in a single reaction. An additional advantage of the real-time PCR is that new sequence variations in the probe regions are also detected. For example, the probe for H63D also detects the S65C mutation.[3] The technique has a wide dynamic range, requires no additional processing, and is rapid. A typical instrument cycle takes 40 minutes. The high cost of the equipment is the main disadvantage. For each assay technique, the appropriate negative and positive controls are required to demonstrate the specificity of the assay.

Interpretation of Test Results

The analytic sensitivity for detecting these mutations should be >99%. False positives and false negatives are rare. However, sequence variations in primers, probes, and restriction sites have the potential to lead to a diagnostic error. A polymorphism reported at the 3′ end of one of the primers described in the original publication[1] was reported to prevent the amplification of wild-type C282Y alleles with some assay methods.[8] Interpretations should be made in the context of all clinical information and account for the reduced penetrance and variable expression observed with HHC. Although genotyping may confirm a diagnosis of HHC in the presence of other signs of iron overload, the utility as a predictive indicator of HHC is less certain. In addition, genotyping of affected individuals within a pedigree may indicate the presence of other mutations in *HFE* or other genes. At this time, genotype alone cannot be used to diagnose, predict, or rule out a diagnosis of HHC. Future studies on the penetrance of these mutations and the benefits of screening for HHC risk should help to clarify these issues.

Laboratory Issues

Proficiency testing for C282Y and H63D mutations in the *HFE* gene is available through the College of American Pathologists (CAP) Molecular Genetics Laboratory (MGL) survey. Cell-line controls can be purchased from Coriell Cell Repositories. (http://ccr.coriell.org/). At the present time, commercial analyte specific reagents (ASRs) are available from Third Wave Technologies, Inc. (Madison, WI), Bio-Rad Laboratories (Hercules, CA), and Nanogen Inc. (San Diego, CA). Due to patents on the *HFE* gene that are currently held by Bio-Rad Laboratories, licensing may be required to perform this test.

HEMOGLOBINOPATHIES

Molecular Basis of Disease

Hemoglobinopathies are the most common single-gene diseases in the world; approximately 5% of the world's population is a carrier for one of the hemoglobin disorders. The extensive study of these disorders has led to much of our knowledge regarding protein structure and function, mutational mechanisms, gene expression during development, and evolution. Many extensive reviews of these disorders are available.[11–13] A list of the numerous characterized mutant alleles is available at the Globin Gene Server (http://globin.cse.psu.edu) and the Human Gene Mutation Database (HGMB) (http://www.uwcm.ac.uk/uwcm/mg/hgmd0.html).

Hemoglobin (Hb) functions as an oxygen carrier in vertebrate red cells. The protein is a tetramer consisting of two α and two β chains. Each subunit contains a heme prosthetic group that is responsible for the oxygen-binding capacity of the protein. These chains are highly homologous to one another at the primary sequence level as well as in three-dimensional structure. Six different globin chains (α, β, γ, δ, ε, ζ) are found in normal human hemoglobins at various times during development.[11,12] The α and β chains are encoded by separate genetic loci (Figure 13-2). The α and α-like chains (ζ, Ψζ, Ψα) are clustered on chromosome 16. In addition, there are two copies of the α-globin genes, α_1 and α_2, on each chromosome 16. The β and β-like chains (ε, Gγ, Aγ, Ψβ, δ) are clustered on chromosome 11.

The developmental regulation of these genes is important to the understanding of many of the hemoglobin disorders. The gene order in each cluster is identical to the order of expression during development. In each

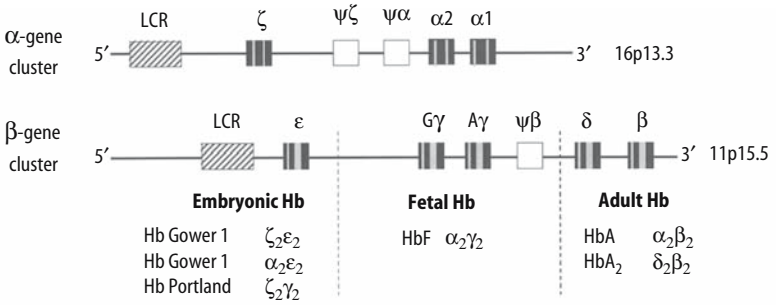


Figure 13-2. Genomic organization and developmental expression of the α- and β-globin gene clusters. The boxes identify the locus control region (LCR; diagonal), pseudogenes (white), exons (black), and introns (gray). The developmental time periods for the expression of the β-globin genes are indicated by the dashed lines. (Adapted from Stamatoyannopoulos G, Nienhuis AW. Hemoglobin switching. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varmus H, eds. *The Molecular Basis of Blood Diseases.* Philadelphia: W.B. Saunders, 1994:107–155.)

developmental stage, there is an equimolar production of α and β chains. Very early in the embryonic stage, $\zeta_2\varepsilon_2$ (Hb Gower 1) is expressed. As the α and γ genes become transcriptionally active, $\alpha_2\varepsilon_2$ (Hb Gower 2) and $\zeta_2\gamma_2$ (Hb Portland) are expressed. In the fetal stage, the Hb is predominantly $\alpha_2\gamma_2$ (Hb F). In normal adults, the Hb is about 98% $\alpha_2\beta_2$ (HbA), and the balance is $\delta_2\beta_2$ (HbA_2). In addition to promoters and enhancer elements, the locus control region (LCR) is responsible for developmental expression of these genes. The LCR for the β locus is approximately 20 kb upstream of the ε-globin gene and the LCR for the α locus is approximately 40 kb upstream from the ζ-globin gene. The LCR is required for expression of all the globin genes at the locus.

The difference in gene dosage and expression patterns between and α and β genes is significant. Each diploid genome encodes four copies of the α gene and two copies of the β gene, making the β globin chain more susceptible to mutations. However, since γ globin is expressed during fetal life and shortly after birth, β globin mutations do not exert their deleterious effects during the prenatal period. In contrast, since α chains are expressed during fetal and postnatal life, mutations in the α chain can cause severe disease during both developmental stages.

The hemoglobinopathies are commonly divided into three categories: structural variants, thalassemias that are characterized by reduced rates of production of α or β chains, and hereditary persistence of fetal hemoglobin (HPFH) characterized by defective β-chain production compensated by γ-chain production.[11] These categories are not exclusive, as structural variants, which result in unstable hemoglobins, can be associated with a thalassemia phenotype. Although an extensive number of abnormal hemoglobins have been described, this chapter focuses on two of the most common structural variants, sickle hemoglobin (HbS) and HbC, and the more common forms of α- and β-thalassemia.

Sickle hemoglobin results from a single nucleotide substitution (GAG→GTG) that changes codon 6 of the β-globin gene from a glutamic acid to a valine (Glu6Val). Homozygosity for this mutation causes sickle cell disease (SD), an autosomal recessive disorder. SD is a major cause of morbidity and mortality in Africa and in populations with individuals of African descent such as the Mediterranean area, the Middle East, and India.[11] Under the condition of low oxygen tension present in the microvasculature, HbS polymerizes into fibers that cause the erythrocytes to sickle. The poor deformability of the sickle red cells and their defective passage in the microcirculation causes the vasoocclusive events that are the hallmark of this disease. The shortened survival of these red cells leads to a chronic hemolytic anemia. The frequency of SD is 1 in 600 in African Americans; the heterozygote frequency in African Americans is approximately 8%, but the frequency in Africa can be much higher.[12] The frequency of the HbS allele, as well as alleles causing thalassemia, is maintained by their protective effect against malaria. The heterozygous state, sickle cell trait, is clinically normal but may be at risk for vasoocclusive events when exposed to low oxygen tension such as flying at high altitude in an airplane with reduced cabin pressure.

The sickling disorders also include compound heterozygous states with the HbS mutation in association with another α- or β-globin variant such as HbC, or α- or β-thalassemia. Similar to HbS, HbC also has an amino acid substitution at position 6 of the β-globin gene (Glu6Lys). HbC is less soluble than HbA and leads to reduced deformability of the red cells and a mild hemolytic anemia. The allele is frequent in persons of African descent; approximately 3% of African Americans are carriers.[12] Individuals with $\beta^S\beta^C$ have SC disease. SC disease is a milder hemolytic disorder than SD and therefore may go unrecognized until a serious complication occurs.

The thalassemias result from the reduced synthesis or stability of either the α- or β-globin chain. The chains in excess due to the α:β imbalance precipitate within the red blood cells, leading to membrane damage and red cell destruction. The defect produces a hypochromic, microcytic anemia. The population distribution of thalassemia includes the Mediterranean, Middle East, portions of Africa, India, and Southeast Asia; however, the migration of these populations has resulted in the worldwide occurrence of these diseases.[11,12]

The most common forms of α-thalassemia are caused by deletion. The genetics of α-thalassemia are complicated by the fact that each normal chromosome has two α genes and that there are haplotypes with either the loss of one (-α) (α^+ thalassemia) or both (—) (α^o thalassemia) copies of the α gene. In addition to the normal genotype, four additional genotypes are possible: silent carrier (αα/-α), α-thalassemia trait (-α/-α or –/αα), HbH (β_4) disease (-α/–), and hydrops fetalis, Hb Barts: γ_4 (–/–). Individuals with HbH disease have a moderately severe hemolytic anemia, while the inheritance of no copies of the α gene (Hb Barts) is incompatible with life. Carriers of α-thalassemia trait have only a mild hypochromic microcytic anemia, but couples with these genotypes are at risk for having a fetus with hydrops fetalis or a child with HbH disease. The (—) genotype, and therefore the risk of hydrops fetalis, is largely restricted to Southeast Asia, while in other groups the (-α) genotype is more common.[11,12] Distinguishing these two carrier states is important for accurate genetic counseling. The most widely occurring single α-globin gene deletions are $-\alpha^{3.7}$ and $-\alpha^{4.2}$. The double α-globin gene deletions $--^{SEA}$, $--^{FIL}$, and $--^{THAI}$ are common in Southeast Asia while $--^{MED}$ and $-(\alpha)^{20.5}$ are more common in the Mediterranean area. Nondeletion α-thalassemia alleles are rare, but the $\alpha^{Constant\ Spring}$, a mutation in the termination codon of the α_2 gene, is frequently present in Southeast Asia.

The β-thalassemias are mainly inherited in an autosomal recessive manner, although there are some dominant forms. Heterozygotes are asymptomatic but with recognizable hematologic parameters including an elevated HbA_2. Homozygotes develop a severe anemia within the first year of life as the switch from γ to β chains occurs. In contrast to α-thalassemia, β-thalassemias are heterogeneous at the molecular level. The molecular mechanism of the mutations is equally varied. Most mutations are single-nucleotide substitutions or frameshift mutations. More than 200 disease-causing mutations have been identified in the β-globin gene; however, within each at-risk population there is a set of four to ten mutations that account for the majority of disease-causing alleles (Table 13-2).[13] HbE is a β-globin variant (Glu26Lys) that can cause a mild thalassemia phenotype. It is one of the most common Hb variants, and the frequency is high in Southeast Asia. Although HbE homozygotes are asymptomatic, genetic compound heterozygotes with another β-thalassemia allele will have an abnormal phenotype. Although deletions are rare, a 619 base pair (bp) deletion involving the 3′ end of the β-globin gene is common in India and Pakistan, accounting for approximately 30% of the β-thalassemia alleles.[11,12] β-thalassemia also may result from deletion of part of the globin gene cluster (e.g., δβ-thalassemia) or from deletions that start 50 to 100 kilobases (kb) upstream from the globin gene cluster and extend 3′ into the cluster (εγδβ-thalassemia). Some of these deletions lead to HPFH. In addition, there are fusion chain variants such as Hb Lepore, a $\delta\beta^+$ thalassemia. Mutations either completely (β^o-thalassemia) or partially (β^+-thalassemia) abolish β-chain production. Milder phenotypes have been recognized that can be explained at the molecular level by homozygosity or compound heterozygosity for mild or silent mutations, coinheritance of α-thalassemia, or coinheritance of a genetic determinant that increases the production of the γ chain.

Clinical Utility of Testing

Sickle cell disease and α- and β-thalassemias and their carrier states can be diagnosed using standard red-cell indices such as Hb, red blood cell number, mean corpuscular volume (MCV), and red cell distribution (RDW) along with Hb electrophoresis or high-performance liquid chromatography (HPLC). The main application of molecular testing is for prenatal diagnosis. When a carrier is identified by abnormal hematologic analysis, the carrier's partner can be screened for carrier status. Molecular testing can be used to identify the disease-causing mutations. Prenatal molecular testing can be offered to couples whose gene mutations have been identified. Testing also can be used to identify mutations in partners with borderline or normal hematologic values when one partner is clearly a carrier, since these couples are at risk for having a fetus with thalassemia. This testing model has been successfully used in population screening to reduce the incidence of thalassemia. In states where hemoglobinopathies are included in the newborn screen, α-thalassemia trait can be detected by the presence of Hb Barts. Since α-thalassemia trait can be asymptomatic or the microcytic

Table 13-2. Population-Specific Mutations in β-Thalassemia

Population	Percentage of Patients	Most Common Mutations
Mediterranean	91–95%	-87 C→G, IVS1-1 G→A, IVS1-6 T→C, IVS1-110 G→A, cd 39 C→T, IVS2-745 C→G
Middle East		Cd 8 –AA, cd 8/9 +G, IVS1-5 G→C, cd 39 C→T, cd 44–C, IVS2-1 G→A
Indian		-619 bp deletion, cd 8/9 +G, IVS1-1 G→T, IVS1-5 G→C, 41/42 –TTCT
Thai		-28 A→G, 17 A→T, 19 A→G, IVS1-5 G→C, 41/42–TTCT, IVS2-654 C→T
Chinese		-28 A→G, 17 A→T, 41/42–TTCT, IVS2-654 C→T
African/African American	75–80%	-88 C→T, –29 A→G, IVS1-5 G→T, cd 24 T→A, IVS11-949 A→G, A→C

Source: Reprinted with permission from Cao A, Galanello R. Beta-thalassemia. GeneReviews [database online]. Seattle, WA: University of Washington. Updated March 18, 2003.

anemia can be mild, molecular testing can help to differentiate trait from iron deficiency. For α-thalassemia, prenatal diagnosis predominantly is used in situations where both members of the couple are carriers of the double α-gene deletion chromosome, as only these couples are at risk for having a fetus with hydrops fetalis. Although the α^o-gene is found almost exclusively in the Southeast Asian population, molecular testing can be used to distinguish the two α-thalassemia carrier states (-α/-α vs –/αα). In addition to prenatal testing for β-thalassemia, molecular testing can be useful for the prediction of the clinical phenotype by identifying mild and silent alleles and for presymptomatic diagnosis of at-risk family members. Molecular testing may also be useful in recognition of complex heterozygotes in populations where multiple hemoglobinopathies are common.

Available Assays

HbS and HbC are caused by single-base substitutions, so a number of assay formats are suitable as long as the mutations can be distinguished from one another. One of the most common methods is PCR-RFLP.[14] Modified primers are used to introduce restriction sites so the β^A, β^S, and β^C alleles can be identified using *Ava* I and *Sty* I. A real-time PCR method using hybridization probes also has been described that can readily identify the β^E allele in addition to the β^A, β^S, and β^C alleles.[15]

The conventional method for detecting α-gene deletions has been Southern blot using probes to the ζ- and α-globin genes. The ζ-globin probe is useful for detecting deletion breakpoint fragments because the α-globin probe will fail to hybridize to a fragment in an α^o-thalassemia fetus.[16] A number of PCR-based methods have been described that detect the most common α^o and α^+ deletions.[17] Since these deletions either partially or completely remove the α2 globin gene, its detection by PCR indicates heterozygosity when a deletion allele is present. The PCR assays are faster, less expensive, and easier to interpret than Southern blot, although the PCR has to be carefully optimized to faithfully amplify the GC-rich sequences at this locus.

Since the mutations causing β-thalassemia are mainly single-nucleotide substitutions, a number of assays using ASO hybridization, AS-PCR, PCR-RFLP,[18] and real-time PCR with hybridization probes have been described.[19] These assays target population-specific mutations. If mutations are not identified, then alternate methods to detect other mutations can be used. The size of the β-globin gene, about 1.6 kb containing three exons, is amenable to DNA sequencing. The advantage of DNA sequencing is that virtually all β-globin mutations can be detected. Unknown mutations also have been identified using denaturing high-performance liquid chromatography (DHPLC). The 619 bp deletion involving the 3′ end of the β-globin gene that is common in India and Pakistan is detectable by PCR or Southern blot analysis.

Interpretation of Test Results

Hematologic screens are more than 95% accurate, with occasional problems detecting carriers with mild phenotypes. The analytic sensitivity for detecting gene mutations should be greater than 99%. False positives and false negatives are rare. However, sequence variations in primers, probes, and restriction sites have the potential to lead to a diagnostic error. The mutation resulting in HbC occurs at 1 bp from the HbS mutation. There is the potential for a diagnostic error with a test that does not clearly distinguish the two mutations.

The mutation detection rate for α- and β-thalassemia will depend on the mutations in the panel and the patient's ethnicity. For α-thalassemia, greater than 95% of the mutations are deletions; the majority of these deletions can be detected in the current assays. For β-thalassemia, if a mutation is not detected in a patient with trait, then DNA sequencing should be considered to identify the mutation. The detection rate using DNA sequencing for β-thalassemia is approximately 99% depending on the actual coverage of the sequence. Deletions larger than the PCR product size or ones that abolish the PCR primer sites will not be detected by DNA sequencing. When interpreting the molecular analysis in relation to the clinical phenotype, the possible coinheritance of genetic determinants that could influence the phenotype must be considered. For example, inheritance of an α-thalassemia allele, silent or mild β-thalassemia allele, HPFH allele, or HbS allele would modify the phenotype. Family studies are useful in sorting out complex genotypes.

As most of the molecular analysis is done for prenatal diagnosis, the possibility of maternal contamination should be considered in interpreting the molecular test result. The contamination with maternal cells of both direct and cultured amniotic fluid and chorionic villus sampling (CVS) is well documented and therefore represents a potential source of error in prenatal diagnosis. Prenatal samples should be tested in parallel with a maternal sample, or both samples should be used to perform a separate analysis for maternal contamination using microsatellite markers, to rule out error due to maternal cell contamination. Laboratories should understand how their testing methods are affected by the presence and the amount of maternal cell contamination. The mutation status of one or both parents, as appropriate, should be tested prior to testing of fetal specimens, preferably within the same laboratory.

Laboratory Issues

Proficiency testing is available for HbS and HbC through the CAP MGL survey. Proficiency testing for α- and β-thalassemia usually is accomplished through a sample exchange with another laboratory that performs the testing. DNA and cell-line controls for many of the hemoglo-

binopathies can be purchased from Coriell Cell Repositories (http://ccr.coriell.org/). Commercial kits for the HbS, HbC, and the common α- and β-thalassemia mutations are available from Bio-Rad Laboratories (Hercules, CA).

RH INCOMPATIBILITY

Molecular Basis of Disease

The Rh antigens are expressed on proteins encoded by two distinct but highly homologous genes, *RHD* and *RHCE*, on chromosome 1p34.3-36.1.[20] The *RHD* and *RHCE* genes are arranged in tandem and likely arose through duplication of a single ancestral gene. The D antigen is expressed from *RHD* and the C/c and E/e antigens are expressed from *RHCE*. The most common *RHD*-negative allele results from the total deletion of *RHD*; however, there are a number of other sequence variations that may result in the loss of expression of the D antigen. Sequence homology and complementation experiments in yeast suggest that the RH proteins are involved in ammonium transport; however, the Rh antigen system is clinically important because antibodies to Rh antigens are involved in hemolytic transfusion reactions, autoimmune hemolytic anemia, and hemolytic disease of the newborn (HDN).

HDN results from the alloimmunization of a mother to a paternally inherited fetal alloantigen after transplacental hemorrhage, unrecognized miscarriage, or invasive procedures. In subsequent pregnancies, these antibodies may destroy the red blood cells of an antigen-positive fetus, leading to hemolytic disease. HDN can occur when fetomaternal incompatibilities exist within any of the different red cell antigen systems, including the RHD, C/c, E/e, Kell, Kidd, Duffy, and M antigen systems. RHD accounts for the majority of HDN cases. The severity of HDN is variable; mild cases require either no treatment or phototherapy, while more severe cases require exchange transfusion at birth, or possibly intrauterine transfusion, and may result in fetal hydrops.[21] Typically, women who are at risk for alloimmunization are screened using antibody titers, but the screen does not accurately predict the severity of hemolytic disease. Serial amniotic fluid ΔOD 450 readings also can be used to estimate the severity of HDN, but this investigation brings not only the additional risks from multiple amniocentesis procedures, but also the risk of further sensitization to the fetus. Using molecular techniques, the Rh status of the fetus can be investigated prenatally to determine whether the fetus is positive for the *RHD* gene and therefore is at risk for HDN. Further invasive diagnostic procedures can be avoided in the RHD-negative fetus of an RHD-negative mother.

The *RH* genes, which are more than 95% homologous at the nucleotide sequence level, both consist of ten exons spanning over 75 kb, and both encode peptides of 417 amino acid residues with a predicted molecular mass of 30 to 35 kilodaltons (kDa)[20] (Figure 13-3). There are a number of sequence variations that can be used to distinguish these two highly similar genes. However, it is important to be aware of a number of variant alleles that exist in this genetic system as the result of gene conversion events between the *RHD* and *RHCE* genes and from point mutations. Databases of these variants are available: Blood Group Antigen Gene Mutation Database (http://www.ncbi.nlm.nih.gov/projects/mhc/xslcgi.fcgi?cmd-bgmut/systems-infoxsystem-rh) and the Rhesus Site (http://www.uni-ulm.de/~wflegel/RH/). These hybrid *RHD-CE-D* alleles may result in either *RHD*-positive or *RHD*-negative haplotypes. Most variant *RHD*-positive alleles encode proteins that do not express all the RHD epitopes. Although an

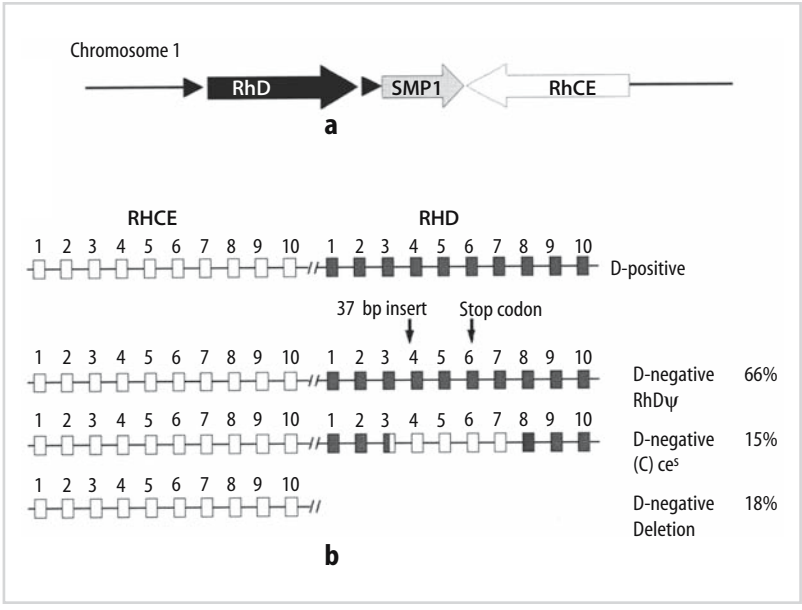


Figure 13-3. (a) Genomic organization of the *RHD* and *RHCE* genes on the short arm of chromosome 1. The *RHD* gene is flanked by two highly homologous Rhesus boxes (black arrowheads). (b) The most common *RHD*-positive and *RHD*-negative haplotypes in Africans. The orientation of the genes is altered to clarify the presentation. (Adapted from Singleton BK, Green CA, Avent ND, et al. The presence of an RHD pseudogene containing a 37 base pair duplication and a nonsense mutation in Africans with the Rh D-negative blood group phenotype. *Blood* 2000;95:12–18. Copyright American Society of Hematology. Used with permission.)

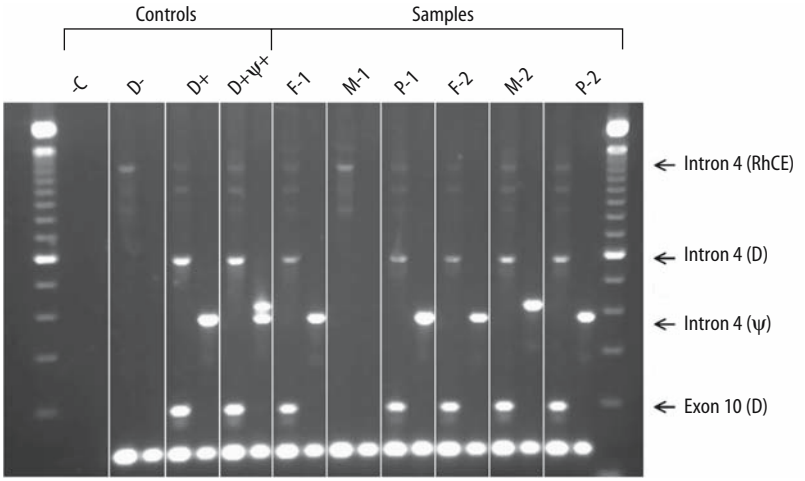


Figure 13-4. *RHD* multiplex genotyping assay. Each sample is tested with two reactions. The first reaction detects exon 10 and intron 4. The second reaction detects the *RHD* gene and the 37bp insertion present in exon 4 of *RHDΨ*. Each reaction contains a control PCR product generated from exon 7 (lowest band in each lane (not marked). The controls are: negative control (-C), *RHD*-negative (D-), *RHD*-positive (D+) and *RHD*-positive/ *RHDΨ*-positive samples. The samples are sets of fetal (F), maternal (M) and paternal (P) samples. In each example, the fetus is *RHD*-positive, the maternal sample is *RHD*-negative and the paternal sample is *RHD*-positive. Note that in M-2, the *RHDΨ* is present. If *RHDΨ* was not specifically detected, the maternal sample would appear to be *RHD*-positive and the genotype of the fetus would not be conclusive.

RHD-negative mother can be alloimmunized by a partial D antigen, hemolytic disease in these cases is rare. The frequency of these variant alleles is low in the white population, but in some ethnic groups these alleles can be common.[20,22] The DAR allele is found in 17% of South African blacks, and DIIIa has been found in 18% of blacks in New York and 28% of blacks from Brazil.

The most common *RHD*-negative allele results from the total deletion of *RHD*; however, there are a number of *RHD*-negative haplotypes that are positive for at least portions of the *RHD* gene (Figure 13-4). In the African population, there are two *RHD*-positive, antigen D-negative alleles that are common and must be considered in any genotyping strategy. *RHDΨ* is present in 66% of RHD-negative Africans and contains a 37 bp insertion at the junction of intron 3 and exon 4 that disrupts the reading frame and leads to a stop codon in exon 6.[23] The Cde^s allele, found in 15% of RHD-negative Africans, is an *RHD-CE-D* hybrid containing exons 1 and 2, part of exon 3, and exons 9 through 10. In addition, a significant percentage of RHD-negative Asians (27%) are positive for the *RHD* gene.

The Rh C/c and E/e antigens also can cause HDN; however, severe cases are usually observed due to anti-c. The Ser103Pro encoded by a 307T→C substitution in exon 2 is responsible for the C/c polymorphism. Typing for the *RHC* allele is straightforward, but the identification of the *RHC* allele is complicated by the fact that exon 2 of the *RHC* allele is identical to exon 2 of *RHD*. A 109 bp insert found in intron 2 of the *RHC* allele can be used for genotyping;[22] however, the Cde^s allele will give a false-negative result with this assay since the intron 2 is of *RHD* origin in the allele found in African Americans. The nt48 cytosine polymorphism in exon 1 (nt 48) is also linked to the *RHC* allele. This polymorphism has a 5% false-positive rate in whites, but in African Americans the false-positive rate is greater than 50%, making this unsuitable for genotyping. The RhE and Rhe antigens are the result of the Pro226Ala polymorphism encoded by the 676C→G substitution in exon 5 of the *RHCE* gene. Rare allelic variants involving exon 5 have been described that would lead to erroneous typing of the E/e polymorphism.

Clinical Utility of Testing

Prenatal blood group determination is routine in pregnancies where the woman has a potentially clinically significant alloantibody and the father is phenotypically heterozygous for the gene encoding the corresponding antigen. Ideally, determining the zygosity of the father allows prediction of the risk for the fetus to inherit the alloallele. A fetus of a heterozygous father has a 50% chance of inheriting the paternal alloallele to which the mother may be immunologically sensitized, while the risk is 100% from a homozygous father. Determining zygosity is straightforward by serological or molecular methods in most antigen systems except for RhD. The molecular basis of the *RHD* deletion has been described for the most common *RHD*-negative haplotype, making it possible to detect the *RHD* deletion.[24] However, in addition to nonpaternity, determination of the zygosity with 100% certainty is difficult due to the allelic heterogeneity of the *RHD*-negative haplotype, so testing of the fetus directly for *RHD* status is important. A fetus that is positive for the gene encoding the alloallele is then identified as a fetus at risk for HDN and can be monitored appropriately. A fetus without the gene is not at risk for HDN within the limits of the assay's sensitivity, and so invasive procedures such as serial amniocentesis, that risk further sensitization, can be avoided.

Genotyping assays for *RH* also can be used to help resolve discrepancies or unusual phenotypic results. Many of the variant alleles display altered phenotypic characteristics that are identified in the immunohematology laboratory or blood bank. Discrepancies in RhD status can occur due to a variety of serological reagents and techniques. Presence of the *RHD* gene can be used to confirm the RhD-positive status of a patient. An RhD-positive individual may develop anti-D after transfusion. Molecular assays can sometimes be used to confirm a partial D phenotype. There are a number of examples in the literature describing the loss or weakening of RhD expression, particularly in cancer patients.[20] In some instances this has been caused by deletion of the *RHD* gene or loss or abnormalities involving chromosome 1. Molecular assays can confirm the loss of genetic material encoding the *RHD* gene.

Available Assays

PCR-based test methods are the most commonly used, since the testing is used for prenatal diagnosis. The assay formats are predominantly AS-PCR, which are well suited for detecting single base pair polymorphisms. Testing for two or more regions of the *RHD* gene is required for identification of allelic variants. Assays must distinguish the *RHD* gene from *RHCE*. A number of assays have been described using the various sequence differences between the genes in exons 3 to 7 and 9, introns 4 and 7, and the 3′ UTR of exon 10.[22,25] Most of the published assays use a combination of amplification of intron 4, exon 7, and exon 10. The use of exon 10 alone is not considered safe because of possible false positives and false negatives; however, this polymorphism is useful in recognizing the Cde^s haplotype. Detection of the 37 bp insertion found in exon 4 of *RHDΨ* also should be included in all genotyping assays.[23] The precise site of the *RHD* deletion has been defined, making detection of the deletion possible with a PCR-based assay.[24] Although this is the most common *RHD*-negative allele in whites and African Americans, this assay will not correctly determine zygosity in the presence of Cde^s or *RHDΨ*. Real-time PCR assays potentially can be used for determining zygosity, as long as more than one region of the *RHD* gene is detected and the assay design takes into account possible genetic variants.

Molecular assays for *RHE/e* take advantage of polymorphisms within exon 5 that are specific to the *RHCE* gene.[22] *RHc* is detected directly using the C307 polymorphism, while the 109 bp intron 2 insertion is used to detect the *RhC* allele. Multiplex genotyping assays for *RHD* and *RHCc* also have been described.[23] Recently, a genotyping method for *RHC* was developed that improves the detection of *RHC* in individuals with the Cde^s haplotype.[22]

These methods are suitable for testing fetal DNA as well as genomic DNA from blood. Fetal DNA sources are typically amniotic fluid or chorionic villi, using either direct or cultured samples. Since fetal DNA preparations can be potentially contaminated with maternal DNA, laboratories must understand how maternal contamination will affect the assay results and have methods to determine the extent of contamination. Testing of variable number of tandem repeats (VNTR) or microsatellite loci in fetal and maternal DNA samples can be used to determine whether maternal cells contaminate the fetal sample.

Another potential source of fetal DNA is maternal plasma. Fetal DNA accounts for approximately 3% of the total cell-free DNA in maternal plasma during the second trimester and increases throughout pregnancy. Since this method avoids invasive techniques that cause risk to the fetus, and prevents further sensitization of the mother, the method is of significant interest. Quantitative real-time PCR methods facilitate molecular testing using fetal DNA from maternal serum. Several studies have demonstrated that fetal DNA in maternal plasma can be used to determine a fetal genotype.[26] However, caution is recommended in using this potentially exciting technology for prenatal testing.[27] The fetal DNA represents a minority of the total DNA, so use of an internal control is important to demonstrate that fetal DNA is present and detectable. Y-chromosome-specific sequences can be used with a male fetus, but other paternal-specific polymorphisms may be required to provide the necessary internal control.

Interpretation of Test Results

By using a multiplex strategy, the sensitivity of testing will be close to 100%, the false-positive rate will be quite low and the positive predictive value close to 1, but these performance characteristics may vary with different testing strategies in different ethnic groups. There is one rare *RHD* allele that would elude most genotyping assays; DHar contains only exon 5 of the *RHD* gene (*RHCE-D$_5$RHCE*). This allele may also cause a false negative for the Rhe antigen. The rate of false-positive results and positive predictive values for different PCR strategies in the European population has been described.[25] However, the diversity of *RHD* alleles is much greater in other ethnic groups, especially Africans. Currently, the majority of variant alleles are accounted for by *Cde^s* and *RHDΨ*. In Asian populations, where the *RHD*-negative haplotype is uncommon, the *RHD* gene has been detected at a frequency of 27.7% among RhD-negative Japanese donors. The major allele appears to be associated with a missense mutation.

These assays should be interpreted in the context of all available information. When conducting prenatal *RH* genotyping, the laboratory must identify discrepancies between parental serotypes and genotypes caused by variant alleles to reduce the risk of false-negative and false-positive results. Variants easily will be identified in RhD-negative mothers that test positive for the *RHD* gene. If the mother is positive for only some parts of the *RHD* gene, it still may be possible to tell if her fetus inherited a paternal *RHD* gene. Allelic variants can be masked in paternal samples by a "normal" *RHD* allele. A masked variant allele could be passed on to the fetus, where it may be undetected without a thorough analysis of the fetal *RHD* gene. If the masked allele is detected, the genotype may not allow certain prediction whether the fetus will have an RhD-positive phenotype. As discussed earlier, allelic variants of the *RHC* are common in individuals of African ancestry. In African Americans, the *Cde^s* allele will result in a 15% to 20% false-negative result since intron 2 is of *RHD* origin. Clearly, in this ethnic group, testing to detect the presence of the paternal *RHC* allele is important, especially in a C-negative fetus. Other rare alleles in the *RH* system may be misidentified in DNA typing assays. When maternal and paternal samples are not available, interpretation should include the possibility of a false-positive or false-negative result.

Laboratory Issues

Rh incompatibility testing requires both phenotypic and genotypic testing of parental samples. Laboratories need to work in conjunction with an immunohematology laboratory, which is able to provide phenotypic data, assist in the analysis of variant alleles, and provide assay controls. *RHD* controls also are available from Coriell Cell Repositories (http://ccr.coriell.org/). The reagents for these tests as well as the tests are typically laboratory developed and validated. Proficiency testing for *RHD* is available in the CAP MGL survey; otherwise, proficiency testing can be accomplished through a sample exchange with another laboratory that performs the testing. In addition, the phenotypic data on each blood sample help identify any proficiency problems in the genotyping assays.

References

1. 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.
2. Hanson EH, Imperatore G, Burke W. HFE gene and hereditary hemochromatosis: a HuGE review. Human Genome Epidemiology. *Am J Epidemiol.* 2001;154:193–206.
3. Lyon E, Frank EL. Hereditary hemochromatosis since discovery of the HFE gene. *Clin Chem.* 2001;47:1147–1156.
4. Imperatore G, Pinsky LE, Motulsky A, Reyes M, Bradley LA, Burke W. Hereditary hemochromatosis: perspectives of public health, medical genetics, and primary care. *Genet Med.* 2003;5:1–8.
5. Camaschella C, Roetto A, Cali A, et al. The gene TFR2 is mutated in a new type of haemochromatosis mapping to 7q22. *Nat Genet.* 2000;25:14–15.
6. Njajou OT, Vaessen N, Joosse M, et al. A mutation in SLC11A3 is associated with autosomal dominant hemochromatosis. *Nat Genet.* 2001;28:213–214.
7. Burke W, Imperatore G, McDonnell SM, Baron RC, Khoury MJ. Contribution of different HFE genotypes to iron overload disease: a pooled analysis. *Genet Med.* 2000;2:271–277.
8. Somerville MJ, Sprysak KA, Hicks M, Elyas BG, Vicen-Wyhony L. An HFE intronic variant promotes misdiagnosis of hereditary hemochromatosis. *Am J Hum Genet.* 1999;65:924–926.
9. Beutler E, Gelbart T. A common intron 3 mutation (IVS3-48c→g) leads to misdiagnosis of the c.845G→A (C282Y) HFE gene mutation. *Blood Cells Mol Dis.* 2000;26:229–233.
10. Pointon JJ, Merryweather-Clarke AT, Carella M, Robson KJ. Detection of C282Y and H63D in the HFE gene. *Genet Test.* 2000;4:115–120.
11. Weatherall DJ, Clegg JB, Higgs DR, Wood WG. The hemoglobinopathies. In: Scriver CR, Beaudet AL, Sly W, Valle D, eds. *The Metabolic and Molecular Basis of Inherited Disease.* 8th ed. New York: McGraw-Hill; 2001:4571–4636.
12. Gay JC, Phillips JA III, Kazazian HH. Hematoglobinopathies and thalassemias. In: Rimoin DL, Connor JM, Pyeritz RE, eds. *Emery and Rimoin's Principles and Practices of Medical Genetics.* New York: Churchill Livingstone; 1997:1599–1626.
13. Cao A, Galanello R. Beta-thalassemia. GeneReviews [database online]. Seattle, WA: University of Washington. Updated March 18, 2003.
14. Hatcher SL, Trang QT, Robb KM, Teplitz RL, Carlson JR. Prenatal diagnosis by enzymatic amplification and restriction endonuclease digestion for detection of haemoglobins A, S and C. *Mol Cell Probes.* 1992;6:343–348.
15. Hermann MG. Genotyping β-globin mutations (HbS, HbC, HbE) by multiplexing probe color and melting temperature. In: Meuer S, Wittwer CT, Nakagawara K, eds. *Rapid Cycle Real-Time PCR: Methods and Applications.* Heidelberg: Springer-Verlag; 2001:119–125.
16. Kattamis AC, Camaschella C, Sivera P, Surrey S, Fortina P. Human alpha-thalassemia syndromes: detection of molecular defects. *Am J Hematol.* 1996;53:81–91.
17. Tan AS, Quah TC, Low PS, Chong SS. A rapid and reliable 7-deletion multiplex polymerase chain reaction assay for alpha-thalassemia. *Blood.* 2001;98:250–251.
18. Tuzmen S, Schechter AN. Genetic diseases of hemoglobin: diagnostic methods for elucidating beta-thalassemia mutations. *Blood Rev.* 2001;15:19–29.
19. Moreno I, Bolufer P, Perez ML, Barragan E, Sanz MA. Rapid detection of the major Mediterranean beta-thalassaemia mutations by real-time polymerase chain reaction using fluorophore-labelled hybridization probes. *Br J Haematol.* 2002;119:554–557.
20. Avent ND, Reid ME. The Rh blood group system: a review. *Blood.* 2000;95:375–387.
21. Bowman JM. Immune hemolytic disease. In: Nathan DG, Oski FA, eds. *Hematology of Infancy and Childhood.* Philadelphia: W.B. Saunders; 1998:53–78.
22. van der Schoot CE, Tax GH, Rijnders RJ, De Haas M, Christiaens GC. Prenatal typing of Rh and Kell blood group system antigens: the edge of a watershed. *Transfus Med Rev.* 2003;17:31–44.
23. Singleton BK, Green CA, Avent ND, et al. The presence of an RHD pseudogene containing a 37 base pair duplication and a nonsense mutation in Africans with the Rh D-negative blood group phenotype. *Blood.* 2000;95:12–18.
24. Wagner FF, Flegel WA. RHD gene deletion occurred in the Rhesus box. *Blood.* 2000;95:3662–3668.
25. Wagner FF, Frohmajer A, Flegel WA. RHD positive haplotypes in D negative Europeans. *BMC Genet.* 2001;2:10.
26. Finning KM, Martin PG, Soothill PW, Avent ND. Prediction of fetal D status from maternal plasma: introduction of a new noninvasive fetal RHD genotyping service. *Transfusion.* 2002;42:1079–1085.
27. Stamatoyannopoulos G, Nienhuis AW. Hemoglobin switching. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varmus H, eds. *The Molecular Basis of Blood Diseases.* Philadelphia: W.B. Saunders; 1994:107–155.

Chapter 14

Cardiovascular Disease

Anant Khositseth and Michael J. Ackerman

Introduction

During the last decade, techniques and advances in molecular genetics and genomics have yielded profound new insights into the fundamental mechanisms and genetic underpinnings for many heritable cardiovascular diseases. The resulting genotype-phenotype correlations facilitate: (1) molecular testing for the preclinical/presymptomatic identification of genetically susceptible individuals, (2) the possibility of gene-based prognosis, and (3) new opportunities for gene-specific or gene-targeted therapy including primary prevention in genotype-positive–phenotype-negative individuals. Cardiology has embraced new genetic discoveries, since sudden cardiac death (SCD) consumes more lives than any other medical condition in developed countries, with 1,000 SCDs occurring each day in the United States. Coronary artery disease (CAD) is the major cause of SCD, while other heritable processes including cardiomyopathies and the channelopathies may also predispose to fatal ventricular arrhythmias.

In this chapter, we review the current understanding with respect to the molecular pathogenic mechanisms and status of molecular testing for these three principle categories: CAD, cardiomyopathies, and channelopathies that render an individual susceptible to SCD. CAD has many risk factors, some of which have a genetic susceptibility, including lipidemias, factor V Leiden mutation, and homocystinuria. Many cardiomyopathies are heritable diseases, and many chromosomal loci of and specific causal genes have been identified and extensively reported, including hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/ARVD), and restrictive cardiomyopathy. Finally, cardiac channelopathies may be responsible for a significant proportion of sudden, unexpected death associated with a structurally normal (negative autopsy) heart and include congenital long QT syndrome (LQTS), Brugada syndrome, idiopathic ventricular fibrillation (IVF), catecholaminergic polymorphic ventricular tachycardia (CPVT), and Andersen syndrome.

CORONARY ARTERY DISEASE

Atherosclerosis, or coronary artery disease, is a complex process involving the interaction of many different key factors, including lipoprotein metabolism, coagulation, and inflammation. Gene mutations underlying these key determinants may lead to over- or underproduction of essential proteins. Other factors, including hypertension, diabetes, smoking, sleep apnea, and obesity, interact to increase risk. Any single gene mutation will provide only a small or modest contribution to overall risk. Because of the effect of gene-environment interaction, a gene mutation may have little effect on CAD in individuals who have low environmental risk but may exert a major effect in the setting of concomitant high environmental risk.

HYPERLIPIDEMIAS (LIPOPROTEIN METABOLISM)

Molecular Basis of Disease

Goldstein and Brown first reported one of the most important molecular mechanisms leading to atherosclerosis: defects of the low-density lipoprotein (LDL) receptor.[1,2] Autosomal dominant hypercholesterolemia (ADH) is one of the most frequent hereditary disorders of lipid metabolism. It is characterized by an isolated elevation of LDL particles leading to premature death from atherosclerosis. ADH is caused mainly by defects in the genes encoding the LDL receptor (*LDLR*, chromosome 19p13.1–13.3),[3,4] apolipoprotein B (*APOB*, chromosome 2p23–p24),[5,6] and other unidentified gene(s) on chromosome 1p34.1-p32.[7] More than 600 different *LDLR* mutations have been identified throughout the 45 kilobase (kb) gene.[8,9] The *APOB* gene spans 43 kb and is divided into 29 exons, and a few mutations in the *APOB* gene have been identified.[10–13]

Autosomal recessive hypercholesterolemia (ARH) is a very rare recessive disorder caused by mutations in a putative adaptor protein called ARH, which is character-

ized by severe hypercholesterolemia (elevation of LDL level), xanthomatosis, and premature CAD.[14] Three mutations in the *ARH* gene on chromosome 1 were identified exclusively from patients in Sardinia, Italy.[15]

Apolipoprotein E is a constituent of several plasma lipoproteins including chylomicrons (VLDL, IDL, and HDL-E), and may be a risk-stratifying biomarker for CAD. The *APOE* alleles (ε2, ε3, and ε4) have a strong and consistent influence on plasma lipids and coronary artery disease, and the ε4 allele may portend the greatest risk for CAD.[16]

Interleukin-6 (IL6) is an inflammatory cytokine associated with the development and severity of CAD.[17] The 174C allele of the *IL6* gene is associated with risk of CAD and high systolic blood pressure in men.[18]

Endothelial nitric oxide synthase (eNOS) is an enzyme catalyzing the synthesis of NO in vascular endothelium and effects vascular relaxation. Alterations in the NOS pathway leading to decreased generation or action of NO are potential mechanisms in the development of premature atherosclerosis.[19] The gene encoding eNOS is located on chromosome 7.

Clinical Utility of Testing

To date, molecular testing of patients with hypercholesterolemia is not widely used clinically, and continues to be used predominantly for research purposes to identify new genes. A recurring theme throughout this chapter is that ADH, like virtually every heritable cardiovascular disease, is genetically heterogeneous. The identification of new genes and variations related to cardiac disease risk may direct the development of new intervention strategies to limit elevation of LDL particles and prevent morbidity and mortality from premature atherosclerosis.

Available Assays

There are no currently available commercial kits. Polymerase chain reaction–single-strand conformation polymorphism (PCR-SSCP) assays on extracted DNA are performed in research laboratories.[20]

Interpretation of Test Results

Test sensitivity and specificity are not known, and the results are simply reported as the presence or absence of specific genetic variants.

HYPERCOAGULABLE STATE (COAGULATION)

Molecular Basis of Disease

The coagulation system is an essential hemostatic mechanism. Imbalances in this system can lead to arterial thrombosis manifested with thrombotic events and development of myocardial infarction. Factor V (F5), a cofactor of factor Xa, plays an important regulatory role in coagulation. A point mutation in the *F5* gene (designated factor V Leiden or *F5* R506Q), accounts for 21% to 64% of cases of hereditary hypercoagulable state[21] and represents one of the most important risk factors for inherited thrombophilia, venous thrombosis, and pulmonary embolus.

The R506Q missense mutation in *F5* (named factor V Leiden because of identification in Leiden, Netherlands) was discovered in 1994 as the cause for activated protein C (APC) resistance. It is believed to have arisen de novo in northern Europe, and all individuals with factor V Leiden share this common ancestor. This mutation is present in 5% of white Americans, in 20% of cases of idiopathic sentinel deep vein thrombosis (DVT), and in 60% of venous thrombosis cases in pregnant women. Individuals heterozygous (5% in whites) for factor V Leiden have a 7-fold increase in relative risk of venous thrombosis, whereas the risk is approximately 80-fold for the 1 in 400 whites homozygous for factor V Leiden.[22,23]

Besides its strong association with risk of venous thrombosis, the presence of factor V Leiden may render significant genetic susceptibility for the development of arterial thrombosis, atherosclerosis, and myocardial infarction.[24–28] However, several studies have failed to demonstrate these associations.[29–31]

Clinical Utility of Testing

Regarding the clinical utility of factor V Leiden testing for CAD, Dunn and colleagues[32] reported that heterozygosity for factor V Leiden is not independently associated with CAD or myocardial infarction. Recently, 2 large meta-analyses[33,34] reported the role of factor V Leiden in myocardial infarction and found that individuals heterozygous for factor V Leiden have an odds ratio of approximately 1.3 for the development of myocardial infarction, but only one of the reports showed that the association was statistically significant. Another meta-analysis in unselected patients with CAD demonstrated no association with factor V Leiden.[35] However, anecdotal reports have shown that homozygosity for factor V Leiden might have a role in development of atherosclerosis and CAD.[25,36] Presently, screening for factor V Leiden in unselected patients at risk for myocardial infarction does not appear to be indicated. However, three independent studies and several case reports revealed a high prevalence of the factor V Leiden allele in patients with myocardial infarction without signs of coronary atherosclerosis.[37] In this selected group (3% to 10% of all myocardial infarction patients), the analysis of factor V Leiden genotype might help with risk assessment. However, there is a growing consensus that factor V Leiden (R506Q) mutational testing represents the standard of care in the evaluation of any venous thrombosis in many settings (see chapter 12 for details).[22]

Available Assays

There are a variety of PCR-based assays including restriction fragment length polymorphism (PCR-RFLP), temperature-modulated heteroduplex analysis by denaturing high-performance liquid chromatography (TMHA-DHPLC), and direct DNA sequencing (see chapter 12).

Interpretation of Test Results

Mutational analysis is sensitive and specific for the presence of factor V Leiden (see chapter 12).

Laboratory Issues

Factor V Leiden (R506Q) mutation testing represents one of the best examples of a functional polymorphism that influences or "modifies" a cardiovascular disease process and is presently a standard and routine molecular test for clinical practice.[22]

HOMOCYSTEINE

Molecular Basis of Disease

Homocysteine is an intermediate formed during the metabolism of methionine, and this metabolite may play a key role in atherosclerosis. Homocystinuria is a rare autosomal recessive disorder characterized by markedly elevated homocysteine levels and excretion of homocysteine in the urine. Patients with homocystinuria can have mental retardation, skeletal abnormalities, and lens dislocations, as well as predisposition to both arterial and venous thromboembolic episodes.[38] In contrast to that in patients with this rare disorder of homocystinuria, the relationship of homocysteine levels in unaffected individuals with premature development of CAD remains unclear. Cleophas and colleagues[39] performed a meta-analysis on the role of homocysteine in CAD and concluded that homocysteine was not an independent risk factor for CAD.

Homocysteine levels are influenced by several factors including folate, vitamin B_6, and vitamin B_{12} intake, as well as genetic factors. Methylenetetrahyodrofolate reductase (MTHFR) is one of the main regulatory enzymes of homocysteine metabolism. The C677T and A1298C mutations in the *MTHFR* gene result in decreased enzymatic activity and contribute to increased homocysteine levels (see chapter 12).[40–42]

Clinical Utility of Testing

Andreassi and colleagues[43] studied the influence of the C677T mutation in the *MTHFR* gene on plasma levels of homocysteine, vitamin B_{12}, and folate in patients with CAD and found an association between the *MTHFR* C677T polymorphism and hyperhomocysteinemia, nutritional deficiencies, and oxidative stress. These findings suggest a plausible molecular mechanism of genetic instability in atherosclerotic lesions.

Available Assays

Genotyping for the *MTHFR* polymorphism, C677T, is performed by PCR amplification and electrophoresis through 6% polyacrylamide gel (see chapter 12).[41]

Interpretation of Test Results

Test sensitivity and specificity are not known.

Laboratory Issues

Both homocysteine levels and molecular testing for the C677T mutation in the *MTHFR* gene are available from clinical laboratories (see chapter 12).

CARDIOMYOPATHIES

HYPERTROPHIC CARDIOMYOPATHY

Hypertrophic cardiomyopathy is a complex cardiovascular disease with a great diversity of morphologic, functional, and clinical features, and affects approximately 1 in 500 individuals.[44] Clinical manifestations are variable, ranging from an asymptomatic or mildly symptomatic course to severe heart failure and SCD.[45] HCM commonly manifests between the second and fourth decades of life but can manifest at the extremes of age as well.[46,47] SCD can be the tragic sentinel event for HCM in children, adolescents, and young adults. Typically, the left ventricle is hypertrophied in the absence of an increased external load (i.e., unexplained hypertrophy) and myocyte hypertrophy, disarray, and interstitial fibrosis are the hallmark histological features of the disease.[48,49]

Molecular Basis of Disease

Hypertrophic cardiomyopathy is a genetic disease, usually with an autosomal dominant mode of inheritance, although autosomal recessive forms and spontaneous germline mutations have been identified (Table 14-1). A familial history of HCM is present in approximately two thirds of all index cases. The first gene for familial HCM was mapped to chromosome 14q11.2–q12 using genomewide linkage analysis in a large Canadian family.[50] Thierfelder and colleagues[51] mapped the second locus to

Table 14-1. Summary of Molecular Basis of Hypertrophic Cardiomyopathies

Gene	Locus	Proteins	Mutations	Frequency (%)
HCM: Genes Coding for Sarcomeric Proteins				
MYH7	14q11.2–q12	β-myosin heavy chain	>70, predominantly missense	35–50
MYBPC3	11p11.2	Cardiac myosin-binding Protein C	>40, predominantly splice junction/deletion	15–20
TNNT2	1q32	Cardiac troponin T	>15, mostly missense	15–20
TPM1	15q22.1	α-tropomyosin	>5 missense	<5
TNNI3	19p13.4	Cardiac troponin I	3 missense, 1 deletion	<1
MYL3	3p21.2–p21.3	Ventricular essential myosin light chain	2 missense	<1
MYL2	12q23–q24.3	Ventricular regulatory myosin light chain	7 missense, 1 truncation	<1
ACTC	15q14	α-cardiac actin	5 missense	<1
TTN	2q24.3	Titin	1 missense	<1
MYH6	14q11.2–q12	α-myosin heavy chain	1 missense, 1 rearrangement	Rare
HCM: Genes Coding for Nonsarcomeric Proteins				
PRKAG2	7q35–q3636	AMP-activated protein kinase	3 point, 1 insertion	<1
MTTI	Mitochondrial	Mitochrondrial DNA	tRNA isoleucin, tRNA glycine	Rare
FRDA	9q13	Frataxin	>200GAA in intron 1	Unknown
DMPK	19q13	Myotonin protein kinase	>50CTG in 3′UTR	Rare

chromosome 15q2, and Watkins and colleagues[52] mapped the third locus to chromosome 1q3. Six other loci were subsequently reported, located on chromosomes 11p11.2,[53] 7q3,[54] 3p21.2–3p21.3,[55] 12q23–q24.3,[56] 15q14,[57] and 2q31.[58]

HCM is a disease of contractile sarcomeric proteins. The R403Q missense mutation in the β-myosin heavy-chain gene (*MYH7*) was the first pathogenic mutation identified for HCM.[59] Subsequently, HCM-related mutations were discovered in genes encoding proteins responsible for the thick and thin filaments of the cardiac sarcomere.[55–58,60–63] Mutations causing familial HCM are primarily missense or frameshift mutations. To date, hundreds of different mutations have been identified in HCM patients in ten genes encoding contractile sarcomeric proteins: β-myosin heavy chain (*MYH7*),[59] myosin-binding protein-C (*MYBPC3*),[61] cardiac troponin T (*TNNT2*),[62] α-tropomyosin (*TPM1*),[62] cardiac troponin I (*TNNI3*),[56] essential myosin light chain (*MYL3*),[55] regulatory myosin light chain (*MYL2*),[55] cardiac α-actin (*ACTC*),[57,60] titin (*TTN*),[58] and α-myosin heavy chain (*MYH6*).[63–65] *MYH7*, *MYBPC3*, and *TNNT2* are cited as the three most common genes, accounting for 35% to 50%, 20% to 25%, and 20% of all HCM cases, respectively.[66]

Recently, HCM has been associated with mutations in genes encoding two nonsarcomeric proteins: *PRKAG2* and *CRP3*.[67–69] *PRKAG2* (chromosome 7q35–q36) encodes the γ2-subunit of the AMP-activated protein kinase and has been implicated in the variant of HCM associated with Wolff-Parkinson-White syndrome (WPW).[67,68] HCM associated with mutations in *PRKAG2* appears to represent a glycogen storage disease.[70]

Geier and colleagues reported three different missense mutations in the human cysteine-rich protein (*CRP3*) gene encoding muscle LIM protein (MLP) in three unrelated patients with familial HCM.[69] In addition, mutations in mitochondrial DNA often cause a complex phenotype that involves multiple organs, including cardiac muscle, with a disease phenotype of HCM.[71]

Clinical Utility of Testing

The specific type of mutation in the causal genes, specifically *MYH7*, *TNNT2*, and *MYBPC3*, may affect the phenotypic expression of HCM, particularly the magnitude of cardiac hypertrophy and the risk of SCD. In the *MYH7* gene, several mutations have been associated with a high risk of SCD (R403Q, R453C, R719W), whereas other mutations have been assigned a "benign" phenotype (G256E, V606M, L908V).[72–74] However, in a cohort of nearly 300 unrelated patients seen in a tertiary referral center for HCM, these particular mutations were present in less than 3% of the cases and did not necessarily adhere to the previously assigned genotype-phenotype classification as "malignant" or "benign" mutations.[75,76] Thus, great caution must be exercised with respect to mutation-specific prognosis and clinical decision making regarding SCD prevention.

HCM caused by mutations in *TNNT2* has been associated with a wide variety of clinical presentations and is usually associated with a low disease penetrance (80%), a mild degree of hypertrophy, but a high incidence of SCD and more extensive myocyte disarray, although some mutations are associated with hypertrophy without risk of arrhythmias.[77–79] HCM caused by mutations in *MYBPC3* has been associated with a low penetrance (60%), relatively mild hypertrophy, a low incidence of SCD, late onset of clinical manifestations, and good prognosis before the age of 40.[80,81]

Genetic studies have revealed that approximately 25% of tested patients have a mutant allele but do not manifest a HCM phenotype (genotype positive–phenotype negative HCM).[82] The variability in phenotypic expression of the mutations could be due to environment or acquired traits (differences in lifestyle, risk factors, and exercise). The existence of modifier genes, which are genes other than the pathogenic genes that affect the phenotypic expression of

genetic disease, could affect the phenotypic expression in HCM.

For example, polymorphisms involving the angiotensin-1 converting enzyme-1 (*ACE1*) gene may be associated with an increased risk of SCD[83] and the severity of left ventricular hypertrophy in HCM.[84] The deletion allele of *ACE1* (DD) is more common in patients with HCM and in patients with a high incidence of SCD compared to a control population.[83,85] Therapeutic intervention using angiotensin II blockade aimed at targeting this potential modifier gene demonstrated that the interstitial fibrosis in a transgenic mouse model of HCM could be reversed.[86] This finding raises the possibility of targeting the angiotensin II receptor to improve diastolic function and perhaps decrease the risk of SCD.

Available Assays

On a research basis, several centers of excellence perform mutational analysis of the HCM-related genes using a variety of methods, usually with initial PCR amplification followed by analysis of the PCR products by DNA sequencing, SSCP, or TMHA-DHPLC. Generally, these laboratories restrict the analysis to the protein-encoding exons and the intron-exon boundaries for the genes implicated in HCM.

Interpretation of Test Results

When and if mutational analysis becomes available from clinical laboratories, prognosis based on the genetic variation must be rendered cautiously. A decision to intervene with an internal cardioverter-defibrillator (ICD) should not be predicated based on the patient's HCM-causing mutation. The greatest impact of genetic testing will likely be realized in the identification of HCM patients with preclinical disease, permitting initiation of proper surveillance and perhaps preventive interventions directed at altering the patient's natural history.[87] However, the profound genetic heterogeneity and current technological limitations challenge the feasibility of a comprehensive HCM gene screen.

Laboratory Issues

Currently, commercial test kits are not available for HCM genetic testing.

DILATED CARDIOMYOPATHY

Molecular Basis of Disease

Dilated cardiomyopathy (DCM) is a heterogeneous heart muscle disease characterized by maladaptive ventricular dilation and impaired systolic function affecting approximately 1 in 2,500 individuals.[88] However, the incidence of DCM may be higher secondary to incomplete or partial penetrance (unidentified asymptomatic patients). A molecular basis for idiopathic DCM was speculated after clinical studies demonstrated that DCM may be inherited in up to 20% to 35% of cases when first-degree relatives are screened carefully.[89–91] Therefore, DCM represents a genetically heterogeneous disease (Table 14-2). Four different modes of inheritance are known: autosomal dominant with only DCM, autosomal dominant with other manifestations (conduction defect, sinus node dysfunction), autosomal recessive, and X-linked. Nearly two thirds of familial DCM has an autosomal dominant inheritance pattern with an isolated cardiac phenotype.[92]

Autosomal dominant familial DCM frequently presents as a mild form of DCM. The phenotype often is characterized by low penetrance. The onset of clinical manifestations is age related, with only 20% of gene mutation carriers under the age of 20 years of age manifesting the disease phenotype. Molecular genetic studies using linkage analysis have identified eight loci for autosomal dominant DCM: 1q32, 2q24.3–31, 2q35, 5q33, 6q12–q16, 9q12–q13, 14q11.2, and 15q11-qter.[93] To date, six genes have been associated with this phenotype: *ACTC* (chromosome 15q11-qter, encoding cardiac actin),[94,95] *DES* (chromosome 2q35, encoding desmin),[96] *TTN* (chromosome 2q24.3–q31, encoding titin),[97] *SGCD* (chromosome 5q33, encoding δ-sarcoglycan),[98] *MYH7* (chromosome 14q11.2, encoding β-myosin heavy chain),[99] and *TNNT2* (chromosome 1q32, encoding cardiac troponin T).[100]

Six loci for autosomal dominant DCM associated with other cardiac or muscle pathologies, or both (conduction defects, sinus node dysfunction, mitral valve prolapse, limb girdle muscular dystrophy, muscular dystrophy, and Emery-Dreifuss muscular dystrophy) or hearing loss have been identified: 1q21, 2q14–q22, 3p22–p25, 10q21–q23, 6q23–q24, and 6q22–q23.[101] So far, only two genes have been identified: *LMNA* (chromosome 1q21, encoding lamin A/C)[102] and *EYA4* (chromosome 6q23–q24, encoding eyes absent 4).[101,103] Missense mutations in the *LMNA* gene are responsible for inherited DCM with conduction system disease. A single nucleotide deletion in *LMNA* causes DCM, frequently associated with conduction system defects and variable skeletal muscle involvement.[102,104]

Autosomal recessive DCM is less frequent and is characterized by a significantly younger age of onset and a worse prognosis compared to the dominant form. This form is typically associated with limb girdle muscular dystrophy and caused by mutations in the genes coding for sarcoglycans. Mutations in three genes have been identified: *SGCA* (chromosome 17q21, encoding α-sarcoglycan), *SGCB* (chromosome 4q12, encoding. β-sarcoglycan), and *SGCD* (chromosome 5q33, encoding δ-sarcoglycan).[105,106]

X-linked familial DCM is a disease of the cytoskeleton caused by mutations in the dystrophin (*DYS*) gene on chromosome Xp21.3.[107,108] Dystrophin is a large cytoskeletal

Table 14-2. Summary of Molecular Basis of Dilated Cardiomyopathies

Gene	Locus	Proteins	Phenotype
DCM: Autosomal Dominant			
MYH7	14q11.2	β-myosin heavy chain	pure DCM, early onset
TNNT2	1q32	Cardiac troponin T	pure DCM, early onset
ACTC	15q11–qter	Actin	pure DCM
SGCD	5q33	δ-sarcoglycan	pure DCM
DES	2q35	Desmin	pure DCM
TTN	2q24.3–q31	Titin	pure DCM
Unknown	6q12–q16		pure DCM
Unknown	1q32		pure DCM
Unknown	9q12–q13		pure DCM
LMNA	1q21	Lamin A/C	DCM + CD
EYA4	6q23–q24	Eye absent 4	DCM + hearing loss
LMNA	1q21	Lamin A/C	DCM + CD + MD (AD-EDMD)
LMNA	1q21	Lamin A/C	DCM + CD + LGMD (LGMD1B)
Unknown	2q14–q22		DCM + CD
Unknown	3p22–p25		DCM + CD + SND
Unknown	10q21–q23		DCM + MVP
Unknown	6q22–q23		DCM + CD + LGMD
DCM: Autosomal Recessive			
SGCA	17q11	α-sarcoglycan	LGMD ± cardiomyopathy
SGCB	4q12	β-sarcoglycan	LGMD + severe cardiomyopathy
SGCD	5q33	δ-sarcoglycan	LGMD + cardiomyopathy
DCM: X-linked			
DYS	Xp21.3	Dystrophin	Pure DCM
TAZ	Xq28	Tafazzin (G4.5)	DCM lethal in infancy
TAZ	Xq28	Tafazzin (G4.5)	DCM + myopathy (Barth syndrome)
EMD	Xq28	Emerin	DCM + CD + MD (XL-EDMD)

CD, conduction defect; SND, sinus node dysfunction; MVP, mitral valve prolapse; LGMD, limb girdle muscular dystrophy; MD, muscular dystrophy; EDMD, Emery-Dreifuss muscular dystrophy.

protein that plays a critical role in membrane organization and stability and in force transduction in skeletal and cardiac myocytes.[107] Mutations in *DYS* are found in patients with Duchenne (severe form and early age of onset) and Becker (mild form and older age of onset) muscular dystrophies.[109]

The *G4.5* gene (chromosome Xq28, encoding tafazzins) is associated with DCM lethal in infancy and DCM with myopathy (Barth syndrome).[110–113] *EMD* mutations (chromosome Xq28, encoding the nuclear membrane protein emerin) are associated with X-linked Emery-Dreifuss muscular dystrophy (DCM, conduction defect, and muscular dystrophy).[114]

Recently, Schmitt and colleagues[115] reported DCM with refractory heart failure caused by a missense mutation in the phospholamban (*PLN*) gene located on chromosome 6q22.1.[116] PLN is an abundant 52-amino-acid transmembrane sarcoplasmic reticulum (SR) phospholipid[117] that regulates the cardiac Ca^{2+} sarcoplasmic reticular Ca^{2+}-adenosine triphosphate (SERCA2a) pump.[118] Haghighi and colleagues[119] also reported a missense mutation in the *PLN* gene, T116G, that results in markedly diminished expression of PLN protein in individuals from two families with inherited DCM. These results indicate that intracellular calcium dysregulation in cardiac myocytes can initiate DCM, as well as perturbations in the cytoskeleton and sarcomere.[120,121]

Clinical Utility of Testing

At the present time, the profound genetic heterogeneity and the lack of any relative hotspots or even "hot genes" precludes a prominent role for molecular testing in DCM.

Available Assays

Mutational analyses are performed in research laboratories. No clinical tests are currently available.

Interpretation of Test Results

The frequency of familial DCM indicates the need for meticulous family screening with at least echocardiographic surveillance of all first-degree relatives. The complexity of the phenotype requires accurate skeletal muscle

investigation that directs the diagnosis of a specific form of familial DCM. The disease can be asymptomatic due to the reduced and age-related penetrance. The profound genetic heterogeneity underlying DCM will challenge greatly the prospects of a molecular test in the next 10 years.

Laboratory Issues

Currently, commercial test kits are not available for DCM genetic testing.

ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY/DYSPLASIA

Arrhythmogenic right ventricular cardiomyopathy/dysplasia is a primary cardiomyopathy characterized by progressive degeneration and fibrous-fatty replacement of right ventricular myocardium, arrhythmias with a left bundle branch block profile, and increased risk of SCD in juveniles.[122–124]

Molecular Basis of Disease

Two modes of inheritance, autosomal dominant and autosomal recessive, are observed with ARVD (Table 14-3). To date, eight forms of autosomal dominant ARVD have been genetically mapped: ARVD1 (14q24.3),[125] ARVD2 (1q42),[126] ARVD3 (14q11–q12),[127] ARVD4 (2q32),[128] ARVD5 (3p23),[129] ARVD6 (10p12–p14),[130] ARVD7 (10q22),[131] and ARVD8 (6p24).[132] Two autosomal recessive forms of ARVD have been genetically mapped: Naxos (17q21)[133] and ARVC/APC (14q24-qterminal).[134] Autosomal dominant ARVD has an incomplete penetrance and variable phenotypic expression.[135] In contrast, Naxos disease is essentially 100% penetrant with relatively early age of onset and homogeneous presentation (i.e., specific hair and skin abnormalities in early infancy and complete penetrance of the cardiac disorder by adolescence).[136]

Table 14-3. Summary of Molecular Basis of Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia

Type	Locus	Gene	Mode of Inheritance	Penetrance
ARVC-1	14q23–q24	Unknown	AD	High
ARVC-2	1q42–q43	*RYR2*	AD	High
ARVC-3	14q12–q22	Unknown	AD	Unknown
ARVC-4	2q32.1–q32.3	Unknown	AD	Unknown
ARVC-5	3p23	Unknown	AD	Unknown
ARVC-6	10p12–p14	Unknown	AD	Unknown
ARVC-7	10q22.3	Unknown	AD	Unknown
ARVC-8	6p24	*DSP*	AD	Unknown
Naxos	17q21	*PG*	AR	100%
ARVD/PC	1424-qter	Unknown	AR	Unknown

AD, autosomal dominant; AR, autosomal recessive.

Clinical Utility of Testing

The variability of phenotype, disease progression, and genetic heterogeneity contribute to the difficulties of diagnosis and risk stratification in ARVD. Although eight chromosomal loci for the autosomal dominant form of ARVD and two loci for the autosomal recessive form have been identified, a specific genetic basis has been identified in only three of the genes: *RYR2* encoding the cardiac ryanodine receptor for ARVD2,[137] *DSP* encoding desmoplakin for ARVD8,[132] and *PG* encoding plakoglobin for Naxos disease.[136]

Tiso and colleagues[137] reported *RYR2* mutations in four ARVD2 families. Two families carried an N2386I mutation, and one family had two different mutations (R176Q and T2504M) on the same allele, and a fourth family had an L433P missense mutation. More recently, Rampazzo and colleagues[132] identified a putative pathogenic mutation in *DSP* responsible for ARVD8 in 26 members of a family in four generations: the missense mutation (S299R) in exon 7 on chromosome 6p24 with no linkage with any DNA marker for ARVD loci known until then. The penetrance was estimated to be 50%. Protonotarios and colleagues[136] reported a mutation in the gene encoding plakoglobin causing Naxos disease fully penetrant by adolescence.

Available Assays

Mutational analyses are performed in research laboratories. No clinical tests currently are available.

Interpretation of Test Results

Routine genetic testing in patients or family members is presently not available.

Laboratory Issues

Currently, commercial test kits are not available for ARVD genetic testing.

RESTRICTIVE CARDIOMYOPATHY

Restrictive cardiomyopathy (RCM) is a myocardial disorder characterized by impaired diastolic filling of the left ventricle and reduced diastolic volume in the presence of normal systolic function and normal myocardial thickness.[45] Although most frequently caused by diseases causing infiltration or fibrosis of the myocardium, RCM

may be found in the absence of a precipitating condition in many patients (idiopathic RCM).[138]

Molecular Basis of Disease

Idiopathic RCM often occurs as a sporadic case, but familial RCM has been reported as an autosomal dominant trait.[139,140] Missense mutations in the *DES* (desmin) gene have been found in several families with desmin-related myopathy with and without RCM.[141,142] Zhang and colleagues[143] described a large family (34 family members) with apparent autosomal dominant inheritance of desmin-associated RCM spanning four generations. This family is relatively unique because no symptom-based skeletal muscle involvement was evident and the known desminopathy and cardiomyopathy genes and loci were ruled out. These data showed that desmin deposition may be associated with different gene defects in RCM. Whether a distinct genetic entity for RCM exists was uncertain until Mogensen and colleagues[144] reported a large family in which family members were affected by either idiopathic RCM or HCM. Using linkage analysis for selected sarcomeric contractile protein genes, cardiac troponin I (*TNNI3*) was identified as the likely gene causing RCM. Subsequent mutation analysis of *TNNI3* by direct sequencing identified a novel missense mutation. Furthermore, *TNNI3* mutations were identified in six of nine unrelated patients with RCM. To date, six missense mutations of *TNNI3* have been identified in RCM patients.[144] These findings suggested that *TNNI3* mutations can lead to not only HCM but also RCM. RCM is part of the spectrum of hereditary sarcomeric contractile protein disease.

ISOLATED LEFT VENTRICULAR NONCOMPACTION

Isolated left ventricular noncompaction (LVNC) is characterized by a hypertrophic left ventricle with deep trabeculations and poor systolic function with or without associated left ventricular dilation. LVNC is a rare disorder thought to be due to an arrest of myocardial morphogenesis.[145,146]

Molecular Basis of Disease

Familial recurrence in LVNC is high and found in approximately 40% of patients.[145] Genetic linkage and mutation analysis have shown that mutations in the *DRP3* gene (chromosome 18q12), encoding the dystrophin-related protein 3 known as α-dystrobrevin, are responsible for LVNC associated with congenital heart defects.[113] In addition, mutations in the *G4.5* gene result in not only Barth syndrome as described in the DCM section above, but also other X-linked infantile cardiomyopathies, including LVNC,[113,147,148] X-linked infantile cardiomyopathy, and X-linked endocardial fibroelastosis.[149,150] Mutations in *G4.5* were identified in two families with isolated LVNC including a missense mutation in exon 4 (C118R) in one family and a splice donor mutation (IVS10+2T→A) in intron 10 in the other family.[113] Recently, Chen and colleagues[151] reported a novel splice acceptor-site mutation of intron 8 of *G4.5* in a family with severe infantile X-linked LVNC without the usual findings of Barth syndrome. The incidence of *G4.5* mutations in LVNC is not high, and specific mutations are not associated with a particular phenotype.

CARDIAC CHANNELOPATHIES/PRIMARY ARRHYTHMIA SYNDROMES

CONGENITAL LONG QT SYNDROME

Molecular Basis of Disease

Congenital long QT syndrome is an inherited cardiac channelopathy characterized by prolongation of the QT interval of the cardiac cycle and increased susceptibility for syncope, seizures, and sudden cardiac death secondary to polymorphic ventricular tachyarrhythmias (torsades de pointes). Congenital LQTS occurs in two main heritable forms: autosomal dominant LQTS, originally described as Romano-Ward syndrome, and autosomal recessive LQTS, originally described as the Jervell and Lange-Nielsen syndrome (JLNS; Table 14-4). LQTS is the first genetically defined type of arrhythmia to be understood at the molecular level as a primary cardiac channelopathy.[152–154] Using genetic linkage and mutational analysis, the genetic heterogeneity seen with the cardiomyopathies is similarly present for this cardiac channelopathy.

Six LQTS genes have been identified: *KVLQT1* (*KCNQ1*, LQT1),[155] *HERG* (*KCNH2*, LQT2),[156] *SCN5A* (LQT3),[157] *ANKB* (Ankyrin-B, LQT4),[158] *KCNE1* (minK, LQT5),[159] and *KCNE2* (MiRP1, LQT6).[160] LQT1 and LQT2 represent the two most common genetic subtypes of LQTS comprising about half of the syndrome. In approximately 30% to 40% of families with LQTS, a genetic defect cannot be identified in the known LQTS-causing genes.[161]

The gene responsible for LQT1 (originally named *KVLQT1*, but more recently named *KCNQ1*), encodes the α-subunit of the slowly activating delayed rectifier potassium ion channel (I_{Ks}). The gene consists of 16 exons, spans approximately 400 kb, and is localized to chromosome 11p15.5. The loss of I_{Ks} channel function decreases the I_{Ks} current, resulting in prolongation of the action potential duration and ventricular repolarization.

LQT2 is due to mutations in the human ether-a-go-go-related gene (*HERG*, recently assigned as *KCNH2*), which codes for the α-subunit of the rapidly activating delayed rectifier potassium ion channel (I_{Kr}) and consists of 16 exons, spans 55 kb of genomic sequence,[162] and is localized

Table 14-4. Summary of Molecular Basis of Cardiac Channelopathies

Type	Locus	Gene	Mode of Inheritance	Current	Mutations	Frequency (%)
Romano-Ward Syndrome: LQTS						
LQT1	11p15.5	*KCNQ1/KVLQT1*	AD	I_{Ks}	mis, non, ins/del	20–25
LQT2	7q35-36	*KCNH2/HERG*	AD	I_{Kr}	mis, non, ins/del	20–25
LQT3	3p21–p24	*SCN5A*	AD	I_{Na}	mis, non, ins/del	<5
LQT4	4q25–q27	*ANKB*	AD	Na/Ca	mis	<1
LQT5	21q22.1	*KCNE1/minK*	AD	I_{Ks}	mis	1
LQT6	21q22.1	*KCNE2/MiRP1*	AD	I_{Ks}	mis	1
Jervell and Lange-Nielsen Syndrome: LQTS						
JLN1	11p15.5	*KCNQ1/KVLQT1*	AR	I_{Ks}	mis, non, ins/del	80
JLN2	21q22.1	*KCNE1/minK*	AR	I_{Ks}	mis	20
Anderson Syndrome						
AS1	17q23	*KCNJ2*	AD	I_{K1}	mis	50
Brugada Syndrome						
BrS1	3p21–p24	*SCN5A*	AD	I_{Na}	mis	20
Catecholaminergic Polymorphic Ventricular Tachycardia (CPVT)						
CPVT1	1q42–q43	*RYR2*	AD	–	mis	50
CPVT2	1p13–21	*CASQ2*	AR	–	mis	Unknown

AD, autosomal dominant; AR, autosomal recessive; mis, missense mutation; non, nonsense mutation; ins/del, insertion/deletion.

to chromosome 7q35–36.[163] Mutations in *HERG* reduce the I_{Kr} current, resulting in prolongation of the action potential duration and repolarization.

LQT3 results from mutations in *SCN5A* (chromosome 3p21–24) that encodes the α-subunit of the cardiac sodium channel.[164,165] In contrast to the other forms of LQTS, in which a reduction of repolarization causes the prolongation of the action potential, *SCN5A* mutations that cause LQT3 cause a "gain of function" in the cardiac sodium channel with an increase in late sodium current. In addition, Ackerman and colleagues identified two decedents with *SCN5A* mutations among 93 cases of sudden infant death syndrome in a statewide population of SIDS autopsies using molecular analysis.[166] *SCN5A* encodes a protein of 2,016 amino acids with a putative structure that consists of four homologous domains (DI to DIV), each of which contains six membrane-spanning segments (S1 to S6) similar to the structure of α-subunits of the potassium channel.

Recently, a mutation in *ANKB* (chromosome 4q25–27), which encodes ankyrin-B, was identified as the pathogenic basis for the elusive LQT4 locus.[158,167] Ankyrin-B (also known as ankyrin 2) is a member of a family of versatile membrane adapters.[168] This study represents the first non-cardiac channel gene to be implicated in LQTS.

The remaining LQTS genotypes include LQT5 secondary to mutations in the β-subunit of the I_{Ks} channel (*KCNE1 or minK*, chromosome 21q22.1)[159] and LQT6 secondary to mutations in *KCNE2* (*MiRP1*).[160] Interestingly, autosomal recessive LQTS or Jervell and Lange-Nielsen syndrome is secondary to homozygous or compound heterozygous mutations involving either *KCNQ1* (JLN1) or *KCNE1* (JLN2) or both.[169–172] By definition, both parents are heterozygous for a pathogenic mutation; however, they are typically asymptomatic with negligible QT interval prolongation.[173] In contrast, their affected offspring not only have a severe cardiac phenotype but also are deaf. The I_{Ks} channel plays a critical role in potassium homeostasis in the endolymph of the inner ear.[174,175]

Clinical Utility of Testing

Molecular testing of the LQTS genes has been performed in research laboratories over the past decade. In addition to bridging the gap from research to clinical practice and providing critical information for families while a clinical assay is being developed, such molecular genetic testing has enabled very significant genotype-phenotype correlations to be delineated. First, cardiac events occur more frequently at rest or during sleep in LQT3, whereas cardiac events are typically related to emotion or exercise (particularly swimming) in LQT1[176,177] and auditory stimuli in LQT2.[177] Second, the recurrence rate of cardiac events while on β-blocker therapy was significantly lower in patients with LQT1 (19%) compared to LQT2 (41%) and LQT3 (50%).[178] Third, LQT3 patients under the age of 40 years have a lower incidence of cardiac events (18%) compared to LQT1 (63%) and LQT2 (46%) patients, but the lethality per cardiac event is greatest in LQT3.[177,179] Fourth, patients with LQT2-causing mutations in the channel pore tend to have a "poorer" prognosis than LQT2 patients with C-

terminal mutations.[180] Fifth, elucidation of the various genetic subtypes also has fostered the development and interpretation of provocative physiological challenges with epinephrine to unmask certain forms of "concealed LQTS."[181,182] In particular, Ackerman and colleagues first demonstrated that patients with LQT1 manifest a paradoxical QT response to low-dose epinephrine, which distinguishes them from other LQTS genotypes and, more important, from normal individuals.[182] Molecular genetic testing is useful in diagnosis of "concealed LQTS," especially in asymptomatic family members of patients with LQTS. These insights further justify the transition of mutational analysis from research laboratories to clinical molecular laboratories.

Available Assays

Mutational analyses of the known LQTS genes are performed in research laboratories by PCR amplification combined with SSCP analysis, TMHA-DHPLC analysis, or direct DNA sequencing of the PCR products. The commercial LQTS gene screen in development involves a single PCR protocol to amplify all the protein-encoding exons for subsequent high-throughput DNA sequencing.

Interpretation of Test Results

Molecular genetic testing for LQTS is likely to play a critical complementary role in the diagnostic evaluation and risk stratification for LQTS. However, the present sensitivity of the genetic test is approximately 60% to 70%. As such, a negative LQTS gene screen will not provide sufficient objective evidence to exclude the presence of LQTS.

Laboratory Issues

The anticipated commercially available clinical diagnostic test for the cardiac channelopathies will test for mutations in the genes responsible for LQT1, LQT2, LQT3, LQT5, and LQT6. Importantly, the gene responsible for LQT4 and the genes implicated in the other primary channelopathies will not be part of the first-generation molecular genetic test. The LQTS channel screen will represent the first available multigene molecular genetic test for any heritable cardiovascular disease.

BRUGADA SYNDROME

Brugada syndrome (BrS) is characterized clinically by syncope and increased risk of SCD, particularly in young men at night, electrocardiographically by a right bundle branch block pattern and ST segment elevation in the precordial leads V1–V3, and echocardiographically by a structurally normal heart.[183,184] In Asia, BrS and sudden nocturnal death syndrome were previously called *lai tai* (Thailand), *pokkuri* (Japan), and *banungut* (Philippines), which mean nocturnal sudden death. BrS has been considered a distinct subgroup of idiopathic ventricular fibrillation (IVF) and may account for 20% to 40% of all patients with IVF.[185] IVF is classified as ventricular fibrillation without demonstrable cardiac or noncardiac causes to account for the episode[186] and accounts for approximately 5% to 12% of all sudden deaths. IVF is associated with a high mortality rate and high recurrence rates of up to 30% during the 5 years after an initial episode of survived cardiac arrest. For both BrS and IVF, the treatment of choice is principally defibrillator therapy.

Molecular Basis of Disease

In early 1998, Chen and colleagues[187] reported the first gene linked to BrS in six families and two sporadic cases of BrS using SSCP and sequencing analysis of *SCN5A*. Mutations in the cardiac sodium channel encoded by *SCN5A* (recall that LQT3 is due to *SCN5A* defects as well) were identified initially in three families: a missense mutation, a two-nucleotide insertion, and a single nucleotide deletion. However, the sites of mutations were distinct from those known to contribute to LQT3, and the QT interval was not prolonged in these patients. In contrast to LQT3, BrS-causing *SCN5A* mutations result in a "loss of function."

SCN5A gene mutations have been identified in approximately 20% of patients with clinical BrS.[188,189] In cases of sporadic BrS, *SCN5A* mutations rarely are found.[190] Recently, Weiss reported a large multigeneration family with an autosomal dominant form of BrS linked to a new locus on chromosome 3p22–25 distinct from *SCN5A*, confirming the suspected genetic heterogeneity of this syndrome.[191] Akai and colleagues[192] recently reported a novel *SCN5A* missense mutation (S1710L) in a symptomatic IVF patient lacking electrocardiographic evidence for BrS.

Clinical Utility of Testing

Considering that only one genotype for BrS exists, and that genotype (BrS1 = *SCN5A*) accounts for 20% of the syndrome, the clinical utility is limited. For families established to have BrS1, genetic testing can be used to distinguish family members whose clinical evaluations are equivocal. Presently, BrS1 versus non-BrS1 genotype does not seem to afford any prognostic information.

Available Assays

Mutational analysis of *SCN5A* for BrS1 screening the known LQTS genes is performed in research laboratories by PCR amplification combined with SSCP analysis, TMHA-DHPLC analysis, or direct DNA sequencing of the PCR products.

Interpretation of Test Results

Sensitivity of current BrS molecular genetic testing is approximately 20%.

Laboratory Issues

The anticipated commercially available clinical genetic test for cardiac channelopathies will include the gene (SCN5A) responsible for both LQT3 and BrS1.

CATECHOLAMINERGIC POLYMORPHIC VENTRICULAR TACHYCARDIA

Catecholaminergic polymorphic ventricular tachycardia is a distinct arrhythmogenic entity in children and adolescents with structurally normal hearts characterized by episodes of syncope, seizures, or sudden death in response to physical activity or emotional stress, resulting in stress-induced ventricular arrhythmias that are predominantly bidirectional ventricular tachycardias.[193]

Molecular Basis of Disease

CPVT can be transmitted in either an autosomal dominant[194] or an autosomal recessive inheritance pattern.[195] Swan and colleagues first reported linkage of CPVT to chromosome 1q42–q43.[194] Subsequently, Priori and colleagues[196] demonstrated that the cardiac ryanodine receptor gene (*RYR2*) mapping on chromosome 1q42-q43 was responsible for this disease. Initially, four *RYR2* missense mutations were found, three sporadic mutations (S2246L, R2474S, and N4104K), and one familial (R4497C). *RYR2* mutations were then identified in 14 of 30 probands (36% of patients with bidirectional ventricular tachycardia, 58% of patients with polymorphic ventricular tachycardia, 50% of patients with catecholaminergic idiopathic ventricular fibrillation, and in nine family members (four silent gene carriers).[197] Lahat and colleagues[195] reported an autosomal recessive form of CPVT mapping to chromosome 1p13–21. Subsequently, this group also described a missense mutation in a highly conserved region of the calsequestrin 2 gene (*CASQ2*) as one potential mechanism for autosomal recessive CPVT.[198]

Clinical Utility of Testing

Although only 50% of CPVT has been elucidated genetically, some important genotype-phenotype relationships have emerged that support a role for molecular genetic testing in addition to the preclinical identification of asymptomatic carriers. Families with CPVT1 (*RYR2* mutations) are more prone to stress-induced ventricular tachycardia at an earlier age than genotype-negative CPVT. In addition, males with CPVT1 may be at greater risk for syncope than males lacking *RYR2* mutations.[197]

Available Assays

Mutation screening is performed in research laboratories on genomic DNA extracts from peripheral blood lymphocytes. Intronic primers amplifying the *RYR2* coding region are used for PCR. PCR products are analyzed by either SSCP or DHPLC.[197]

Interpretation of Test Results

The sensitivity of molecular testing for the two CPVT-causing genes is approximately 50%.

Laboratory Issues

Currently, commercial test kits are not available for CPVT genetic testing.

ANDERSEN SYNDROME

Molecular Basis of Disease

Andersen syndrome (AS) is a rare inherited disorder characterized by periodic paralysis, prolongation of the QT interval with ventricular arrhythmias, and characteristic dysmorphic features including low-set ears, micrognathia, short stature, scoliosis, hypertelorism, broad forehead, and clinodactyly.[199,200] AS is inherited in an autosomal dominant manner,[201] although many cases are sporadic. Plaster and colleagues[202] demonstrated that mutations in the *KCNJ2* gene (chromosome 17q23) encoding the inward-rectifying K^+ channel (Kir2.1) are responsible for AS. Subsequently, several additional heterozygous missense mutations were identified in *KCNJ2* in families with AS.[203–205] These *KCNJ2* mutations produce loss of function and dominant negative suppression of Kir2.1 channel function.[205]

Clinical Utility of Testing

Because of the prolonged QT interval manifest in most patients with *KCNJ2* mutations, this gene also has been implicated in the pathogenesis of LQT7. However, no families with isolated, autosomal dominant LQTS have been reported to have mutations in *KCNJ2*. Because *KCNJ2* mutations were not found in every family with AS, this multiorgan syndrome is certain to be genetically heterogeneous as well, and *KCNJ2* mutations may be more accurately designated as the molecular basis for AS1.

Available Assays

The entire coding region of Kir2.1 is amplified from the genomic DNA by PCR and then analyzed by DNA sequencing.[205] Again, analysis of *KCNJ2* is performed only in research laboratories at the present time.

Interpretation of Test Results

Test sensitivity and specificity are not known.

Laboratory Issues

Currently, commercial test kits are not available for AS genetic testing.

Conclusions

Over the past ten years, profound molecular breakthroughs and advances related to the Human Genome Project have facilitated the discovery of fundamental pathogenic mechanisms for several heritable cardiovascular diseases. Important genotype-phenotype insights have been identified and have been applied in the clinical setting for proper identification and treatment of preclinical, asymptomatic individuals and for prognostic considerations. Already, genetic testing of critical polymorphisms such as factor V Leiden that predispose to venous thrombosis and pulmonary embolism has made the transition from a test performed solely in research laboratories to a readily available clinical molecular test. However, in contrast to clinical molecular testing for risk-modifying polymorphisms, the development of clinical molecular genetic tests for the disease-causing genes of such processes as cardiomyopathies and channelopathies has been hindered by the immense genetic heterogeneity for these diseases. Undoubtedly, technological advances will be made over the next decade that will enable comprehensive gene screening for diseases such as LQTS and HCM to join the ranks of existing clinical genetic tests such as *CFTR* screening for cystic fibrosis and *BRCA1/2* screening for breast cancer.

References

1. Goldstein JL, Brown MS. Binding and degradation of low density lipoproteins by cultured human fibroblasts: comparison of cells from a normal subject and from a patient with homozygous familial hypercholesterolemia. *J Biol Chem.* 1974;249:5153–5162.
2. Brown MS, Goldstein JL. A receptor-mediated pathway for cholesterol homeostasis. *Science.* 1986;232:34–47.
3. Lindgren V, Luskey KL, Russell DW, Francke U. Human genes involved in cholesterol metabolism: chromosomal mapping of the loci for the low density lipoprotein receptor and 3-hydroxy-3-methylglutaryl-coenzyme A reductase with cDNA probes. *Proc Natl Acad Sci U S A.* 1985;82:8567–8571.
4. Yamamoto T, Davis CG, Brown MS, et al. The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell.* 1984;39:27–38.
5. Law SW, Lackner KJ, Hospattankar AV, et al. Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of the gene to chromosome 2. *Proc Natl Acad Sci U S A.* 1985;82:8340–8344.
6. Innerarity TL, Weisgraber KH, Arnold KS, et al. Familial defective apolipoprotein B-100: low density lipoproteins with abnormal receptor binding. *Proc Natl Acad Sci U S A.* 1987;84:6919–6923.
7. Varret M, Rabes JP, Saint-Jore B, et al. A third major locus for autosomal dominant hypercholesterolemia maps to 1p34.1–p32. *Am J Hum Genet.* 1999;64:1378–1387.
8. Heath KE, Gahan M, Whittall RA, Humphries SE. Low-density lipoprotein receptor gene (LDLR) world-wide website in familial hypercholesterolaemia: update, new features and mutation analysis. *Atherosclerosis.* 2001;154:243–246.
9. Wilson DJ, Gahan M, Haddad L, et al. A World Wide Web site for low-density lipoprotein receptor gene mutations in familial hypercholesterolemia: sequence-based, tabular, and direct submission data handling. *Am J Cardiol.* 1998;81:1509–1511.
10. Soria LF, Ludwig EH, Clarke HR, Vega GL, Grundy SM, McCarthy BJ. Association between a specific apolipoprotein B mutation and familial defective apolipoprotein B-100. *Proc Natl Acad Sci U S A.* 1989;86:587–591.
11. Gaffney D, Reid JM, Cameron IM, et al. Independent mutations at codon 3500 of the apolipoprotein B gene are associated with hyperlipidemia. *Arterioscler Thromb Vasc Biol.* 1995;15:1025–1029.
12. Pullinger CR, Hennessy LK, Chatterton JE, et al. Familial ligand-defective apolipoprotein B. Identification of a new mutation that decreases LDL receptor binding affinity. *J Clin Invest.* 1995;95:1225–1234.
13. Nissen H, Hansen PS, Faergeman O, Horder M. Mutation screening of the codon 3500 region of the apolipoprotein B gene by denaturing gradient-gel electrophoresis. *Clin Chem.* 1995;41:419–423.
14. Goldstein JL, Hobbs H, Brown M. Familial hypercholesterolaemia. In: Scriver C, Beauder A, Sly W, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease.* 8th ed. New York: McGraw-Hill; 2001:2863–2913.
15. Arca M, Zuliani G, Wilund K, et al. Autosomal recessive hypercholesterolaemia in Sardinia, Italy, and mutations in ARH: a clinical and molecular genetic analysis. *Lancet.* 2002;359:841–847.
16. Wilson PW, Schaefer EJ, Larson MG, Ordovas JM. Apolipoprotein E alleles and risk of coronary disease. A meta-analysis. *Arterioscler Thromb Vasc Biol.* 1996;16:1250–1255.
17. Ridker PM, Rifai N, Stampfer MJ, Hennekens CH. Plasma concentration of interleukin-6 and the risk of future myocardial infarction among apparently healthy men. *Circulation.* 2000;101:1767–1772.
18. Humphries SE, Luong LA, Ogg MS, Hawe E, Miller GJ. The interleukin-6-174 G/C promoter polymorphism is associated with risk of coronary heart disease and systolic blood pressure in healthy men. *Eur Heart J.* 2001;22:2243–2252.
19. Soufi M, Sattler AM, Maisch B, Schaefer JR. Molecular mechanisms involved in atherosclerosis. *Herz.* 2002;27:637–648.
20. Garcia-Garcia AB, Real JT, Puig O, et al. Molecular genetics of familial hypercholesterolemia in Spain: ten novel LDLR mutations and population analysis. *Hum Mutat.* 2001;18:458–459.
21. Griffin JH, Evatt B, Wideman C, Fernandez JA. Anticoagulant protein C pathway defective in majority of thrombophilic patients. *Blood.* 1993;82:1989–1993.
22. Grody WW, Griffin JH, Taylor AK, Korf BR, Heit JA. American College of Medical Genetics consensus statement on factor V Leiden mutation testing. *Genet Med.* 2001;3:139–148.
23. Press RD, Bauer KA, Kujovich JL, Heit JA. Clinical utility of factor V Leiden (R506Q) testing for the diagnosis and management of

thromboembolic disorders. *Arch Pathol Lab Med.* 2002;126:1304–1318.
24. Rosendaal FR, Siscovick DS, Schwartz SM, et al. Factor V Leiden (resistance to activated protein C) increases the risk of myocardial infarction in young women. *Blood.* 1997;89:2817–2821.
25. Holm J, Zoller B, Svensson PJ, Berntorp E, Erhardt L, Dahlback B. Myocardial infarction associated with homozygous resistance to activated protein C. *Lancet.* 1994;344:952–953.
26. Holm J, Hillarp A, Zoller B, Erhardt L, Berntorp E, Dahlback B. Factor V Q506 (resistance to activated protein C) and prognosis after acute coronary syndrome. *Thromb Haemost.* 1999;81:857–860.
27. Marz W, Seydewitz H, Winkelmann B, Chen M, Nauck M, Witt I. Mutation in coagulation factor V associated with resistance to activated protein C in patients with coronary artery disease. *Lancet.* 1995;345:526.
28. Holm J, Zoller B, Berntorp E, Erhardt L, Dahlback B. Prevalence of factor V gene mutation amongst myocardial infarction patients and healthy controls is higher in Sweden than in other countries. *J Intern Med.* 1996;239:221–226.
29. Ridker PM, Hennekens CH, Lindpaintner K, Stampfer MJ, Eisenberg PR, Miletich JP. Mutation in the gene coding for coagulation factor V and the risk of myocardial infarction, stroke, and venous thrombosis in apparently healthy men. *N Engl J Med.* 1995;332:912–917.
30. Prohaska W, Mannebach H, Schmidt M, Gleichmann U, Kleesiek K. Evidence against heterozygous coagulation factor V 1691 G→A mutation with resistance to activated protein C being a risk factor for coronary artery disease and myocardial infarction. *J Mol Med.* 1995;73:521–524.
31. Jeffery S, Leatham E, Zhang Y, Carter J, Pratel P, Kaski JC. Factor V Leiden polymorphism (FV Q506) in patients with ischaemic heart disease, and in different populations groups. *J Hum Hypertens.* 1996;10:433–434.
32. Dunn ST, Roberts CR, Schechter E, Moore WE, Lee ET, Eichner JE. Role of factor V Leiden mutation in patients with angiographically demonstrated coronary artery disease. *Thromb Res.* 1998;91:91–99.
33. Boekholdt SM, Bijsterveld NR, Moons AH, Levi M, Buller HR, Peters RJ. Genetic variation in coagulation and fibrinolytic proteins and their relation with acute myocardial infarction: a systematic review. *Circulation.* 2001;104:3063–3068.
34. Juul K, Tybjaerg-Hansen A, Steffensen R, Kofoed S, Jensen G, Nordestgaard BG. Factor V Leiden: the Copenhagen City Heart Study and 2 meta-analyses. *Blood.* 2002;100:3–10.
35. Wu AH, Tsongalis GJ. Correlation of polymorphisms to coagulation and biochemical risk factors for cardiovascular diseases. *Am J Cardiol.* 2001;87:1361–1366.
36. Heinrich J, Budde T, Assmann G. Mutation in the factor V gene and the risk of myocardial infarction. *N Engl J Med.* 1995;333:881.
37. Endler G, Mannhalter C. Polymorphisms in coagulation factor genes and their impact on arterial and venous thrombosis. *Clin Chim Acta.* 2003;330:31–55.
38. Mudd SH, Skovby F, Levy HL, et al. The natural history of homocystinuria due to cystathionine beta-synthase deficiency. *Am J Hum Genet.* 1985;37:1–31.
39. Cleophas TJ, Hornstra N, van Hoogstraten B, van der Meulen J. Homocysteine, a risk factor for coronary artery disease or not? A meta-analysis. *Am J Cardiol.* 2000;86:1005–1009, A8.
40. Jacques PF, Bostom AG, Williams RR, et al. Relation between folate status, a common mutation in methylenetetrahydrofolate reductase, and plasma homocysteine concentrations. *Circulation.* 1996;93:7–9.
41. Frosst P, Blom HJ, Milos R, et al. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet.* 1995;10:111–113.
42. Castro R, Rivera I, Ravasco P, et al. 5,10-Methylenetetrahydrofolate reductase 677C→T and 1298A→C mutations are genetic determinants of elevated homocysteine. *QJM.* 2003;96:297–303.
43. Andreassi MG, Botto N, Cocci F, et al. Methylenetetrahydrofolate reductase gene C677T polymorphism, homocysteine, vitamin B12, and DNA damage in coronary artery disease. *Hum Genet.* 2003;112:171–177.
44. Maron BJ, Bonow RO, Cannon RO 3rd, Leon MB, Epstein SE. Hypertrophic cardiomyopathy. Interrelations of clinical manifestations, pathophysiology, and therapy (2). *N Engl J Med.* 1987;316:844–852.
45. Richardson P, McKenna W, Bristow M, et al. Report of the 1995 World Health Organization/International Society and Federation of Cardiology Task Force on the Definition and Classification of cardiomyopathies. *Circulation.* 1996;93:841–842.
46. Popjes ED, St John Sutton M. Hypertrophic cardiomyopathy: pathophysiology, diagnosis, and treatment. *Geriatrics.* 2003;58:41–46; quiz 47.
47. Maron BJ. Hypertrophic cardiomyopathy: a systematic review. *JAMA.* 2002;287:1308–1320.
48. Maron BJ. Hypertrophic cardiomyopathy. *Lancet.* 1997;350:127–133.
49. Maron BJ, Anan TJ, Roberts WC. Quantitative analysis of the distribution of cardiac muscle cell disorganization in the left ventricular wall of patients with hypertrophic cardiomyopathy. *Circulation.* 1981;63:882–894.
50. Jarcho JA, McKenna W, Pare JA, et al. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. *N Engl J Med.* 1989;321:1372–1378.
51. Thierfelder L, MacRae C, Watkins H, et al. A familial hypertrophic cardiomyopathy locus maps to chromosome 15q2. *Proc Natl Acad Sci U S A.* 1993;90:6270–6274.
52. Watkins H, MacRae C, Thierfelder L, et al. A disease locus for familial hypertrophic cardiomyopathy maps to chromosome 1q3. *Nat Genet.* 1993;3:333–337.
53. Carrier L, Hengstenberg C, Beckmann JS, et al. Mapping of a novel gene for familial hypertrophic cardiomyopathy to chromosome 11. *Nat Genet.* 1993;4:311–313.
54. MacRae CA, Ghaisas N, Kass S, et al. Familial hypertrophic cardiomyopathy with Wolff-Parkinson-White syndrome maps to a locus on chromosome 7q3. *J Clin Invest.* 1995;96:1216–1220.
55. Poetter K, Jiang H, Hassanzadeh S, et al. Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nat Genet.* 1996;13:63–69.
56. Kimura A, Harada H, Park JE, et al. Mutations in the cardiac troponin I gene associated with hypertrophic cardiomyopathy. *Nat Genet.* 1997;16:379–382.
57. Mogensen J, Klausen IC, Pedersen AK, et al. Alpha-cardiac actin is a novel disease gene in familial hypertrophic cardiomyopathy. *J Clin Invest.* 1999;103:R39–R43.
58. Satoh M, Takahashi M, Sakamoto T, Hiroe M, Marumo F, Kimura A. Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene. *Biochem Biophys Res Commun.* 1999;262:411–417.
59. Geisterfer-Lowrance AA, Kass S, Tanigawa G, et al. A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. *Cell.* 1990;62:999–1006.
60. Olson TM, Doan TP, Kishimoto NY, Whitby FG, Ackerman MJ, Fananapazir L. Inherited and de novo mutations in the cardiac actin gene cause hypertrophic cardiomyopathy. *J Mol Cell Cardiol.* 2000;32:1687–1694.
61. Bonne G, Carrier L, Bercovici J, et al. Cardiac myosin binding protein-C gene splice acceptor site mutation is associated with familial hypertrophic cardiomyopathy. *Nat Genet.* 1995;11:438–440.
62. Thierfelder L, Watkins H, MacRae C, et al. Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell.* 1994;77:701–712.
63. Freeman K, Lerman I, Kranias EG, et al. Alterations in cardiac adrenergic signaling and calcium cycling differentially affect the progression of cardiomyopathy. *J Clin Invest.* 2001;107:967–974.

64. Carrier L, Jongbloed R, Smeets H, Doevendans PA. Hypertrophic cardiomyopathy. In: Doevendans P, Wilde, AA, eds. *Cardiovascular Genetics for Clinicians.* Kluwer Academic Publishers: Dordrecht, The Netherlands; 2001:139–154.
65. Seidman C. Genetic causes of inherited cardiac hypertrophy: Robert L. Frye lecture. *Mayo Clin Proc.* 2002;77:1315–1319.
66. Marian AJ. Modifier genes for hypertrophic cardiomyopathy. *Curr Opin Cardiol.* 2002;17:242–252.
67. Gollob MH, Seger JJ, Gollob TN, et al. Novel PRKAG2 mutation responsible for the genetic syndrome of ventricular preexcitation and conduction system disease with childhood onset and absence of cardiac hypertrophy. *Circulation.* 2001;104:3030–3033.
68. Blair E, Redwood C, Ashrafian H, et al. Mutations in the gamma(2) subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis. *Hum Mol Genet.* 2001;10: 1215–1220.
69. Geier C, Perrot A, Ozcelik C, et al. Mutations in the human muscle LIM protein gene in families with hypertrophic cardiomyopathy. *Circulation.* 2003;107:1390–1395.
70. Arad M, Benson DW, Perez-Atayde AR, et al. Constitutively active AMP kinase mutations cause glycogen storage disease mimicking hypertrophic cardiomyopathy. *J Clin Invest.* 2002;109:357–362.
71. Simon DK, Johns DR. Mitochondrial disorders: clinical and genetic features. *Annu Rev Med.* 1999;50:111–127.
72. Watkins H, Rosenzweig A, Hwang DS, et al. Characteristics and prognostic implications of myosin missense mutations in familial hypertrophic cardiomyopathy. *N Engl J Med.* 1992;326:1108–1114.
73. Marian AJ, Roberts R. Molecular genetic basis of hypertrophic cardiomyopathy: genetic markers for sudden cardiac death. *J Cardiovasc Electrophysiol.* 1998;9:88–99.
74. Enjuto M, Francino A, Navarro-Lopez F, Viles D, Pare JC, Ballesta AM. Malignant hypertrophic cardiomyopathy caused by the Arg723Gly mutation in beta-myosin heavy chain gene. *J Mol Cell Cardiol.* 2000;32:2307–2313.
75. Van Driest SL, Ackerman MJ, Ommen SR, et al. Prevalence and severity of "benign" mutations in the beta-myosin heavy chain, cardiac troponin T, and alpha-tropomyosin genes in hypertrophic cardiomyopathy. *Circulation.* 2002;106:3085–3090.
76. Ackerman MJ, VanDriest SL, Ommen SR, et al. Prevalence and age-dependence of malignant mutations in the beta-myosin heavy chain and troponin T genes in hypertrophic cardiomyopathy: a comprehensive outpatient perspective. *J Am Coll Cardiol.* 2002;39: 2042–2048.
77. Watkins H, McKenna WJ, Thierfelder L, et al. Mutations in the genes for cardiac troponin T and alpha-tropomyosin in hypertrophic cardiomyopathy. *N Engl J Med.* 1995;332:1058–1064.
78. Forissier JF, Carrier L, Farza H, et al. Codon 102 of the cardiac troponin T gene is a putative hot spot for mutations in familial hypertrophic cardiomyopathy. *Circulation.* 1996;94:3069–3073.
79. Moolman JC, Corfield VA, Posen B, et al. Sudden death due to troponin T mutations. *J Am Coll Cardiol.* 1997;29:549–555.
80. Charron P, Dubourg O, Desnos M, et al. Clinical features and prognostic implications of familial hypertrophic cardiomyopathy related to the cardiac myosin-binding protein C gene. *Circulation.* 1998;97:2230–2236.
81. Niimura H, Bachinski LL, Sangwatanaroj S, et al. Mutations in the gene for cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy. *N Engl J Med.* 1998;338:1248–1257.
82. Maron BJ, Niimura H, Casey SA, et al. Development of left ventricular hypertrophy in adults in hypertrophic cardiomyopathy caused by cardiac myosin-binding protein C gene mutations. *J Am Coll Cardiol.* 2001;38:315–321.
83. Marian AJ, Yu QT, Workman R, Greve G, Roberts R. Angiotensin-converting enzyme polymorphism in hypertrophic cardiomyopathy and sudden cardiac death. *Lancet.* 1993;342:1085–1086.
84. Lechin M, Quinones MA, Omran A, et al. Angiotensin-I converting enzyme genotypes and left ventricular hypertrophy in patients with hypertrophic cardiomyopathy. *Circulation.* 1995;92:1808–1812.
85. Yoneya K, Okamoto H, Machida M, et al. Angiotensin-converting enzyme gene polymorphism in Japanese patients with hypertrophic cardiomyopathy. *Am Heart J.* 1995;130:1089–1093.
86. Lim DS, Lutucuta S, Bachireddy P, et al. Angiotensin II blockade reverses myocardial fibrosis in a transgenic mouse model of human hypertrophic cardiomyopathy. *Circulation.* 2001;103:789–791.
87. Van Driest SL, Will ML, Atkins DL, Ackerman MJ. A novel TPM1 mutation in a family with hypertrophic cardiomyopathy and sudden cardiac death in childhood. *Am J Cardiol.* 2002;90:1123–1127.
88. Durand JB, Abchee AB, Roberts R. Molecular and clinical aspects of inherited cardiomyopathies. *Ann Med.* 1995;27:311–317.
89. Michels VV, Pastores GM, Moll PP, et al. Dystrophin analysis in idiopathic dilated cardiomyopathy. *J Med Genet.* 1993;30:955–957.
90. Keeling PJ, Gang Y, Smith G, et al. Familial dilated cardiomyopathy in the United Kingdom. *Br Heart J.* 1995;73:417–421.
91. Grunig E, Tasman JA, Kucherer H, Franz W, Kubler W, Katus HA. Frequency and phenotypes of familial dilated cardiomyopathy. *J Am Coll Cardiol.* 1998;31:186–194.
92. Mestroni L, Rocco C, Gregori D, et al. Familial dilated cardiomyopathy: evidence for genetic and phenotypic heterogeneity. Heart Muscle Disease Study Group. *J Am Coll Cardiol.* 1999;34:181–190.
93. Keller DI, Carrier L, Schwartz K. Genetics of familial cardiomyopathies and arrhythmias. *Swiss Med Wkly.* 2002;132:401–407.
94. Olson TM, Michels VV, Thibodeau SN, Tai YS, Keating MT. Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science.* 1998;280:750–752.
95. Olson TM, Kishimoto NY, Whitby FG, Michels VV. Mutations that alter the surface charge of alpha-tropomyosin are associated with dilated cardiomyopathy. *J Mol Cell Cardiol.* 2001;33:723–732.
96. Li D, Tapscoft T, Gonzalez O, et al. Desmin mutation responsible for idiopathic dilated cardiomyopathy. *Circulation.* 1999;100:461–464.
97. Gerull B, Gramlich M, Atherton J, et al. Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nat Genet.* 2002;30:201–204.
98. Tsubata S, Bowles KR, Vatta M, et al. Mutations in the human delta-sarcoglycan gene in familial and sporadic dilated cardiomyopathy. *J Clin Invest.* 2000;106:655–662.
99. Abraham WT, Gilbert EM, Lowes BD, et al. Coordinate changes in Myosin heavy chain isoform gene expression are selectively associated with alterations in dilated cardiomyopathy phenotype. *Mol Med.* 2002;8:750–760.
100. Li D, Czernuszewicz GZ, Gonzalez O, et al. Novel cardiac troponin T mutation as a cause of familial dilated cardiomyopathy. *Circulation.* 2001;104:2188–2193.
101. Keller DI, Carrier L, Schwartz K. Genetics of familial cardiomyopathies and arrhythmias. *Swiss Med Wkly.* 2002;132:401–407.
102. Fatkin D, MacRae C, Sasaki T, et al. Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N Engl J Med.* 1999;341:1715–1724.
103. Schonberger J, Levy H, Grunig E, et al. Dilated cardiomyopathy and sensorineural hearing loss: a heritable syndrome that maps to 6q23–24. *Circulation.* 2000;101:1812–1818.
104. Arbustini E, Pilotto A, Repetto A, et al. Autosomal dominant dilated cardiomyopathy with atrioventricular block: a lamin A/C defect-related disease. *J Am Coll Cardiol.* 2002;39:981–990.
105. Bies RD, Maeda M, Roberds SL, et al. A 5′ dystrophin duplication mutation causes membrane deficiency of alpha-dystroglycan in a family with X-linked cardiomyopathy. *J Mol Cell Cardiol.* 1997;29: 3175–3188.
106. Melacini P, Fanin M, Duggan DJ, et al. Heart involvement in muscular dystrophies due to sarcoglycan gene mutations. *Muscle Nerve.* 1999;22:473–479.
107. Franz WM, Cremer M, Herrmann R, et al. X-linked dilated cardiomyopathy: novel mutation of the dystrophin gene. *Ann N Y Acad Sci.* 1995;752:470–491.

108. Towbin JA, Bowles NE. Genetic abnormalities responsible for dilated cardiomyopathy. *Curr Cardiol Rep.* 2000;2:475–480.
109. Matsuo M. Duchenne and Becker muscular dystrophy: from gene diagnosis to molecular therapy. *IUBMB Life.* 2002;53:147–152.
110. Bione S, D'Adamo P, Maestrini E, Gedeon AK, Bolhuis PA, Toniolo D. A novel X-linked gene, G4.5. is responsible for Barth syndrome. *Nat Genet.* 1996;12:385–389.
111. Barth PG, Wanders RJ, Vreken P, Janssen EA, Lam J, Baas F. X-linked cardioskeletal myopathy and neutropenia (Barth syndrome) (MIM 302060). *J Inherit Metab Dis.* 1999;22:555–567.
112. Cantlay AM, Shokrollahi K, Allen JT, Lunt PW, Newbury-Ecob RA, Steward CG. Genetic analysis of the G4.5 gene in families with suspected Barth syndrome. *J Pediatr.* 1999;135:311–315.
113. Ichida F, Tsubata S, Bowles KR, et al. Novel gene mutations in patients with left ventricular noncompaction or Barth syndrome. *Circulation.* 2001;103:1256–1263.
114. Morris GE, Manilal S. Heart to heart: from nuclear proteins to Emery-Dreifuss muscular dystrophy. *Hum Mol Genet.* 1999;8: 1847–1851.
115. Schmitt JP, Kamisago M, Asahi M, et al. Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science.* 2003;299:1410–1413.
116. Otsu K, Fujii J, Periasamy M, et al. Chromosome mapping of five human cardiac and skeletal muscle sarcoplasmic reticulum protein genes. *Genomics.* 1993;17:507–509.
117. Frank K, Kranias EG. Phospholamban and cardiac contractility. *Ann Med.* 2000;32:572–578.
118. Brittsan AG, Kranias EG. Phospholamban and cardiac contractile function. *J Mol Cell Cardiol.* 2000;32:2131–2139.
119. Haghighi K, Kolokathis F, Pater L, et al. Human phospholamban null results in lethal dilated cardiomyopathy revealing a critical difference between mouse and human. *J Clin Invest.* 2003;111:869–876.
120. Chen J, Chien KR. Complexity in simplicity: monogenic disorders and complex cardiomyopathies. *J Clin Invest.* 1999;103:1483–1485.
121. Towbin JA. The role of cytoskeletal proteins in cardiomyopathies. *Curr Opin Cell Biol.* 1998;10:131–139.
122. Corrado D, Thiene G, Nava A, Rossi L, Pennelli N. Sudden death in young competitive athletes: clinicopathologic correlations in 22 cases. *Am J Med.* 1990;89:588–596.
123. Nava A, Bauce B, Basso C, et al. Clinical profile and long-term follow-up of 37 families with arrhythmogenic right ventricular cardiomyopathy. *J Am Coll Cardiol.* 2000;36:2226–2233.
124. Basso C, Thiene G, Corrado D, Angelini A, Nava A, Valente M. Arrhythmogenic right ventricular cardiomyopathy. Dysplasia, dystrophy, or myocarditis? *Circulation.* 1996;94:983–991.
125. Rampazzo A, Nava A, Danieli GA, et al. The gene for arrhythmogenic right ventricular cardiomyopathy maps to chromosome 14q23–q24. *Hum Mol Genet.* 1994;3:959–962.
126. Rampazzo A, Nava A, Erne P, et al. A new locus for arrhythmogenic right ventricular cardiomyopathy (ARVD2) maps to chromosome 1q42–q43. *Hum Mol Genet.* 1995;4:2151–2154.
127. Severini GM, Krajinovic M, Pinamonti B, et al. A new locus for arrhythmogenic right ventricular dysplasia on the long arm of chromosome 14. *Genomics.* 1996;31:193–200.
128. Rampazzo A, Nava A, Miorin M, et al. ARVD4, a new locus for arrhythmogenic right ventricular cardiomyopathy, maps to chromosome 2 long arm. *Genomics.* 1997;45:259–263.
129. Ahmad F, Li D, Karibe A, et al. Localization of a gene responsible for arrhythmogenic right ventricular dysplasia to chromosome 3p23. *Circulation.* 1998;98:2791–2795.
130. Li D, Ahmad F, Gardner MJ, et al. The locus of a novel gene responsible for arrhythmogenic right-ventricular dysplasia characterized by early onset and high penetrance maps to chromosome 10p12–p14. *Am J Hum Genet.* 2000;66:148–156.
131. Melberg A, Oldfors A, Blomstrom-Lundqvist C, et al. Autosomal dominant myofibrillar myopathy with arrhythmogenic right ventricular cardiomyopathy linked to chromosome 10q. *Ann Neurol.* 1999;46:684–692.
132. Rampazzo A, Nava A, Malacrida S, et al. Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy. *Am J Hum Genet.* 2002;71:1200–1206.
133. Coonar AS, Protonotarios N, Tsatsopoulou A, et al. Gene for arrhythmogenic right ventricular cardiomyopathy with diffuse nonepidermolytic palmoplantar keratoderma and woolly hair (Naxos disease) maps to 17q21. *Circulation.* 1998;97:2049–2058.
134. Frances R, Rodriguez Benitez AM, Cohen DR. Arrhythmogenic right ventricular dysplasia and anterior polar cataract. *Am J Med Genet.* 1997;73:125–126.
135. Nava A, Thiene G, Canciani B, et al. Familial occurrence of right ventricular dysplasia: a study involving nine families. *J Am Coll Cardiol.* 1988;12:1222–1228.
136. Protonotarios N, Tsatsopoulou A, Anastasakis A, et al. Genotype-phenotype assessment in autosomal recessive arrhythmogenic right ventricular cardiomyopathy (Naxos disease) caused by a deletion in plakoglobin. *J Am Coll Cardiol.* 2001;38:1477–1484.
137. Tiso N, Stephan DA, Nava A, et al. Identification of mutations in the cardiac ryanodine receptor gene in families affected with arrhythmogenic right ventricular cardiomyopathy type 2 (ARVD2). *Hum Mol Genet.* 2001;10:189–194.
138. Kushwaha SS, Fallon JT, Fuster V. Restrictive cardiomyopathy. *N Engl J Med.* 1997;336:267–276.
139. Zachara E, Bertini E, Lioy E, Boldrini R, Prati PL, Bosman C. Restrictive cardiomyopathy due to desmin accumulation in a family with evidence of autosomal dominant inheritance. *G Ital Cardiol.* 1997;27:436–442.
140. Aroney C, Bett N, Radford D. Familial restrictive cardiomyopathy. *Aust N Z J Med.* 1988;18:877–878.
141. Goldfarb LG, Park KY, Cervenakova L, et al. Missense mutations in desmin associated with familial cardiac and skeletal myopathy. *Nat Genet.* 1998;19:402–403.
142. Dalakas MC, Park KY, Semino-Mora C, Lee HS, Sivakumar K, Goldfarb LG. Desmin myopathy, a skeletal myopathy with cardiomyopathy caused by mutations in the desmin gene. *N Engl J Med.* 2000;342:770–780.
143. Zhang J, Kumar A, Stalker HJ, et al. Clinical and molecular studies of a large family with desmin-associated restrictive cardiomyopathy. *Clin Genet.* 2001;59:248–256.
144. Mogensen J, Kubo T, Duque M, et al. Idiopathic restrictive cardiomyopathy is part of the clinical expression of cardiac troponin I mutations. *J Clin Invest.* 2003;111:209–216.
145. Ichida F, Hamamichi Y, Miyawaki T, et al. Clinical features of isolated noncompaction of the ventricular myocardium: long-term clinical course, hemodynamic properties, and genetic background. *J Am Coll Cardiol.* 1999;34:233–240.
146. Chin TK, Perloff JK, Williams RG, Jue K, Mohrmann R. Isolated noncompaction of left ventricular myocardium: a study of eight cases. *Circulation.* 1990;82:507–513.
147. Bleyl SB, Mumford BR, Thompson V, et al. Neonatal, lethal noncompaction of the left ventricular myocardium is allelic with Barth syndrome. *Am J Hum Genet.* 1997;61:868–872.
148. Bleyl SB, Mumford BR, Brown-Harrison MC, et al. Xq28-linked noncompaction of the left ventricular myocardium: prenatal diagnosis and pathologic analysis of affected individuals. *Am J Med Genet.* 1997;72:257–265.
149. D'Adamo P, Fassone L, Gedeon A, et al. The X-linked gene G4.5 is responsible for different infantile dilated cardiomyopathies. *Am J Hum Genet.* 1997;61:862–867.
150. Gedeon AK, Wilson MJ, Colley AC, Sillence DO, Mulley JC. X linked fatal infantile cardiomyopathy maps to Xq28 and is possibly allelic to Barth syndrome. *J Med Genet.* 1995;32:383–388.
151. Chen R, Tsuji T, Ichida F, et al. Mutation analysis of the G4.5 gene in patients with isolated left ventricular noncompaction. *Mol Genet Metab.* 2002;77:319–325.

152. Ackerman MJ, Clapham DE. Ion channels—basic science and clinical disease. *N Engl J Med.* 1997;336:1575–1586.
153. Ackerman MJ. The long QT syndrome: ion channel diseases of the heart. *Mayo Clin Proc.* 1998;73:250–269.
154. Keating MT, Sanguinetti MC. Molecular and cellular mechanisms of cardiac arrhythmias. *Cell.* 2001;104:569–580.
155. Wang Q, Curran ME, Splawski I, et al. Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nat Genet.* 1996;12:17–23.
156. Curran ME, Splawski I, Timothy KW, Vincent GM, Green ED, Keating MT. A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell.* 1995;80:795–803.
157. Wang Q, Shen J, Splawski I, et al. SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. *Cell.* 1995;80: 805–811.
158. Mohler PJ, Schott JJ, Gramolini AO, et al. Ankyrin-B mutation causes type 4 long-QT cardiac arrhythmia and sudden cardiac death. *Nature.* 2003;421:634–639.
159. Splawski I, Tristani-Firouzi M, Lehmann MH, Sanguinetti MC, Keating MT. Mutations in the hminK gene cause long QT syndrome and suppress IKs function. *Nat Genet.* 1997;17:338–340.
160. Abbott GW, Sesti F, Splawski I, et al. MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell.* 1999;97:175–187.
161. Splawski I, Shen J, Timothy KW, et al. Spectrum of mutations in long-QT syndrome genes. KVLQT1, HERG, SCN5A, KCNE1, and KCNE2. *Circulation.* 2000;102:1178–1185.
162. Curran ME, Splawski I, Timothy KW, Vincent GM, Green ED, Keating MT. A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell.* 1995;80:795–803.
163. Jiang C, Atkinson D, Towbin JA, et al. Two long QT syndrome loci map to chromosomes 3 and 7 with evidence for further heterogeneity. *Nat Genet.* 1994;8:141–147.
164. Wang Q, Shen J, Splawski I, et al. SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. *Cell.* 1995; 80:805–811.
165. Wang Q, Shen J, Li Z, et al. Cardiac sodium channel mutations in patients with long QT syndrome, an inherited cardiac arrhythmia. *Hum Mol Genet.* 1995;4:1603–1607.
166. Ackerman MJ, Siu BL, Sturner WQ, et al. Postmortem molecular analysis of SCN5A defects in sudden infant death syndrome. *JAMA.* 2001;286:2264–2269.
167. Schott JJ, Charpentier F, Peltier S, et al. Mapping of a gene for long QT syndrome to chromosome 4q25–27. *Am J Hum Genet.* 1995;57: 1114–1122.
168. Mohler PJ, Gramolini AO, Bennett V. Ankyrins. *J Cell Sci.* 2002;115: 1565–1566.
169. Tyson J, Tranebjaerg L, Bellman S, et al. IsK and KvLQT1: mutation in either of the two subunits of the slow component of the delayed rectifier potassium channel can cause Jervell and Lange-Nielsen syndrome. *Hum Mol Genet.* 1997;6:2179–2185.
170. Chen Q, Zhang D, Gingell RL, et al. Homozygous deletion in KVLQT1 associated with Jervell and Lange-Nielsen syndrome. *Circulation.* 1999;99:1344–1347.
171. Splawski I, Timothy KW, Vincent GM, Atkinson DL, Keating MT. Molecular basis of the long-QT syndrome associated with deafness. *N Engl J Med.* 1997;336:1562–1567.
172. Schulze-Bahr E, Wang Q, Wedekind H, et al. KCNE1 mutations cause Jervell and Lange-Nielsen syndrome. *Nat Genet.* 1997;17:267–268.
173. Schulze-Bahr E, Wedekind H, Haverkamp W, et al. The LQT syndromes—current status of molecular mechanisms. *Z Kardiol.* 1999;88:245–254.
174. Romey G, Attali B, Chouabe C, et al. Molecular mechanism and functional significance of the MinK control of the KvLQT1 channel activity. *J Biol Chem.* 1997;272:16713–16716.
175. Drici MD, Arrighi I, Chouabe C, et al. Involvement of IsK-associated K+ channel in heart rate control of repolarization in a murine engineered model of Jervell and Lange- Nielsen syndrome. *Circ Res.* 1998;83:95–102.
176. Ackerman MJ, Tester DJ, Porter CJ. Swimming, a gene-specific arrhythmogenic trigger for inherited long QT syndrome. *Mayo Clin Proc.* 1999;74:1088–1094.
177. Schwartz PJ, Priori SG, Spazzolini C, et al. Genotype-phenotype correlation in the long-QT syndrome: gene-specific triggers for life-threatening arrhythmias. *Circulation.* 2001;103:89–95.
178. Moss AJ, Zareba W, Hall WJ, et al. Effectiveness and limitations of beta-blocker therapy in congenital long-QT syndrome. *Circulation.* 2000;101:616–623.
179. Zareba W, Moss AJ, Schwartz PJ, et al. Influence of genotype on the clinical course of the long-QT syndrome. International Long-QT Syndrome Registry Research Group. *N Engl J Med.* 1998;339:960–965.
180. Moss AJ, Zareba W, Kaufman ES, et al. Increased risk of arrhythmic events in long-QT syndrome with mutations in the pore region of the human ether-a-go-go-related gene potassium channel. *Circulation.* 2002;105:794–799.
181. Shimizu W, Tanabe Y, Aiba T, et al. Differential effects of beta-blockade on dispersion of repolarization in the absence and presence of sympathetic stimulation between the LQT1 and LQT2 forms of congenital long QT syndrome. *J Am Coll Cardiol.* 2002;39:1984–1991.
182. Ackerman MJ, Khositseth A, Tester DJ, Hejlik JB, Shen WK, Porter CB. Epinephrine-induced QT interval prolongation: a gene-specific paradoxical response in congenital long QT syndrome. *Mayo Clin Proc.* 2002;77:413–421.
183. Martini B, Nava A, Thiene G, et al. Ventricular fibrillation without apparent heart disease: description of six cases. *Am Heart J.* 1989;118:1203–1209.
184. Brugada P, Brugada J. Right bundle branch block, persistent ST segment elevation and sudden cardiac death: a distinct clinical and electrocardiographic syndrome. A multicenter report. *J Am Coll Cardiol.* 1992;20:1391–1396.
185. Remme CA, Wever EF, Wilde AA, Derksen R, Hauer RN. Diagnosis and long-term follow-up of the Brugada syndrome in patients with idiopathic ventricular fibrillation. *Eur Heart J.* 2001;22:400–409.
186. Viskin S, Belhassen B. Idiopathic ventricular fibrillation. *Am Heart J.* 1990;120:661–671.
187. Chen Q, Kirsch GE, Zhang D, et al. Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. *Nature.* 1998;392: 293–296.
188. Priori SG, Napolitano C, Gasparini M, et al. Natural history of Brugada syndrome: insights for risk stratification and management. *Circulation.* 2002;105:1342–1347.
189. Priori SG, Napolitano C, Gasparini M, et al. Clinical and genetic heterogeneity of right bundle branch block and ST-segment elevation syndrome: a prospective evaluation of 52 families. *Circulation.* 2000;102:2509–2515.
190. Wichter T, Schulze-Bahr E, Eckardt L, et al. Molecular mechanisms of inherited ventricular arrhythmias. *Herz.* 2002;27:712–739.
191. Weiss R, Barmada MM, Nguyen T, et al. Clinical and molecular heterogeneity in the Brugada syndrome: a novel gene locus on chromosome 3. *Circulation.* 2002;105:707–713.
192. Akai J, Makita N, Sakurada H, et al. A novel SCN5A mutation associated with idiopathic ventricular fibrillation without typical ECG findings of Brugada syndrome. *FEBS Lett.* 2000;479:29–34.
193. Leenhardt A, Lucet V, Denjoy I, Grau F, Ngoc DD, Coumel P. Catecholaminergic polymorphic ventricular tachycardia in children: a 7-year follow-up of 21 patients. *Circulation.* 1995;91:1512–1519.
194. Swan H, Piippo K, Viitasalo M, et al. Arrhythmic disorder mapped to chromosome 1q42–q43 causes malignant polymorphic ventricular tachycardia in structurally normal hearts. *J Am Coll Cardiol.* 1999;34:2035–2042.
195. Lahat H, Eldar M, Levy-Nissenbaum E, et al. Autosomal recessive catecholamine- or exercise-induced polymorphic ventricular tachycardia: clinical features and assignment of the disease gene to chromosome 1p13–21. *Circulation.* 2001;103:2822–2827.

196. Priori SG, Napolitano C, Tiso N, et al. Mutations in the cardiac ryanodine receptor gene (hRyR2) underlie catecholaminergic polymorphic ventricular tachycardia. *Circulation.* 2001;103:196–200.
197. Priori SG, Napolitano C, Memmi M, et al. Clinical and molecular characterization of patients with catecholaminergic polymorphic ventricular tachycardia. *Circulation.* 2002;106:69–74.
198. Eldar M, Pras E, Lahat H. A missense mutation in the CASQ2 gene is associated with autosomal-recessive catecholamine-induced polymorphic ventricular tachycardia. *Trends Cardiovasc Med.* 2003;13:148–151.
199. Andersen ED, Krasilnikoff PA, Overvad H. Intermittent muscular weakness, extrasystoles, and multiple developmental anomalies: a new syndrome? *Acta Paediatr Scand.* 1971;60:559–564.
200. Sansone V, Griggs RC, Meola G, et al. Andersen's syndrome: a distinct periodic paralysis. *Ann Neurol.* 1997;42:305–312.
201. Canun S, Perez N, Beirana LG. Andersen syndrome autosomal dominant in three generations. *Am J Med Genet.* 1999;85:147–156.
202. Plaster NM, Tawil R, Tristani-Firouzi M, et al. Mutations in Kir2.1 cause the developmental and episodic electrical phenotypes of Andersen's syndrome. *Cell.* 2001;105:511–519.
203. Andelfinger G, Tapper AR, Welch RC, Vanoye CG, George AL Jr, Benson DW. KCNJ2 mutation results in Andersen syndrome with sex-specific cardiac and skeletal muscle phenotypes. *Am J Hum Genet.* 2002;71:663–668.
204. Ai T, Fujiwara Y, Tsuji K, et al. Novel KCNJ2 mutation in familial periodic paralysis with ventricular dysrhythmia. *Circulation.* 2002;105:2592–2594.
205. Tristani-Firouzi M, Jensen JL, Donaldson MR, et al. Functional and clinical characterization of KCNJ2 mutations associated with LQT7 (Andersen syndrome). *J Clin Invest.* 2002;110:381–388.

Chapter 15

Neurodegenerative Disorders

Nicholas T. Potter

HUNTINGTON DISEASE

Molecular Basis of Disease

Huntington disease (HD) is an autosomal dominant neurodegenerative disorder clinically characterized by the presence of choreiform movements, psychiatric sequelae, and dementia. While the majority (>90%) of HD patients become clinically symptomatic in adulthood, 5% to 10% of patients present with the juvenile-onset form of the disease, which is almost invariably associated with inheritance of the mutant allele from a symptomatic father. Unlike the adult-onset form of the disease, juvenile HD is generally characterized by the presence of progressive rigidity, seizures, ataxia, and dystonia.

The HD gene (*IT15*) is located on chromosome 4p16. The gene spans more than 200 kilobases (kb) and contains 67 exons. The encoded protein, huntingtin, is a 3,136 amino acid protein with an approximate molecular weight of 350 kilodaltons (kDa) that shares no known homologies with other characterized proteins. HD is caused by a toxic gain-of-function mechanism associated with the expansion of a polyglutamine tract within the protein that is translated from a CAG repeat region within exon 1 of the gene. The CAG repeat length is highly polymorphic in the general population,[1] with the largest normal allele currently defined as carrying ≤26 CAG repeats. Alleles carrying ≥40 CAG repeats, identified in symptomatic patients with a presumptive clinical diagnosis of HD, are considered diagnostic. Alleles with 27 to 35 CAG repeats are defined as "mutable normal or intermediate alleles." While alleles in this size range have yet to be convincingly associated with an HD phenotype, they can be meiotically unstable in sperm and result in the expansion of paternally derived alleles.[2] Alleles found to carry 36 to 39 CAG repeats are defined as "HD alleles with reduced penetrance." Alleles in this size range have been found in both clinically as well as neuropathologically confirmed HD patients and in elderly asymptomatic individuals.[2]

Clinical Utility of Testing

The direct quantitative analysis of HD CAG repeat number is clinically available in a large number of clinical molecular laboratories.[3] As HD appears to be a genetically homogeneous disease, molecular testing has been routinely utilized with confidence for diagnostic, predictive, and prenatal purposes. In the case of predictive testing, professional standards strongly encourage that this test be offered only to individuals who have reached the statutory age of majority (generally 18 years of age). For adults, a formal multidisciplinary predictive testing protocol is offered at many sites for those desiring determination of their carrier status before the onset of symptoms. The recommended presymptomatic protocol includes pretest counseling and evaluation of the individual being tested by a clinical geneticist, genetic counselor, psychologist or psychiatrist, and neurologist over several visits, with the option to not receive test results even after testing has been completed. With few exceptions, interpretation of laboratory results is unambiguous and the resulting genetic counseling consultations follow those for other autosomal dominant adult-onset disorders.

Available Assays

Detection of HD CAG repeat expansions can be determined by both PCR and Southern blot methods. For PCR, several sets of primers, assay conditions, amplicon separation, and detection techniques have been published.[2] Regardless of the particular PCR test method employed, optimization of the assay conditions and post-PCR analyses are essential to ensure accurate and unambiguous quantitation of the HD CAG repeat length (Figure 15-1). CAG sizing anomalies have been observed in both comparative studies of ^{32}P incorporation methods[2] and comparative post-PCR analyses utilizing agarose, capillary, and denaturing polyacrylamide gel electrophoretic methods.

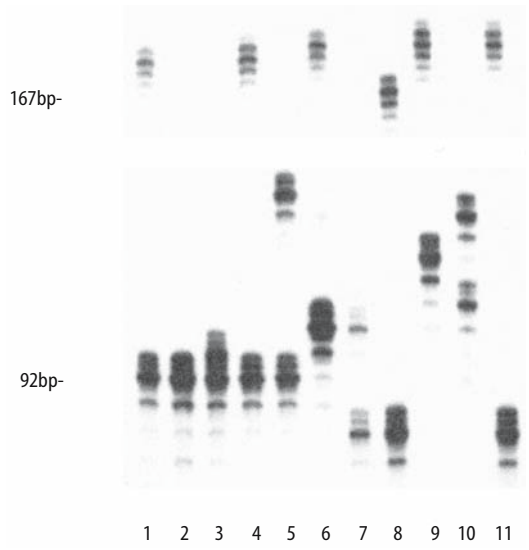


Figure 15-1. PCR genotyping of the Huntington disease polymorphic CAG repeat utilizing ^{32}P-dCTP incorporation and the AI/C2 primer pair.[4,5] Separation and sizing of alleles was performed on a 35 cm × 43 cm 6% denaturing polyacrylamide gel followed by autoradiography. The numbers of CAG repeats in each of the two alleles are 17/44, 17/17, 17/18, 17/44, 17/25, 19/45, 15/19, 15/40, 22/45, 20/24, and 15/45 for patients 1 to 11, respectively.

While PCR testing has a sensitivity of approximately 99%, the fidelity of amplification of very large alleles is not empirically known. However, conventional PCR protocols reproducibly amplify alleles in the 115 to 120 CAG repeat range,[2] and modified PCR protocols (long-template PCR) have enabled detection of an approximately 265 CAG repeat allele in a patient with juvenile-onset disease.[6] As such, accurate quantitation of patient amplicon sizes should be empirically determined by comparison to appropriate external or internal standards. These generally include M13 sequencing ladders, cloned reference standards, and appropriate normal and abnormal patient controls with independently verified repeat sizes, usually by sequence analysis of the CAG repeat.

As the length of the polymorphic CAG repeat alone is associated with HD, patient genotyping based on the use of a single primer pair that amplifies both the CAG and adjacent CCG repeat[2] should be discouraged. The CCG repeat, which lies 12 base pairs (bp) 3′ of the HD CAG repeat, has been shown to be polymorphic and as a result may lead to diagnostic inaccuracies for both normal and HD allele sizing. Furthermore, polymorphisms surrounding or within the CAG tract have been identified with a collective frequency of approximately 1%.[2] These nucleotide substitutions can be categorized into two groups: (1) those that modify primer-annealing sites and (2) those that result in the loss of sequence interruption between the CAG and CCG tracts. With the first category, nucleotide changes may result in the misinterpretation of genotyping data due to an allele-specific amplification failure associated with improper primer annealing. In the second category, rare A-to-G substitutions within the intervening 12 bp segment between the CAG and CCG tracts can result in increased meiotic instability of the tract as well as miscalculation of uninterrupted CAG repeat length based on conventional calculation formulas.[2]

Southern blot methods, although not generally used for routine diagnostic testing, are often essential for the identification or confirmation of very large expansions typically associated with juvenile-onset HD, because these alleles are typically refractory to robust amplification by conventional PCR. Furthermore, this method may be useful for the confirmation of patients (generally children) who appear to be homozygous for two normal-sized alleles.[2]

Interpretation of Test Results

PCR testing for the HD CAG repeat has a reported sensitivity of 99%. The remaining 1% of patients represent HD phenocopies, and to date at least two distinct genetic loci (*HDL1* and *HDL2*) have been identified.[2] Testing specificity is 100%.[1] In summary, the interpretative challenges in HD testing are primarily related to the need to recognize that CAG ranges and descriptors may be modified over time. For reporting of HD results, the American College of Medical Genetics (ACMG) has recently established updated guidelines regarding the definitions of CAG repeat range descriptors and interpretative guidelines for genetic counseling purposes.[2]

Laboratory Issues

There is currently no commercially available in vitro diagnostic test kit for HD. The College of American Pathologists (CAP) offers HD proficiency challenges twice a year as part of the CAP/ACMG Molecular Genetics Laboratory (MGL) Survey. Genotyped HD patient DNA for use as controls can be obtained from the Coriell Cell Repositories (Camden, NJ; http://ccr.coriell.org/).

AUTOSOMAL DOMINANT CEREBELLAR ATAXIAS AND FRIEDREICH ATAXIA

Molecular Basis of Disease

The autosomal dominant cerebellar ataxias (ADCAs) collectively represent a clinically heterogeneous group of neurological disorders. To date, at least 27 distinct loci have been identified, and clinical testing is presently available for 12 of these (Table 15-1). Because the collective incidence of these disorders is high in both unselected and selected ataxia cohorts[7,8] and the degree of clinical overlap among the various ADCAs is extensive, molecular genetic testing is a valuable diagnostic tool. Most of the molecularly characterized ADCAs (dentatorubro pallidoluysian atrophy or DRPLA and SCA1, 2, 3, 6, 7, 8, 10, 12, and 17) are

Table 15-1. Molecular Classification of the Autosomal Dominant Cerebellar Ataxias (ADCA) and Friedreich Ataxia (FRDA)

Locus	Inheritance	Chromosome	Gene Product	Clinical Test[1]
DRPLA	AD	12p13	Atrophin-1	Yes
SCA-1	AD	6p23	Ataxin-1	Yes
SCA-2	AD	12q24	Ataxin-2	Yes
SCA-3	AD	14q24	Ataxin-3	Yes
SCA-4	AD	16q22	Puratrophin-1	No
SCA-5	AD	11p11	Spectrin beta chain, brain 2	Yes
SCA-6	AD	19p13	Voltage-dependent P/Q-type calcium channel a-1A subunit	Yes
SCA-7	AD	3p21	Ataxin-7	Yes
SCA-8	AD	13q21	Not named	Yes
SCA-9 (not assigned)	AD			
SCA-10	AD	22q13	Ataxin-10	Yes
SCA-11	AD	15q14	Unknown	No
SCA-12	AD	5q31	Protein phospatase 2, regulatory subunit B	Yes
SCA-13	AD	19q13	Potassium voltage-gated channel subfamily 3, mumber 3	No
SCA-14	AD	19q13	Protein kinase C, gamma type	Yes
SCA-15	AD	3p26	Unknown	No
SCA-16	AD	8q22	Unknown	No
SCA-17	AD	6q27	TATA-box binding protein	Yes
SCA-18	AD	7q22	Unknown	No
SCA-19	AD	1q21	Unknown	No
SCA-20	AD	11cen	Unknown	No
SCA-21	AD	7p21	Unknown	No
SCA-22	AD	1p21	Unknown	No
SCA-23	AD	20p13	Unknown	No
SCA-25	AD	2p21	Unknown	No
SCA-26	AD	19p13	Unknown	
SCA-27	AD	13p34	Fibroblast growth factor 14	No
SCA-28	AD	18p11	Unknown	No
FRDA	AR	9p13	Frataxin	Yes

Sources: See GeneTests (www.genetests.org), The Human Gene Nomenclature Database (www.gene.ucl.ac.uk/nomenclature/), and The Neuromuscular Disease Center, Washington University, St Louis, MO (www.neuro.wustl.edu/neuromuscular/index.html).
[1]DNA-based testing.
AD, autosomal dominant; AR; autosomal recessive.

associated with reiterated repeat expansion mutations. DRPLA and SCA1, 2, 3, 6, 7, 12, and 17 are associated with CAG trinucleotide repeat expansion mutations. SCA8 is associated with a CTG expansion mutation, and SCA10 with an ATTCT pentanucleotide repeat expansion.

In the majority of these disorders the location of the repeat is exonic (DRPLA and SCA1, 2, 3, 6, 7, and 17), and the mechanism of disease pathogenesis (with the exception of SCA6) is thought to involve a toxic gain of function in the protein. With some exceptions, SCA5 is associated with the spectrin beta chain,[3] the function of the normal gene product is not known. SCA6 is associated with a CAG expansion in the *CACNA1A* gene, which encodes the voltage-dependent P/Q-type calcium channel α-1A subunit. SCA12 is associated with a CAG expansion in the 5′ UTR of the *PPP2R2B* gene, which encodes a brain-specific regulatory subunit of the protein phosphatase PP2A.[9] SCA13 with the protein voltage-gated channel subfamily C member 3,[3] SCA14 with the protein kinase C, gamma type,[3] SCA17 is associated with a CAG repeat expansion in the TATA-binding protein (*TBP*) gene, which encodes a general transcription initiation factor.[10] SCA8 is associated with a CTG expansion within the 3′ untranslated exon of a fully processed transcript of a gene of unknown function[11] and SCA10 is associated with an intronic expansion of a pentanucleotide (ATTCT) sequence within the *SCA10* gene encoding a protein of unknown function.[12]

Friedreich ataxia (FRDA) is the most common form of hereditary ataxia, with a prevalence of approximately 2 to 4 per 100,000 and a carrier frequency of approximately 1 in 90. The disease has an autosomal recessive mode of inheritance. FRDA is clinically characterized by the onset of gait and limb ataxia before the age of 25 years, decreased deep tendon reflexes, dysarthria, pyramidal signs, Babinski responses, and decreased position or vibration sense (or both) in the lower limbs. Approximately 65% of patients also develop a cardiomyopathy, while approximately 10% of patients develop diabetes. Atypical clinical presentations have been reported in approximately 25% of patients and generally have included later ages of onset (after 25 years of age), retained reflexes, and slower disease progression. In addition, genetically confirmed patients with more complex phenotypes have recently been reported and include individuals with idiopathic spastic paraparesis, motor and sensory neuropathy, limb and axial dystonia, and chorea with myoclonus.[13]

FRDA is caused by mutations in the *FRDA* gene that encodes the protein frataxin. Frataxin is a 210 amino acid protein that localizes to the mitochondrial inner membrane, where it is required for mitochondrial iron

homeostasis. Approximately 96% of patients are homozygous for GAA expansion mutations within intron 1 of the *FRDA* gene. However, the remaining 4% of patients are compound heterozygotes for a GAA expansion on one allele and a point mutation on the other. Both mutation types, however, are gene inactivating.[13]

Clinical Utility of Testing

The utility of offering testing for the molecularly characterized ADCAs and FRDA is very high. For the ADCAs, diagnostic, predictive, as well as prenatal testing can be performed with confidence. Although allelic heterogeneity has been reported in FRDA (i.e., point mutations in compound heterozygotes), the vast majority of FRDA patients have homozygous expansions of the GAA repeat; hence, diagnostic, predictive, and prenatal testing can also be offered with confidence. As is the case with HD, predictive testing for the ADCAs should be offered only to individuals who have reached the statutory age of majority (generally 18 years age). For adults, a formal multidisciplinary predictive testing protocol, similar to the one utilized for HD predictive testing, should be offered to those desiring determination of their carrier status. Again, with few exceptions, interpretation of the laboratory result is unambiguous, and the resulting genetic counseling consultations follow those for other autosomal dominant adult-onset disorders. For FRDA families, appropriate counseling consistent with an autosomal recessive pattern of inheritance is required.

Available Assays

Detection and quantitation of the repeat expansions associated with the ADCAs can be determined by both PCR and Southern blot methods, and the choice of methodology is dependent on the characteristic size range for the disease-specific expansion mutation. For example, very large expansions, such as those routinely seen in SCA10, require the use of Southern blot, whereas those seen in other ADCAs (SCA6, for example) are amenable to detection by routine PCR analysis. There are some notable exceptions, however, which are described below. As already alluded to for HD, regardless of the particular PCR method employed, assay conditions and post-PCR analyses should be optimized to ensure accurate and unambiguous quantitation of repeat length. Furthermore, CAG sizing anomalies for ADCA testing also have been observed in comparative studies of capillary and denaturing polyacrylamide gel electrophoretic methods.[14] As such, accurate quantitation of patient amplicon sizes should be empirically determined by comparison to appropriate external or internal standards.

One of the important methodological considerations for quantitation of expansion mutations in general is the recognition of the existence of "extreme expansions" associated with several of the SCAs. For example, both the infantile and juvenile-onset forms of SCA2 and SCA7 have been associated with large, or "extreme," expansions generally greater than 200 CAG repeats.[15] In a situation that is analogous to the large expansions of juvenile-onset HD, these alleles can be refractory to efficient PCR or difficult to separate using conventional polyacrylamide gel electrophoresis (PAGE). Test results that are apparently "homozygous normal" genotypes in infants or children with a high index of clinical suspicion of SCA2 or SCA7 should be confirmed using a Southern blot test that can detect large expansions (Figure 15-2a and b).[17] This testing strategy also is important for SCA8, since repeat lengths of greater than 800 have been described in adults with SCA8.[18] Detailed current information regarding repeat ranges, clinical descriptors, and disease-specific commentaries for the ADCAs is available on the web.[19]

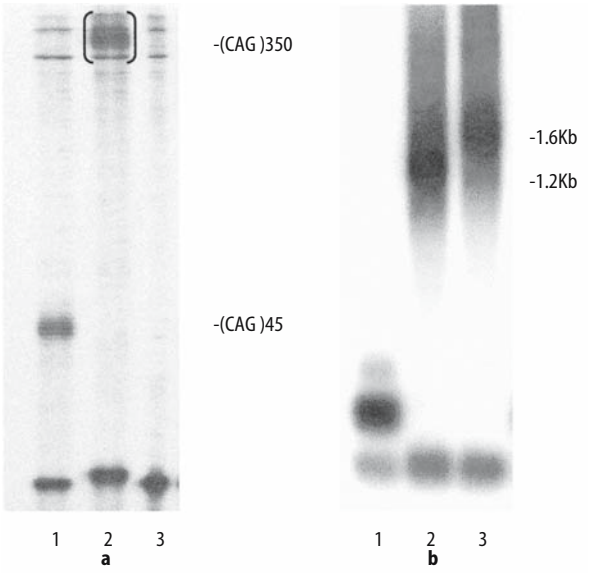


Figure 15-2. Molecular testing for the SCA2 CAG repeat. (a) PCR genotyping of the SCA2 polymorphic CAG repeat utilizing the UH13/UH10 primer pair[16] is shown for two related patients (lane 1, mother; lane 2, son) with adult-onset and juvenile-onset SCA2, respectively.[15] Separation and sizing of alleles was performed on a 35 cm × 43 cm 6% denaturing polyacrylamide gel followed by autoradiography. The numbers of CAG repeats in each of the two alleles are 22/45 (lane 1, affected mother run in duplicate), 23/~350 (lane 2, affected child run in duplicate), and 22/22 (lane 3, normal unaffected control). (b) PCR-Southern blot analysis[17] of the same mother-child samples shown in Figure 15-2a, confirming the presence of the extreme expansion in the affected son. Allele sizes are 22/45 (lane 1, affected mother), 23/~350 (lane 2, affected child), and 23/~400 (lane 3, positive extreme expansion control). (Figure 15-2b courtesy of Dr. Karen Snow-Bailey, Auckland District Health Board, Auckland, New Zealand.)

Analysis of the FRDA GAA expansion mutation is generally performed by long-template PCR (Figure 15-3a and b) or by Southern blot analysis. Although the fidelity of this assay is high, PCR artifacts resembling expansions can be seen, particularly in samples with two normal-sized alleles (Figure 15-3a, lane 4). The expansion artifacts are thought to be heteroduplexes, as they do not appear when amplicons are analyzed under denaturing conditions. These artifacts should not result in significant diagnostic errors for patients who are clearly heterozygous for two normal-sized

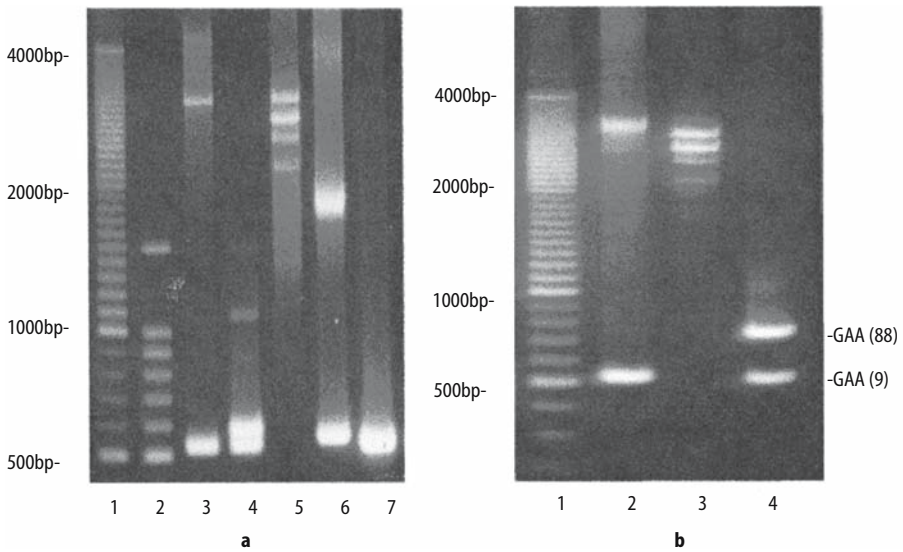


Figure 15-3. PCR analysis of the FRDA polymorphic GAA expansion mutation utilizing the 104F/629R primer pair[20] and the Expand Long Template PCR System (Roche Diagnostics, Mannheim, Germany). (a) Lane 1, 100 bp step ladder G6951 (Promega, Madison, WI); lane 2, 100 bp ladder G2101 (Promega); lane 3, FRDA carrier with ~850 GAA repeats; lane 4, normal individual with allele sizes of 9 and 26 GAA repeats; lane 5, FRDA patient with allele sizes of ~800/~900 GAA repeats; lane 6, FRDA carrier with ~450 GAA repeats; lane 7, normal control. (b) Lane 1, 100 bp step ladder G6951 (Promega); lane 2, FRDA compound heterozygote (see Figure 15-4) with ~1000 GAA repeats; lane 3, FRDA patient with allele sizes of ~800/~900 GAA repeats; lane 4, FRDA carrier with allele sizes of 9 and 88 GAA repeats (confirmed by DNA sequencing). PCR products were separated through 1% agarose on a 10.5 cm × 14 cm horizontal gel and the alleles visualized after ethidium bromide staining.

alleles, as the intensity of the true alleles is usually much greater than that of the heteroduplexes. Ambiguities may remain, however, because the resolving power of 1% agarose gels for amplicons in the 500 bp range is limited. These can be clarified, if necessary, by either Southern blot or denaturing PAGE.

Interpretation of Test Results

PCR tests for the ADCAs have a sensitivity of >99%, and complete testing including Southern blot has a sensitivity of close to 100%. For the reporting of results, interpretation of the findings requires the integration of available clinical information, repeat size, and its clinical descriptor (i.e., normal, intermediate, or abnormal). Again, the subtleties of effective genetic counseling for the ADCAs require an understanding of the significance of repeat sizes and an appreciation that repeat ranges and descriptors may change over time.[19] PCR tests for FRDA have a sensitivity of approximately 96%, but the presence of a frataxin point mutation must be considered in patients with a high degree of clinical suspicion and only one expanded GAA allele (Figure 15-3b, lane 2). In such situations, referral to a laboratory that offers point mutation analysis is strongly recommended. See also Figure 15-4.

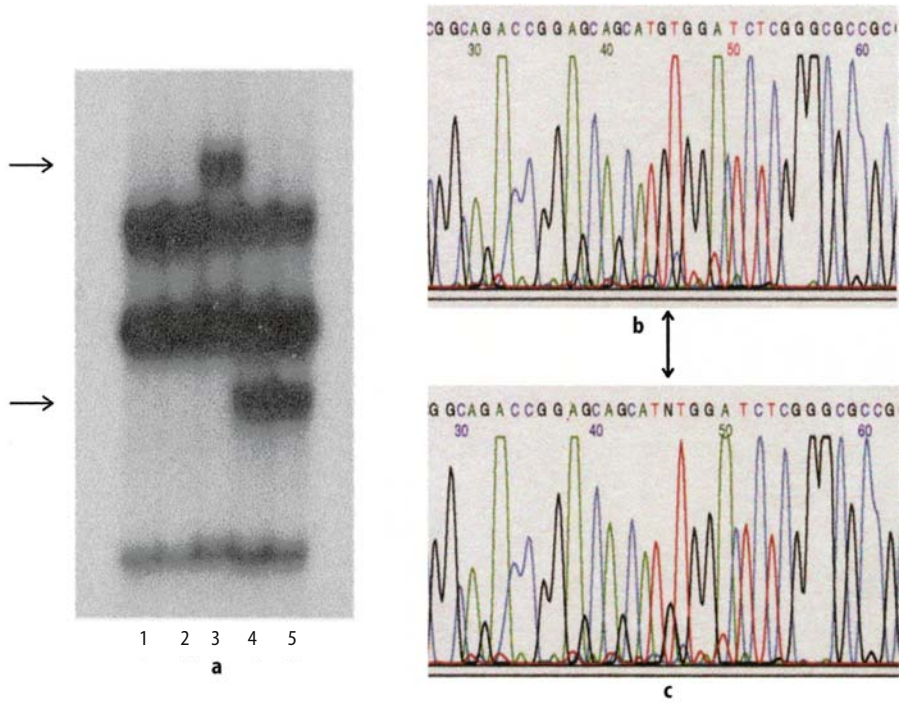


Figure 15-4. (a) SSCP analysis of the *FRDA* gene exon 1 amplicon revealing the presence of abnormal conformers (→) in two FRDA compound heterozygotes. Lanes 1 and 2, exon 1 amplicon from a normal control; lane 3, FRDA compound heterozygote carrying a G→A transition at nucleotide 3 of the *FRDA* gene;[21] lanes 4 and 5 (duplicate), FRDA compound heterozygote carrying an A→C transversion at nucleotide 1 of the *FRDA* gene. (b, c) Representative exon 1 electropherograms from a normal individual (b) and the 3G→A/GAA_{1000} FRDA compound heterozygote (c) shown in Figure 15-4a, lane 3, confirming the presence of the mutation in the heterozygous state (↕).

Laboratory Issues

There are currently no in vitro diagnostic test kits commercially available for the ADCAs or FRDA. The CAP offers proficiency testing twice a year as part of the CAP/ACMG MGL Survey that includes challenges for SCA1, 2, 3, 6, and 7 as well as FRDA.

ALZHEIMER DISEASE

Molecular Basis of Disease

Alzheimer disease is an adult-onset slowly progressive dementia with gradual erosion of intellectual function. Alzheimer disease is the most common form of dementia in the elderly, accounting for approximately 50% of the dementia in patients over age 85 years. Alzheimer disease can be categorized as either sporadic (75% of cases) or familial (25%). The 25% of familial cases are associated with either late- (20%) or early-onset disease (5%) with autosomal dominant/multifactorial or classic autosomal dominant inheritance patterns (Table 15-2). Of the five identified Alzheimer disease loci, four have been characterized at the gene and protein levels (AD1-4) and clinical (DNA) testing is available for three (AD2-4).

Late-onset Alzheimer disease (AD2), representing approximately 20% of the familial cases, has been associated with the presence of the apolipoprotein E (APOE) ε4 allele in many large studies.[22] APOE is a plasma protein involved in the transport of cholesterol and is a component of the very low-density lipoprotein (VLDL). The protein exists in three isoforms (apoE2, E3, and E4) encoded for by three alleles (ε2, ε3, and ε4). In the general population, the ε3 allele is the most common allele, found in approximately 77% of individuals. As stated, while many studies have corroborated the observation that the presence of an ε4 allele in an individual with dementia increases the probability that the patient has Alzheimer disease, this association is not complete. In patients with a clinical diagnosis of Alzheimer disease, the presence of an ε4/ε4 genotype is virtually diagnostic (sensitivity of approximately 97%); however, since approximately 45% of Alzheimer disease patients do not have an ε4 allele, APOE genotyping is not specific for Alzheimer disease.[22,23]

In contrast to late-onset Alzheimer disease, pathologic mutations in at least three genes have been associated with the development of early-onset familial Alzheimer disease (EOFAD). AD1, comprising approximately 10% to 15% of EOFAD, is associated with mutations in the gene encoding the amyloid precursor protein (APP), a 110 to 130 kDa ubiquitously expressed protein. The small proteolytic fragment of APP, $A\beta_{1\text{-}42}$, is found as a major component of amyloid plaques, the neuropathological hallmark of Alzheimer disease. Over a dozen APP missense mutations have been described, with one mutational hotspot (V717) described in multiple kindreds. AD3 (accounting for as much as 70% of EOFAD) and AD4 (accounting for less than 5% of EOFAD) are both associated with mutations in two highly homologous proteins, presenilin 1 and presenilin 2 (PSEN1 and PSEN2). The specific functions of these proteins are not completely understood. Expression of both PSEN1 and PSEN2 mutations in transfected cells has demonstrated an upregulatory effect on $A\beta_{1\text{-}42}$ secretion, suggesting a role in the processing and secretion of APP and its amyloidogenic peptide. To date, more than 40 missense mutations have been identified in AD3 kindreds, whereas only three mutations have been reported in AD4 families.[22,23]

Table 15-2. Molecular Classification of Alzheimer Disease (AD)

Locus	Inheritance	Chromosome	Gene Product	Clinical Test[1]
AD1	AD	21q21.3	APP[2]	Yes
AD2	AD/Multifactorial[3]	19q13	APOE[4]	Yes
AD3	AD	14q24	PSEN1[5]	Yes
AD4	AD	1q31	PSEN2[6]	Yes
AD5	AD/Multifactorial	12q11	Unknown	No

Source: See GeneTests (www.genetests.org).
[1] DNA-based testing.
[2] APP (amyloid precursor protein).
[3] risk-factor/genetic susceptibility.
[4] APOE (apolipoprotein E).
[5] PSEN1 (presenilin 1).
[6] PSEN2 (presenilin 2).
AD, autosomal dominant

Clinical Utility of Testing

Although APOE represents a risk factor for late-onset familial and sporadic AD, the presence of an ε4 allele is neither necessary nor sufficient for the development of Alzheimer disease. As such, the use of *APOE* genotyping in the clinical setting remains controversial. As a diagnostic adjunct in the clinical evaluation of a patient with late-onset dementia, the test may have some demonstrable clinical utility; however, there is general agreement that this test should not be used for predictive purposes and should never be considered for prenatal testing. In contrast, within the context of a high degree of clinical suspicion and the documentation of a family history of Alzheimer disease, DNA diagnostic testing for both AD3 and AD4 has demonstrable utility. Confirmation of a pathologic mutation in a proband will establish a diagnosis and clarify the EOFAD subtype such that both presymptomatic testing for appropriate family members and prenatal testing for at-risk pregnancies could be offered.

Available Assays

APOE genotyping can be performed by traditional polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP) analysis or by the Invader method (Third Wave Technologies, Madison, WI).[24] Detection of pathologic mutations in *APP*, *PSEN1* and *PSEN2* requires direct DNA sequencing.

Interpretation of Test Results

Interpretation of an *APOE* genotype requires the concurrent evaluation of other clinical information available at the time of testing and should never be interpreted in the absence of this information. As such, testing is limited to situations where the index of clinical suspicion for AD is quite high. In contrast, the identification of a pathologic *PSEN1* or *PSEN2* mutation in a symptomatic proband is diagnostic. Testing sensitivity for *PSEN1* is estimated to be between 30% and 60%; therefore, a negative result needs to be interpreted with caution. While both presymptomatic as well as prenatal testing are theoretically possible in mutation-positive kindreds, the possible existence of incomplete or reduced penetrance must be discussed during genetic counseling.[25]

Laboratory Issues

No commercial in vitro diagnostic test kits are currently available for Alzheimer disease testing, although Third Wave Technologies (Madison, WI) does offer *APOE* genotyping reagents. Several *APOE* genotyped cell lines from Alzheimer disease pedigrees can be obtained from the Coriell Cell Repositories (Camden, NJ; http://ccr.coriell.org/). Formal proficiency testing for Alzheimer disease is not currently available.

PARKINSON DISEASE

Molecular Basis of Disease

Parkinson disease (PD) is a chronic, progressive, idiopathic neurodegenerative disorder of late onset. Clinically, the disease is characterized by the development of rigidity, bradykinesia, resting tremor, and postural instability. Neuropathologic changes are notable for the presence of Lewy bodies and the selective degeneration of dopaminergic neurons in the pars compacta and substantia nigra. Although most patients present with a late-onset sporadic form of the disease, the identification of familial forms of PD has clearly established the role of genetic factors in disease etiology and pathogenesis. To date, at least 11 forms of PD can be distinguished on a genetic basis, and genes for seven of these (PARK1, PARK2, PARK5, PARK6, PARK7, PARK8, and PARK11) have been identified (Table 15-3).

PARK1 is associated with early-onset autosomal dominant PD. Despite an earlier age of onset, clinically PARK1 patients have typical dopa-responsive PD as well as classic neuropathologic findings at autopsy. Two missense mutations in the α-synuclein (*SCNA*) gene at 4q21 have been identified to date. The first, A53T, was found in 13 families of Italian-Greek descent[26] and the second, A30P, identified in a single German kindred.[27] Both mutations appear to be highly penetrant. α-synuclein is a 140-amino-acid protein predominantly expressed in neuronal tissues. Although the function of α-synuclein is not well characterized, this protein is one of the major protein constituents of the Lewy body. One attractive hypothesis proposes that the mutant forms of the protein have a propensity to oligomerize and form toxic neuronal aggregates[28] that contribute to the formation of insoluble fibrils through the disruption of the cellular ubiquitin-proteosome pathway and ultimately neuronal cell death via an apoptotic mechanism.

PARK2, or juvenile-onset autosomal recessive PD, is associated with mutations in the *PARKIN* gene that maps to chromosome 6q25. The gene consists of 12 exons and encodes a 465 amino acid protein. Functionally, parkin contains a ubiquitin-like domain in the N-terminus and two RING finger domains in the C-terminus. Similar to other proteins containing RING finger domains, parkin has been shown to have an E3 ubiquitin ligase activity. The association of parkin mutations with an autosomal recessive form of PD suggests that it is the loss of E3 ubiquitin ligase activity that directly contributes to the pattern of neurodegeneration seen in PD. Like PARK1, this mechanism is thought to involve disruption of the ubiquitin-proteosome system, resulting in the abnormal accumulation of substrate proteins. To date, more than 35 pathologic loss-of-function *PARKIN* mutations have been identified.

Table 15-3. Molecular Classification of Parkinson Disease (PD)

Locus	Inheritance	Chromosome	Gene Product	Clinical Test[1]
PARK1	AD	4q21	SNCA[2]	Yes
PARK2	AR	6q25	Parkin	Yes
PARK3	AD	2p13	Unknown	No
PARK4	AD	4p15	Unknown	No
PARK5	AD	4p14	UCHL1[3]	No
PARK6	AR	1p35	Serine-threonine protein kinase	Yes
PARK7	AR	1p36	DJ-1	No
PARK8	AD	12p11	Leucine rich repeat kinase 2	Yes
PARK9 (Kufor-Rakeb Syndrome)	AR	1p36	Unknown	No
PARK10	Idiopathic	1p32	Unknown	No
PARK11 (?)	AD	2q22	NR4A2	No

Source: Adapted from the Human Gene Nomenclature Database (www.gene.ucl.ac.uk/nomenclature/).
[1] DNA-based testing.
[2] SNCA, α-synuclein.
[3] UCHL1, ubiquitin carboxyl-terminal esterase L1 (ubiquitin thiolesterase).
AD, autosomal dominant; AR, autosomal recessive.

Mutations are generally exon rearrangements leading to deletions, duplications, triplications, or point mutations. While current estimates indicate that >90% of all *PARKIN* mutations can be identified by molecular methods,[29,30] testing is presently available only on a research basis.

PARK5, a rare autosomal dominant form of PD, may be associated with mutations in the *UCHL1* gene, although to date only a single partially penetrant missense mutation in one family has been identified, raising the possibility that this sequence alteration might be a rare genetic polymorphism.[31] Ubiquitin C terminal hydrolase L1 (UCHL1) is an enzyme of the ubiquitin C terminal hydrolase family. Abundantly expressed in brain, this enzyme is directly involved in the protein ubiquitination pathway responsible for the normal clearance and recycling of proteins from neurons. The mutant form of UCHL1 has been reported to have diminished enzyme activity, resulting in impaired protein clearance through the ubiquitin-proteosome pathway.[28] PARK6 is associated with mutations in the serine-threonine protein kinase gene (PINK1).[3]

Mutations (deletion and point mutations) in the recently identified *DJ1* gene are associated with PARK7, another autosomal recessive form of PD. Although the specific function of this newly identified gene is unknown, preliminary data suggest involvement in the cellular oxidative stress response.[32] PARK8 is associated with mutations in the leucine rich repeat kinase 2 gene.[3]

The most recently characterized PD gene, *NR4A2*, is associated with yet another autosomal dominant form of the disease (PARK11). *NR4A2*, encoding a member of nuclear receptor superfamily, is involved in the differentiation of nigral dopaminergic neurons, suggesting that mutations within the gene can result in dopaminergic dysfunction. Both a deletion and a missense mutation in exon 1 of *NR4A2* have been identified in multiple symptomatic members of ten kindreds segregating with an autosomal dominant PD phenotype.[33]

Clinical Utility of Testing

Testing for PARK1, 2, 6, and 8 are available clinically. Within the context of a high degree of clinical suspicion and the documentation of a family history of PD, molecular testing has demonstrable utility. Confirmation of a pathologic mutation in a proband will establish a diagnosis and clarify the mode of inheritance such that, theoretically, both presymptomatic testing for appropriate family members and prenatal testing for at-risk pregnancies could be offered.

Available Assays

Testing for patients with suspected *SCNA* gene mutations can be performed by conventional PCR-RFLP analysis of the mutation-containing exons. Testing for *PARKIN* mutations generally necessitates the use of a semiquantitative PCR methodology to detect heterozygous exonic rearrangements (deletions, duplications, and triplications), which would not be detected by conventional nonquantitative PCR or direct sequencing.[30]

Interpretation of Results

The identification of a pathologic mutations in a symptomatic proband is diagnostic.

Laboratory Issues

No commercial in vitro diagnostic test kits nor formal proficiency testing are currently available for PARK1.

AMYOTROPHIC LATERAL SCLEROSIS

Molecular Basis of Disease

Amyotrophic lateral sclerosis (ALS) is an adult-onset neurodegenerative disease characterized by progressive loss of function of both upper and lower motor neurons (UMN, LMN) of the brain and spinal cord. The disease is characterized clinically by the presence of deficits in UMN and LMN function. UMN deficit signs generally include hyperreflexia, extensor plantar response, and increased muscle tone. LMN deficit signs include weakness, muscle wasting and cramping, and fasciculations. While 90% of ALS cases occur sporadically, the remaining 10% of patients have familial forms of the disease, which can be classified based on the mode of inheritance (Table 15-4).

Only ALS1 and ALS2 have been characterized at the molecular and protein levels. The copper-zinc superoxide dysmutase gene (*SOD1*) encodes superoxide dysmutase (SOD1), a metalloenzyme consisting of 153 amino acids with both copper and zinc binding sites. SOD1 catalyzes the conversion of superoxide anions to hydrogen peroxide and

Table 15-4. Molecular Classification of Amyotrophic Lateral Sclerosis (ALS)

Locus	Inheritance	Chromosome	Gene Product	Clinical Test[1]
ALS1	AD	21q22.1	SOD1[2]	Yes
ALS2	AR	2q33	Aslin	Yes
ALS3	AD	18q31	Unknown	No
ALS4	AD	9q34	Unknown	No
ALS5	AR	15q15	Unknown	No
ALS-FTD[3]	AD	9q22	Unknown	No

Source: See GeneTests (www.genetests.org).
[1] DNA-based testing.
[2] SOD1, superoxide dysmutase (Cu-Zn).
[3] ALS-FTD, ALS with frontotemporal dementia.
AD, autosomal dominant. AR, autosomal recessive.

molecular oxygen and, as such, is thought to have a protective function in the cell by preventing oxidative damage caused by the accumulation of free radicals. More than 90 mutations have been reported in the *SOD1* gene, with one, A4V, accounting for approximately 50% of all mutations found in North American families. Approximately 20% of patients with familial ALS and approximately 3% of patients with sporadic ALS have *SOD1* mutations (associated with ALS1).[34]

The gene product for ALS2, or autosomal recessive juvenile-onset ALS, recently has been identified and characterized. The protein, alsin, is predicted to encode a 184 kDa protein containing 1,657 amino acids. Analysis of protein sequence homologies suggests that alsin may belong to a family of GTPase regulator proteins. All ALS2 families studied to date carry deletion mutations that lead to frameshifts, generating premature stop codons.[35,36]

Clinical Utility of Testing

Within the context of a high degree of clinical suspicion and the documentation of a positive family history, DNA diagnostic testing for ALS1 and ALS2 has demonstrable utility. Confirmation of a pathologic mutation in a proband will establish a diagnosis and clarify the mode of inheritance such that both presymptomatic testing for appropriate family members and prenatal testing for at-risk pregnancies could be offered.

Available Assays

Due to allelic heterogeneity for both ALS1 and ALS2, clinical testing for ALS1 requires the use of direct DNA sequencing.

Interpretation of Test Results

The identification of a pathologic *SOD1* mutation in a symptomatic proband is diagnostic. However, as test sensitivity is estimated to be no greater than 50%, negative results must be interpreted with caution. While both presymptomatic as well as prenatal testing are theoretically possible in mutation-positive kindreds, the existence of incomplete or reduced penetrance makes pretest counseling difficult.

Laboratory Issues

No commercial in vitro diagnostic test kits are available for ALS. Several cell lines and at least one purified genomic DNA sample from a family with autosomal dominant ALS can be obtained from the Coriell Cell Repositories (Camden, NJ; http://ccr.coriell.org/); however, the mutation segregating in the family has not been characterized. Formal proficiency testing for ALS1 currently is not available.

DYSTONIA

Molecular Basis of Disease

Although they are not strictly considered a neurodegenerative disease, clinical and molecular descriptions of the dystonias have historically been found in texts describing neurodegenerative disorders. With the recent neuropathological descriptions associated with DYT3 (X-linked dystonia-parkinsonism, or "Lubag" form of the disease) and DYT14,[37,38] the discussion of the dystonias in this chapter becomes even more justified. The dystonias represent a clinically heterogeneous group of disorders characterized by sustained involuntary muscle contractions leading to repetitive twisting movements and abnormal postures. At least 14 forms of dystonia (Table 15-5) can be distinguished on a genetic basis,[37,39] and genes for three of the dystonias (DYT1, DYT5, and DYT11) have been identified.[37,39] DYT1, also known as early-onset primary dystonia, is inherited in an autosomal dominant manner and is usually associated with reduced penetrance. Essentially all patients with DYT1 (>99.9%) have a 3 bp deletion (GAG) in the *DYT1* gene. A second mutation (an 18 bp in-frame deletion in exon 5 of *DYT1*) has been described in one patient with early-onset dystonia and myoclonic features.[40] The *DYT1* gene localizes to chromosome 9q34 and encodes a 332 amino acid protein called torsinA. Although its function remains unknown, torsinA shares sequence homologies with the AAA super-

Table 15-5. Molecular Classification of the Dystonias

Locus	Inheritance	Chromosome	Gene Product	Clinical Test[1]
DYT1	AD	9q34	TorsinA	Yes
DYT2	AR	Unknown	Unknown	No
DYT3	X-linked	Xq13.1	Putative DYT3 protein	Yes
DYT4	AD	Unknown	Unknown	No
DYT5	AD	14q22.1	GCH1[2]	Yes
	AR	11p15	TH[3]	No
DYT6	AD	8p21	Unknown	No
DYT7	AD	18p11.3	Unknown	No
DYT8	AD	2q25	Unknown	No
DYT9	AD	1p21	Unknown	No
DYT10	AD	16p11	Unknown	No
DYT11	AD	7q21	ε-Sarcoglycan	No
	AD	11q23	Dopamine receptor D2	No
DYT12	AD	19q13	Unknown	No
DYT13	AD	1p36.13	Unknown	No
DYT14	AD	14q13	Unknown	No

Sources: Adapted from [31, 33].
[1] DNA-based testing.
[2] GCH1 (GTP cyclohydrolase I).
[3] TH (tyrosine hydroxylase).
AD, autosomal dominant; AR, autosomal recessive.

family of ATPases[37] and is thought to have a role in several cellular functions including protein folding and degradation, membrane trafficking, vesicle fusion, cytoskeletal dynamics, and physiological responses to stress.[37]

DYT5, or dopamine-responsive dystonia (DRD), is associated with mutations in two different genes. Homozygous as well as compound heterozygous mutations in the tyrosine hydroxylase (*TH*) gene at 11p15.5 are associated with the rare autosomal recessive form of DRD. The more common autosomal dominant form of DRD is associated with mutations in the GTP cyclohydrolase I (*GCH1*) gene, which maps to 14q22.1. Both enzymes are involved in the bio-synthetic pathway for the neurotransmitter dopamine. Recently a third DRD locus (DYT14) has been mapped to chromosome 14q13 but the specific gene and its protein product are not known.[37]

DYT11, or myoclonus-dystonia, is associated with loss-of-function mutations in the ε-sarcoglycan gene at 7q21 in a large number of families and represents a major locus for myoclonus-dystonia. However, the concurrent finding of a single myoclonus-dystonia family segregating a missense mutation in the D2 dopamine receptor (*DRD2*) gene that maps to 11q23, and the identification of linkage of another large myoclonus-dystonia family to a locus on chromosome 18p11, suggests some degree of genetic heterogeneity for this form of dystonia.[37]

Clinical Utility of Testing

Clinical molecular testing is presently available for only DYT1, DYT3, and the *GCH1*-linked form of DYT5.

Available Assays

Molecular testing for DYT1 is routinely performed by PCR-RFLP analysis of the GAG-containing *DYT1* exon (Figure 15-5). Determination of pathologic mutations in *GCH1* requires direct DNA sequencing of the gene.

Interpretation of Test Results

Essentially all patients with typical early-onset dystonia (DYT1) carry the GAG deletion on one allele, thus the test sensitivity approaches 100%. In contrast, the specificity of the assay is 60% to 70%, because DYT1 is inherited as an autosomal dominant trait with reduced penetrance (30% to 40%). Counseling of the parents of the proband should include an evaluation of their clinical status and consideration of molecular testing to determine parental origin of the mutation. The risk of inheriting a *DYT1* allele from a proband is 50%; however, the probability that the gene carrier will become symptomatic is estimated to be 30% to 40%. Prenatal testing is clinically available for fetuses at 50% risk of inheriting a *DYT1* allele, once the presence of the mutation has been confirmed in the family. New mutations, although rare, have been reported, but the mutation rate is unknown.

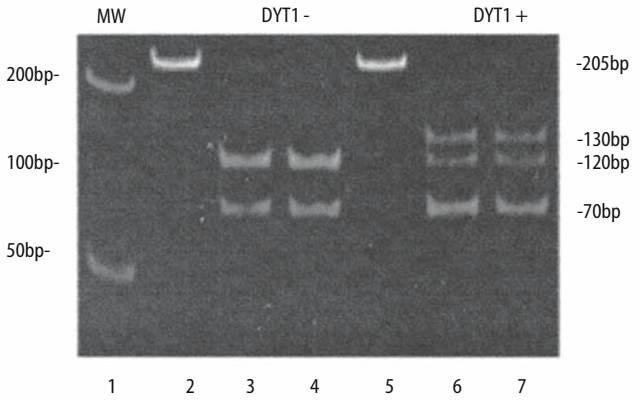


Figure 15-5. PCR-RFLP analysis of the *DYT1* GAG deletion mutation associated with autosomal dominant early-onset torsion dystonia (DYT1). DNA was amplified with the 6419/H48 primer pair[21] and the resulting 205 bp amplicon digested with *Bse* RI. The presence of the GAG deletion mutation results in the loss of a *Bse* RI site on the mutant allele and the generation of a novel 130 bp fragment in addition to the 120 bp and 70 bp fragments generated from the normal allele. Lane 1, 100 bp ladder G2101 (Promega, Madison, WI); lane 2, undigested amplicon; lanes 3 and 4, *Bse* RI digests from two *DYT1* negative patients; lane 5, undigested amplicon; lanes 6 and 7, *Bse* RI digests from two *DYT1* positive patients (note the presence of the novel 130 bp fragment).

Unlike in DYT1, more than 85 mutations have been identified in the *GCH1* form of DYT5 and, on average, approximately 30% to 40% of clinically symptomatic patients will not carry a pathologic mutation in the coding region or intron-exon boundaries of the gene. As such, clinical testing (direct sequencing) is usually limited to diagnostic testing in a proband with the knowledge that testing sensitivity may reach only 60%. If a mutation is detected, carrier testing or prenatal testing or both can be considered for family members. Counseling for the offspring and siblings of a molecularly-confirmed proband is also complicated by the influence of gender-related reduced penetrance and the existence of new mutations.[42]

Laboratory Issues

Currently no commercial in vitro diagnostic test kits are available for DYT1, DYT3, or DYT5. DYT1 DNA can be obtained from the Coriell Cell Repositories (Camden, NJ; http://ccr.coriell.org/). Currently, no formal proficiency testing is available.

References

1. Hug M, Hayden M. Huntington disease. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 7, 2003.
2. Potter NT, Spector EB, Prior TW. Technical standards and guidelines for Huntington disease. In: The American College of Medical Genetics Standards and Guidelines for Clinical Genetics Laboratories, Bethesda MD. Available at: http://www.acmg.net. Accessed February 17, 2003.

3. GeneTests [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 7, 2003.
4. Andrew SE, Goldberg YP, Theilman J, et al. A CCG repeat polymorphism adjacent to the CAG repeat in the Huntington disease gene: implications for diagnostic accuracy and predictive testing. *Hum Mol Genet.* 1994;3:65–67.
5. Margolis RL, Stine OC, Callahan C, et al. Two novel single-base-pair-substitutions adjacent to the CAG repeat in the Huntington disease gene (IT15): implications for diagnostic testing. *Am J Hum Genet.* 1999;64:323–326.
6. Maher TA, Ito M, Loose BA, et al. Implementation of long PCR for the detection of very large trinucleotide repeat expansion in a 3-year-old with juvenile Huntington disease. *Am J Hum Genet.* 2002;71:552.
7. Bird TD. Hereditary ataxia overview. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 7, 2003.
8. Potter NT, Nance MA, for the Ataxia Molecular Diagnostic Testing Group. Genetic testing for ataxia in North America. *Mol Diagn.* 2000;5:91–99.
9. Holmes SE, O'Hearn EE, McInnis MG, et al. Expansion of a novel CAG trinucleotide repeat in the 5′ region of *PPP2R2B* is associated with SCA12. *Nat Genet.* 1999;4:391–392.
10. Nakamura K, Jeong SY, Uchihara T, et al. SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum Mol Genet.* 2001;10:1441–1448.
11. Koob MD, Moseley ML, Schut LJ, et al. An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat Genet.* 1999;21:379–384.
12. Matsuura T, Yamagata T, Burgess DL, et al. Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. *Nat Genet.* 2000;26:191–194.
13. Bidichandani SI, Ashizawa T. Friedreich ataxia. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 7, 2003.
14. Dorschner MO, Barden D, Stephens K. Diagnosis of five spinocerebellar ataxia disorders by multiplex amplification and capillary electrophoresis. *J Mol Diagn.* 2002;4:108–113.
15. Mao R, Aylsworth AS, Potter NT, et al. Childhood-onset ataxia: testing for large CAG-repeats in SCA2 and SCA7. *Am J Med Genet.* 2002;110:338–345.
16. Imbert G, Saudou F, Yvert G, et al. Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat Genet.* 1996;14:285–291.
17. Snow K, Mao R. Extreme expansion detection in spinocerebellar ataxia type 2 and type 7. In: Potter NT, ed. *Neurogenetics: Methods and Protocols.* Vol. 217. Totowa, NJ: Humana Press; 2002:41–50.
18. Dalton JC, Day JW, Ranum LPW, et al. Spinocerebellar ataxia type 8. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 7, 2003.
19. Potter NT. Update on the genetics of ataxia. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed February 17, 2003.
20. Filla A, De Michelle G, Cavalcanti F, et al. The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. *Am J Hum Genet.* 1996;59:554–560.
21. Potter NT, Miller CA, Anderson IJ. Mutation detection in an equivocal case of Friedreich ataxia. *Pediatr Neurol.* 2000;22:413–415.
22. Bird TD. Alzheimer overview. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 7, 2003.
23. McKusick VA. Alzheimer disease (OMIM #104300). Online Mendelian Inheritance in Man [database online]. Baltimore: Johns Hopkins University. Available at: http://www.ncbi.nlm.nih.gov/omim/. Accessed January 7, 2003.
24. Third Wave Technologies. Apolipoprotein E (ApoE). L-0008. Madison, WI; 2002. Available at: http://www.twt.com. Accessed February 17, 2003.
25. Janssen JC, Beck JA, Campbell TA, et al. Early onset familial Alzheimer disease: mutation frequency in 31 families. *Neurology.* 2003;60:235–239.
26. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the α-synuclein gene identified in families with Parkinson's disease. *Science.* 1997;276:2045–2047.
27. Krüger R, Kuhn W, Müller T, et al. Ala30Pro mutation in the gene encoding α-synuclein in Parkinson's disease. *Nat Genet.* 1998;18:106–108.
28. Mouradian MM. Recent advances in the genetics and pathogenesis of Parkinson disease. *Neurology.* 2002;58:179–185.
29. Brice A, Dürr A, Lücking C, et al. Parkin type of juvenile Parkinson disease. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 7, 2003.
30. Lücking CB, Brice A. Semiquantitative PCR for the detection of exon rearrangements in the parkin gene. In: Potter NT, ed. *Neurogenetics: Methods and Protocols.* Vol. 217. Totowa, NJ: Humana Press; 2002:13–26.
31. Leroy E, Boyer R, Auburger G, et al. The ubiquitin pathway in Parkinson's disease. *Nature.* 1998;395:451–452.
32. Bonifati V, Rizzu P, van Baren MJ, et al. Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *Science.* 2003;299:256–259.
33. Le W-D, Xu P, Jankovic J, et al. Mutations in NR4A2 associated with familial Parkinson disease. *Nat Genet.* 2003;33:85–87.
34. Gaudette M, Saddique T. Amyotrophic lateral sclerosis overview. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 7, 2003.
35. Yang Y, Hentati A, Deng H-X, et al. The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis. *Nat Genet.* 2001;29:160–165.
36. Hadano S, Hand CK, Osuga H, et al. A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nat Genet.* 2001;29:166–173.
37. deCarvalho Aguiar PM, Ozelius LJ. Classification and genetics of dystonia. *Lancet Neurol.* 2002;1:316–325.
38. Grotzsch H, Pizzolato G-P, Ghika J, et al. Neuropathology of a case of dopa-responsive dystonia associated with a new genetic locus, DYT14. *Neurology.* 2002;58:1839–1842.
39. Klein C, Ozelius LJ. Dystonia: clinical features, genetics, and treatment. *Curr Opin Neurol.* 2002;15:491–497.
40. Leung JC, Klein C, Friedman J, et al. Novel mutation in the Tor1A (DYT1) gene in atypical early onset dystonia and polymorphisms in dystonia and early onset parkinsonism. *Neurogenetics.* 2001;3:133–143.
41. Ozelius LJ, Hewett JW, Page CE, et al. The early-onset torsion dystonia gene (DYT1) encodes an ATP-binding protein. *Nat Genet.* 1997;17:40–48.
42. Furukawa Y. Dopa-responsive dystonia. GeneReviews [database online]. Seattle: University of Washington. Available at: http://www.genetests.org. Accessed January 21, 2003.

Chapter 16

Skin and Connective Tissue Disorders

James C. Hyland

OSTEOGENESIS IMPERFECTA

Molecular Basis of Disease

Osteogenesis imperfecta (OI) is a clinically heterogeneous disorder resulting from the reduced synthesis or the accumulation of abnormal type I collagen, which is a major structural component of many connective tissues including bone. Type I collagen is a trimeric molecule composed of two α1(I) chains and one α2(I) chain, encoded by the *COL1A1* and *COL1A2* genes, respectively. Fibrillar collagens have a long triple helical domain composed of Gly-X-Y repeats. Glycine residues at every third position are required for the formation of the triple helical domain.[1] OI is predominantly inherited in an autosomal dominant fashion and is caused by a wide variety of mutational events in either the *COL1A1* or *COL1A2* gene.

Phenotypic characteristics and clinical manifestations of OI may include fractures, blue or dark sclera, bowing of the femurs, short stature, dentinogenesis imperfecta, hearing loss, and additional assorted skeletal abnormalities.[2] Four subtypes are recognized based on clinical findings (Table 16-1).

Type I is the mild variant of OI. Type I patients have blue sclera and increased bone fragility, but usually obtain a near normal stature. Fractures involve the long bones, ribs, and bones of the hands, and may be few or numerous. The fractures heal without deformity and the fracture rate decreases following puberty.

Type II is the perinatal lethal variant of OI with up to 60% dying on the day of birth. Many infants are premature with low birth weight. Infants have beaded ribs, a soft calvarium, dark sclera, and short extremities. The legs are bowed and hips flexed. Fractures are rare since the bones are soft.

Type III is known as the progressive deforming variant of OI. Fractures often occur in utero, and bones are undermineralized at birth. Distal metaphyseal cystic structures disrupt the growth plates. Short stature is inevitable and many patients develop severe kyphoscoliosis, leading to cardiopulmonary in-sufficiency that contributes to premature death. Dentinogenesis imperfecta also is characteristic.

Type IV is the mildly deforming variant of OI. Fractures may occur during birth or in utero. The patients are usually short, often have dentinogenesis imperfecta, and have mild to moderate skeletal deformities. The fracture rate may decrease at puberty. Scoliosis or kyphoscoliosis may compromise cardiopulmonary function.[3]

OI Type I Phenotype Mutations

Mutational events resulting in *COL1A1* null alleles are a major cause of the type I variant. Null alleles may result from point mutations directly producing stop codons, or insertions, deletions, and splice-site alterations causing frame shifts with the generation of downstream stop codons. Splice-site mutations also may lead to the direct inclusion of intronic sequences containing a stop codon. The unifying theme is that these mutations lead to nonsense-mediated messenger RNA (mRNA) decay.[3,4] Other *COL1A1* gene mutations in type I OI result in specific glycine substitutions and interference with procollagen chain association. *COL1A2* gene mutations infrequently cause OI type I. Defined mutations include those altering consensus splice sites, resulting in exon skipping, or mutations causing specific glycine substitutions.[3]

OI Type II Phenotype Mutations

Numerous types of mutations are associated with type II perinatal lethal OI. Glycine substitutions in the triple helical domain of the proα(1) and proα(2) chains are the most common.[3] Glycine-to-stop codon mutations also occur. Rare, multiexon rearrangements involving regions of the *COL1A1* gene that code for the triple helical domain are lethal. These rearrangements include a deletion and a

Table 16-1. Osteogenesis Imperfecta

Type	Genes	Observed Mutations	Available Tests
I Mild	*COL1A1*	Mutations resulting in null alleles, splice-site mutations, and glycine substitutions	*COL1A1* null allele analysis; heteroduplex analysis with DNA sequencing
	COL1A2 (rare)	Splice-site mutations and glycine substitutions	Heteroduplex analysis with DNA sequencing
II Perinatal Lethal	*COL1A1* *COL1A2*	Glycine and other substitutions, splice-site mutations with exon skipping, small in-frame insertions or deletions, and frame shifts	*COL1A1* and *COL1A2* heteroduplex analysis with DNA sequencing
III Progressive Deforming	*COL1A1* *COL1A2*	Glycine substitutions, splice-site mutations, and rare out-of-frame deletions	*COL1A1* and *COL1A2* heteroduplex analysis with DNA sequencing
IV Mild Deforming	*COL1A1* *COL1A2*	Glycine substitutions, splice-site mutations resulting in insertions and deletions, and exon skipping	*COL1A1* and *COL1A2* heteroduplex analysis with DNA sequencing

recombination resulting in duplication.[3] A large multiexon deletion in the *COL1A2* gene has been defined. Splice-site mutations in both the *COL1A1* and *COL1A2* genes lead to the skipping of exons coding for triple helical domains.[3] Deletions of single Gly-X-Y triplets have also been observed.[3] Mutations in the carboxy-terminal propeptide of proα1(I) interfere with proα1(I) and proα2(I) chain assembly. These include substitutions, frameshifts, and small in-frame deletions.[3]

The pathogenesis of type II OI involves the synthesis of aberrant proα1(I) or proα2(I) chains. These chains either are incorporated into trimeric molecules, ultimately resulting in reduced thermal stability, or, alternatively, interfere with trimer assembly. In most instances autosomal dominant inheritance is observed. Some cases of mosaicism have been defined. One compound heterozygote has been observed. One perinatal lethal case proved to be a homozygote.[3]

OI Type III Phenotype Mutations

Type III OI is associated with dominantly inherited mutations in the *COL1A1* gene, including those causing glycine substitutions and splice-site mutations,[3] as well as mutations in the *COL1A2* gene causing glycine substitutions, single glycine deletions, and splice-site mutations.[3] Rare cases of OI type III are associated with an autosomal recessive pattern of inheritance in consanguineous families.[3] In one family, the proband inherited two copies of the same 4 base pair deletion in the *COL1A2* gene, which results in a frame shift that ultimately inhibits the incorporation of the proα2(I) chains into procollagen.[3] In a second family, two siblings were homozygous for a specific glycine-to-serine substitution in the triple helix; the heterozygous parents displayed a mild OI phenotype including generalized osteopenia.[3]

OI Type IV Phenotype Mutations

Type IV OI inheritance is autosomal dominant and associated with specific point mutations in both the *COL1A1* and *COL1A2* genes, resulting in glycine substitutions. Mutations leading to the skipping of specific *COL1A1* and *COL1A2* exons have been defined. Certain in-frame *COL1A2* gene insertions and deletions also occur.[3]

Clinical Utility of Testing

Molecular testing has proved very useful for prenatal diagnosis of OI. Analysis can be completed within 2 to 3 weeks of receipt of cultured amniocytes or chorionic villus sampling (CVS). If a submitted sample is derived from a family with a predetermined mutation, the turnaround time is compressed to a few days. Molecular testing is indicated in cases where protein-based testing is equivocal. Certain mutations may not affect the quantity or quality of collagen synthesized by fibroblasts or CVS specimens, or may do so at levels not detectable by relatively insensitive protein gel electrophoresis methods. Molecular testing also does not require establishing fibroblast cultures, significantly reducing turnaround times. Molecular testing is used to predict the risk of OI recurrence with subsequent pregnancies. The possibility of very rare germline mosaicism must be kept in mind when families are counseled. Testing may be indicated for the evaluation of relatives at risk for disease. Increasingly, testing is used to predict clinical severity.

Genomic DNA testing for OI has proved beneficial in cases where questions of abuse arise. Mutations have been defined in children who have been removed from parents due to suspected abuse. Alternatively, failure to define a mutation is useful for child welfare and potential prosecution. Results must be interpreted with caution and with attention to assay sensitivity.

Available Assays

Commercial molecular testing for OI is available in the United States and Europe. The two US laboratories offering molecular testing essentially utilize identical approaches for genomic analysis. Oligonucleotide primer sets are used to generate a series of polymerase chain reaction (PCR) products corresponding to all exons and exon-intron boundaries of both the *COL1A1* and *COL1A2* genes. The PCR products (approximately 100 amplicons) are then analyzed for the presence of heteroduplexes by conforma-

tion sensitive gel electrophoresis (CSGE) in a screening test. PCR products displaying heteroduplexes are sequenced.[4] The major advantages of this testing approach are that results are generated in a timely fashion (procollagen analysis is not required), and many specimen types are adequate as a source of DNA for testing. The major disadvantage is that this approach will not directly detect large deletions or the complete deletion of an allele. Careful analysis of the sequencing results will indicate a lack of polymorphic markers, which should be an indication that further analysis is required.

In some instances, genomic analysis may be preceded by procollagen synthetic studies. These protein studies may be used to establish a diagnosis or, alternatively, may be utilized to determine the methodology followed for mutation detection.

A European laboratory offering testing performs *COL1A1* null allele analysis on fibroblasts displaying either decreased type I procollagen synthesis or a normal type I procollagen synthetic profile on sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE). Null allele analysis consists of determining the heterozygosity of the two *COL1A1* alleles. If heterozygosity is evident, expression of both alleles is assessed by sequence analysis of complementary DNA (cDNA) reverse transcribed from *COL1A1* mRNA. If a *COL1A1* null allele is suggested, genomic screening is performed. Cases lacking a *COL1A1* null allele and all cases displaying structurally abnormal type I procollagen on SDS-PAGE are subjected to *COL1A1* and *COL1A2* cDNA mutation screening either by single-strand conformation polymorphism (SSCP) gels or by CSGE.[4] Suspicious products are then sequenced. While this is a comprehensive approach, a biopsy is required, as well as procollagen studies using fibroblast cultures. Consequently, the turnaround time is in the range of 6 to 12 months.

Interpretation of Test Results

Sequencing results are compared to previously defined mutations and polymorphisms. New findings require clinical correlation and possibly analysis of family members of the proband. The CSGE screening method is quite sensitive for most types of mutations. However, it is insensitive to large deletions since PCR products may not be generated, and it occasionally does not detect certain other mutations. In cases with a strong clinical history and negative screening by CSGE, complete genomic sequencing may be indicated.

Laboratory Issues

Individual laboratories are responsible for quality control since the number of laboratories testing for OI is limited. The US laboratories performing OI testing are accredited by the College of American Pathologists (CAP). Internal proficiency testing programs have been approved by the CAP.

EPIDERMOLYSIS BULLOSA

Molecular Basis of Disease

Epidermolysis bullosa (EB) is a clinically and genetically heterogeneous group of mechanobullous disorders. Defects in at least ten genes are known to cause the clinical syndrome (Table 16-2).[5,6] EB is characterized by blistering of the skin and certain other tissues following mild trauma. Severity ranges from mild blistering to severe bulla formation with resultant erosions, scarring, and ultimately mutilation. Patients may exhibit nail, hair, and dental abnormalities, as well as corneal erosions, pyloric atresia, esophageal strictures, and neuromuscular features.[5–7]

EB classification is based on the precise level of blister formation in the dermal-epidermal junction and has been divided into three general categories: EB simplex, junctional EB, and dystrophic EB. Molecular characterization has more recently led to a proposed fourth group, the hemidesmosomal variants.[6]

EB Simplex

In EB simplex (EBS), blistering is intraepidermal and occurs as a consequence of the degeneration of the basal epidermal cells. Blistering decreases with age. Blisters heal without scarring and the prognosis is favorable. Several clinical subtypes are currently recognized. Several subtypes are caused by mutations in the *KRT5* and *KRT14* genes coding for keratins 5 and 14, respectively.[5,6] Inheritance is usually autosomal dominant, although some keratin 14 null allele mutations display recessive inheritance.[5] Mutations in different portions of the genes are correlated with different subtypes of EBS.[5]

EBS Associated with Muscular Dystrophy

EBS associated with muscular dystrophy is inherited in an autosomal recessive manner.[5,6] Cases associated with homozygous mutations are often consanguineous. Others represent compound heterozygotes.[6] In EBS with muscular dystrophy, mutations occur in the *PLEC1* gene encoding plectin, a cytoskeleton-associated protein with intermediate filament, β4 integrin, and actin-binding domains.[5,6] In skin, plectin mediates the attachment of keratins to the hemidesmosome. In muscle, plectin also functions to bridge intermediate filaments and actin filaments. Plectin deficiencies in the skin lead to defective anchorage of the cytoskeleton to the plasma membrane. In muscle, disorganized structural filaments result in diminished mechanical resistance of muscle cells.[5] In most instances, small

Table 16-2. Epidermolysis Bullosa

Subtype	Affected Genes	Mutations	Affected Molecules	Available Tests
EB Simplex				
EBS generalisata	*KRT5/KRT14*	M/rare SS	Keratin5/14	Direct sequencing of PCR products
EBS localisata	*KRT5/KRT14*		Keratin5/14	
EBS herpetiformis	*KRT5/KRT14*		Keratin5/14	
EBS with "mottled pigmentation"	*KRT5/KRT14*		Keratin5/14	
EBS with muscular dystrophy	*PLEC1*	N/I/D	Plectin	Protein truncation and heteroduplex analysis by DHPLC with sequencing
EBS Ogna	*PLEC1*	N/I/D	Plectin	
Junctional EB				
JEB Herlitz	*LAMA3/LAMB3/LAMC2*	N	Laminin 5	Heteroduplex analysis of PCR products and DNA sequencing
LEB localisata	*COL17A1*	M	Collagen XVII	
JEB with pyloric atresia	*ITGA6*	N	Integrin α6 chain	
	ITGB4	N/M	Integrin β4 chain	
Generalized atrophic benign EB	*COL17A1*	N	Collagen XVII	
	LAMB3	M	Lamin 5	
Dystrophic EB				
DEB generalisata	*COL7A1*	D/F/I/M/N/ SS	Type VII collagen	Heteroduplex analysis of PCR products and DNA sequencing
DEB localisata	*COL7A1*		Type VII collagen	
DEB generalisata mutilans	*COL7A1*		Type VII collagen	
DEB generalisata nonmutilans	*COL7A1*		Type VII collagen	
DEB localisata	*COL7A1*		Type VII collagen	
DEB inversa	*COL7A1*		Type VII collagen	
Transient bullous dermolysis of the newborn	*COL7A1*		Type VII collagen	

D, deletion; F, frameshift; I, insertion; M, missense; N, nonsense; SS, splice site.

insertions, deletions, or nonsense mutations result in premature chain termination codons and protein truncation.[6]

Junctional EB

Junctional EB (JEB) is composed of a series of clinically distinct subtypes. In JEB, defective structural molecules of the hemidesmosomal anchoring filament complex lead to detachment of basal keratinocytes from the lamina densa. The result is epidermal dysadhesion and blister formation within the lamina lucida.[6] Extreme clinical variability exists and ranges from a lethal subtype, JEB Herlitz, to a milder variant, JEB progressiva. Mutations in six genes have been identified, and all are inherited in an autosomal recessive manner. Many patients are compound heterozygotes. The involved genes are *LAMA3*, *LAMB3,* and *LAMC2* encoding the three chains of laminin 5, *COL17A1* encoding collagen XVII, and *ITG6A* and *ITGB4* encoding the integrin α6 chain and the integrin β4 chain, respectively.[5,6] With the exception of *LAMB3*, most mutations are private and distributed throughout the entire length of the genes.

Laminin 5 Defects

The heterotrimer laminin 5 is composed of the α3, β3, and γ2 chains, encoded by the genes *LAMA3*, *LAMB3*, and *LAMC2,* respectively. Laminin 5 links the hemidesmosomal protein α6β4 integrin and collagen VII, an anchoring fibril component. Laminin 5 also is cross-linked to other basement membrane networks. Homozygous or compound heterozygous mutations in any of the individual genes encoding laminin 5 chains can lead to nonsense-mediated mRNA decay that is manifested as null alleles. Absence of any single chain results in the failure of assembly and secretion of laminin 5. Patients with JEB Herlitz, the most severe form of JEB, lack laminin 5.[5,8] However, some patients lacking laminin 5 present with mild to moderate JEB.[5] *LAMB3* gene defects account for 80% of all laminin gene mutations in JEB. Two common *LAMB3* nonsense mutations, R42X and R635X, account for 50% of laminin 5 mutations. Missense mutations in one or both alleles of the individual laminin genes are associated with mild JEB phenotypes.[5–7]

α6β4 Integrin Gene Defects

JEB with pyloric atresia is associated with a marked reduction or absence of α6β4 integrin, a molecule normally expressed in skin and the gastrointestinal tract. This integrin binds to other hemidesmosomal components including collagen type XVII and laminin 5. Most lethal cases are due to nonsense mutations in the *ITGA6* or *ITGB4* genes. Missense mutations in the *ITGB4* gene are associated with mild to lethal phenotypes.[5,6]

COL17A1 Gene Defects

Defects in the *COL17A1* gene underlie generalized atrophic EB. This mild variant of JEB is characterized by lifelong skin blistering, nail dystrophy, enamel hypoplasia, and alopecia localized to the eyelids. Collagen XVII is a transmembrane protein playing a role in anchoring the basal keratinocytes to the underlying basement membrane. Most *COL17A1* mutations resulting in generalized atrophic EB result in nonsense-mediated mRNA decay. Some missense mutations have been identified. Most patients are compound heterozygotes. Some have both nonsense and missense mutations.[5,6]

Dystrophic (Severely Blistering) EB

In dystrophic EB (DEB), dermal blistering occurs beneath the basement membrane with the potential for scarring. DEB is subdivided into at least seven discrete subtypes (Table 16-2). Extracutaneous manifestations may include the absence or dystrophy of the nails, contractures, oral, corneal, and conjunctival blistering, mucosal involvement of the pharynx, larynx, and esophagus, dystrophy of the teeth, and mutilation. Inheritance may be autosomal recessive or autosomal dominant and is subtype dependent.[5,6] *COL7A1* gene defects underlie all variants. Collagen VII, a homotrimer, is the major structural component of anchoring fibrils. Collagen VII has a triple helical domain interrupted by a "hinge" region, and a large noncollagenous domain composed of submodules with homology to specific known proteins. The anchoring fibrils form networks with other collagen types and provide stability to the lower portion of the dermal basement membrane.[5]

DEB Autosomal Dominant Mutations

Missense mutations resulting in triple helical domain glycine substitutions or leading to substitutions in the conserved submodules are often manifested in a dominant negative manner. Deletions of portions of the triple helical domain also are dominantly inherited. Phenotypes are variable.[5,6]

DEB Recessive Mutations

Recessively inherited forms of DEB are due to premature translation termination codons and result from point mutations or from frameshifts due to insertions or deletions. Homozygous or compound heterozygous nonsense mutations result in a severe phenotype.[5,6]

Clinical Utility of Testing

Results of testing are useful for genetic counseling and to estimate risk of recurrence in subsequent pregnancies. For example, some cases of DEB previously considered sporadic are due to recessive mutations, and the tests demonstrate either homozygous or compound heterozygous mutations.

Prenatal testing is available for all forms of EB and may be performed as early as 10 weeks for CVS and 12 weeks for amniotic fluid. In isolated instances, preimplantation genetic testing has been accomplished for previously defined parental mutations.[8] Test results have been utilized to predict clinical severity, phenotype, and prognosis, and have contributed toward developing an evolving molecularly based classification system for EB.[7]

Available Assays

Molecular testing is available for all of the mentioned EB subtypes in a single laboratory.

EBS: *KRT5* and *KRT14* Genes

PCR products derived from the *KRT5* and *KRT14* genes are directly sequenced (Ellen Pfendner, personal communication).

EBS Associated with Muscular Dystrophy: *PLEC1* Gene

The *PLEC1* gene is assessed for mutations by a protein truncation assay. Genomic DNA is used to generate PCR products from two large exons constituting about 75% of the coding sequence of *PLEC1*. The sequences are amplified using a forward primer containing a phage promoter, the codon for the first methionine, and sequences homologous to *PLEC1*. Reverse primers contain an in-frame stop codon. The PCR products are utilized in a transcription-translation system with radiolabeled methionine. Radiolabeled peptides are analyzed on polyacrylamide gels. Truncated products reflect nonsense mutations and are verified by direct sequencing of the corresponding PCR product. Most *PLEC1* mutations cause stop codons in exon 32.[6] If a mutation is not defined, heteroduplex analysis is utilized to scan PCR products generated from the remaining exons[6] (Ellen Pfendner, personal communication).

JEB: *LAMA3*, *LAMB3* and *LAMC2* Genes; JEB with Pyloric Atresia: *ITGA6* and *ITGB4* Genes; Generalized Atrophic EB: *COL17A1* Gene; and DEB: *COL7A1* Gene

The *LAMA3*, *LAMB3, LAMC2, ITGA6, ITGB4, COL17A1,* and *COL7A1* genes are analyzed in a similar manner. PCR products generated from exons are analyzed by denaturing high-performance liquid chromatography (DHPLC). Those exons displaying anomalous elution

profiles are directly sequenced[6] (Ellen Pfendner, personal communication).

Interpretation of Test Results

DHPLC is a sensitive screening technique for heteroduplex analysis. The protein-truncation test method utilized for a portion of the *PLEC1* gene analysis is less sensitive compared to DHPLC but has proven useful in defining mutations. The PCR-based tests have a high specificity since oligonucleotide primers used for amplification generate unique products with appropriate sequences.

Laboratory Issues

No in vitro diagnostic test kits are commercially available. All described testing is performed in a large university-based laboratory that is accredited by the CAP. An internal proficiency testing program has been approved by the CAP.

EHLERS-DANLOS SYNDROME

The Ehlers-Danlos syndrome (EDS) is a connective tissue disorder that is divided into numerous subtypes with distinct genetic and clinical findings (Table 16-3). In general, EDS is characterized by joint hypermobility, skin hyperextensibility and tissue fragility. The inheritance pattern is type specific. In most instances, mutations are in collagen genes or in genes coding for molecules involved with the processing or posttranslational modification of collagens.[9]

EDS TYPE I AND EDS TYPE II (CLASSIC TYPES)

Molecular Basis of Disease

The classic types of EDS are considered EDS type I and EDS type II. These types constitute about 90% of the cases and differ only in the degree of organ system involvement.[9] EDS type I is characterized by marked skin involvement with generalized joint hypermobility and musculoskeletal deformities. Hypotonia, delayed motor development, easy bruisability, and prematurity may occur. Tissue fragility and extensibility may be manifested by hiatal hernia, bowel rupture, aortic aneurysm rupture, or postoperative hernia.[9] EDS type II shares manifestations with EDS type I, but the phenotypic presentation is less severe and patients may remain undiagnosed or be diagnosed later in life. The skin is less involved and joint laxity may be confined to the hands and feet.[9]

EDS types I and II are associated with defects in several genes. The majority of cases are due to mutations in either the *COL5A1* or *COL5A2* gene.[9] Most mutations cause dominant negative effects. In some instances, classic EDS appears to be due to haploinsufficiency.[9,10] In one family, a more severely affected member was determined to be compound heterozygous for *COL5A1* gene mutations.[9] The *TNXA* gene also has been implicated in classic EDS with an autosomal recessive inheritance pattern.[9,11] This gene encodes tenascin-X, a collagen-fibril-associated protein of undetermined function. In a few isolated cases, mutations in both the *COL1A1* and *COL1A2* genes have been associated with the classic EDS phenotype.[9]

Clinical Utility of Testing

Molecular *COL5A1* and *COL5A2* gene testing is used to confirm the diagnosis of classic EDS and characterize mutations. Testing is indicated for first-degree relatives of

Table 16-3. The Ehlers-Danlos Syndrome

Type	Genes	Affected Molecules	Available Tests
EDS I and II: Classic types	*COL5A1*, *COL5A2*	Type V collagen	Null allele screening; heteroduplex analysis with DNA sequencing
	COL1A1, *COL1A2*	Type I collagen	Not offered
	TNXA	Tenascin X	Not offered
EDS III: Hypermobile type	? *COL5A3*, ?*COL3A1*	? Type V collagen	Not offered
		? Type III collagen	Not offered
EDS IV: Vascular type	*COL3A1*	Type III collagen	Null allele screening; heteroduplex analysis with DNA sequencing
EDS VIA: Kyphoscoliotic type	*PLOD1*	Lysyl hydroxylase	Sequencing of cDNA
EDS VIB: Kyphoscoliotic type	Unknown	Unknown	Not offered
EDS VIIA: Arthrochalasic type	*COL1A1*	Type I collagen	Procollagen studies
EDS VIIB: Arthrochalasic type	*COL1A2*	Type I collagen	Procollagen studies
EDS VIIC: Dermatosparactic type	Procollagen N-proteinase gene	Procollagen N-proteinase	Procollagen studies

individuals with EDS, since considerable phenotypic heterogeneity exists. Molecular testing has revealed compound heterozygosity in a more severely affected member of one family displaying a classic EDS phenotype. Test results are useful for genetic counseling. Prenatal mutation testing of the *COL5A1* and *COL5A2* genes has not been reported[9] but could be accomplished in an approach identical to that described for OI.

Classic EDS infrequently is associated with defects in the *COL1A1* or *COL1A2* genes. Testing of these genes may be indicated following negative findings on *COL5A1* and *COL5A2* analysis. Prenatal testing for *COL1A1* and *COL1A2* gene defects could be accomplished in an approach identical to that described for OI.

Diagnostic confirmation of EDS type I or II may benefit the patient in several ways. Physiotherapeutic programs may be initiated in children with hypotonia and delayed motor development. Patients may also be counseled to avoid activities inducing heavy joint strain. As in most types of EDS, wounds should be closed without tension and deep stitches should be applied and left in place for twice the normal length of time.

Available Assays

Research-based *COL5A1* and *COL5A2* gene testing is available. The initial step is screening both genes for null alleles. Genomic DNA is first analyzed for the presence of polymorphic markers. Extracted mRNA is then used to generate cDNA that is sequenced to determine whether both alleles are expressed.[11] Approximately 20% to 25% of classic EDS patients will have a null allele (Peter Byers, personal communication). If a null allele is not evident, *COL5A1* or *COL5A2* genes are analyzed in a manner similar to that described for the *COL1A1* and *COL1A2* genes under OI (Peter Byers, personal communication).

COL1A1 and *COL1A2* gene defects in classic EDS are rare. Molecular testing is not routinely offered but could be performed in an approach identical to that outlined for OI above. Molecular testing for the *TNXA* gene is not clinically available.

EDS TYPE III (HYPERMOBILITY TYPE)

Molecular Basis of Disease

Patients with EDS type III have generalized joint hypermobility. They suffer frequent dislocations, effusions, and precocious arthritis. They may have variably hyperextensible or smooth, velvety skin. Tissue fragility is not a characteristic, but mitral valve prolapse is prevalent. Inheritance is autosomal dominant and the candidate gene in at least some cases is *COL5A3*.[9] Another candidate gene is the *COL3A1* gene.[9] Molecular testing is not available.

EDS TYPE IV (VASCULAR TYPE)

Molecular Basis of Disease

Patients with EDS type IV have the worst prognosis of all subtypes.[9] Arterial complications include spontaneous hemorrhages, aneurysms, dissections, and arteriovenous fistulas. Inguinal hernia is common. Rupture of the colon, gravid uterus, and other organs occurs.[9,12,13] Complications of pregnancy also include an incompetent cervix, a prolapsed uterus, and fragility of the membranes. The skin is thin, translucent, and fragile, but not hyperelastic. Marked bruising and delayed wound healing is seen. Slight hypermobility of the joints of the hand and feet as well as congenital hip dislocation can occur. Additional findings may include keratoconus, talipes equinovarus, gingival recession, thrombophlebitis, venous varicosities, keloid formation, tendon or muscle rupture, and (hemato-) pneumothorax.[9] Patients have a characteristic facial appearance with large eyes, a thin nose and thin lips. Complications are rare before puberty but increase with age.[9,12,13] Survival is foreshortened, with a median age at death of 48 years.[12,13]

EDS type IV is caused by defects in the *COL3A1* gene.[9,12] The *COL3A1* gene encodes proα1(III) collagen chains. These chains combine and are processed to form trimeric type III collagen molecules with a large, central, triple helical domain similar to type I collagen. Type III collagen is found in the dermis, blood vessel walls, and connective tissue scaffold of organs.

Inheritance of EDS type IV is autosomal dominant, although a few instances of mosaicism have been documented.[13] Approximately 50% of the cases represent de novo mutations and 50% are inherited.[13] The majority of the reported *COL3A1* gene mutations are private.[9,12] Several types of mutations have been documented, including glycine substitutions, mutations leading to single exon skipping, and large in-frame deletions or intronic inclusions.[9,12,14] These mutations are manifested in a dominant negative manner. Some splice-site acceptor mutations may result in null alleles by the introduction of premature chain-termination codons and subsequent mRNA-mediated decay.[14,15] Direct introduction of a stop codon due to a point mutation also has been observed.[9,13]

Clinical Utility of Testing

Due to the severity and progressive nature of EDS type IV, mutational analysis is indicated. Genetic testing is useful for confirmation of the diagnosis, genetic counseling, and prenatal diagnosis. Diagnosis confirmation is important so that patients can, if possible, avoid elective or invasive surgical procedures and other high-risk activities. Vascular manipulations are particularly problematic. In some instances, prophylactic measures are used, such as total colectomy following bowel rupture. Pregnant women with

EDS type IV should be followed in a high-risk clinic.[9,13] Testing of asymptomatic children in an EDS type IV family may eliminate concern if the child does not carry the family-specific mutations or alternatively lead to increased surveillance if testing is positive. Mutational analysis is also used to facilitate genotype-phenotype correlations.

Available Assays

Clinical molecular testing is available for EDS type IV. Genetic analysis is preceded by procollagen synthetic studies utilizing cultured dermal fibroblasts. A series of overlapping, type III collagen cDNA products are generated from mRNA from those cultures displaying decreased procollagen synthesis or aberrant procollagen. The cDNA products are then sequenced.[12] For those cases with a high index of suspicion for EDS type IV, but with negative protein studies, *COL3A1* null allele analysis is performed. The approach is similar to that described for the *COL5A1* and *COL5A2* genes (see above under EDS types I and II).[15] Additional analysis includes sequencing the region of the gene coding for the C-propeptides, since this region is critical for procollagen chain association.

EDS TYPE VI (KYPHOSCOLIOTIC TYPE)

Molecular Basis of Disease

EDS type VI is divided into two subtypes. Most cases are classified as EDS VIA and are due to a deficiency of the lysyl hydroxylase (LH) enzyme. Those cases displaying normal LH activity are designated as EDS VIB.[9] The genetic defect underlying EDS VIB is unknown.

Newborns with EDS type VI usually present with severe hypotonia and kyphoscoliosis at birth. The kyphoscoliosis is progressive and results from both ligamentous laxity and muscular hypotonia.[9] Patients may have generalized joint laxity and scleral fragility. Rupture of the globe of the eye or retinal detachment may occur. Additional findings may include tissue fragility, easy bruising, arterial rupture, microcornea, and a marfanoid habitus.[9,13] Inheritance is autosomal recessive and a history of affected siblings is a diagnostic criterion.[9]

Lysyl hydroxylase 1 (procollagen-lysine, 2-oxyoglutarate 5-dioxygenase) is encoded by the *PLOD1* gene. This enzyme catalyzes the hydroxylation of lysyl residues in Gly-X-Lys sequences in procollagen molecules. Hydroxylysines and lysines are substrates in a series of extracellular reactions ultimately resulting in the formation of intermolecular collagen cross-links. These cross-links are essential for collagen stability. Hydroxylysines also serve as attachment sites for mono- and disaccharides.[9]

Confirmed *PLOD1* gene mutations include those leading to amino-acid substitutions, exon skipping, premature termination codons (deletions and insertions resulting in frameshifts and directly introduced nonsense codons), and a common 7-exon duplication.[9,16,17] This duplication, involving exons 10 through 16 of *PLOD1*, is due to Alu-Alu repeat-mediated recombination in introns 9 and 16. The duplication is present in at least 20% of the EDS type VI cases.[9,17]

Clinical Utility of Testing

PLOD1 gene analysis is useful for confirmation of the diagnosis of EDS type VI. Results of genetic testing may be used to define the risk of recurrence from obligate heterozygous parents of an affected child.[9] Since severe muscular hypotonia at birth is a major diagnostic criterion for EDS type VI, *PLOD1* gene analysis may be indicated when EDS is considered in the differential diagnosis of neonatal hypotonia. Confirmation of diagnosis is important so that routine ophthalmic examinations to screen for retinal detachment and echocardiograms for aortic root enlargement may be initiated. Prenatal diagnosis is available.[17]

Available Assays

Molecular testing for EDS VIA is available only on a research basis. The approach initially involves the quantification of *PLOD1* mRNA by northern blot analysis. If sufficient mRNA is available, a full-length cDNA transcript is generated by reverse transcription (RT)-PCR. The cDNA is directly sequenced. Mutations are verified by sequencing genomic DNA. If a null allele is suspected, the testing proceeds directly to sequencing of genomic DNA, without the intervening cDNA steps.[17] Alternatives to DNA analysis exist.[9]

EDS TYPE VIIA AND EDS TYPE VIIB (ARTHROCHALASIC TYPE)

Molecular Basis of Disease

Arthrochalasic type EDS is characterized by severe, generalized joint hypermobility with recurrent subluxations, and congenital hip dislocation. Muscular hypotonia is evident at birth. Short stature due to kyphoscoliosis can occur. Patients also may be affected by osteopenia and bone fractures similar to mild OI.[9]

The inheritance pattern for arthrochalasic type EDS is autosomal dominant. Some sporadic mutations are recorded. The underlying defects are mutations leading to the deletion of exon 6 in either the *COL1A1* (EDS VIIA) or *COL1A2* (EDS VIIB) genes.[9,18] Exon 6 in both genes encodes the procollagen N-propeptide cleavage site and a lysine residue involved in collagen cross-linking. Mutations resulting in the skipping of exon 6, or direct genomic deletions in either gene, result in aberrant procollagen processing. The results are increased solubility of mutant procollagen and decreased collagen deposition in tissues.[9]

Clinical Utility of Testing

A procollagen protein test is available to confirm the diagnosis. Prenatal diagnosis is not available.

Available Assays

Molecular testing is not currently offered for arthrochalasic type EDS. Heteroduplex analysis as applied to the *COL1A1* and *COL1A2* genes in OI could be used to define exon 6 splice-site point mutations. Analysis of procollagens synthesized in the absence and presence of dextran sulfate, an agent promoting the activity of procollagen N-proteinase, is available.

EDS TYPE VIIC (DERMATOSPARACTIC TYPE)

Molecular Basis of Disease

The rare dermatosparactic type EDS is due to deficiency of procollagen N-proteinase, which is responsible for cleaving the amino-propeptides of procollagen types I and II. Procollagens containing the retained amino-propeptides accumulate, interfere with fibrillogenesis, and cannot form inter- and intramolecular cross-links through a critical lysyl residue.[9] Patients have lax, extremely fragile, redundant skin that heals with minimal scarring. Micrognathia and joint laxity are prominent. The sclera may be blue and fingers and limbs short. Premature rupture of the fetal membranes often occurs. Inheritance is autosomal recessive. Most patients have homozygous mutations.[9]

Clinical Utility of Testing

Procollagen testing is useful to confirm the diagnosis and for genetic counseling.

Available Assays

Molecular testing currently is not available. A procollagen protein analysis identical to that offered for EDS types VIIA and VIIB is available.[18]

ACHONDROGENESIS

Strictly speaking, achondrogenesis (ACG) refers to lethal chondrodysplasias presenting with a short trunk and very short extremities. Various classification schemes have identified subgroups of clinically distinct patients. Clinically recognized subtypes currently include types 1A, 1B, and 2.[19,20]

ACHONDROGENESIS TYPE 1B

Molecular Basis of Disease

Achondrogenesis type1B (ACG1B) lies at the severe end of a spectrum of disorders, all caused by defects in the diastrophic dysplasia sulfate transporter (*DTDST*) gene.[20,21] Death occurs before or shortly after birth. Affected fetuses are short, have a normal- or near-normal-sized head, with a flat face, narrow thorax, and protuberant abdomen. The limbs are severely shortened. Fingers and toes are short and stubby and the feet are rotated inward. The neck is also short and the infant appears hydropic. Cartilage is brown tinted, translucent, and friable. Histological studies reveal reduced extracellular matrix, coarse collagen fibers, and disorganized growth plates. Biochemical analysis reveals markedly decreased sulfate content of the proteoglycans. Undersulfated proteoglycans result from defects in the DTDST protein. The DTDST protein is a transmembrane protein involved in cellular sulfate uptake and functions as an anion exchanger in fibroblasts and chondrocytes.[20]

ATELOSTEOGENESIS TYPE 2

Atelosteogenesis type 2 (AO2) is a second, usually lethal chondrodysplasia, due to defects in the *DTDST* gene.[21,22,23] The phenotype is intermediate between ACG1B and the less-severe forms of chondrodysplasias due to defects in the *DTDST* gene.[20,21] Inheritance is autosomal recessive.

Numerous mutations have been defined in the *DTDST* gene. Certain mutations are associated with the lethal ACG1B or AO2 phenotypes, and some are seen with less severe phenotypes.[20,21] Mutations include short deletions causing frameshifts resulting in the downstream generation of stop codons, point mutations leading directly to stop codons, and short in-frame deletions leading to loss of an amino acid. Point mutations resulting in substitutions in the transmembrane domains are also associated with ACG1B.[20,21] Both homozygous and compound heterozygous patients have been identified.

Clinical Utility of Testing

Mutation analysis is useful for prenatal diagnosis and may aid in discrimination between the less severe chondrodysplasias caused by other *DTDST* gene defects. Carrier detection may be indicated in relatives of those with confirmed mutations. Genotype and phenotype correlations exist and potentially can be used to predict outcomes of various mutation combinations in families. The test can be used to confirm the diagnosis and may be used to discriminate between ACG1B and ACG2.

Available Assays

Exons and exon-intron boundaries of the *DTDST* gene are amplified with ten sets of PCR primers to generate a series of overlapping products. The discrete PCR fragments then are analyzed by SSCP. PCR products with an abnormal SSCP pattern are sequenced in both directions. If the screening test fails to reveal any suspicious products in a patient with a strong clinical history, the entire set of PCR products is sequenced.[21]

Interpretation of Test Results

The sensitivity of SSCP is reported as 80% to 90%. Overall test sensitivity, including DNA sequencing, exceeds 90% for identification of both mutant alleles.[21] In those instances where no mutations or only a single mutation is found, the mutation may occur in a 3′ or 5′ region of the gene not covered by the analytic approach.

Laboratory Issues

Molecular testing of the *DTDST* gene is available in one European laboratory. No in vitro diagnostic test kits are commercially available.

ACHONDROGENESIS TYPE 2

Molecular Basis of Disease

Achondrogenesis type 2 (ACG2) also is known as the Langer-Saldino type.[19] ACG2 and hypochondrogenesis represent the severe end of a spectrum of disorders with defects in a common gene. In ACG2 and hypochondrogenesis, the mutated gene directly encodes a defective form of the cartilage matrix protein, type II collagen.[19] Type II collagen is the major fibrillar collagen in cartilaginous tissues. ACG2 is more severe than hypochondrogenesis, but both may be associated with hydrops fetalis. Affected infants often are premature, stillborn, or die shortly following birth. They have large heads with soft cranial bones and flat faces. The neck, ribs, and long bones are short. The long bones are also broad, with cup-shaped metaphyses. The iliac wings are hypoplastic, and vertebral bodies are variably ossified. The ribs are often fractured.[19] The quantity of type II collagen present is case dependent. Biochemical studies indicate overmodification of type II collagen. This finding is consistent with the synthesis of abnormal type II collagen molecules.[19]

Essentially all cases of ACG2 and hypochondrogenesis arise as de novo mutations in the *COL2A1* gene; however, the possibility of germline mosaicism exists.[19] Mutations cause a dominant negative effect. Substitutions of the invariant glycines in Gly-X-Y repeat sequences are found scattered throughout the length of the triple helical domain.[19]

In addition to ACG2 and hypochondrogenesis, specific *COL2A1* gene mutations have been defined in spondyloepiphyseal dysplasia congenital, spondyloepiphyseal dysplasia Strudwick or spondyloepimetaphyseal dysplasia, spondyloepiphyseal dysplasia late onset, Kniest dysplasia, and certain cases of Stickler and Marshall syndromes. These disorders are distinguished by the phenotypic presentation and to some extent by defined mutations.[19]

Clinical Utility of Testing

Prenatal testing has been utilized to confirm the diagnosis in suspected ACG2 and hypochondrogenesis cases. Testing may discriminate between suspected ACG2 and hypochondrogenesis due to defects in the *COL2A1* gene, and ACG1B and AO2 due to defects in the *DTDST* gene. More importantly, the test may discriminate between ACG2 and hypochondrogenesis, and other, less-severe chondrodysplasias in prenatal cases. Testing also may confirm the diagnosis of ACG2 and hypochondrogenesis in an infant dying shortly after birth. Finally, molecular testing may be useful in the analysis of a parent suspected of being a mosaic.

Available Assays

COL2A1 gene mutation analysis is available in the United States and Europe. Testing is based on the generation of PCR products corresponding to all exons including the exon-intron boundaries[22] (Paul Couche, personal communication). The generated PCR products are screened for heteroduplexes by DHPLC or CSGE. Heteroduplexes are sequenced for mutation confirmation. The loss of polymorphisms in the *COL2A1* gene region is suggestive of complete or partial gene deletion. Genomic DNA isolated from specimens displaying loss of polymorphic markers may be evaluated for deletions by Southern blot analysis. Alternatively, mRNA studies may be conducted.

Interpretation of Test Results

Mutation detection is dependent on PCR product heteroduplex formation and ultimately recognizing an underlying DNA sequence variation. Technologies dependent on heteroduplex formation for mutation detection with other genes have sensitivities exceeding 90%.[22]

Laboratory Issues

Commercial testing for *COL2A1* gene mutation detection is available from at least two large academic laboratories, one located in Europe and one located in the United States.

The US laboratory is accredited by the CAP. An internal proficiency testing program has been approved by the CAP.

MARFAN SYNDROME

Molecular Basis of Disease

Marfan syndrome is a serious disorder of connective tissue with an autosomal dominant inheritance pattern and an incidence estimated at 1 in 5,000 to 10,000.[22,23] Organ systems involved may include the skin, eyes, skeleton, lung, fascia, adipose tissue, central nervous system, and cardiovascular system. Life expectancy is considerably foreshortened. Diagnosis is dependent on the presence of a series of major or minor criteria and may include molecular test results.[24] Marfan syndrome patients are long-limbed, thin individuals with arachnodactyly and elongated heads, and often with pectus deformities (excavatum or cariatum). Positive wrist (Walker-Murdoch sign) and thumb (Steinberg sign) signs reflect long fingers and lax joints. Pes planus is a common finding. Abnormal spinal curvature is frequent. Ectopia lentis is a major diagnostic criterion, and additional ocular findings exist. Patients are at increased risk for spontaneous pneumothorax. Mitral valve prolapse and severe regurgitation can lead to congestive heart failure and pulmonary hypertension. The mitral annulus may be calcified. Patients are subject to ventricular and supraventricular dysarrhythmia, and cardiac electrical disturbances may result in death. Aortic root dilation and the potential for dissecting aneurysm also represent life-threatening cardiovascular-system manifestations. Although aortic dilation may begin in utero, the dilation rate is variable and may not be apparent in childhood.[23]

Marfan syndrome is due to defects in fibrillin-1 (Table 16-4).[22,23,25–28] Fibrillin-1 is a cysteine-rich protein that assembles into microfibrils. Microfibrils are important constituents of elastic fibers. Perturbed elastogenesis is a direct consequence of defective fibrillin-1.[23] Some additional, nonstructural functions have been suggested for fibrillin-1.[23]

The *FBN1* gene encodes fibrillin-1. Fibrillin-1 is composed of several types of repeated motifs. These repeats include epidermal growth factor-like domains, calcium-binding epidermal growth factor-like (cbEGF-like) domains, latent transforming growth factor β-binding proteins (LTBP-1), also known as eight cys domains, and two hybrid motifs.[23,26] A proline-rich region is predicted to form a polyproline II helix.[23] Polyproline regions are hypothesized to function in intermolecular interactions. The EGF-like and cbEGF-like domains contain highly conserved amino acids, including six cysteine residues that form intradomain disulfide bonds.[23,26]

FBN1 mutations characterized in patients with Marfan syndrome include those leading to substitutions, in-frame or out-of-frame insertions and deletions, duplications, the introduction of stop codons, and the skipping of individual or multiple exons.[27,28] Mutations affecting the EGF-like and particularly the cbEGF-like motifs are common.[27]

Overlap Syndromes

FBN1 mutations have been defined in disorders sharing but not meeting the diagnostic criteria of Marfan syndrome.[23,25] Findings in the autosomal dominant form of ectopia lentis may be restricted to lens dislocation or may include skeletal or cardiovascular abnormalities. Familial aortic dissection or isolated aortic aneurysm and associated aortic dissection are inherited in an autosomal dominant manner and present with aortic aneurysm, dissection, or both. Some families have deformity of the thoracic cage.[23]

Clinical Utility of Testing

A disease causing *FBN1* mutation fulfills a major diagnostic criterion for Marfan syndrome diagnosis; hence, testing is indicated in suspected cases. Since Marfan syndrome is clinically heterogeneous and can display considerable intrafamily variability, testing can identify additional family members at risk. Mutation confirmation is important for initiation of a surveillance program consisting of periodic echocardiography. This is particularly important for children since they may not exhibit aortic root dilation. Prenatal diagnosis for previously defined mutations is available. Results of testing also are useful for genetic counseling. DNA testing is contributing to the rapid expansion of the international *FBN1* mutational database, which is contributing to a better understanding of the disease and to improving genotype-phenotype correlations.

Table 16-4. Achondrogenesis/Hypochondrogenesis and Marfan Syndrome

Type	Genes	Affected Molecules	Available Tests
Achondrogenesis 1A (ACG1A)	Unknown	Unknown	Not offered
Achondrogenesis 1B (ACG1B)	*DTDST*	Diastrophic sulfate transporter	SSCP analysis and sequencing of *DTDST*-derived PCR products
Atelosteogenesis 2 (AO2) and hypochondrodrogensis	*COL2A1*	Type II collagen	Heteroduplex analysis and sequencing of *COL2A1*-derived PCR products
Marfan syndrome	*FBN1*	Fibrillin-1	Heteroduplex analysis and sequencing of *FBN1*-derived PCR products

Available Assays

Genetic testing is available for Marfan syndrome in a few laboratories. An mRNA-based approach involves generation of *FBN1* cDNA. The cDNA is sequenced for mutation detection. Potential mutations are verified in genomic DNA using *FBN1* oligonucleotide primer sets. An alternative method uses genomic DNA to generate PCR products corresponding to all individual exons, including the exon-intron boundaries. The PCR products are analyzed for heteroduplex formation either by DHPLC or CSGE. Alternatively, PCR products can be analyzed for a migration shift by temperature gradient gel electrophoresis (TGGE). PCR products displaying heteroduplex formation or altered mobility are sequenced.[22,27,28] Both the genomic DNA and mRNA methods have advantages and disadvantages. The genomic DNA CSGE method has a sensitivity of greater than 90% when tested against patients with a confirmed clinical history of Marfan syndrome. The genomic DNA method will not detect large or complete allelic deletions. The mRNA method has the advantage of detecting large deletions but will fail to detect any mutations leading to unstable mRNA. This is a significant problem since null alleles are observed with stop codon mutations in certain genes and result from mRNA-mediated decay. The approaches should be viewed as complementary with awareness of the advantages and disadvantages of both.

Interpretation of Test Results

The sensitivity of the CSGE genomic DNA test method exceeds 90% for point mutations in patients with a confirmed diagnosis of Marfan syndrome.[22] The reported sensitivity of the TGGE test method is 51%.[28] Adoption of DHPLC by the laboratory utilizing the TGGE approach may increase sensitivity (Peter Robinson, personal communication).

Laboratory Issues

The US laboratory performing the genomic DNA testing method is accredited by the CAP. An internal proficiency testing program has been accepted by the CAP.

References

1. Kielty CM, Grant ME. The collagen family: structure, assembly, and organization in the extracellular matrix. In: Royce PM, Steinmann B, eds. *Connective Tissue and Its Heritable Disorders*. 2nd ed. New York: Wiley-Liss; 2002:159–221.
2. Silence D. Osteogenesis imperfecta: an expanding panorama of variants. *Clin Orthop.* 1981;159:11–25.
3. Byers PH, Cole WG. Osteogenesis imperfecta. In: Royce PM, Steinmann B, eds. *Connective Tissue and Its Heritable Disorders.* 2nd ed. New York: Wiley-Liss; 2002:358–430.
4. Körkkö J, Ala-Kokko L, De Paepe AD, et al. Analysis of the COL1A1 and COL1A2 genes by PCR amplification and scanning by conformational-sensitive gel electrophoresis identifies only COL1A1 mutations in 15 patients with osteogenesis imperfecta type I: identification of common sequences of null allele mutations. *Am J Hum Genet.* 1998;62:98–110.
5. Bruckner-Tuderman L. Epidermolysis bullosa. In: Royce PM, Steinmann B, eds. *Connective Tissue and Its Heritable Disorders.* 2nd ed. New York: Wiley-Liss; 2002:687–725.
6. Pulkkinen L, Uitto J. Mutational analysis and molecular genetics of epidermolysis bullosa. *Matrix Biol.* 1999;18:29–42.
7. Uitto J, Pulkkinen L. Molecular genetics of heritable blistering disorders. *Arch Dermatol.* 2001;137:1458–1462.
8. Cserhalmi-Friedman PB, Tang Y, Adler A, et al. Preimplantation genetic testing in two families at risk for recurrence of Herlitz junctional epidermalolysis bullosa. *Exp Dermatol.* 2000;9:290–297.
9. Steinmann B, Royce PM, Superti-Furga A. The Ehlers-Danlos syndrome. In: Royce PM, Steinmann B, eds. *Connective Tissue and Its Heritable Disorders.* 2nd ed. New York: Wiley-Liss; 2002:431–523.
10. Schalkwijk J, Zweers MC, Steijlen PM, et al. A recessive form of the Ehlers-Danlos syndrome caused by tenascin-X deficiency. *N Engl J Med.* 2001;345:1167–1175.
11. Schwarze U, Atkinson M, Hoffman GG, et al. Null alleles of the COL5A1 gene of type V collagen are a cause of the classic forms of Ehlers-Danlos syndrome (types I and II). *Am J Hum Genet.* 2000;66:1757–1765.
12. Pepin M, Schwarze U, Superti-Furga A, et al. Clinical and genetic features of Ehlers-Danlos syndrome type IV, the vascular type. *N Engl J Med.* 2000;342:673–680.
13. Pepin MG, Byers PH. Ehlers-Danlos syndrome, vascular type. GeneReviews [database online]. 1999. Available at: http://www.genetests.org. Updated 2002.
14. Schwarze U, Goldstein JA, Byers PH. Splicing defects in the COL3A1 gene: marked preference for 5′ (donor) splice-site mutations in patients with exon-skipping mutations and Ehlers-Danlos syndrome type IV. *Am J Hum Genet.* 1997;61:1276–1278.
15. Schwarze U, Schievink WI, Petty E, et al. Haploinsufficiency for one COL3A1 allele of type III procollagen results in a phenotype similar to the vascular form of Ehlers-Danlos syndrome type IV. *Am J Hum Genet.* 2001;69:989–1001.
16. Hyland J, Ala-Kokko L, Royce P, et al. A homozygous stop codon in the lysyl hydroxylase gene in a family with Ehlers-Danlos syndrome type VI. *Nat Genet.* 1992;2:228–231.
17. Yeowell HN, Walker LC. Mutations in the lysyl hydroxylase 1 gene that result in enzyme deficiency and the clinical phenotype of Ehlers-Danlos syndrome type VI. *Mol Genet Metab.* 2000;71:212–224.
18. Byers PH, Duvic M, Atkinson M. Ehlers-Danlos syndrome typeVIIA and type VIIB result from splice-junction mutations or genomic deletions that involve exon 6 in the COL1A1 and COL1A2 genes of type I collagen. *Am J Med Genet.* 1997;72:94–105.
19. Horton WA, Hecht JT. Chondrodysplasias: disorders of cartilage matrix proteins. In: Royce PM, Steinmann B, eds. *Connective Tissue and Its Heritable Disorders.* 2nd ed. New York: Wiley-Liss; 2002:909–937.
20. Superti-Furga A. Skeletal dysplasias related to defects in sulfate metabolism. In: Royce PM, Steinmann B, eds. *Connective Tissue and Its Heritable Disorders.* 2nd ed. New York: Wiley-Liss; 2002:939–960.
21. Rossi A, Superti-Furga A. Mutations in the diastrophic sulfate transporter (DTDST) gene (SLC26A2): 22 novel mutations, mutation review, associated skeletal phenotypes, and diagnostic relevance. *Hum Mutat.* 2001;17:159–171.
22. Körkkö J, Kaitila I, Lönnqvist L, et al. Sensitivity of conformational sensitive gel electrophoresis in detecting mutations in Marfan-syndrome and related conditions. *J Med Genet.* 2002;39:34–41.

23. Pyeritz RE, Dietz HC. Marfan syndrome and other microfibrillar disorders. In: Royce PM, Steinmann B, eds. *Connective Tissue and Its Heritable Disorders*. 2nd ed. New York: Wiley-Liss; 2002:585–626.
24. De Paepe A, Devereux RB, Dietz HC, et al. Revised diagnostic criteria for the Marfan syndrome. *Am J Hum Genet.* 1996;62:417–426.
25. Milewicz DM, Zsolt U, Boyd C. Genetic disorders of the elastic fiber system. *Matrix Biol.* 2000;19:471–480.
26. Handford PA, Downing AK, Reinhardt DP, et al. Fibrillin: from domain to supramolecular assembly. *Matrix Biol.* 2000;19:457–470.
27. Robinson PN, Booms P, Katzke S, et al. Mutations of FBN1 and genotype-phenotype correlations in Marfan syndrome and related fibrillinopathies. *Hum Mutat.* 2002;20:153–161.
28. Katzke S, Booms P, Tiecke F, et al. TGGE screening of the entire FBN1 coding sequence in 126 individuals with Marfan syndrome and related fibrillinopathies. *Hum Mutat.* 2002;20:197–208.

Section II

Inherited Cancers

Section Editors: Karen L. Kaul and Debra G.B. Leonard

Chapter 17

Inherited Breast Cancer

Wendy S. Rubenstein

Molecular Basis of Disease

Breast Cancer

Breast cancer is the most common cancer among women in Western countries, with about 180,000 new cases and 40,000 deaths occurring annually in the United States. Epidemiologic factors consistently associated with breast cancer risk include a family history of breast cancer, breast biopsy features, and hormonal risk factors such as age at menarche, parity, and age at first live birth. After female gender and age, family history of breast cancer is the most significant risk factor. In a meta-analysis of family history of breast cancer as a risk factor, the relative risk ranged from 1.5 for a second-degree relative to 3.6 for a mother and sister with breast cancer.[1] Relative risks are significantly influenced by the degree of relationship of affected relatives and their age of breast cancer onset, with closer degrees of relationship and younger age of onset conveying higher risks. An analysis of family history as a risk factor using data from the Swedish Family-Cancer Database showed a population-attributable fraction of about 11%.[2]

Breast cancer is a complex disease, resulting from an incompletely characterized interplay of genetic and environmental factors. In the majority of cases, the level of genetic analysis currently available cannot be used to stratify risk. However about 5% to 10% of breast cancer is hereditary, that is, due to the transmission of highly penetrant mutations in breast cancer–predisposing genes. Within hereditary breast cancer families, mutation status is the overriding risk factor, and genetic analysis can be used to clarify risk and guide medical management in a highly effective way. Genetic risk assessment consists of evaluating the pattern of cancers in the family, judging which of the known hereditary breast cancer syndromes fits the pattern, and pursuing genetic analysis.[3]

A specific genetic syndrome can be elucidated in about half of hereditary breast cancer families. Additional genes remain to be described.[4] Risk-conferring alleles are conceptualized as high-penetrance genes with low prevalence (e.g., *BRCA1* and *BRCA2* [hereditary breast-ovarian cancer (HBOC) syndrome], *TP53* [Li-Fraumeni syndrome], *PTEN* [Cowden syndrome], *LKB1*[Peutz-Jeghers syndrome]) or low-penetrance genes with high prevalence (possibly *CHEK2*,[5] *ATM*,[6] and the *TGFBR1*6A* allele[7]). The latter category may have small effects in individuals, but large aggregate effects in populations because they are common.

Using data from the Anglian Breast Cancer Study, Pharoah et al.[8] found that the best-fitting genetic model hypothesized that susceptibility to breast cancer is due to several loci, each conferring a modest independent risk. Assuming that all the susceptibility genes could be identified, they showed that the half of the population at highest risk would account for 88% of all affected individuals. Clinical testing for one or a few low-penetrance genes at a time would be unsatisfying in many respects. Mutations are found in many individuals but convey small risks, and gene-gene and gene-environment interactions are unknown, limiting clinical utility. Ultimately, whole genome screening might be used in combination with knowledge about such interactions to achieve higher predictive power and allow for efficient breast cancer risk stratification.

BRCA1 and *BRCA2*

Newman et al. published the first study providing quantitative evidence for an autosomal dominant breast cancer susceptibility allele, accounting for an estimated 4% of breast cancer families and conveying an 82% lifetime risk of breast cancer.[9] Following the report by Hall et al. of linkage to chromosome 17q21 for early-onset hereditary breast cancer,[10] *BRCA1* was isolated using a positional cloning strategy in 1994.[11] Subsequently, *BRCA2*, a second breast-ovarian cancer susceptibility gene, was localized to chromosome 13q12-q13 and cloned.[12,13] GenBank, the National Institutes of Health genetic sequence database, lists entries for *BRCA1* and *BRCA2* as U14680 and U43746, respectively (http://www.ncbi.nlm.nih.gov/Genbank/).

Elucidation of the functions of BRCA1 and BRCA2 has lagged behind the technical capability of carrier detection, delineation of the clinical syndrome, and demonstration of the efficacy of medical management strategies. The manifold functions of BRCA1 and BRCA2 are incompletely characterized. *BRCA1* and *BRCA2* encode very large proteins with 1863 and 3418 amino acids, respectively; each bears little homology to other known proteins or to each other. BRCA1 appears to play a role in numerous cellular functions including transcriptional regulation and influence of estrogen receptor activity, chromatin remodeling, DNA damage repair (homologous recombination and repair of transcription-coupled oxidation-induced DNA damage), centrosome duplication, cell growth, apoptosis, and cell cycle checkpoint control.[14] BRCA1 contains an N-terminal RING domain that interacts with BARD1. Two BRCA1 C-terminal (BRCT) domains are present, which are found in proteins involved in DNA repair and control of the cell cycle. *BRCA2* contains eight highly conserved BRC repeats of 30 to 40 residues in exon 11, which bind to RAD51, a key recombinational repair protein. After exposure of cells to DNA damage, BRCA1 relocalizes from nuclear foci to sites of DNA synthesis and becomes hyperphosphorylated. BARD1, BRCA2, and RAD51 all relocalize with BRCA1.[15]

Available Assays

Overview of Mutation Types

The *BRCA1* locus spans 100 kilobases (kb), encompasses 22 coding exons, and encodes an open reading frame of 5.5 kb (of 24 exons, exons 1 and 4 are noncoding). *BRCA2* has 26 coding exons and an open reading frame of 10 kb. While the size of these genes presents an enormous challenge to mutational analysis, the problem is compounded by the distribution of mutations throughout the coding regions and introns.[16] Interpretation of results is further complicated by the occurrence of numerous polymorphisms of uncertain significance. All types of mutations have been found, including truncating mutations, frameshift mutations, and missense mutations. Furthermore, loss of gene function also occurs through large duplications, deletions, and rearrangements, including promoter deletions. Because the complete genetic characterization of *BRCA1* and *BRCA2* is an ongoing process, the technique(s) selected for mutation detection must be comprehensive in order to provide an accurate clinical result. In general, both genes must be assayed because the *BRCA1* and *BRCA2* clinical syndromes cannot be readily distinguished in individual families.

BRCA1 and *BRCA2* mutations and polymorphisms are catalogued in the Breast Cancer Information Core (BIC), an open-access online database hosted by the National Human Genome Research Institute (http://research.nhgri.nih.gov/bic/). At the time of this writing, the BIC database contained 6672 entries for *BRCA1*, with 1236 distinct mutations, polymorphisms, and variants, 709 of which had been reported only once. For *BRCA2*, BIC contains 5624 entries, including 1380 distinct mutations, polymorphisms, and variants, 871 of which had been reported only once. Among 7461 consecutive full-sequence analyses of *BRCA1* and *BRCA2* performed at a commercial reference laboratory, 689 (61%) of the mutations identified occurred in *BRCA1* and 440 (39%) in *BRCA2*. Of the 424 different mutations detected, 256 (60%) were frameshifts, 106 (25%) were nonsense, 9 (2.1%) were missense, and 53 (12.5%) occurred in the analyzed regions of introns.[16] Genomic rearrangements are discussed below. The technical and financial limitations imposed by the complexity of genetic analysis impacts the availability of clinical genetic testing and the ability to perform genetic epidemiologic studies.

Functional and Phenotypic Screening

Given the multifunctional nature of *BRCA1* and *BRCA2*, the development of functional assays amenable to clinical testing is not readily anticipated. An immunoassay based on diminished anticarboxy vs antiamino immunoreactivity to *BRCA1* and *BRCA2* mutations has been applied to buccal cells[17] and ovarian cancers.[18] This assay would be useful for screening truncating mutations, which comprise the majority of mutations, but has not been further characterized.

There is no simple pathologic correlate of the presence of *BRCA1* or *BRCA2* mutations in breast tumors (akin to microsatellite instability for colorectal cancer and mismatch repair gene deficiency) to serve as an effective screen. An immunohistochemical tumor phenotype (typically negative for estrogen receptor, progesterone receptor, and HER2 and positive for TP53) and distinctive morphology (high mitotic count, lymphocytic infiltrate, smooth noninfiltrative pushing borders, and an excess of medullary and atypical medullary types) have been discerned for breast cancers with *BRCA1* mutations.[19] However, these features are neither sufficiently sensitive nor specific to reliably predict germline mutation status.

Initial reports on expression profiling in breast tumors have found expression patterns indicative of *BRCA1* loss of function.[20,21] If these results can be replicated and if they are widely applicable to a variety of mutation types, then an expression-profiling-based screen for breast tumors could be developed.

Sequence Analysis, Gene Scanning Techniques, and Linkage Analysis

A comparison of analytical methods for *BRCA1* and *BRCA2* is presented in Table 17–1. Direct sequence analysis is considered the gold standard test. The BIC Web site provides extensive information on methodologies includ-

Table 17-1. A Comparison of Analytical Methods for *BRCA1* and *BRCA2*

	Sensitivity					
	Frameshift, Nonsense, and Missense Mutations; Analyzed Regions of Introns	Large Deletions, Duplications, and Rearrangements	Specificity	Cost	Labor	Variants of Uncertain Significance Detected
Direct nucleotide sequence analysis	>98%	No	99%	Expensive; ~$3000	Intensive	Yes
SSCP	65%	No	<95%?	Low (excluding sequence confirmation)	Moderate	Yes
CSGE	60%	No	<95%?	Low (excluding sequence confirmation)	Moderate	Yes
TDGS	91%	No	Lower than other mutation-scanning techniques	Low (excluding sequence confirmation)	Moderate	Yes
DHPLC	>99%	No	>95%	Low (excluding sequence confirmation) but high equipment costs	Moderate	Yes
DOVAM-S	High but unknown for BRCA1/2	No	Unknown	Moderate	Moderate	Yes
PTT	<<80%	No	Unknown	Moderate	Moderate	No
Southern blot, exon 13 duplication, fluorescence DNA microarray, color bar coding	No	Unknown; ~3–12%	Unknown	Expensive if comprehensive	Intensive	No
MLPA	No	High but unknown for BRCA1/2	Unknown	Moderate	Moderate	No
Linkage analysis	Theoretically up to 100% but limited to suitable families		Dependent on family LOD score	Expensive	High	No
Ashkenazi Jewish founder mutation analysis	78–96% in Ashkenazi Jews	No	>99%	Low; ~$450	Low	No

ing a genomic primer database (intronic sequence databases), group-specific primer sequences from different laboratories, and a test method protocol database. Due to the high complexity of testing for gene mutations in *BRCA1* and *BRCA2*, a highly automated approach to analysis is worthwhile, consisting of sample tracking, robotic assay implementation, and data analysis (described in greater detail below; see Quality Control and Laboratory Issues). Sample DNA is amplified by polymerase chain reaction (PCR); then the PCR products are directly sequenced in forward and reverse directions using M13 forward and reverse tagged primer pairs, generating 36,000 nucleotides of sequence information per patient.[22] Thus, analysis of the *BRCA1* and *BRCA2* genes and management of the multitudinous data generated is a challenging and expensive feat.

Gene-scanning technologies hold appeal as an alternative to sequencing, particularly for large-scale epidemiological studies, because sequence analysis of wild-type exons is avoided. Gene-scanning techniques that have been applied to *BRCA1* and *BRCA2* analysis include single-strand conformational polymorphism analysis (SSCP), conformational sensitive gel electrophoresis (CSGE),[23] two-dimensional gene scanning (TDGS),[24] and denaturing high-performance liquid chromatography (DHPLC).[25]

A number of other mutation-detection techniques have been used or proposed as applicable to *BRCA1* and *BRCA2* gene analysis. A technique called DOVAM-S, an acronym for "detection of virtually all mutations-SSCP,"[26] is a variation of SSCP that scans multiple exons in a single lane under several different nondenaturing electrophoresis conditions, resulting in improved sensitivity (e.g., for analysis of the factor IX and *ATM* genes). Another approach, multiplex ligation-dependent probe amplification (MLPA)[27] with diploid-to-haploid conversion, has been applied to analysis of the *MSH2* and *MLH1* genes, resulting in detection of rearrangements not amenable to detection by standard sequencing. Deletions of sizes that fall between the detection limits of sequence analysis and Southern blot analysis are a possible target area for newer methodologies. For example, a method to identify unusual haplotypes that may represent small intragenic deletions

has been described.[28] Due to the high proportion of truncating mutations, protein truncation testing also is a viable way to screen for mutations[29–31] but is relatively insensitive.

The accuracy of gene-scanning techniques was evaluated in a blinded comparison by Eng et al.[31] They compared the sensitivity, specificity, and cost of SSCP, CSGE, TDGS, and DHPLC for detecting 58 distinct mutations in *BRCA1* among 50 samples. An additional 15 samples were included that had no mutations detected by sequence analysis. Exons 2 to 24 had been previously analyzed at a commercial reference laboratory. Mutation types included protein truncating and missense mutations as well as intronic sequence alterations occurring within 20 bp proximal or 10 bp distal to the ends of the exons. Sensitivity was evaluated by the ability to detect an abnormality initially and also to confirm the mutation by sequence analysis and report the result correctly. DHPLC was the only method that correctly identified each of the 58 mutations. A prior study also had reported a sensitivity of 100% for DHPLC.[32] The sensitivities of the other methods in the Eng et al. study were as follows: SSCP 65%, CSGE 60%, and TDGS 91%. Further aspects of this study are discussed in the section "Quality Control and Laboratory Issues."

The sensitivity of testing in the clinical setting reflects several factors, beyond the sensitivity of the specific laboratory techniques, which require clarification. Given particular constellations of family history, there is a certain expectation or probability that a germline mutation in *BRCA1* or *BRCA2* is causative.[33] Failure to identify a mutation can arise due to several reasons. First, the syndrome is not due to *BRCA1* or *BRCA2* mutations but due to another known syndrome for which genetic testing could be offered. Second, the syndrome is not due to *BRCA1* or *BRCA2* mutations but due to mutations in an undiscovered gene or genes.[34] Third, the wrong person in the family was tested; other family members have HBOC syndrome. Fourth, the family does not have hereditary breast cancer. And finally, there is a *BRCA1* or *BRCA2* mutation that cannot be identified using current methodologies. Cancer genetic consultation helps to distinguish these possibilities and to guide genetic testing strategies.[3] We focus here on the last problem of false-negative results and the use of alternative test methods to identify mutations that are not detectable by sequence analysis.

A study of 237 families collected by the Breast Cancer Linkage Consortium (BCLC), each with at least four cases of breast cancer diagnosed by age 60 years or male breast cancer at any age, provided insight into the genetic epidemiology of *BRCA1* and *BRCA2* and the sensitivity of mutation detection methods.[35] Disease was linked to *BRCA1* in 52% of families and to *BRCA2* in 32% of families. Families with female breast cancer only (no ovarian cancer or male breast cancer) showed the highest rate of nonlinkage (67%); that is, the breast cancers in these families were attributable to other genes. This study used estimated odds of linkage and found an overall sensitivity of genetic testing methods (reflecting several different methods) of 63% (95% confidence interval 51–77%).

In a study that assessed the clinical validity of a computer program called BRCAPRO, the sensitivity of genetic testing was much higher.[36] BRCAPRO implements a statistical model for calculating the probability that an individual carries a *BRCA1* or *BRCA2* mutation (or neither). Age of breast and ovarian cancer onset in the proband and affected relatives and the age of cancer-free unaffected relatives are important components. The model incorporates published penetrance and prevalence estimates, autosomal dominant inheritance, and Bayesian updating. Genetic testing for subjects in this study was performed at several different laboratories but involved either complete DNA sequencing, or analysis limited to the three common Ashkenazi Jewish mutations in those subjects of Ashkenazi Jewish ancestry. The estimated sensitivity of genetic testing in this study was at least 85%. This was interpreted to reflect a trade-off between estimated sensitivity and the estimated prevalence of other susceptibility genes; if genetic testing has perfect sensitivity, then negative sequencing results are attributable to other genes.

These two studies highlight the continued utility of linkage analysis for families with high pretest probabilities and unidentifiable mutations. Intragenic markers are available[35] and linkage studies are as reliable as direct gene testing for those families that are suitable for analysis. For some families with breast cancer linked to *BRCA1* and *BRCA2*, current analysis cannot yet close the gap.

Genomic Rearrangements

Several genomic rearrangements within *BRCA1* and its coding region have been described. Rearrangements account for a significant proportion of mutations, and many are population specific. The first deleterious genomic rearrangement was reported by Puget et al.[37] and involves deletion of exon 17 due to Alu element recombination. They subsequently reported four additional new deletions,[38] two of which involved Alu elements as well as a recombination hotspot due to a *BRCA1* pseudogene upstream of *BRCA1*.[39]

The Puget group was the first to report a 6 kb duplication of *BRCA1* exon 13, which was evaluated in a multicenter study involving collaborators from 19 countries.[40] Based on haplotype analysis, the *BRCA1* exon 13 duplication is most likely a founder mutation and is seen primarily in English-speaking countries or in countries with ancestral links to Britain. The mutation was observed among 10 of 1831 affected families from English-speaking countries.

Three *BRCA1* genomic rearrangements in the Dutch population represent founder mutations and account for as much as 36% of *BRCA1* mutations for families of Dutch ancestry.[41] The three deletions comprise exon 22, exon 13,

and exons 13–16. Analysis of these rearrangements for families of Dutch ancestry therefore is indispensable.

Unger et al.[42] systematically evaluated the frequency of *BRCA1* genomic rearrangements in a cohort of American women seeking genetic testing for a family history of breast and ovarian cancer whose results were negative for *BRCA1* coding-region mutations. Using Southern blot analysis and exon 13 duplication screening, they found germline genomic rearrangements in 11.9% (5/42) of families. These results confirm that a significant proportion of *BRCA1* mutations can be discerned by genomic techniques but also provide evidence for the role of additional genes. In this study, a higher prior probability was associated with a higher likelihood of detecting a rearrangement, suggesting that a high pretest probability (e.g., ≥30%) could be used to guide further analysis for families with negative genetic testing by sequence analysis.

BRCA1 and *BRCA2* genetic testing performed by Myriad Genetic Laboratories, Inc (Salt Lake City, UT), incorporates an automated high-throughput multiplexed assay for five recurrent mutations: deletions of exons 8–9, 13, 14–20, and 22, and the exon 13 duplication. In a study of 7570 patients, 1036 deleterious mutations were identified, including 31 (3.0%) with one of these rearrangements.[43] The exon 13 duplication represented 83.9% of the rearrangement mutations and 2.5% of the total mutations, making this the most prevalent non-Ashkenazi mutation. European ancestry was reported for more than 90% of the families with the exon 13 duplication.

Analysis of genomic rearrangements has been accomplished using additional sophisticated techniques. For example, a fluorescence DNA microarray assay has been developed for the detection of large gene rearrangements spanning several kilobases of genomic DNA.[44] In addition, color bar coding has been used for analysis of *BRCA1*,[45] which involves DNA combing, done by homogeneous stretching of DNA molecules at a constant rate of 2 kb/μm. Fluorescence in situ hybridization is performed on the DNA using a probe that covers the entire *BRCA1* gene, in addition to long-range PCR products. The PCR products cover variable lengths of exons, providing the bar code.

Founder Mutations

Genetic analysis of *BRCA1* and *BRCA2* may be simplified in cases where founder mutations comprise a high proportion of mutations within a particular ethnic group. Numerous ethnic-specific mutations have been described across the globe.[46] About 2.5% of Ashkenazi Jews that are unselected for a family history of cancer carry one of three founder mutations: *BRCA1* 187delAG (also called 185delAG), *BRCA1* 5385insC (also called 5382insC), and *BRCA2* 6174delT.[47] Various methods have been described that specifically target the founder mutations. About 78% to 96% of Ashkenazi Jews with detectable mutations carry one of the founder mutations; thus, further analysis is required in high-risk families that do not carry one of the founder mutations.[16,48] These highly prevalent mutations account for a large proportion of breast and ovarian cancer in Ashkenazi Jews: about 20% of breast cancer diagnosed by age 40, and 30% to 40% of ovarian cancer.

In Dutch populations, genomic rearrangements in *BRCA1* account for as much as 36% of mutations.[41] In Icelandic populations, the *BRCA2* 999del5 founder mutation is present in about 0.6% of the population.[49] The Icelandic founder mutation accounts for about 40% of male breast cancer cases and most of the excess ovarian and prostate cancer in breast cancer families. Additional founder mutations have been reported in Norway,[50] Sweden,[51] Poland,[52] and numerous other countries, as well as in African and African American families.[53]

Interpretation

Truncating mutations comprise the largest proportion of mutation types identifiable with sequencing. All truncating mutations are associated with loss of gene function and clinical expression of breast and ovarian cancer, with the exception of those that prematurely terminate the protein product of *BRCA1* less than 10 amino acids from the C-terminus or the protein product of *BRCA2* less than 110 amino acids from the C-terminus. The major problems in interpretation of the clinical significance of sequence variations apply to missense variants and intronic mutations, chain-terminating mutations that truncate *BRCA1* and *BRCA2* distal to amino acid positions 1853 and 3308, respectively, and mutations that eliminate the normal stop codons for these proteins.

Frank et al. described the largest study of *BRCA1* and *BRCA2* genetic testing results and clinical characteristics, for consecutive tests on 10,000 individuals.[16] One or more variants of uncertain clinical significance in the absence of deleterious mutations were observed in 970 (13%) of the 7461 individuals sequenced in this study.[16] Epidemiological and biological criteria can be applied to distinguish functional from benign variants with some success.[54] For example, the prevalence of each variant in a control population, cosegregation of the variant with cancer within families, location of the variant within the gene, functional assays, demonstration of abnormal mRNA transcript processing, type of the amino acid substitution, and degree of conservation among species[55] provide clues as to whether the mutation is deleterious.[16] Variant *BRCA* test results pose a commonly encountered and highly problematic issue,[56] which can possibly lead to the inappropriate use of medical interventions such as prophylactic surgery;[57] thus having genetic counseling and competent test interpretation is paramount.

It should be pointed out that false-positive results (as reflected by specificities <100%) pose the most potential

harm since "positive" results are used to make decisions about prophylactic breast and ovarian surgery, which are irreversible decisions. False-negative results are far more common but are dealt with by reverting to pedigree information and maintaining a high level of suspicion where indicated.

Clinical Significance

BRCA1 (Online Mendelian Inheritance in Man [OMIM; database online] #113705) and *BRCA2* (OMIM #600185) mutations predispose female carriers to a high lifetime risk of breast cancer (>80%) and ovarian cancer (40–65% for *BRCA1* carriers and 20% for *BRCA2* carriers). The clinical features and management of HBOC syndrome have been reviewed.[57] Average ages of breast and ovarian cancer onset are generally younger for *BRCA1* carriers than for *BRCA2* carriers, but each can manifest as breast cancer in the 20s. *BRCA2* contains an ovarian cancer cluster region (OCCR) in exon 11 defined by nucleotides 4075–6503 that appears to convey higher ovarian cancer risks; this region includes the 6174delT Ashkenazi Jewish founder mutation. Penetrance estimates have varied widely among studies, reflecting differences in populations and in study methods.[58] Generally, lower penetrance estimates have been seen in studies of the three Ashkenazi Jewish founder mutations. However, in a study of Ashkenazi Jewish probands selected on the basis of having incident primary invasive breast cancer, King et al. found penetrance estimates (to age 80 years) similar to that described in other groups: 82% for breast cancer for *BRCA1* and *BRCA2* carriers, 54% for ovarian cancer for *BRCA1* carriers, and 23% for ovarian cancer for *BRCA2* carriers.[59]

Male breast cancer is seen in excess in *BRCA1* and *BRCA2* families, with about two thirds of positive cases involving *BRCA2* and one third involving *BRCA1*.[16] Lifetime risk of breast cancer is about 6% for male *BRCA2* carriers and is probably lower for male *BRCA1* carriers. *BRCA2* carriers also are at excess risk for pancreatic cancer.

Many effective cancer risk-management strategies are available for *BRCA1* and *BRCA2* carriers, as well as for families with a high clinical suspicion of genetic predisposition.[60] The chief value of genetic testing is to confirm the need for medical interventions, particularly those that are irreversible such as prophylactic mastectomy and prophylactic oophorectomy. As well, a true negative result (i.e., in the setting of a known familial mutation) obviates the need for aggressive surveillance and prevention measures, and provides reassurance to the person tested as well as to their offspring. Surveillance and management for HBOC syndrome is summarized by Lynch et al.[57] and includes consideration of chemoprevention of breast (e.g., tamoxifen) and ovarian (e.g., oral contraceptives) cancers, MRI surveillance for breast cancer under a research protocol, and early breast cancer surveillance (age 25 years) in at-risk female relatives.

Quality Control and Laboratory Issues

The study of Eng et al.[31] highlights many of the quality control issues facing laboratories performing analysis of *BRCA1* and *BRCA2*. For the laboratory performing the SSCP technique, a false-positive result was reported that occurred due to switching of samples; this also accounted for one of the false-negative results. There appeared to be variability in the ability to detect subtle changes by SSCP. For the laboratory performing CSGE, false-negative results were in part due to erroneous interpretation of CSGE results, and in part due to failure to confirm abnormal gel mobility results using sequence analysis. Clerical errors led to three additional false-negative results. However, CSGE is readily applied to the analysis of known mutations at a low cost. A caveat regarding SSCP and CSGE is that these test methods can systematically fail to identify single-nucleotide substitutions and specific mutations such as the common *BRCA1* 5385insC mutation.

Although TDGS is appealing as a particularly high-throughput method, five mutations were missed due to incorrect interpretation of the two-dimensional gel, and two mutations were not successfully confirmed using sequence analysis. In addition, three false-positive results were reported, two of which might have been introduced by the long-distance PCR method. The TDGS method is prone to preferential amplification of the wild-type allele. DHPLC, which showed the best performance among the DNA scanning methods, may be further improved by the use of proofreading DNA polymerases to increase the fidelity of the PCR.[61]

The high-throughput DNA platform developed by Myriad Genetic Laboratories, Inc (Salt Lake City, UT), addresses many of the quality control issues that surfaced in the above study (personal communication, T. Scholl, December 2002). The test requisition and blood specimen are matched by a bar code that is scanned into a database. Pre-PCR procedures are carried out in positive pressure clean laboratories where genomic DNA is isolated from mononuclear cells. Next, 82 test PCR inoculations are performed, in conjunction with two reactions that detect M13 priming sites to control for contamination by amplified products. The PCR amplification products are thermal cycle-sequenced and the sequencing reaction products analyzed on slab gels using high-throughput sequencers (ABI 377). Data analysis also is automated, and a proprietary data review application analyzes the sequence for mutations and polymorphisms. Algorithms are applied to the data to assess signal-to-noise ratios, alignment scores, and signal intensities. A data reviewer who reviews electropherograms from sense and antisense DNA strands confirms sequence variants; antisense DNA information is reverse color complemented to facilitate review. Background subtraction of the composite wild-type sequence also can be done. The involved amplicon then is confirmed by repeat sequence analysis. The

reported sensitivity of this platform is 98%, with a specificity of 99%, for mutations detectable by sequence analysis.

Conclusion

Mutation detection in the *BRCA1* and *BRCA2* genes is of great clinical import but is extremely challenging from a technical standpoint. The large variety of mutation types mandates the use of multiple mutation detection methods if a sensitivity approaching 100% is desired. Methods for assessing pretest probability based on clinical presentation[33] can be used to determine which cases merit the highest degree of mutational analysis. Further improvements in mutation detection methods are needed to clarify the true sensitivity of testing versus the proportion of families with mutations in other hereditary breast cancer genes. Clinical interpretation is beleaguered by the high prevalence of genetic variants of uncertain significance. Development of functional assays is highly desirable and may help classify variants into benign or disease-causing categories.

References

1. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. *Int J Cancer.* 1997;71:800–809.
2. Hemminki K, Granstrom C, Czene K. Attributable risks for familial breast cancer by proband status and morphology: a nationwide epidemiologic study from Sweden. *Int J Cancer.* 2002;100:214–219.
3. Offit K. *Clinical Cancer Genetics. Risk Counseling and Management.* New York: Wiley-Liss; 1998.
4. de Jong MM, Nolte IM, te Meerman GJ, et al. Genes other than BRCA1 and BRCA2 involved in breast cancer susceptibility. *J Med Genet.* 2002;39:225–242.
5. Meijers-Heijboer H, van den OA, Klijn J, et al. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet.* 2002;31:55–59.
6. Thorstenson YR, Roxas A, Kroiss R, et al. Contributions of ATM mutations to familial breast and ovarian cancer. *Cancer Res.* 2003;63:3325–3333.
7. Kaklamani VG, Hou N, Bian Y, et al. TGFBR1*6A and cancer risk: a meta-analysis of seven case-control studies. *J Clin Oncol.* 2003;21:3236–3243.
8. Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet.* 2002;31:33–36.
9. Newman B, Austin MA, Lee M, King MC. Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci U S A.* 1988;85:3044–3048.
10. Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science.* 1990;250:1684–1689.
11. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science.* 1994;266:66–71.
12. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature.* 1995;378:789–792.
13. Tavtigian SV, Simard J, Rommens J, et al. The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat Genet.* 1996;12:333–337.
14. Deng CX, Brodie SG. Roles of BRCA1 and its interacting proteins. *Bioessays.* 2000;22:728–737.
15. Scully R, Livingston DM. In search of the tumour-suppressor functions of BRCA1 and BRCA2. *Nature.* 2000;408:429–432.
16. Frank TS, Deffenbaugh AM, Reid JE, et al. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J Clin Oncol.* 2002;20:1480–1490.
17. Byrne TJ, Reece MT, Adams LA, Hoffman DE, Lane MA, Cohn GM. A rapid immunoassay predicts BRCA1 and BRCA2 mutations in buccal cells. *Oncol Rep.* 2000;7:1203–1207.
18. Byrne TJ, Reece MT, Adams LA, Lane MA, Cohn GM. An antibody assay predictive of BRCA1 mutations in ovarian tumors and normal tissue. *Oncol Rep.* 2000;7:949–953.
19. Lakhani SR, Van De Vijver MJ, Jacquemier J, et al. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol.* 2002;20:2310–2318.
20. Hedenfalk I, Duggan D, Chen Y, et al. Gene-expression profiles in hereditary breast cancer. *N Engl J Med.* 2001;344:539–548.
21. Van 't Veer LJ, Dai H, Van De Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* 2002;415:530–536.
22. Frank TS, Manley SA, Olopade OI, et al. Sequence analysis of BRCA1 and BRCA2: correlation of mutations with family history and ovarian cancer risk. *J Clin Oncol.* 1998;16:2417–2425.
23. Ganguly A, Rock MJ, Prockop DJ. Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes. *Proc Natl Acad Sci U S A.* 1993;90:10325–10329.
24. van Orsouw NJ, Dhanda RK, Rines RD, et al. Rapid design of denaturing gradient-based two-dimensional electrophoretic gene mutational scanning tests. *Nucleic Acids Res.* 1998;26:2398–2406.
25. Wagner T, Stoppa-Lyonnet D, Fleischmann E, et al. Denaturing high-performance liquid chromatography detects reliably BRCA1 and BRCA2 mutations. *Genomics.* 1999;62:369–376.
26. Buzin CH, Wen CY, Nguyen VQ, et al. Scanning by DOVAM-S detects all unique sequence changes in blinded analyses: evidence that the scanning conditions are generic. *Biotechniques.* 2000; 28:746–750, 752–753.
27. Nakagawa H, Hampel H, de la Chapelle A. Identification and characterization of genomic rearrangements of MSH2 and MLH1 in Lynch syndrome (HNPCC) by novel techniques. *Hum Mutat.* 2003;22:258–263.
28. Hendrickson BC, Pruss D, Lyon E, Scholl T. Application of haplotype pair analysis for the identification of hemizygous loci. *J Med Genet.* 2003;40:346–347.
29. Hogervorst FB, Cornelis RS, Bout M, et al. Rapid detection of BRCA1 mutations by the protein truncation test. *Nat Genet.* 1995;10:208–212.
30. Hardy CA. The protein truncation test. *Methods Mol Biol.* 2002; 187:87–108.
31. Eng C, Brody LC, Wagner TM, et al. Interpreting epidemiological research: blinded comparison of methods used to estimate the prevalence of inherited mutations in BRCA1. *J Med Genet.* 2001;38:824–833.
32. Gross E, Arnold N, Goette J, Schwarz-Boeger U, Kiechle M. A comparison of BRCA1 mutation analysis by direct sequencing, SSCP and DHPLC. *Hum Genet.* 1999;105:72–78.
33. Rubinstein WS, O'Neill SM, Peters JA, Rittmeyer LJ, Stadler MP. Mathematical modeling for breast cancer risk assessment: state of the art and role in medicine. *Oncology (Huntington).* 2002;16:1082–1094.
34. Antoniou AC, Pharoah PD, McMullan G, Day NE, Ponder BA, Easton D. Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. *Genet Epidemiol.* 2001;21:1–18.

35. Ford D, Easton DF, Stratton M, et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet.* 1998;62:676–689.
36. Berry DA, Iversen ES Jr, Gudbjartsson DF, et al. BRCAPRO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. J *Clin Oncol.* 2002;20: 2701–2712.
37. Puget N, Torchard D, Serova-Sinilnikova OM, et al. A 1-kb Alu-mediated germ-line deletion removing BRCA1 exon 17. *Cancer Res.* 1997;57:828–831.
38. Puget N, Stoppa-Lyonnet D, Sinilnikova OM, et al. Screening for germ-line rearrangements and regulatory mutations in BRCA1 led to the identification of four new deletions. *Cancer Res.* 1999;59: 455–461.
39. Puget N, Gad S, Perrin-Vidoz L, et al. Distinct BRCA1 rearrangements involving the BRCA1 pseudogene suggest the existence of a recombination hot spot. *Am J Hum Genet.* 2002;70:858–865.
40. The BRCA1 Exon 13 Duplication Screening Group. The exon 13 duplication in the BRCA1 gene is a founder mutation present in geographically diverse populations. *Am J Hum Genet* 2000;67(1): 207–212.
41. Petrij-Bosch A, Peelen T, van Vliet M, et al. BRCA1 genomic deletions are major founder mutations in Dutch breast cancer patients. *Nat Genet.* 1997;17:341–345.
42. Unger MA, Nathanson KL, Calzone K, et al. Screening for genomic rearrangements in families with breast and ovarian cancer identifies BRCA1 mutations previously missed by conformation-sensitive gel electrophoresis or sequencing. *Am J Hum Genet.* 2000;67:841–850.
43. Hendrickson B, Deffenbaugh AM, Gaglio CA, et al. Prevalence results for five recurrent *BRCA1* rearrangement mutations in 7570 analyses. *Am J Hum Genet.* 2003;73(5):487.
44. Frolov A, Prowse AH, Vanderveer L, Bove B, Wu H, Godwin AK. DNA array-based method for detection of large rearrangements in the BRCA1 gene. *Genes Chromosomes Cancer.* 2002;35:232–241.
45. Gad S, Scheuner MT, Pages-Berhouet S, et al. Identification of a large rearrangement of the BRCA1 gene using colour bar code on combed DNA in an American breast/ovarian cancer family previously studied by direct sequencing. *J Med Genet.* 2001;38:388–392.
46. Szabo CI, King MC. Population genetics of BRCA1 and BRCA2. *Am J Hum Genet.* 1997;60:1013–1020.
47. Roa BB, Boyd AA, Volcik K, Richards CS. Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2. *Nat Genet.* 1996;14:185–187.
48. Kauff ND, Perez-Segura P, Robson ME, et al. Incidence of non-founder BRCA1 and BRCA2 mutations in high risk Ashkenazi breast and ovarian cancer families. *J Med Genet.* 2002;39:611–614.
49. Thorlacius S, Sigurdsson S, Bjarnadottir H, et al. Study of a single BRCA2 mutation with high carrier frequency in a small population. *Am J Hum Genet.* 1997;60:1079–1084.
50. Moller P, Heimdal K, Apold J, et al. Genetic epidemiology of BRCA1 mutations in Norway. *Eur J Cancer.* 2001;37:2428–2434.
51. Einbeigi Z, Bergman A, Kindblom LG, et al. A founder mutation of the BRCA1 gene in Western Sweden associated with a high incidence of breast and ovarian cancer. *Eur J Cancer.* 2001;37:1904–1909.
52. Menkiszak J, Gronwald J, Gorski B, et al. Hereditary ovarian cancer in Poland. *Int J Cancer.* 2003;106:942–945.
53. Olopade OI, Fackenthal JD, Dunston G, Tainsky MA, Collins F, Whitfield-Broome C. Breast cancer genetics in African Americans. *Cancer.* 2003;97(suppl):236–245.
54. Deffenbaugh AM, Frank TS, Hoffman M, Cannon-Albright L, Neuhausen SL. Characterization of common BRCA1 and BRCA2 variants. *Genet Test.* 2002;6:119–121.
55. Fleming MA, Potter JD, Ramirez CJ, Ostrander GK, Ostrander EA. Understanding missense mutations in the BRCA1 gene: an evolutionary approach. *Proc Natl Acad Sci U S A.* 2003;100:1151–1156.
56. Petrucelli N, Lazebnik N, Huelsman KM, Lazebnik RS. Clinical interpretation and recommendations for patients with a variant of uncertain significance in BRCA1 or BRCA2: a survey of genetic counseling practice. *Genet Test.* 2002;6:107–113.
57. Lynch HT, Snyder CL, Lynch JF, Riley BD, Rubinstein WS. Hereditary breast-ovarian cancer at the bedside: role of the medical oncologist. *J Clin Oncol.* 2003;21:740–753.
58. Narod SA. Modifiers of risk of hereditary breast and ovarian cancer. *Nat Rev Cancer.* 2002;2:113–123.
59. King MC, Marks JH, Mandell JB. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. *Science.* 2003;302:643–646.
60. Scheuer L, Kauff N, Robson M, et al. Outcome of preventive surgery and screening for breast and ovarian cancer in BRCA mutation carriers. *J Clin Oncol.* 2002;20:1260–1268.
61. Muhr D, Wagner T, Oefner PJ. Polymerase chain reaction fidelity and denaturing high-performance liquid chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2002;782:105–110.

Chapter 18

Familial Adenomatous Polyposis and Turcot and Peutz-Jeghers Syndromes

Holly L. Neibergs and Amy T. Massey

FAMILIAL ADENOMATOUS POLYPOSIS

Familial adenomatous polyposis (FAP) is an autosomal dominantly inherited disorder that predisposes affected individuals to colon cancer through the early development of hundreds to thousands of adenomatous polyps (Figure 18-1). Florid polyposis throughout the colon will develop in 50% of affected individuals by age 16, and 95% will have polyposis by age 35.[1] If left untreated, colorectal cancer is inevitable in those with FAP, with an average age at diagnosis of 39 years. The incidence of FAP is estimated to be 1 in 8,300 to 1 in 14,025 live births and represents less than 1% of all colon cancers.[1] FAP is clinically diagnosed when an individual has greater than 100 colorectal adenomatous polyps.

Individuals affected with FAP may exhibit a host of benign extracolonic maladies. Gastric lesions identified in individuals with FAP are generally of two types: fundic gland hamartomatous polyps (occur in approximately 50% of cases) or adenomatous polyps (found in approximately 10%). The risk for gastric cancer from both types of polyps is small. Adenomatous polyps of the duodenum are much more common (approximately 90% of FAP patients). Adenomatous polyps of the periampullary region occur in 50% of cases, and may result in obstruction of the pancreatic duct, leading to pancreatitis. Visualization of periampullary polyps may be difficult due to their size, and so the use of a side-viewing endoscope is recommended.

Benign extraintestinal manifestations of FAP include cutaneous lesions (lipomas, fibromas, and sebaceous and epidermoid cysts), desmoid tumors, osteomas, dental abnormalities, congenital hypertrophy of the retinal pigment epithelium (CHRPE), and nasopharyngeal angiofibromas.[1] Desmoid tumors are identified in approximately 10% of individuals with FAP and result from a proliferation of myofibroblasts. The vast majority of desmoids arise after surgery and are present in the small intestine mesentery, retroperitoneum, and abdominal wall. Desmoids are usually slow growing but may become large enough to compress or obstruct organs, resulting in morbidity or mortality for approximately 5% of patients. Mesenteric and retroperitoneal desmoids often are not amenable to surgical or medical therapy, although some desmoid tumors have been successfully treated with antiestrogen medications. Desmoid tumors and duodenal carcinoma are the most common cause of death in FAP patients following prophylactic colectomy. Osteomas are frequently seen in the long bones, mandible, and skull. Dental abnormalities have been reported in approximately 17% of individuals with FAP and may involve unerupted teeth, absence of teeth, or supernumerary teeth. CHRPE are flat pigmented lesions of the retina believed to be present at birth. CHRPE is a benign condition that may be detected by fundoscopic examination.

Extracolonic malignancies may be present in individuals with FAP. Duodenum, periampullary area, stomach, pancreas, papillary thyroid, liver, and brain carcinomas are known to occur with greater frequency in FAP patients. Adenocarcinomas of the bile ducts and adrenal glands also have been reported, but the risk of these cancers in individuals with FAP remains low. Carcinomas of the small intestine, predominantly of the duodenum or periampullary area, are seen in 4% to 12% of FAP patients. Gastric adenocarcinoma occurs in 0.5% of FAP patients. Pancreatic and papillary thyroid carcinomas are present in approximately 2% of individuals with FAP. One in 150 children under the age of 5 years with a mutation for FAP will develop hepatoblastoma.[1] Tumors of the central nervous system (most commonly medulloblastoma, as seen in Turcot syndrome) occur in less than 1% of individuals with FAP. Individuals with FAP who also have an osteoma, fibroma, or epidermoid cyst have, in the past, been diagnosed with Gardner syndrome. However, since the identification of *APC* germline mutations in both FAP and Gardner syndrome, it is now recognized that FAP and Gardner syndrome are the same entity with different clinical presentations.

Attenuated FAP is characterized by an average of 30 to 100 cumulative colonic adenomas, although the number of adenomatous polyps varies greatly. Attenuated FAP

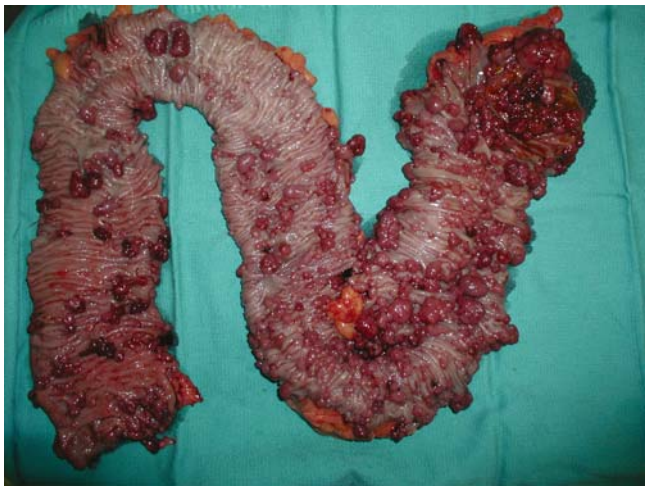


Figure 18-1. Adenomatous polyps present in the colon of an individual affected with familial adenomatous polyposis. (Courtesy of Dr. Wayne B. Tuckson.)

patients also have germline mutations in the *APC* gene, but the location of the mutations often differs from those of classic FAP.[2] Polyps tend to present at a later age and in a more proximal location in individuals with attenuated FAP. There may be considerable difficulty in distinguishing attenuated FAP from hereditary nonpolyposis colorectal cancer (HNPCC) syndromes based on clinical criteria alone. Endoscopic surveillance may be more challenging in individuals with attenuated FAP because the adenomas tend to be more flat and plaquelike as opposed to polypoid.[3] The identification of these lesions may require stains, such as indigo carmine, during colonoscopy for adequate visualization. Colon cancer risk remains high, but with a 10 to 15 year later age of onset than observed in classic FAP. Adenomas and carcinomas of the upper gastrointestinal tract also may be present.[4] Attenuated FAP is clinically diagnosed when an individual presents with numerous colonic adenomatous polyps under the age of 60 years. Extracolonic features characteristic of FAP may aid in the diagnosis but are not required.

Molecular Basis of Disease

The identification of the *APC* (adenomatous polyposis coli) gene, responsible for FAP, was greatly facilitated by a patient with polyposis and an interstitial deletion of 5q visible by cytogenetic analysis. Subsequent linkage analysis confirmed the localization of the FAP gene to 5q21. Individuals affected with FAP contain one germline *APC* mutation and acquire an additional somatic mutation of the normal (wild-type) allele during their lifetime. Additional somatic mutations must occur before transformation occurs. Approximately one third of all FAP cases are due to de novo mutations with no family history of polyposis or colorectal cancer.

The *APC* gene has several alternatively spliced forms. When translation begins at the first methionine, a 2,843 amino acid protein results. The majority (75%) of the gene is contained in a 6,579 base pair (bp) open reading frame within exon 15. The APC protein functions in cell adhesion, signal transduction, and transcriptional activation,[2] and is localized to the nucleus and membrane/cytoskeleton in human epithelial cells. It binds to GSK3β, β-catenin, γ-catenin, tubulin, and EB1. β-catenin functions in cell adhesion by binding to E-cadherin. The cytoplasmic level of β-catenin is suppressed by APC-mediated degradation, preventing activation of oncogenes, such as MYC and cyclin D1 by a T cell transcription factor (Wingless-*Wnt* signal mediator). Failure of the APC protein to maintain low levels of β-catenin results in activation of these growth-promoting genes. APC also accumulates at the kinetochore during mitosis, contributes to kinetochore-microtubule attachment, and may play a role in chromosome segregation.

An extensive number of FAP families have been examined for mutations in the *APC* gene. Most deleterious mutations cause loss or disruption of the central β-catenin binding region of the APC protein. Of the FAP families studied, approximately 80% have been demonstrated to contain intragenic mutations that result in the truncation of the APC protein. Truncating mutations were the result of nonsense mutations (33%), small insertions (6%), and deletions (55%), largely in the first half of exon 15.[5] Several hundred mutations of the *APC* gene have been reported, with the two most common mutations found in codons 1061 and 1309. Individuals with mutations in codon 1309 tend to exhibit a great number of colonic adenomas at an earlier age.

A missense mutation at codon 1307 (I1307K) of *APC* is found in the Ashkenazi Jewish population. Individuals with the I1307K mutation have an estimated odds ratio of 1.85 or a 10% to 20% lifetime risk of developing colon cancer. This mutation results in a lysine substitution of an isoleucine and is believed to increase the risk of colorectal cancer by creating a hypermutable site. Of Ashkenazi Jewish families with a family history of colorectal cancer, 28% are estimated to have this variant. I1307K is found in approximately 6% of the Ashkenazi Jewish population but is extremely rare in the general population. This mutation does not result in the polyposis phenotype. Colorectal cancers associated with I1307K are identical to sporadic colorectal tumors.

A second missense variant in the *APC* gene has been described that may be associated with a predisposition to colon cancer. A consensus has not been reached regarding the role of the E1317Q mutation in colon cancer or adenomas. At this time, clinical mutation testing specific for E1317Q is not available.

Clinical Utility of Testing

Genetic testing is the standard of care for management of individuals with FAP.[1] Diagnostic testing establishes the FAP genotype in the family and provides the opportunity

for conclusive genetic testing for at-risk asymptomatic family members. Early detection of FAP and attenuated FAP may allow for early intervention and prevention of colorectal cancer. The use of genetic testing for recognition of presymptomatic family members improves their diagnosis of FAP and circumvents costly colorectal cancer surveillance in those who have not inherited the FAP-associated mutation. Not only is genetic testing more cost effective than repeated sigmoidoscopy screening for at-risk family members, but life expectancy of individuals with FAP is extended as well. Germline mutations in the Mut Y homologue (*MYH*) gene also may be evaluated concurrently with *APC* mutations due to the overlap of phenotypes (see Interpretation of Test Results below).

Genetic testing is appropriate for presymptomatic at-risk family members of an individual diagnosed with FAP, and is ideally offered in adolescence, prior to the age when colon screening by sigmoidoscopy or colonoscopy would begin. Prenatal testing also is available for FAP. DNA extracted from fetal cells obtained by amniocentesis or chorionic villus sampling may be used for genetic testing. Requests for prenatal diagnosis of adult-onset diseases remain controversial, and provide a challenging situation that mandates careful and thoughtful genetic counseling.

APC mutations are less frequently identified in those diagnosed with FAP without a family history than in those individuals with a family history.[6] Children of an affected individual have a 50% chance of inheriting their parent's mutation and should be offered genetic testing at an appropriate age. Those that test negative for an *APC* mutation (within the context of a known mutation in the family) remain at the general population risk for colorectal cancer and may forego colorectal cancer screening until age 50. Those testing positive will require heightened cancer surveillance and prevention at an earlier age.

Genetic testing may be offered to individuals of Ashkenazi Jewish ancestry who have a personal or family history of colon cancer or polyps, or those that have a heightened concern for colon cancer. Genetic testing for I1307K in the Ashkenazi Jewish population has been met with some controversy, as the value of the testing may be limited if the management of individuals positive for the variant is unchanged. A good proportion of the individuals with this variant have a close relative affected with colon cancer or personal risk factors that already increase the need for colorectal cancer screening regardless of whether the person carries the I1307K variant.

Identification of the location of the mutation in the *APC* gene may provide clinical insight into the management and surveillance of FAP patients. FAP manifestations may be associated with the location of the mutation within the *APC* gene. Mutations associated with classic FAP are found from codon 168 to codon 1600, in which the full range of phenotypes may be seen. Truncating mutations between codons 1403 and 1578 are associated with extracolonic manifestations such as desmoid tumors and osteomas, but not CHRPE.[5,7] CHRPE-associated mutations are associated with mutations between codons 463 and 1387. A particularly severe form of FAP with several thousand colonic polyps is associated with mutations between codons 1350 and 1464. Desmoid tumors are more frequently (61%) associated with mutations between codons 1444 and 1580 than mutations prior to codon 1444.[8]

In attenuated FAP, mutations of the *APC* gene are generally identified in three distinct regions. The first region is at the 5′ end of the *APC* gene and extends to codon 157 in exon 4. The second region consists of an alternatively spliced area of exon 9, and the third region begins at the distal 3′ end of the *APC* gene beyond codon 1595 in exon 15.[2] Attenuated FAP patients with mutations at the amino terminus demonstrate the greatest variation in the number of colorectal adenomas. In addition, a more severe phenotype of upper gastrointestinal lesions is seen in these individuals than in those with mutations in either exon 9 or the 3′ region of the *APC* gene. Mutations in the extreme 5′ end of the *APC* gene may result in a nearly full-length functional protein by the initiation of translation downstream of the mutated site.[9] These proteins appear to have the ability to downregulate both β-catenin and Wnt signaling, and to induce cell cycle arrest, suggesting that the amino terminus may not be required for APC function. Although genotype-phenotype correlations aid in predicting possible clinical manifestations, these correlations are not certain.

Available Assays

Protein truncation assays (see chapter 2) of the *APC* gene have been commonly used as a commercial test for mutations in FAP families. This assay capitalizes on the frequency of truncating mutations as a cause of classic FAP. An advantage of this method is that a mutation resulting in a truncated protein is almost certainly deleterious. One drawback of this assay is that it is not reliable in detecting *APC* mutations in the extreme 5′ or 3′ ends of the gene. Missense variants and alterations outside the coding region are not detected with this technique. Protein truncation assays do not provide information regarding the specific mutation responsible for FAP in the family, so direct sequencing of the truncated region is necessary to identify the specific mutation.

Linkage analysis is used to identify FAP carriers. Linkage analysis requires the participation of affected and unaffected family members, and provides a clinical testing alternative for families that do not have *APC* mutations identifiable by protein truncation or by direct DNA sequencing. Linkage analysis is informative in 90% to 95% of families who are able to submit the appropriate samples for testing. However, not all families have a sufficient number of members who are willing or able to participate in genetic testing for linkage analysis to be informative. Additionally, individuals with de novo mutations may not

be aided by linkage analysis, unless their children become affected. Linkage analysis also may prove problematic if the markers used are not informative (i.e., are homozygous) in the affected and unaffected family members.

Direct sequence analysis of the *APC* coding, flanking, and intronic regions is available. This method has the advantage of directly identifying the specific familial mutation. Mutations may be detected by automated methods utilizing gel or capillary electrophoresis. While DNA sequencing is the most comprehensive genetic test, it may miss mutations due to rearrangements or large insertions or deletions, or that result in altered expression. Sequencing also will reveal variants of uncertain clinical significance, which go undetected when indirect methods such as protein truncation and linkage analysis are used. Denaturing high-performance liquid chromatography (DHPLC), single-strand conformation polymorphism (SSCP) and denaturing gradient gel electrophoresis (DGGE) are mutation screening methods that may be used in conjunction with sequencing for the detection of an *APC* mutation.

To detect large rearrangements in the *APC* gene, a quantitative polymerase chain reaction (PCR) assay may be used to amplify all exons of the APC gene. Genes that exhibit rearrangements will differ in the quantity of PCR product produced. Gene regions showing differences in the quantity of amplified PCR products are subsequently confirmed by additional methods such as long-range PCR with the use of intragenic single-nucleotide polymorphic (SNP) marker analysis, or Southern blot analysis.

Genetic testing for the I1307K variant is available by allele-specific oligonucleotide hybridization (ASOH) or by sequencing.

Interpretation of Test Results

The sensitivity or accuracy of genetic tests varies by the method used. Common genetic testing methods for *APC* mutation detection include protein truncation, linkage analysis, sequencing, and ASOH. Protein truncation assays will detect a mutation in approximately 80% of FAP families. When used in combination with other mutation detection methods such as SSCP, the mutation detection rate may improve to 90%. In linkage analysis, the linkage of a specific haplotype with FAP in a family provides 95% to 99% accuracy in identifying carriers. The sensitivity for DNA sequencing is reported to range from 90% to 99%. The sensitivity of genetic analysis for the I1307K variant of the *APC* gene by ASOH or sequencing approaches 100%.

Unfortunately, not all individuals that meet the clinical criteria for FAP have an identifiable mutation. In approximately 30% of affected families, an identifiable *APC* mutation is not found. This may be the result of large rearrangements of the *APC* gene, an upstream mutation that reduces the expression of the gene, or perhaps other susceptibility loci such as *MYH*. Many individuals with polyposis that test negative for mutations in *APC* may have an autosomal recessive disorder caused by the inheritance of mutations in the *MYH* gene.[10] For those with multiple adenomas (15 to 100), about one third will have biallelic mutations in *MYH*. Due to the high incidence of *MYH* mutations in individuals with polyposis, *MYH* often is evaluated in conjunction with *APC* genetic testing.

For classic FAP, DNA testing usually is suggested for children at 10 years of age or older, as colon surveillance is recommended to begin at 10 to 12 years of age.[1] For individuals at risk for attenuated FAP, colon screening should begin at 18 to 20 years of age and so genetic testing should be offered at that time. There is some debate as to the adequacy of using flexible sigmoidoscopy or if full colonoscopy is required for surveillance. Prophylactic colectomy should be performed to prevent colon cancer after adenomas are identified, but controversy remains regarding the timing and extent of the procedure (ileorectal anastomosis or ileoanal pouch anastomosis). If a rectal segment remains after surgery, routine endoscopic screening is required on a biannual basis. The use of oral nonsteroidal antiinflammatory Cox-2 inhibitors to prevent or eliminate polyps in the retained rectum has been reported to be effective in the short term but has not been recommended as a means of primary care.[1]

In addition to colorectal surveillance, patients need to be examined for extracolonic manifestations throughout their life at regular intervals, before and after surgery. Upper endoscopic screenings of the stomach, duodenum, and periampullary region is suggested every 6 months to 3 years. Cox-2 inhibitors have been reported to regress duodenal adenomas in some, but not all studies.[1] Small bowel x-ray is suggested every 1 to 3 years when duodenal adenomas are detected or before colectomy. Annual examination of the thyroid is warranted and ultrasonography may be considered. Children under the age of 5 years whose parents are affected with FAP may benefit from screening of serum alpha-fetoprotein levels and imaging of the liver to detect hepatoblastoma. Detection of this rapidly growing, often fatal tumor at an early stage provides the best opportunity for cure. For families with FAP but without an identified *APC* mutation, members must participate in intensive, long-term colorectal cancer surveillance.

Laboratory Issues

Genetic counseling is recommended before genetic testing is ordered (chapter 3). Referral of individuals affected or at risk for FAP to a genetic counselor or hereditary cancer center that routinely tests FAP patients is recommended.

TURCOT SYNDROME

Molecular Basis of Disease

Turcot syndrome is a rare complex genetic disorder that results in primary colonic neoplasms and concurrent primary malignant tumors of the central nervous system. This syndrome has been reported as a dominant and a recessive disorder as homozygous and compound heterozygous mutations in *APC*, *MLH1*, *MSH2*, and *PMS2* have been described. *MSH6* also may be involved, as one individual with homozygous *MSH6* mutations was reported with Turcot syndrome.[11]

Turcot syndrome represents an intricate interplay between the characteristics of the germline mutation and additional somatic events in target tissues. Central nervous system tumors associated with this syndrome are typically either a medulloblastoma or a glioblastoma multiforme. Turcot syndrome was at one time thought to be a separate disease entity, but now it is known that two thirds of the cases are associated with mutations in the *APC* gene, while one third are associated with mutations in the mismatch repair genes, *MLH1*, *MSH2*, *MSH6*, or *PMS2*.

Individuals with Turcot syndrome who present with a medulloblastoma, astrocytoma, or ependymoma tend to have a germline mutation in the *APC* gene.[12] Polyposis and extracolonic manifestations are often present. Approximately 80% of families with Turcot syndrome and polyposis will have a detectable *APC* mutation. Polyposis may be identified before or after the manifestation of the brain tumor. The age of onset and type of brain tumor may differ greatly even between affected family members. The medulloblastomas and colorectal cancers in these families do not exhibit microsatellite instability. As is the case with FAP, the *APC* mutations encode protein products that are truncated. Analysis of individuals with FAP and brain tumors suggests that mutations in codons 457 to 1309 of the *APC* gene may account for the majority of cases.[12,13] It is not understood why some individuals will develop brain tumors and others will not when the mutations in the *APC* gene are similar. The lifetime relative risk for those affected with FAP to develop medulloblastoma is estimated to be 92 times that of the general population, but the absolute lifetime risk remains below 1%.[12]

Individuals who present with glioblastomas and colorectal tumors likely have mismatch repair gene germline mutations in *MLH1*, *MSH2*, or *PMS2*. Café au lait spots also may be seen. Both the glioblastomas and colorectal tumors may exhibit microsatellite instability that characterizes HNPCC (see chapter 19). This phenomenon is characteristic of an inability to repair interstrand mismatch errors during DNA replication and often is associated with defects in the mismatch repair genes. Microsatellite instability is more prevalent or present to a greater degree in colorectal cancers than in central nervous system cancers.[12,14] Examination of repetitive sequences of coding and 5′ upstream regions of several loci revealed frequent alterations that increase in prevalence with the advancement of colorectal cancer according to Dukes categories.[15] The colorectal adenomas in patients with *MLH1*, *MSH2*, or *PMS2* mutations may not be clinically distinguishable from those resulting from an *APC* mutation. Large numbers of colorectal adenomas may be present, and these individuals may not meet the clinical criteria for HNPCC.

The tumorigenic pathway leading to colorectal cancer and glioblastomas in mismatch repair genes that are associated Turcot syndrome has been shown to be somewhat heterogeneous. Leung et al. reported that colorectal carcinomas frequently contained somatic frameshift mutations of *TGFBRII*, while glioblastomas did not.[14] Microsatellite instability–high glioblastomas overexpress the TP53 protein, whereas colorectal carcinomas had normal TP53 levels. Glioblastomas exhibit chromosomal instability with aneuploid DNA content, but colorectal tumors are diploid or contain near-diploid DNA content.[14] Loss of 9p is frequently observed in glioblastomas, but 9p is not commonly lost in colorectal cancers. These differences suggest that the underlying germline mutation present in mismatch repair genes may facilitate malignant processes in the central nervous system that are divergent from those in the colon.[14]

Clinical Utility of Testing

Diagnostic and presymptomatic genetic testing is available for Turcot syndrome. Genetic testing of individuals affected with Turcot syndrome may help clarify their diagnosis. The number of polyps, type of brain cancer, and occurrence of extracolonic cancers aid in directing the choice of genetic testing of the *APC* gene or the mismatch repair genes. For individuals choosing to proceed with mutation analysis, genetic testing may establish the disease genotype in a family and provide the opportunity for informative presymptomatic testing for other family members. Carriers may be offered intense surveillance of appropriate organs, and noncarriers may feel a relief of anxiety and stress. Microsatellite instability testing may be used from tumor specimens, although not all tumors in individuals with a DNA mismatch repair gene mutation will present with DNA replication errors. The benefits and limitations of genetic testing would be the same as those discussed for FAP and HNPCC.

Available Assays

Assays available are the same as those described for FAP (*APC* gene) and HNPCC (for *MLH1* and *PMS2)* in chapter 19.

Interpretation of Test Results

Approximately 80% of Turcot syndrome families with polyposis will have a detectable *APC* mutation, similar to that of FAP. Turcot syndrome families associated with

tumors with replication errors or HNPCC have a lower mutation detection rate in the DNA mismatch repair genes.

PEUTZ-JEGHERS SYNDROME

Peutz-Jeghers syndrome is a rare autosomal dominant disorder characterized by mucocutaneous melanin pigmentation, gastrointestinal hamartomatous polyposis, and benign and malignant tumors of the gastrointestinal tract, breast, ovary, cervix, and testis. The incidence of Peutz-Jeghers syndrome is unknown, but estimates range from 1 in 8,300 to 200,000 live births.[16,17] An estimated 25% of Peutz-Jeghers cases are not familial.

Hamartomas with a smooth muscle core are the pathognomonic feature of Peutz-Jeghers (Figure 18-2a and b). Peutz-Jeghers polyps exhibit a central branching smooth muscle cell structure that extends to the lamina propria and is covered with epithelial cells of normal appearance.[17] Polyposis usually becomes symptomatic early in adolescence, though intestinal obstruction has been reported in infancy.[16] Hamartomatous polyps may be present throughout the entire gastrointestinal tract, but are commonly present in the small intestine (90%) or the colon and rectum (33%). Polyps may be of mixed histologic types (hyperplastic, adenomatous) but are mostly hamartomatous and may number from one to dozens.

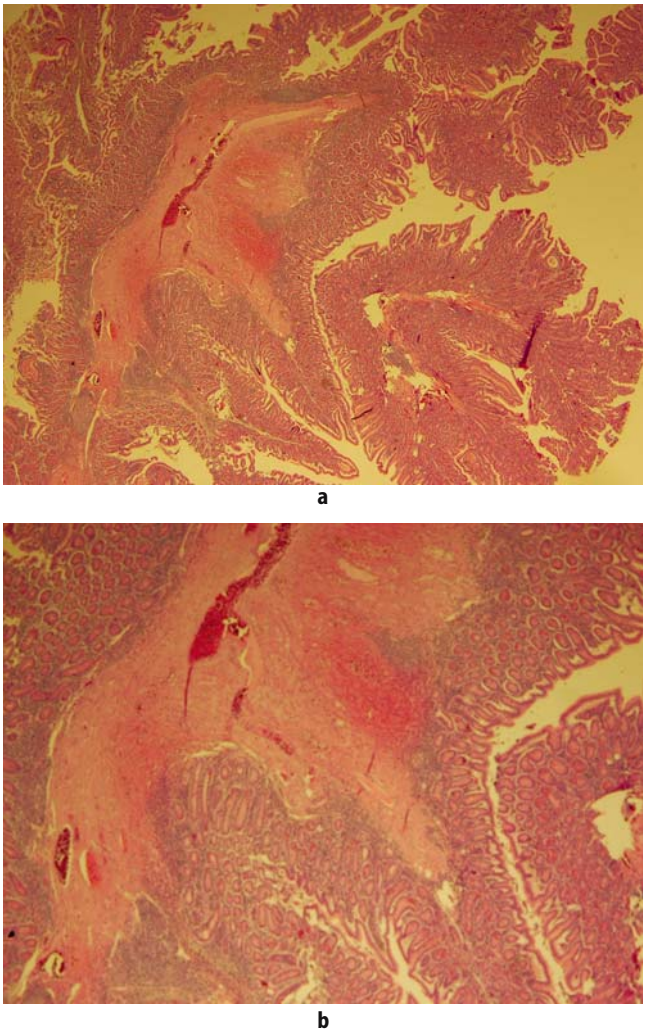


Figure 18-2. (a) The harmatomatous polyp of a Peutz-Jeghers patient is distinguished by its smooth branching muscle core and normal-appearing epithelium that covers its surface. (b) Magnified view of a harmatomatous Peutz-Jegher polyp. (Courtesy of Dr. Catherine Sewell.)

Melanin pigmentation may develop in Peutz-Jeghers patients under the age of 5 years as dark blue to dark brown macules around the mouth, eyes, nostrils, perianal area, buccal mucosa, hands, feet, or axillary areas. Pigmentation may vary greatly between affected individuals of a family and between families. If pigmentation does develop, it may fade or disappear with age. The presence of pigmentation is helpful in the diagnosis of Peutz-Jeghers but may not be used as a sole clinical indication, as similar pigmentation may be found in up to 15% of the general population.[17]

The clinical presentation of Peutz-Jeghers is variable. The most common presentations of Peutz-Jeghers are small bowel intussusception, colon obstruction, and gastrointestinal bleeding. Severe digestive obstruction may be identified in affected individuals throughout life, but the average age of onset is 29 years.[18] Affected individuals also may exhibit intermittent abdominal pain or rectal prolapse. For an individual to be clinically diagnosed with Peutz-Jeghers, one of the following criteria must be met: (*a*) three or more histologically confirmed Peutz-Jeghers polyps, (*b*) any number of Peutz-Jeghers polyps with a family history of Peutz-Jeghers, (*c*) characteristic mucocutaneous pigmentation with a family history of Peutz-Jeghers, or (*d*) any number of Peutz-Jeghers polyps and characteristic mucocutaneous pigmentation.[15]

The estimated incidence of cancer among Peutz-Jeghers patients is 18-fold higher than in the general population.[19] No correlation has been found with the incidence of cancer and the severity of polyposis or the presence of pigmentation.[17] Cancer is most frequently identified in the colon (40%), small intestine (13%), stomach (29%), pancreas (36%), lung (15%), breast (54%), ovary (21%), and uterus (9%).[19] A rare form of cervical cancer that is seen more frequently in women with Peutz-Jeghers is adenoma malignum. This is an aggressive form of cervical cancer that originates from glandular cells of the cervix.[16]

Benign ovarian sex-cord tumors with annular tubules and estrogen-producing ovarian tumors resulting in precocious puberty are found more commonly in females affected with Peutz-Jeghers.[16] Two cases of Sertoli cell tumor of the ovary were reported in sisters with the syndrome. In young males with Peutz-Jeghers, estrogen-producing testicular Sertoli cell tumors can occur and may lead to precocious puberty and gynecomastia.

Molecular Basis of Disease

Germline mutations in *LKB1* (also known as *STK11*) on 19p13.3 were initially identified through positional cloning. *LKB1* mutations have subsequently been identified in most, but not all, Peutz-Jeghers families, suggesting that additional loci may be involved.[17] Families that do not harbor an *LKB1* mutation do not differ by clinical manifestations or ethnicity from families with *LKB1* mutations. Individuals with nonfamilial (de novo) Peutz-Jeghers often display the same germline *LKB1* mutations as those found in individuals with a family history of this syndrome.[17]

LKB1 functions as a tumor suppressor gene whose protein physically associates with TP53 to regulate TP53-dependent apoptosis pathways.[20,21] *LKB1* also interacts with PTEN, which is responsible for other hereditary hamartoma syndromes, and STRAD and MO25α, which regulate epithelial polarity.[22,23] *LKB1* encodes a 433 amino acid protein that is present in both the cytoplasm and the nucleus. The kinase domain resides within codons 50 to 337 of *LKB1* and shares limited homology with other serine-threonine and tyrosine kinase family members. Most germline mutations in *LKB1* result in protein truncation; however, rare missense mutations within the kinase domain have been observed.[21]

Inactivation of this tumor suppressor gene is a critical early event in the development of hamartomas and adenocarcinomas.[24] Adenocarcinomas in patients with Peutz-Jeghers syndrome demonstrate altered TP53 expression and loss of heterozygosity (LOH) in 17p and 18q. Microsatellite instability, LOH near the *APC* gene, or *KRAS* mutations were identified in these tumors.[24]

Clinical Utility of Testing

Genetic testing is useful to confirm the diagnosis of Peutz-Jeghers syndrome or to identify presymptomatic first-degree relatives. Diagnostic and presymptomatic testing is offered from buccal swabs or blood specimens. Prenatal testing is available using chorionic villus samples or fresh or cultured amniocytes.

Cancer screenings for affected and at-risk individuals vary from the general population. Colonoscopy should be conducted every 1 to 2 years beginning in adolescence. Identification of a Peutz-Jeghers polyp requires routine endoscopic surveillance every 1 to 2 years throughout the life of the individual. Upper gastrointestinal endoscopy, small intestinal double-contrast radiology, or push enteroscopy also has been suggested at 2-year intervals due to the risk for small bowel cancers.[18]

Men with Peutz-Jeghers syndrome should perform regular testicular examinations (beginning at age 10 years) that should be followed by ultrasound if a mass or clinical symptoms occur.[18] It is recommended that women with this syndrome perform monthly self breast examinations and undergo annual clinical breast examinations beginning at age 18 to 20 years. Mammography should ensue at age 25 to 35 years and be conducted every 2 years until age 40, at which time they can follow general population guidelines for breast cancer screening. Annual pelvic examinations and Pap smears should begin in the teens. Surveillance for ovarian cancer is complicated by a lack of efficient screening tools that avoid unnecessary laparoscopic surgeries, but nevertheless, annual abdominal and transvaginal ultrasound alone or in combination with CA-125 testing is recommended.[18] Although these surveillance practices are recommended, there have been no published reports proving their efficacy. Additional studies are needed to determine the appropriate cancer screening modalities in individuals with Peutz-Jeghers syndrome.

Available Assays

Complete DNA sequencing of *LKB1* is available for individuals with an unknown mutation. Benefits and limitations of DNA sequencing are the same as those discussed for sequencing for identification of *APC* mutations.

Interpretation of Test Results

The mutation identification rate for *LKB1* mutations in individuals with Peutz-Jeghers syndrome has ranged from 58% to 100%, depending on the study.[16] For affected families without an identifiable mutation in *LKB1*, interpretation of the test results must be made with caution. Several possible explanations exist for a negative test result in an individual with this syndrome. One possibility is that a locus other than *LKB1* is responsible for the syndrome in this individual, as families have been reported that exhibit linkage to regions other than 19p. Another possibility is that the testing method was not able to detect the mutation present in the family. For those families without an identifiable mutation, all first-degree relatives should participate in heightened cancer surveillance.

Site-specific testing should be offered to family members of affected individuals with a known *LKB1* mutation. Individuals that test positive for a deleterious mutation are at increased risk for cancer and should alter their cancer surveillance accordingly. For those who have not inherited the mutation, their risk for malignancy is that of the general population.

Laboratory Issues

Testing for Peutz-Jeghers syndrome is available from a number of laboratories following genetic counseling.

References

1. American Gastroenterological Association. AGA technical review on hereditary colorectal cancer and genetic testing. *Gastroenterology.* 2001;121:198–213.
2. Soravia C, Berk T, Madlensky L, et al. Genotype-phenotype correlations in attenuated adenomatous polyposis coli. *Am J Hum Genet.* 1998;62:1290–1301.
3. Lynch HT, Smyrk TC, Watson P, et al. Hereditary flat adenoma syndrome: a variant of familial adenomatous polyposis? *Dis Colon Rectum.* 1992;35:411–421.
4. Spirio L, Olschwang S, Groden J, et al. Alleles of the APC gene: an attenuated form of familial polyposis. *Cell.* 1993;75:951–957.
5. Kinzler KW, Vogelstein B. Colorectal tumors. In: Vogelstein B, Kinzler KW, eds. *The Genetic Basis of Human Cancer.* New York: McGraw-Hill; 2002:583–612.
6. Truta B, Allen BA, Conrad PG, et al. A comparison of the phenotype and genotype in adenomatous polyposis patients with and without a family history. *Fam Cancer.* 2005;4:127–133.
7. Moisio A-L, Jarvinen H, Peltomaki P. Genetic and clinical characterisation of familial adenomatous polyposis: a population based study. *Gut.* 2002;50:845–850.
8. Friedl W, Caspari R, Sengteller M, et al. Can APC mutation analysis contribute to therapeutic decisions in familial adenomatous polyposis? Experience from 680 FAP families. *Gut.* 2001;48:515–521.
9. Heppner Goss K, Trzepacz C, Tuohy TM, Groden J. Attenuated APC alleles produce functional protein from internal translation initiation. *Proc Natl Acad Sci USA.* 2002;99:8161–8166.
10. Sieber OM, Lipton L, Crabtree M, et al. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. *N Engl J Med.* 2003;348:791–799.
11. Hedge MR, Chong B, Blazo ME, et al. A homozygous mutation in MSH6 causes Turcot syndrome. *Clin Cancer Res.* 2005;11:4689–4693.
12. Hamilton SR, Liu B, Parsons RE, et al. The molecular basis of Turcot's syndrome. *N Engl J Med.* 1995;332:839–847.
13. Wallis YL, Morton DG, McKeown CM, Macdonald F. Molecular analysis of the APC gene in 205 families: extended genotype-phenotype correlation in FAP and evidence of the role of APC amino acid changes in colorectal cancer predisposition. *J Med Genet.* 1999;36:14–20.
14. Leung SY, Yuen ST, Chan TL, et al. Chromosomal instability and p53 inactivation are required for genesis of glioblastoma but not for colorectal cancer in patients with germline mismatch repair gene mutation. *Oncogene.* 2000;19:4079–4083.
15. Miyaki M, Iijima T, Shiba K, et al. Alterations of repeated sequences in 5′ upstream and coding regions in colorectal tumors from patients with hereditary nonpolyposis colorectal cancer and Turcot syndrome. *Oncogene.* 2001;20:5215–5218.
16. Boardman LA. Heritable colorectal cancer syndromes: recognition and preventive management. *Gastroenterol Clin North Am.* 2002;31:1107–1131.
17. Aaltonen L. Peutz-Jeghers syndrome. In: Vogelstein B, Kinzler KW, eds. *The Genetic Basis of Human Cancer.* New York: McGraw-Hill; 2002:337–341.
18. Hampel H, Peltomaki P. Hereditary colorectal cancer: risk assessment and management. *Clin Genet.* 2000;58:89–97.
19. Giardiello FM, Brensinger JD, Tersmette AC, et al. Very high risk of cancer in familial Peutz-Jeghers syndrome. *Gastroenterology.* 2000;119:1447–1453.
20. Ylikorkala A, Avizienyte E, Tomlinson IP, et al. Mutations and impaired function of LKB1 in familial and non-familial Peutz-Jeghers syndrome and a sporadic testicular cancer. *Hum Mol Genet.* 1999;8:45–51.
21. Karuman P, Gozani O, Odze RD, et al. The Peutz-Jegher gene product LKB1 is a mediator of p53-dependent cell death. *Mol Cell.* 2001;7:1307–1319.
22. Mehenni H, Lin-Marq N, Buchet-Poyau K, et al. LKB1 interacts with and phosphorylates PTEN: a functional link between two proteins involved in cancer predisposing syndromes. *Hum Mol Genet.* 2005;14:2209–2219.
23. Forcet C, Etienne-Manneville S, Gaude H, et al. Functional analysis of Peutz-Jeghers mutations reveals that the LKB1 C-terminal region exerts a crucial role in regulating both the AMPK pathway and the cell polarity. *Hum Mol Genet.* 2005;14:1283–1292.
24. Gruber SB, Entius MM, Petersen GM, et al. Pathogenesis of adenocarcinoma in Peutz-Jeghers syndrome. *Cancer Res.* 1998;58:5267–5270.

Chapter 19

Hereditary Nonpolyposis Colorectal Cancer

Erik C. Thorland and Stephen N. Thibodeau

Clinical Characteristics

Hereditary nonpolyposis colorectal cancer (HNPCC) is an autosomal dominant colon cancer syndrome. The first description of a cancer-prone family with HNPCC dates back to the late 1800s.[1] However, it was not until the work of Lynch in the 1970s that a more complete clinical picture of this disorder began to emerge.[2,3] The diagnosis of HNPCC has, until recently, been based primarily on family history. As a result, reliably differentiating patients with HNPCC from those with sporadic cancer has been difficult. However, the constellation of several clinical characteristics, in addition to family history, may raise suspicion of HNPCC.

One of the primary clinical characteristics of HNPCC is the early age at which tumors develop. The average age of colon cancer diagnosis in individuals with HNPCC is in the early to mid-40s, although many tumors may occur in the 20s or even in teenage years. This is in contrast to a mean age of diagnosis of the mid-60s for sporadic colorectal cancer (CRC).[4] In addition to CRC, several other tumor types are also observed at an increased frequency in families with HNPCC. These include endometrial, ovarian, gastric, small intestine, brain, and urinary and biliary tract tumors.[5] In addition, the Muir-Torre variant is associated with the typical extracolonic features of HNPCC in addition to sebaceous gland tumors and keratoacanthomas.[6]

Individuals with HNPCC also have a high rate of synchronous and metachronous tumors. Although HNPCC patients do not develop colon polyps more frequently than the general population, their polyps tend to be larger, to be more villous, and to contain more high-grade dysplasia.[7] These observations are consistent with an "aggressive adenoma" theory, in which, although polyps are not more common, they progress to carcinoma more frequently and more quickly than sporadic polyps.[8] Thus, the increased rate of tumor progression may be the major factor accounting for the cancer predisposition in HNPCC.

Several histologic features of the colon tumors are characteristic of a particular subgroup of HNPCC patients (those with defective DNA mismatch repair). These features include high histologic grade, mucinous or signet ring cell differentiation, medullary features, and a host immune response characterized by a Crohn's-like lymphoid reaction and an excess of tumor-infiltrating lymphocytes.[9] Although many of the histologic features of HNPCC would be considered aggressive (e.g., high grade and signet ring differentiation), HNPCC is paradoxically associated with a more favorable prognosis. In one study, the 5-year survival rate after CRC in individuals with HNPCC was reported to be 86% versus 59% for those with sporadic CRC.[10] Possible biologic explanations for the favorable prognosis include a decreased propensity of these tumors to metastasize, the high frequency of diploid DNA content, wild-type *TP53* status, and enhanced immune surveillance.[11–14]

In 1991, a set of guidelines (the Amsterdam criteria) was established to aid in the identification of families with HNPCC and to make the process of defining families for research purposes more uniform (Table 19-1a).[15] These criteria were based primarily on an extended family history of CRC and were instrumental in helping to identify the genes responsible for a majority of families with HNPCC. However, from a clinical perspective, several shortcomings of the Amsterdam criteria were subsequently identified. First, although many individuals with HNPCC have affected relatives in multiple generations, this is not always the case, because of small family size, variable penetrance, new mutations, or incomplete pedigree information. Second, these criteria did not consider the many extracolonic tumors frequently associated with HNPCC. Thus, use of the Amsterdam criteria for clinical purposes likely leads to an underdiagnosis of HNPCC. As a result, the true incidence of HNPCC and the underlying molecular defects have been difficult to establish. The estimates of the incidence of HNPCC vary widely, but population-based studies indicate that the diagnosis accounts for approximately 2% of all CRC.[16,17]

To address the limitations of the initial Amsterdam criteria, the Amsterdam criteria II[18] and the Bethesda guidelines[19] were proposed (Table 19-1b and c). These criteria

Table 19-1. Minimum Criteria for Clinical Diagnosis of HNPCC

a. Amsterdam criteria I
1. At least three relatives should have histologically verified CRC; one of them should be a first-degree relative to the other two.
2. At least two successive generations should be affected.
3. In one of the relatives colorectal cancer should be diagnosed before age 50.
4. Familial adenomatous polyposis should be excluded.

b. Amsterdam criteria II
1. At least three relatives should have an HNPCC-associated cancer (CRC; cancer of the endometrium, small bowel, ureter, or renal pelvis).
2. One should be a first-degree relative of the other two.
3. At least two successive generations should be affected.
4. At least one should be diagnosed before age 50.
5. Familial adenomatous polyposis should be excluded in the CRC case(s), if any.
6. Tumors should be verified by pathologic examination.

c. Bethesda criteria
1. Individuals with cancer in families that meet the Amsterdam criteria.
2. Individuals with two HNPCC-related cancers including synchronous and metachronous CRC, or associated extracolonic cancers.*
3. Individuals with CRC and a first-degree relative with CRC and/or HNPCC-related extracolonic cancer and/or a colorectal adenoma; one of the cancers diagnosed before age 45, and the adenoma before age 40.
4. Individuals with CRC or endometrial cancer diagnosed before age 45.
5. Individuals with right-sided CRC showing undifferentiated pattern on histopathology.†
6. Individuals with signet ring cell type CRC.‡
7. Individuals with adenomas diagnosed before age 40.

*Endometrial, ovarian, gastric, hepatobiliary, small bowel, transitional cell carcinoma of the renal pelvis or ureter.
†Solid cribriform defined as a poorly undifferentiated carcinoma composed of irregular solid sheets of large eosinophilic cells and containing small glandlike spaces.
‡Composed of more than 50% signet ring cells.

allow for a more liberal definition of HNPCC and include individuals with HNPCC-associated tumors as well as patients with early onset or multiple primary tumors, or both, whose family history is either incomplete or for whom the pedigree is small. These guidelines also include criteria based on the common histologic presentations of the tumors and make recommendations for germline DNA testing.

Molecular Basis

In 1993, three independent groups reported a novel type of genomic instability in a subset of colorectal tumors.[20–22] This instability was termed microsatellite instability (MSI) and was characterized by expansions or contractions in the number of tandem repeats at numerous microsatellite loci in the DNA. In these studies, MSI was identified in both a subset of sporadic colon tumors and nearly all tumors derived from patients diagnosed with HNPCC. Microsatellites are tandem repeats of a simple DNA sequence (1–6 base pairs [bp] in length) that are abundant and randomly distributed throughout the genome. They comprise approximately 3% of the human genome and are found, on average, every 2 kb.[23] These tandem repeats are thought to arise by slippage events during DNA replication. Slippage events have been frequent enough during human evolution that there are usually multiple alleles at any given microsatellite sequence.

The molecular basis for the observed instability and its role in HNPCC was subsequently elucidated in a remarkably short period of time due to two additional critical observations. First, investigators working in the field of yeast genetics recognized that the pattern of MSI observed in the colorectal tumors was similar to that seen in bacteria and yeast with defects in DNA mismatch repair (MMR).[24] This finding suggested that the genes involved in human MMR might be responsible for the MSI observed in a subset of CRC. The second important observation was the discovery of linkage in a subset of HNPCC families to chromosome 2p.[25] Linkage analysis, along with the clues provided by the yeast MMR work, subsequently led to the cloning of a human MutS homolog (*hMSH2*) and the identification of germline mutations in the same subset of HNPCC kindreds with linkage to chromosome 2p.[26,27] These critical observations set the stage for the tremendous amount of work accomplished over the past decade to gain a molecular understanding of HNPCC.

MMR is one of several mechanisms involved in the correction of mutations that occur as a result of exogenous or endogenous mutagens or misincorporation during DNA replication. This system is highly conserved through evolution, and much of what is known about MMR is based on work performed in bacteria and yeast (reviewed in Reference 28). MMR requires the concerted action of several proteins. In humans, at least six different proteins are directly involved in MMR. Deformities in the DNA helix caused by single base pair mismatches or insertion or deletion loop mutations are recognized and bound by a heterodimer called MutS consisting of hMSH2 and either hMSH6 or hMSH3. Each of these heterodimers has specific but partially redundant recognition specificity for the two different types of mutations. hMSH2/hMSH3 preferentially binds to insertion/deletion loops, while hMSH2/hMSH6 preferentially recognizes single base pair mismatches. A second heterodimer called MutL, composed of hMLH1 and either hPMS2 or hMLH3 (or possibly hPMS1), is responsible for differentiating between the template and newly synthesized strand (destined for repair) and for coordinating the interplay between the recognition complex and other proteins necessary for MMR. The hMLH1/hPMS2 heterodimer provides the bulk of this function, while hMLH1/hMLH3 plays a minor role. Following mismatch recognition and assembly of the repair proteins, the error-containing strand is excised, resynthesized, and relegated to complete the repair process (reviewed in Reference 29).

With the identification of the human MMR genes, it has become apparent that a significant fraction of clinically

defined HNPCC is due to alterations in these genes. In mutation carriers, MMR function is phenotypically normal, although every somatic cell carries one inactivated and one normal copy of the MMR gene mutated in the specific HNPCC family. However, cells in individuals who are heterozygous for a mutation in an MMR gene are susceptible to loss of the wild-type allele through large-scale chromosomal losses or deletions manifesting as loss of heterozygosity [LOH] or other somatic mutations within the gene. Inactivation of both alleles leads to the loss of functional MMR and a hypermutable phenotype characterized by a large increase in single-base changes and insertion or deletion mutations at microsatellite loci and other loci.[30,31] This phenotype is observed in the laboratory as MSI, which is an outcome of defects in the repair process, not of increased replication errors. The resulting hypermutable cell, however, is also susceptible to the accumulation of mutations in cellular protooncogenes and tumor suppressor genes, conferring a growth advantage that results in clonal expansion and drives the oncogenic process. It is important to note that the vast majority of microsatellites do not occur within critical regions of genes. Therefore, instability at these noncoding loci does not appear to have a phenotypic effect. However, MSI provides a very useful phenotypic marker for identifying those tumors with defective MMR.

Since the initial cloning of *hMSH2* and the discovery of germline mutations in a subset of families with HNPCC, the human homologues of most, if not all, MMR genes have been cloned and characterized. In addition to *hMSH2*, germline mutations have been identified in four other MMR genes in families with HNPCC: *hMLH1*, *hMSH6*, *hPMS1*, and *hPMS2*.[32–35] Mutations also have been observed in *hMLH3*, but the putative pathogenic role of these mutations remains to be determined.[36]

Hundreds of germline mutations have been reported and are cataloged in two databases (http://www.insight-group.org/ and http://www.med.mun.ca/MMRvariants/). Germline mutations in *hMSH2* and *hMLH1* account for approximately 40% each of the reported mutations in families with defective MMR as the underlying cause of disease. This likely reflects experimental data indicating that the protein products of these genes are indispensable for MMR function.[29] Nearly 10% of germline mutations occur in the *hMSH6* gene. It has been reported that *hMSH6* mutations often occur in HNPCC families characterized by a less-typical clinical presentation, including a later onset of cancer, a relatively higher occurrence of endometrial cancer, and a lower degree of MSI in tumor tissue.[37–39] A smaller number of families have been reported that carry *hPMS2* mutations.

Mutation analysis of MMR genes has provided estimates of the proportion of families that meet clinical criteria for HNPCC and also have identifiable germline mutations.[40,41] Depending on the techniques used and the number of genes examined, between 40% and 90% of families that fit the Amsterdam criteria have identifiable germline mutations.[42] If the less-restrictive Amsterdam II criteria are utilized, this proportion decreases to 5% to 50%.[43,44] There are several possible explanations for this finding. The manner in which the population of HNPCC families under study is ascertained plays an important role in the proportion of families that are found to be positive for mutations in MMR genes. Large, well-defined families with many affected individuals are much more likely to harbor recognizable mutations than are families that barely meet the minimum clinical criteria. A portion of undetected mutations in these families also may be due to technical issues related to the mutation-screening methods, which detect only certain types of mutations and generally have limited mutation-detection rates that range from 60% to 90%. In addition, most laboratories have traditionally examined only the *hMSH2* and *hMLH1* genes for mutations, thus biasing against families that carry mutations in other MMR genes. Also, families that fit the clinical criteria for HNPCC may do so because of environmentally induced aggregates or may occur by chance as a result of the high frequency of CRC in the general population. Finally, another portion of these clinically defined HNPCC families may be due to yet undiscovered genes that predispose to HNPCC. These unidentified genes may or may not be involved in MMR. In support of this hypothesis, linkage analyses suggest that HNPCC families exist that do not involve known loci.[45,46] More important, a significant fraction of HNPCC families do not have evidence of defective MMR when tumors from family members are tested.[17] Families which fit the Amsterdam criteria I for HNPCC but do not have defective MMR have recently been designated as "familial colorecteal cancer type X" to distinguish them from families with hereditary MMR deficiency. These families have a lower incidence of colorectal cancer and may not have any increased incidence of other cancers.[42]

Because of the heterogeneity noted above, it is now possible to make a distinction between HNPCC and a hereditary defect in an MMR gene which is now termed Lynch syndrome.[43] HNPCC is a clinical diagnosis primarily based on family history and other clinical information. As discussed above, current data indicate that there is a subset of families with HNPCC in which tumors do not exhibit any phenotypic evidence of defects in MMR (MSI or the absence of protein expression), and germline mutations in MMR genes have not been identified. Therefore, it is likely that other genes not involved in MMR are responsible for the diagnosis of some patients with HNPCC. Conversely, not all patients with germline defects in MMR genes meet the clinical criteria for HNPCC.[17] Thus, Lynch syndrome is a genetic diagnosis based on the finding of defective MMR in the tumor and a germline mutation in one of the MMR genes. Not all patients with clinically defined HNPCC will have Lynch syndrome, and not all patients with Lynch syndrome will fulfill the clinical criteria for HNPCC. Lynch syndrome is likely to be a relatively homogenous diagnosis despite the locus heterogeneity, whereas HNPCC is more heterogeneous and includes both Lynch syndrome and other hereditary genetic causes of HNPCC. This important distinction has implications for the identification of patients with CRC who should undergo tumor

screening and potentially germline testing (see Strategy for HNPCC testing, below). This distinction also has important implications as it relates to the natural history of the disease.[42]

Available Assays

MSI

MSI is the phenotypic reflection of an underlying MMR defect in a tumor. For this analysis, both tumor and normal tissue or cells are required. Typically, paraffin-embedded tissue sections (10 μm thick) are microdissected to obtain relatively pure tumor cells. Areas containing >70% tumor cells are typically used, although regions containing as little as 50% can provide adequate results. Normal cells are derived from adjacent normal mucosa, sections from archival normal surgical specimens for the same patient, if available, or peripheral blood. DNA is extracted separately from the tumor and normal cells. A number of extraction methods are now available that provide good-quality DNA from paraffin-embedded tissue. Isolated DNA is analyzed by the polymerase chain reaction (PCR) using radioactive or fluorescent-labeled primers to amplify and visualize multiple segments of DNA containing microsatellite loci. The PCR products are subjected to size-based separation on polyacrylamide gels or, more recently, through capillary electrophoresis.

MSI can be observed by comparing the electrophoretic patterns of amplified DNA from both tumor and normal tissue and is scored as the presence of novel fragments in tumor DNA compared to normal DNA (Figure 19-1a). For dinucleotide repeats, these novel fragments can be seen as higher- or lower-molecular-weight fragments (expansion or contraction of repeat). For mononucleotide repeats, however, lower-molecular-weight fragments are typically detected. Allelic imbalance, or LOH, is not scored as MSI since this alteration is the result of an entirely different mechanism (Figure 19-1, panel b). Overall, we have found that the most important aspect of this test is to have a well-optimized PCR assay. Because of PCR artifact, results will be far more difficult to interpret if each of the assays is not sufficiently optimized.

Historically, a major obstacle to understanding the tumor phenotype associated with defective MMR was a lack of consensus on the definition of MSI. The number of microsatellite loci and the proportion that demonstrate instability in order to be classified as MSI positive were highly variable between studies. The likelihood that any particular microsatellite will be susceptible to instability in a tumor probably relates to the inherent instability at the locus. Thus, differences in marker susceptibility to instability may lead to differences in the number of tumors scored as MSI positive.

In 1998, guidelines proposed for MSI testing at a National Cancer Institute (NCI) workshop were published.[19] These guidelines provide a reference panel of five validated microsatellite loci to ensure uniformity of diagnosis between laboratories. Laboratories could use the panel directly or use this panel as a basis of comparison to validate their own set of markers. These loci consist of two mononucleotide (BAT 25 and BAT 26) and three dinucleotide (D2S123, D5S346, and D175250) repeat markers. Mononucleotide markers are relatively stable and tend to be monomorphic. Mononucleotide markers were included in the panel because reports indicate that cells with mutant *hMSH6* were mainly associated with instability at mononucleotide markers, while cells with mutations in *hMLH1*, *hMSH2*, or *hPMS2* lead to pronounced instability at both mono- and dinucleotide repeats.[48] Thus, analysis of only dinucleotide markers could lead to false-negative MSI results for tumors caused by mutations in *hMSH6*.

Several studies have reported that the fraction of markers demonstrating instability within tumors have a bimodal distribution with a breakpoint at approximately 30% to 40%.[49] Therefore, tumors analyzed for MSI can be divided into three discrete groups: MSI-H (high-frequency MSI), MSI-L (low-frequency MSI), and MSS (microsatellite stable). For the NCI panel of five markers, MSI-H is defined as two or more unstable markers, MSI-L as one unstable marker, and MSS as no unstable markers. Although this group of five markers (or a comparable group) often is adequate to identify tumors with the MSI-H phenotype, it is not particularly useful for distinguishing between MSS and MSI-L. A larger number of markers is necessary to make this distinction. When only one marker is unstable or when technical difficulties limit the number of informative markers, an additional five markers may be necessary to help differentiate between MSI-H and MSI-L/MSS. In these cases, the absolute number may not be as important as defining the overall frequency of unstable microsatellites. When more than five markers are used, instability at ≥30% of markers is considered MSI-H, while MSI-L is defined as instability in <30% of the markers tested. More recent data suggest that the use of a greater number of mononucleotide markers (two to four per set), by themselves, may have a greater sensitivity and specificity for the detection of the MSI-H phenotype.[50] Additionally, because these mononucleotide markers tend to be monomorphic, investigators have suggested that testing can be performed with tumor only (i.e., without the need to compare to normal tissue from that patient).[51] However, caution should be used with this approach since other alleles can be detected in some patients, especially in certain racial groups. Without the use of normal tissue from that patient, such results could be scored as false positives. An example of a patient that is heterozygous for BAT 26 is shown in Figure 19-1c. In addition, the quasimonomorphic marker BAT26, which is located immediately downstream of exon 5 in the *MSH2* gene, is widely regarded the most sensitive and specific marker of MSI. Pastrello and colleagues recently demonstrated that in individuals with large germline deletions of *MSH2* which span exon 5, *BAT26* is more likely to show no

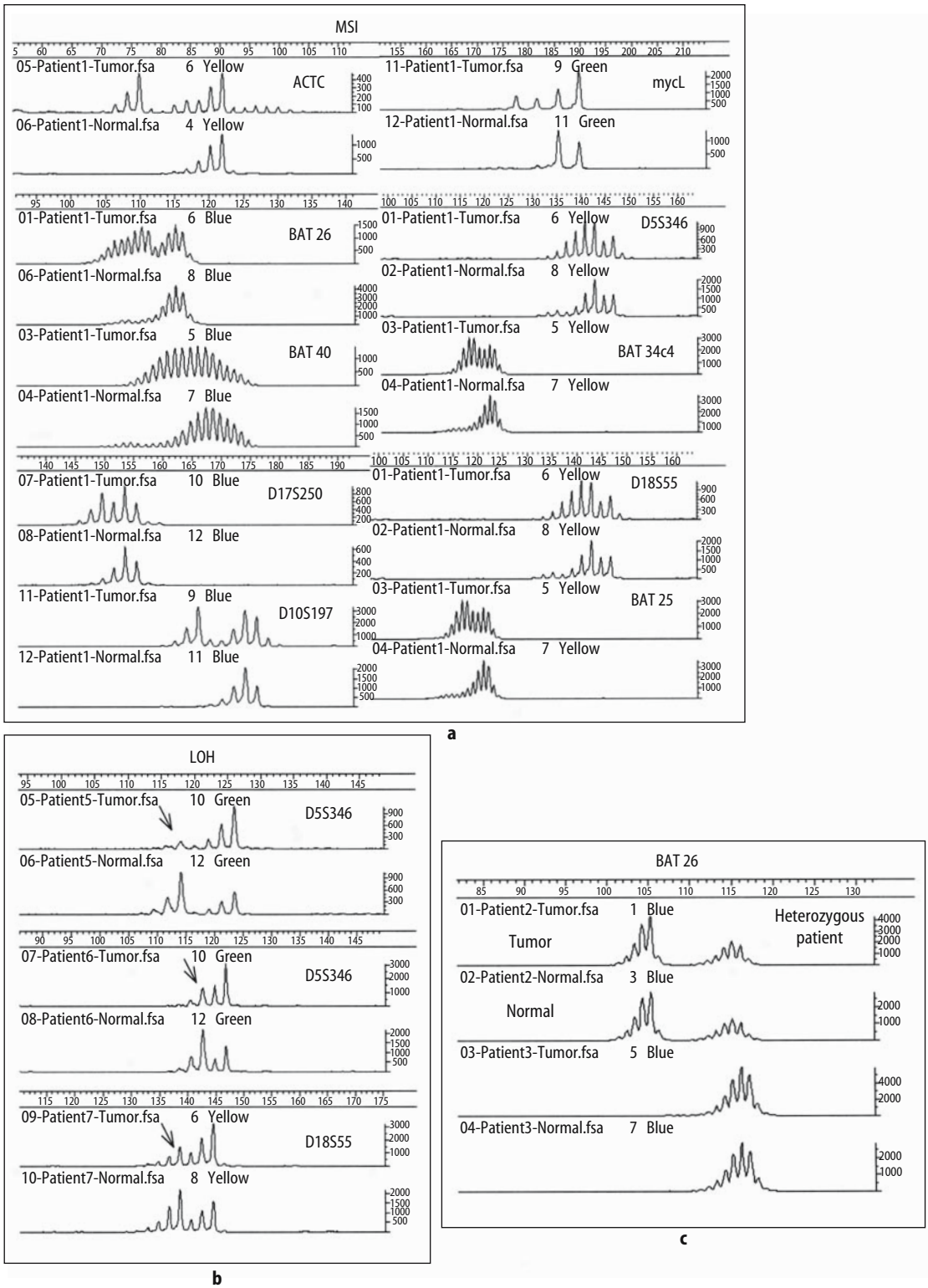


Figure 19-1. MSI analysis of normal and tumor DNA from a CRC patient using four mononucleotide repeat markers (BAT 25, BAT 26, BAT 40, BAT 34c4) and six dinucleotide repeat markers (D17S250, ACTC, D10S197, D18S55, D5S346, mycL) with the use of an ABI3100 genetic analyzer. (a) Example of MSI for each of the ten markers. (b) Example of microsatellite changes that can be misinterpreted as MSI: allelic imbalance or LOH. (c) Example of a patient that is heterozygous for the marker BAT 26.

MSI when tested on tumor tissue, due to the subsequent somatic deletion of the corresponding region of the wild-type allele.[52] This work has highlighted a limitation of using *BAT26* alone for the detection of MMR deficiency, a limitation of considerable magnitued considering that large deletions of *MSH2* comprise 15% of all known deleterious MMR mutations of which up to one-half span exon 5. A third NCI workshop was convened in December 2002 to further develop guidelines for the clinical diagnosis of HNPCC and MSI testing. Results of this workshop advocated the

continued use of the original NCI panel of five markers with recognition that in the future additional mononucleotide markers may be added to the panel.[51]

The proper differentiation between MSI-H and MSI-L/MSS tumors is important because these two tumor subsets have very different biological characteristics with different molecular etiologies. Current data indicate that the MSI-H phenotype is due almost exclusively to defective MMR. MSI-L and MSS tumors, so far, appear to share the same phenotypic characteristics with little or no involvement of defective MMR. MSI-L tumors show a genetic profile similar to sporadic MSS CRC (with typical *KRAS* and *TP53* mutations and LOH at 5q, 17p, and 18q), but not to MSI-H CRC.[53,54] In addition, several histopathologic studies demonstrate that MSI-H tumors reflect the clinical and pathological features typically associated with HNPCC (reviewed above), while MSI-L/MSS tumors are not associated with these features.[49] Together, these data indicate that MSI-L/MSS CRC are fundamentally different from MSI-H CRC. This assay, therefore, provides a sensitive test for detecting the presence of an abnormality in the MMR pathway.

As noted earlier, MSI is not a specific characteristic of CRC associated with HNPCC. Large prospective studies of unselected CRC patients have determined the frequency of MSI-H in both sporadic and HNPCC-derived tumors. These studies indicated that 12% to 28% of unselected CRC display MSI.[16,17] However, mutation analysis of the samples with MSI-H revealed germline MMR gene mutations in only approximately 2% of the patients. MSI also has been described in other sporadic component tumors within the HNPCC spectrum.[55] These data led to an interesting paradox: although most CRC derived from individuals with a clinical diagnosis of HNPCC demonstrate MSI-H, only a minority of unselected tumors with MSI-H are from HNPCC families. These data indicate that two groups of phenotypically indistinguishable MSI-H tumors exist: those that are due to an underlying germline mutation in an MMR gene and those that are due to another mechanism. Insight into this paradox was gained, in part, by the analysis of MMR proteins in colorectal tumor specimens by immunohistochemistry (IHC). Nearly all tumors that displayed MSI-H, but without germline MMR gene mutations, demonstrated a loss of hMLH1 protein expression. Subsequent analysis of tumor tissue for *hMLH1* mutations indicated that somatic mutation was a relatively infrequent event. However, analysis of the methylation status of the *hMLH1* promoter region indicated that aberrant promoter methylation was the most likely explanation for the loss of hMLH1 expression in these tumors.[56,57] Thus, it appears that MMR inactivation can account for nearly all CRC demonstrating the MSI-H phenotype.

Immunohistochemistry

The examination of CRC tissue for loss of expression of MMR proteins by IHC complements MSI analysis in screening for defective MMR in HNPCC. IHC is a relatively fast, easily interpretable, and inexpensive method that is performed on slides from paraffin-embedded tumor tissue. Tumor cells that demonstrate an absence of nuclear staining in the presence of positive staining in surrounding normal cells (lymphocytes and normal colonic epithelium) are interpreted as having an absence of protein expression (Figure 19-2b). Loss of MMR protein expression is highly correlated with an MSI-H phenotype, but not with MSI-L/MSS.[58]

In addition to defining the presence of defective MMR, an important feature of IHC analysis lies in the finding that loss of expression of a particular MMR protein is predictive of the gene most likely mutated. IHC analysis of tumor from moderate- and high-risk individuals is routinely available for hMSH2, hMLH1, hMSH6, and hPMS2. Interestingly, loss of expression of hMSH2 is accompanied by loss of hMSH6, and similarly, loss of expression of hMLH1 is accompanied by loss of hPMS2. However, loss of hMSH6 and hPMS2 can occur alone. The mechanistic reasons for the concomitant loss of these proteins are not entirely clear. However, since these combinations of proteins also make up the MutS and MutL MMR complexes, it is possible that

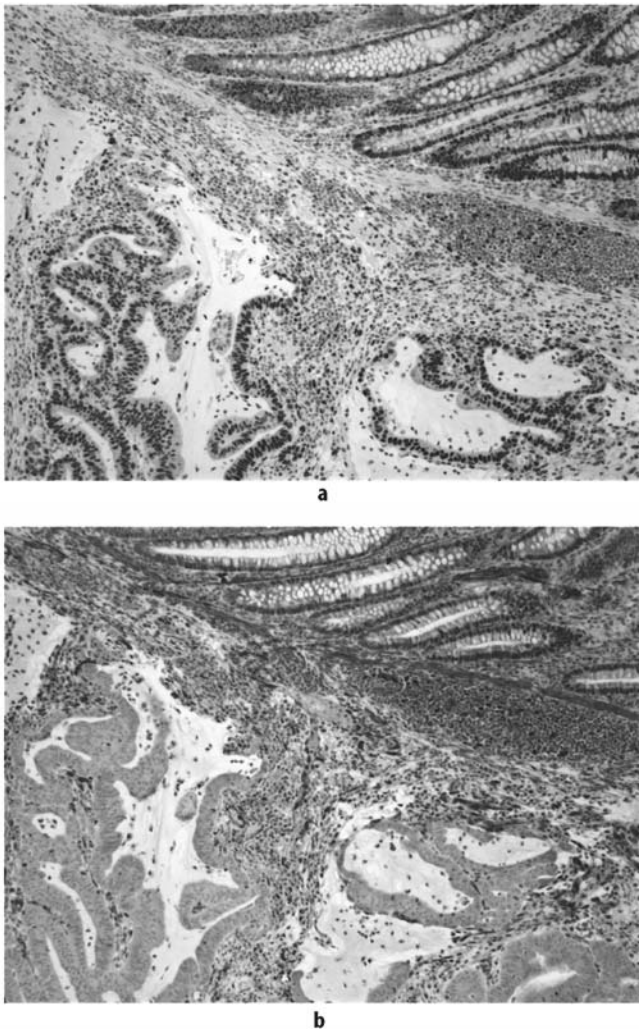


Figure 19-2. Immunohistochemical analysis with anti-hMLH1 antibody. (a) Tumor and surrounding normal tissue showing nuclear staining for hMSH6 expression. (b) Normal tissue with positive staining, but tumor tissue shows an absence of staining, indicating a loss of hMSH6 protein expression.

the loss of one partner of a complex has effects on the expression, protein folding, or proper cellular localization of the other partner. In practice, the loss of expression of hMSH2, loss of hMSH6 alone, or loss of hPMS2 alone is strong evidence for the presence of a germline mutation in that respective gene. Thus, efforts to determine the precise germline mutation responsible for HDMMR in an individual can be focused on a single gene. In addition, absence of MMR protein expression may indicate that an identified MMR gene mutation has functional significance.

Like MSI analysis, however, IHC analysis cannot distinguish between somatic and germline alterations. Sporadic tumors with *hMLH1* promoter hypermethylation also demonstrate a loss of hMLH1 (and hPMS2) protein expression; hence, loss of expression of hMLH1 does not establish the presence of a heritable germline gene mutation. In addition, cases have been described in which tumors with an MSI-H phenotype did not reveal a concomitant loss of protein expression.[59] Thus, MSI analysis (which indicates a phenotypic defect in MMR) in conjunction with IHC (which indicates which of the MMR genes is inactivated to produce that phenotype) is likely the most effective method for screening tumors in order to identify individuals at high risk for HNPCC, and in whom mutation screening should be considered.

Mutation Analysis

Molecular genetic testing in an individual with a suspected diagnosis of HNPCC has important implications for both the affected individual and at-risk family members. The identification of a deleterious germline mutation constitutes a genetically defined diagnosis of Lynch syndrome and is important for the treatment and subsequent surveillance of the disease. In addition, the identification of a germline mutation allows predictive testing to be performed on unaffected relatives at a relatively small cost and with essentially 100% accuracy. Thus, the identification of a germline mutation in families with suspected HNPCC is of paramount importance.

Germline mutations have been identified in any one of five MMR genes in families with Lynch syndrome (http://www.insight-group.org/ and http://www.med.mun.ca/MMRvariants/). Many different types of mutations have been detected, and they are spread throughout the length of the genes with no apparent mutational hotspots. Therefore, genetic testing is a complex, time-consuming, and expensive process. Multiple laboratories offer genetic testing for HNPCC using a variety of methods. Some laboratories begin with a more general mutation-screening test to identify the region harboring a mutation. Mutation-screening techniques are designed to screen relatively large segments of DNA for sequence changes but cannot identify the precise mutation. Additionally, the mutation-detection rate for the majority of these techniques is less than 100%, typically 80% to 90%. However, these techniques are designed to be faster and less labor-intensive than direct sequencing. Common mutation-screening techniques include conformation sensitive gel electrophoresis (CSGE),[60] single-strand conformation polymorphism (SSCP),[61] denaturing gradient gel electrophoresis (DGGE),[62] and denaturing high-performance liquid chromatography (DHPLC).[63] All these methods involve PCR amplification of the region to be screened (typically a single exon with a portion of adjacent intronic sequence) followed by electrophoretic or chromatographic separation of the mutant alleles from the wild-type alleles. Once a region of a gene is identified as positive by a mutation-screening method, direct sequencing of this specific region is performed.

Since the majority of mutations in the MMR genes result in premature truncation of the respective protein product, the protein truncation test (PTT) which utilizes an in vitro transcription/translation assay also has been utilized to detect mutations.[64] The size of the resultant protein products analyzed by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) indicates the relative position of the mutation within the gene for subsequent targeted sequencing. However, PTT results may be complicated by the general difficulty of isolating and manipulating RNA; the alternative splicing of *hMSH2* and *hMLH1* transcripts, which may yield multiple protein products; and nonsense-mediated decay, which may preferentially degrade mRNA species with premature stop codons.[65]

Direct sequencing of the entire coding regions of MMR genes has been another strategy for mutation detection. With the advent of automated fluorescent sequencing and the increased ability to detect heterozygous sequence changes, however, direct sequencing is now feasible and is performed in most laboratories.

Both mutation-screening techniques and direct sequencing are limited by the fact that they do not detect large deletions, duplications, or other genomic rearrangements. Recent studies indicate that these mutations make up a significant proportion of mutations observed in *hMSH2* (approximately 30%) and *hMLH1* (approximately 5–10%)[66–68]. Due to the preponderance of deletions and duplications, analysis by Southern blot or quantitative PCR-based methods such as multiplex ligation-dependent probe amplification (MLPA) should be part of a routine mutation-detection protocol for Lynch syndrome.[69] An example of a Southern blot test for all 16 exons of *hMSH2* is shown in Figure 19-3. For this specific assay, three in vitro synthesized probes (G1, G2, and G3) were constructed to clearly delineate each individual exon.

Although traditional mutation analysis should detect the large majority of deleterious mutations in MMR genes, some mutations can be detected only when the two alleles are studied separately. Analysis of expression from each allele independently through conversion of diploidy to haploidy, involving separation of the two alleles in somatic cell hybrids, permits the detection of mutations in promoter regions or introns that may affect transcription or mRNA splicing.[70]

All genomic sequence changes are potentially deleterious. However, unlike nonsense, frameshift, splicing, or deletion mutations (which account for the majority of

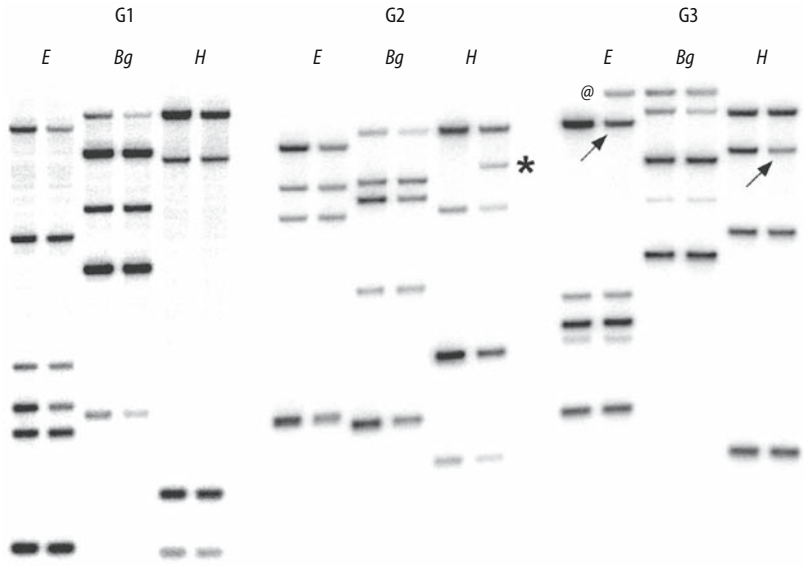


Figure 19-3. Southern blot screen for all 16 exons of the *hMSH2* gene. Quantitative intensity measurements are taken for all bands and are compared to a normal control. Arrows indicate some of the detected changes in band intensity due to hemizygosity. @, a novel restriction fragment generated by the deletion; *, an exon 8 *Hind* III polymorphism; G1, G2, and G3, the probe sets used with each blot; Bg, H, and E, the restriction endonucleases *Bgl* II, *Hind* III, and *EcoR* I, respectively.

mutations in the MMR genes), it is inherently difficult to determine the pathogenicity of missense changes. IHC analysis, evolutionary conservation, and the segregation of the mutation with disease within the family may all provide clues as to the pathogenicity of such mutations. However, functional assays are likely to be the most useful in the interpretation of these alterations. These assays have traditionally been performed in bacterial systems. Clinical utility of functional assays awaits the development of fast, inexpensive, and robust mammalian systems. Clearly the identification of deleterious mutations in families with HNPCC is an extremely difficult task due to genetic and clinical heterogeneity.

Strategy for HNPCC Testing

The ultimate goal of screening for HNPCC is to identify all cases of Lynch syndrome or other hereditary defects present in a population. In view of limited resources, identification of individuals who would benefit most from genetic testing is essential. To most efficiently identify germline mutations in patients identified through family history and clinical criteria to be at risk for HNPCC, a stepwise diagnostic procedure is recommended (Figure 19-4). The utility of this tiered approach has been demonstrated in studies that have successfully screened an unselected

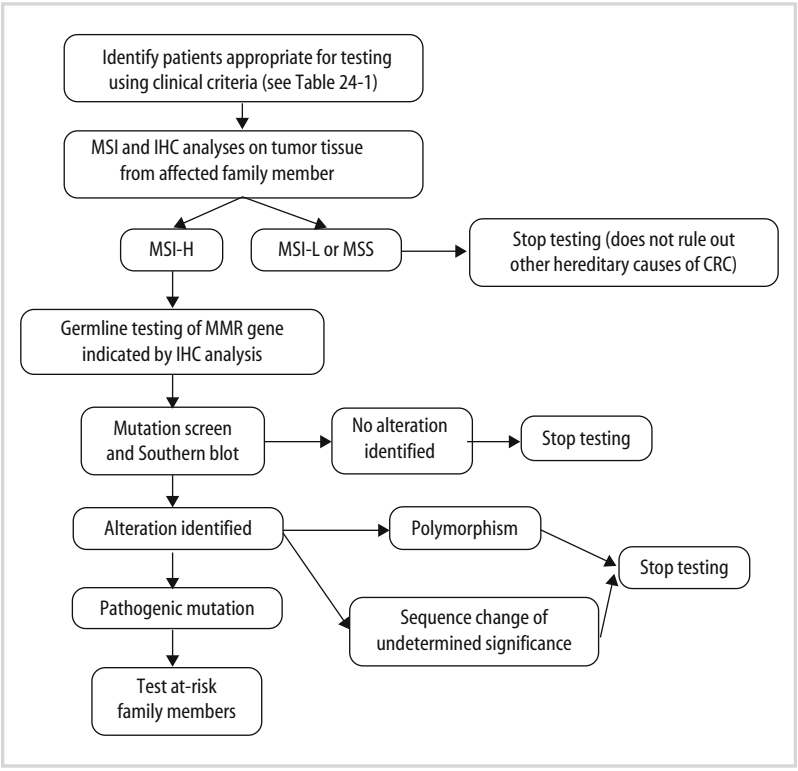


Figure 19-4. Flowchart for recommended HNPCC diagnostic procedure.

population with CRC to identify those who might benefit from genetic testing.[16,17]

Utilizing this approach, a combination of MSI and IHC analysis is first performed on tumor tissue as an initial screen for individuals at increased risk to have germline MMR defects. Since MSI is found in virtually all tumors derived from individuals with Lynch syndrome, it is generally unproductive to search for germline mutations in MMR genes in patients whose tumors do not demonstrate a mutator phenotype. Depending on the prior risk of the patient being tested, a large fraction of cases may not have evidence of defective MMR. This is especially the case for individuals that have a moderate risk for having HNPCC. Data from our Molecular Diagnostic Laboratory, which includes both moderate-risk (young age of onset only, one additional family member with CRC only, etc.) and high-risk (Amsterdam criteria) referral cases, has demonstrated that up to 70% of cases do not have evidence of defective MMR. Direct sequencing of several MMR genes in such cases would generally not be productive. The combination of MSI and IHC testing is a relatively inexpensive prescreen that can eliminate unnecessary and expensive germline testing when performed as the first step in a testing protocol.

If the tumor demonstrates an MSI-H phenotype, then IHC provides important information about the specific MMR gene that is most likely mutated. The subset of patients identified then would be considered for germline testing with appropriate genetic counseling. Germline testing of the appropriate MMR gene (identified by IHC) by mutation screening or direct sequencing in conjunction with Southern blot (or other assays capable of detecting germline deletions or other genomic rearrangements) would follow in those patients who provide informed consent. In this case, only the relevant gene is analyzed for a germline mutation, again minimizing the need to test several MMR genes. This approach is designed to be the most cost-effective and judicious use of resources at the present time. However, as technological advancements continue, this testing algorithm may be amended after further studies have proven the effectiveness of other approaches.

References

1. Lynch HT, Smyrk TC, Watson P, et al. Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review. *Gastroenterology.* 1993;104:1535–1549.
2. Lynch HT, Shaw MW, Magnuson CW, Larsen AL, Krush AJ. Hereditary factors in cancer. Study of two large midwestern kindreds. *Arch Intern Med.* 1966;117:206–212.
3. Lynch HT, Krush AJ. Cancer family "G" revisited: 1895–1970. *Cancer.* 1971;27:1505–1511.
4. Lynch HT, de la Chapelle A. Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet.* 1999;36:801–818.
5. Watson P, Lynch HT. The tumor spectrum in HNPCC. *Anticancer Res.* 1994;14:1635–1639.
6. Lynch HT, Fusaro RM, Roberts L, Voorhees GJ, Lynch JF. Muir-Torre syndrome in several members of a family with a variant of the Cancer Family Syndrome. *Br J Dermatol.* 1985;113:295–301.
7. Jass JR, Stewart SM. Evolution of hereditary non-polyposis colorectal cancer. *Gut.* 1992;33:783–786.
8. Jass JR. Colorectal adenoma progression and genetic change: is there a link? *Ann Med.* 1995;27:301–306.
9. Alexander J, Watanabe T, Wu TT, Rashid A, Li S, Hamilton SR. Histopathological identification of colon cancer with microsatellite instability. *Am J Pathol.* 2001;158:527–535.
10. Aarnio M, Mustonen H, Mecklin JP, Jarvinen HJ. Prognosis of colorectal cancer varies in different high-risk conditions. *Ann Med.* 1998;30:75–80.
11. Frei JV. Hereditary nonpolyposis colorectal cancer (Lynch syndrome II). Diploid malignancies with prolonged survival. *Cancer.* 1992;69:1108–1111.
12. Kouri M, Laasonen A, Mecklin JP, Jarvinen H, Franssila K, Pyrhonen S. Diploid predominance in hereditary nonpolyposis colorectal carcinoma evaluated by flow cytometry. *Cancer.* 1990;65:1825–1829.
13. Dolcetti R, Viel A, Doglioni C, et al. High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. *Am J Pathol.* 1999;154:1805–1813.
14. Gryfe R, Kim H, Hsieh ET, et al. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. *N Engl J Med.* 2000;342:69–77.
15. Vasen HF, Mecklin JP, Khan PM, Lynch HT. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). *Dis Colon Rectum.* 1991;34:424–425.
16. Aaltonen LA, Salovaara R, Kristo P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med.* 1998;338:1481–1487.
17. Cunningham JM, Cheong-Yong K, Christensen E, et al. The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. *Am J Hum Genet.* 2001;69:780–790.
18. Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology.* 1999;116:1453–1456.
19. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 1998;58:5248–5257.
20. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon [comments]. *Science.* 1993;260:816–819.
21. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature.* 1993;363:558–561.
22. Aaltonen LA, Peltomaki P, Leach FS, et al. Clues to the pathogenesis of familial colorectal cancer. *Science.* 1993;260:812–816.
23. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature.* 2001;409:860–921.
24. Strand M, Prolla TA, Liskay RM, Petes TD. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature.* 1993;365:274–276.
25. Peltomaki P, Aaltonen LA, Sistonen P, et al. Genetic mapping of a locus predisposing to human colorectal cancer. *Science.* 1993;260:810–812.
26. Fishel R, Lescoe MK, Rao MR, et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell.* 1993;75:1027–1038.
27. Leach FS, Nicolaides NC, Papadopoulos N, et al. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell.* 1993;75:1215–1225.
28. Fishel R, Kolodner RD. Identification of mismatch repair genes and their role in the development of cancer. *Curr Opin Genet Dev.* 1995;5:382–395.
29. Buermeyer AB, Deschenes SM, Baker SM, Liskay RM. Mammalian DNA mismatch repair. *Annu Rev Genet.* 1999;33:533–564.

30. Malkhosyan S, McCarty A, Sawai H, Perucho M. Differences in the spectrum of spontaneous mutations in the hprt gene between tumor cells of the microsatellite mutator phenotype. *Mutat Res.* 1996;316:249–259.
31. Ohzeki S, Tachibana A, Tatsumi K, Kato T. Spectra of spontaneous mutations at the hprt locus in colorectal carcinoma cell lines defective in mismatch repair. *Carcinogenesis.* 1997;18:1127–1133.
32. Bronner CE, Baker SM, Morrison PT, et al. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature.* 1994;368:258–261.
33. Papadopoulos N, Nicolaides NC, Wei YF, et al. Mutation of a mutL homolog in hereditary colon cancer. *Science.* 1994;263:1625–1629.
34. Miyaki M, Konishi M, Tanaka K, et al. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. *Nat Genet.* 1997;17:271–272.
35. Nicolaides NC, Papadopoulos N, Liu B, et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature.* 1994;371:75–80.
36. Wu Y, Berends MJ, Sijmons RH, et al. A role for MLH3 in hereditary nonpolyposis colorectal cancer. *Nat Genet.* 2001;29:137–138.
37. Akiyama Y, Sato H, Yamada T, et al. Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. *Cancer Res.* 1997;57:3920–3923.
38. Wijnen J, de Leeuw W, Vasen H, et al. Familial endometrial cancer in female carriers of MSH6 germline mutations. *Nat Genet.* 1999;23:142–144.
39. Wu Y, Berends MJ, Mensink RG, et al. Association of hereditary nonpolyposis colorectal cancer-related tumors displaying low microsatellite instability with MSH6 germline mutations. *Am J Hum Genet.* 1999;65:1291–1298.
40. Beck NE, Tomlinson IP, Homfray T, et al. Use of SSCP analysis to identify germline mutations in HNPCC families fulfilling the Amsterdam criteria. *Hum Genet.* 1997;99:219–224.
41. Wagner A, Barrows A, Wijnen JT, et al. Molecular analysis of hereditary nonpolyposis colorectal cancer in the United States: high mutation detection rate among clinically selected families and characterization of an American founder genomic deletion of the MSH2 gene. *Am J Hum Genet.* 2003;72:1088–1100.
42. Lindor NM, Rabe MS, Peterson GM, et al. Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency. *JAMA*. 2005;293:1979–1985.
43. Lynch HT, Smyrk T, Lynch J. An update of HNPCC (Lynch syndrome). *Cancer Genet Cytogenet.* 1997;93:84–99.
44. Nystrom-Lahti M, Wu Y, Moisio AL, et al. DNA mismatch repair gene mutations in 55 kindreds with verified or putative hereditary nonpolyposis colorectal cancer. *Hum Mol Genet.* 1996;5:763–769.
45. Lewis CM, Neuhausen SL, Daley D, et al. Genetic heterogeneity and unmapped genes for colorectal cancer. *Cancer Res.* 1996;56:1382–1388.
46. Huang J, Kuismanen SA, Liu T, et al. MSH6 and MSH3 are rarely involved in genetic predisposition to nonpolypotic colon cancer. *Cancer Res.* 2001;61:1619–1623.
47. Baudhuin LM, Burgart LJ, Leontovich O, et al. Use of microsatellite instability and immunohistochemistry testing for the identification of individuals at risk for Lynch syndrome. *Familial Cancer*. 2005;4:255–265.
48. Bhattacharyya NP, Skandalis A, Ganesh A, Groden J, Meuth M. Mutator phenotypes in human colorectal carcinoma cell lines. *Proc Natl Acad Sci U S A.* 1994;91:6319–6323.
49. Thibodeau SN, French AJ, Cunningham JM, et al. Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. *Cancer Res.* 1998;58:1713–1718.
50. Zhou XP, Hoang JM, Li YJ, et al. Determination of the replication error phenotype in human tumors without the requirement for matching normal DNA by analysis of mononucleotide repeat microsatellites. *Genes Chromosomes Cancer.* 1998;21:101–107.
51. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Nat Cancer Inst*. 2004;96:261–267.
52. Pastrello C, Baglioni S, Tibiletti MG, Papi L, Fornasarig M, Morabito A, Agostini M, Genuardi M, Viel A: Stability of BAT26 in tumours of hereditary nonpolyposis colorectal cancer patients with MSH2 intragenic deletion, *Eur J Hum Genet*. 2006,14:63–68.
53. Konishi M, Kikuchi-Yanoshita R, Tanaka K, et al. Molecular nature of colon tumors in hereditary nonpolyposis colon cancer, familial polyposis, and sporadic colon cancer. *Gastroenterology.* 1996;111:307–317.
54. Jass JR, Biden KG, Cummings MC, et al. Characterisation of a subtype of colorectal cancer combining features of the suppressor and mild mutator pathways. *J Clin Pathol.* 1999;52:455–460.
55. Cunningham JM, Boardman LA, Burgart LJ, Thibodeau SN. Microsatellite instability in colon cancers. In: Wells R, Warren S, eds. *Genetic Instabilities and Hereditary Neurological Diseases*. Academic Press, San Diego, CA; 1998:791–807.
56. Kane MF, Loda M, Gaida GM, et al. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res.* 1997;57:808–811.
57. Cunningham JM, Christensen ER, Tester DJ, et al. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res.* 1998;58:3455–3460.
58. Dietmaier W, Wallinger S, Bocker T, Kullmann F, Fishel R, Ruschoff J. Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression. *Cancer Res.* 1997;57:4749–4756.
59. Thibodeau SN, French AJ, Roche PC, et al. Altered expression of hMSH2 and hMLH1 in tumors with microsatellite instability and genetic alterations in mismatch repair genes. *Cancer Res.* 1996;56:4836–4840.
60. Ganguly A, Rock MJ, Prockop DJ. Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes [published erratum appears in *Proc Natl Acad Sci U S A*. 1994;91:5217]. *Proc Natl Acad Sci U S A*. 1993;90:10325–10329.
61. Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci U S A*. 1989;86:2766–2770.
62. Lerman LS, Silverstein K, Grinfeld E. Searching for gene defects by denaturing gradient gel electrophoresis. *Cold Spring Harb Symp Quant Biol.* 1986;51:285–297.
63. Underhill PA, Jin L, Lin AA, et al. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.* 1997;7:996–1005.
64. Powell SM, Petersen GM, Krush AJ, et al. Molecular diagnosis of familial adenomatous polyposis. *N Engl J Med.* 1993;329:1982–1987.
65. Frischmeyer PA, Dietz HC. Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet.* 1999;8:1893–1900.
66. Wijnen J, van der Klift H, Vasen H, et al. MSH2 genomic deletions are a frequent cause of HNPCC. *Nat Genet.* 1998;20:326–328.
67. Wang Y, Friedl W, Lamberti C, et al. Hereditary nonpolyposis colorectal cancer: frequent occurrence of large genomic deletions in MSH2 and MLH1 genes. *Int J Cancer.* 2003;103:636–641.
68. Baudhuin LM, Mai M, French AJ, et al. Analysis of hMLH1 and hMSH2 gene dosage alterations in hererditary nonpolyposis colorectal cancer patients by novel methods. *J Mol Diagn*. 2005;7:26–235.
69. Gille JJ, Hogervorst FB, Pals G, et al. Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach. *Br J Cancer.* 2002;87:892–897.
70. Yan H, Papadopoulos N, Marra G, et al. Conversion of diploidy to haploidy. *Nature.* 2000;403:723–724.

Chapter 20

Multiple Endocrine Neoplasia Syndromes

Barbara Zehnbauer

Introduction

Multiple endocrine neoplasia (MEN) syndromes include several types of autosomal dominant inherited familial cancer syndromes, each characterized by a different pattern of endocrine gland tumors in affected individuals. The two major types are MEN1 (Wermer syndrome) and MEN2 (Sipple syndrome). MEN1 is an autosomal dominant disorder characterized by a high frequency of peptic ulcer disease and primary endocrine abnormalities involving the parathyroids (90–97% of patients), pancreatic islets (30–80% of patients; including adenoma, prolactinoma, insulinoma, glucagonoma, gastrinoma, etc.), and anterior pituitary (15–50% of patients).[1] MEN2 includes subtypes MEN2A, MEN2B, and familial medullary thyroid carcinoma (FMTC, non-MEN), with the primary clinical features of medullary thyroid carcinoma (MTC; 95% of patients), pheochromocytoma (pheo; 50% of MEN2A and MEN2B), parathyroid hyperplasia (15–30% of MEN2A and rarely in MEN2B), plus mucosal neuromas (lips and tongue), ganglioneuromas of the gastrointestinal tract, and marfanoid habitus in MEN2B only.[1] The MEN2A diagnostic category characterizes approximately 60% to 90% of patients with MEN2, FMTC accounts for 5% to 35%, and MEN2B for about 5%.[2] In addition, MTC and pheo may be bilateral or multifocal with an earlier age of onset than sporadic occurrence of the same tumor type.

Penetrance of both MEN1 and MEN2 is nearly complete by 60 years of age.[1,3] While these syndromes are uncommon (1 in 10,000 to 100,000 for MEN1 and 1 in 25,000 to 30,000 for MEN2), diagnosis of these autosomal dominant disorders has important implications for other family members because first-degree relatives have a 50% risk of inheriting the causative mutation and developing the disease. Early detection is critical for the most effective intervention with the goal of treatment prior to tumor metastasis. However, these two syndromes and their underlying molecular pathology present very different challenges for the clinical molecular laboratory as a study in contrasts of allelic heterogeneity and detection sensitivity.

MEN1

Molecular Basis of Disease

The *MEN1* gene maps to human chromosome 11q13 and was cloned in 1997.[4] *MEN1* is a tumor suppressor gene[5] that contains 10 exons and encodes a ubiquitously expressed 2.8 kilobase (kb) transcript. The predicted 610 amino acid protein product, termed menin, exhibits no apparent similarities to any previously known proteins or consensus protein motifs, and its exact function is unknown. Menin is a nuclear protein that interacts with proteins involved in transcription and cell growth regulation. At least 400 different *MEN1* mutations (nonsense, frameshift, insertion, deletion, missense, and splicing defects) have been described and are distributed throughout the 9.8 kb of genomic DNA comprising the *MEN1* gene (Figure 20-1).[3] At least 75% of all reported *MEN1* mutations will produce a truncated menin protein primarily by nonsense and frameshift mutations. Missense and in-frame genomic alterations have been described in the domains of menin that interact with *JUND*, *SMAD3*, and *NAKB1*, three major effectors in transcription and cell growth regulation pathways.[6] These and other alterations that predict a loss of menin function are consistent with a role of *MEN1* as a tumor suppressor gene.[3] In mouse models of MEN1, loss of both *MEN1* alleles in the mouse germline is lethal in the embryo but does not result in tumor formation.[7] Conditional gene knockouts targeted to pancreatic islet β-cells or parathyroid cells result in growth of insulinomas or parathyroid adenomas, respectively.[8,9] No correlation has been observed between *MEN1* genotype and MEN1 phenotype.[1,3,10] In about 20% of MEN1 families no mutation in the *MEN1* gene has been identified.[11] Extensive listing and descriptions of *MEN1* mutations are available online from the Human Gene Mutation Database Cardiff

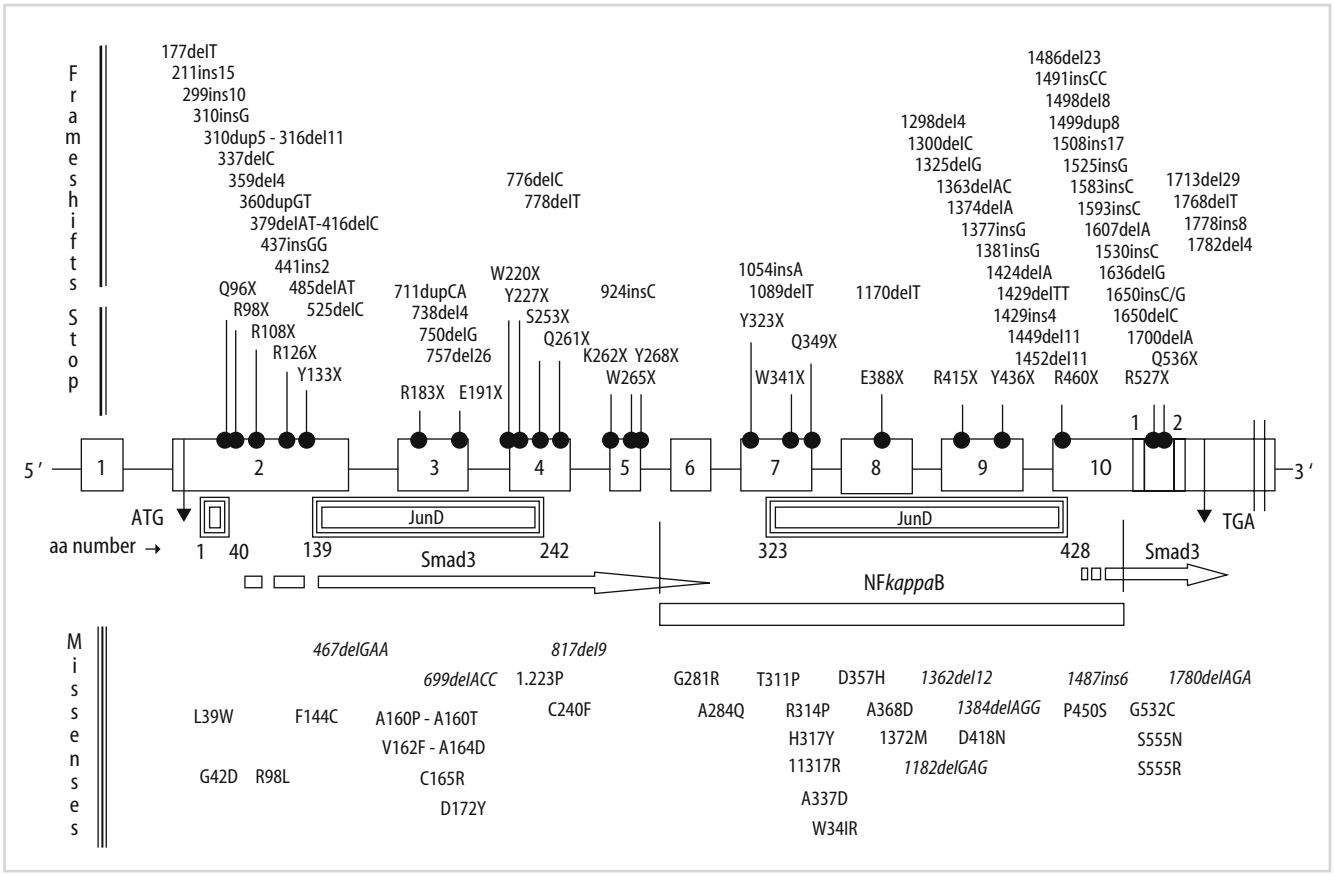


Figure 20-1. Schematic representation of the *MEN1* gene indicating the intron-exon organization, as well as the relative location, type, and distribution of germline mutations characterized in patients with MEN1. Functional domains of menin that interact with JUND, SMAD3, and NFKB are indicated below the gene map. (From Wautot V, Vercherat C, Lespinasse J, et al. Germline mutation profile of MEN1 in multiple endocrine neoplasias type 1: search for correlations between phenotype and the functional domains of the MEN1 protein. *Hum Mutat* 2002;20:35–47, copyright © 2002. Reprinted by permission of Wiley-Liss, Inc, a subsidiary of John Wiley and Sons, Inc.)

(http://uwcmml1s.uwcm.ac.uk/uwcm/mg/search/120173.html) and the Weizmann Institute of Science, GeneCards (http://www.genecards.org/cgi-bin/carddisp.pl?gene=MEN1).

Available Assays

The mutations in the 9.8 kb genomic DNA of the *MEN1* gene that are associated with the MEN1 syndrome are diverse in both type and distribution. This creates a considerable challenge in the design of a DNA diagnostic test. The initial molecular pathology screening of affected members of MEN1 families should include analysis of the entire gene. Most approaches incorporate polymerase chain reaction (PCR) amplification of each exon with a survey for sequence variants by dideoxy fingerprinting,[4] heteroduplex analysis (HA), single-strand conformation polymorphism (SSCP), or direct DNA sequencing.[3,10,12] PCR products from exons showing sequence variants by conformational changes are usually sequenced to identify the specific mutation. These strategies will identify mutations in 75% to 80% of MEN1 families. The majority of mutations are nonsense or frameshift mutations that predict expression of a truncated menin protein with loss of function. Missense mutations in *MEN1* are primarily clustered in the domains that interact with the transcription factors, JUND (codons 1–40, 139–242, 323–428), SMAD3 (distal to codon 478), and NFKB (codons 276–479), of key cell growth pathways (Figure 20-1).[13] Protein truncation testing (PTT) could be useful in detecting the effects of such mutations, but PTT has been complicated for clinical molecular laboratory implementation. Newer versions of PTT technology may yield more promising approaches for a comprehensive *MEN1* mutation detection strategy.[14,15]

As in most PCR-based screening procedures, some types of mutations can be missed, including large deletions or insertions, point mutations in the 5′ regulatory or untranslated regions, nucleotide changes in the introns, or at the sites of PCR primer annealing, which may adversely affect PCR amplification efficiency. Large germline DNA deletions can be identified by Southern blot analysis and have been described by several investigators[16,17] as abnormal restriction endonuclease fragments. These were observed in MEN1 families following negative findings using dideoxy fingerprinting and direct DNA sequencing approaches. Linkage analysis also may be used to define familial haplotypes that confer genetic predisposition to develop MEN1.

Interpretation

As for any clinical molecular genetic test, the ability to identify a specific gene mutation to correlate with the disease status of an affected individual depends on the sensitivity of the assay. A well-defined clinical phenotype is essential for accurate interpretation and reporting of molecular test results to the healthcare professional. Genotype-phenotype correlations are either imperfect or nonexistent for these disorders. Both MEN1 and MEN2 are inherited in autosomal dominant patterns; thus, a single germline mutation confers predisposition for these cancer syndromes. At the cellular level, MEN1-associated tumors are caused by inactivation of both copies of this tumor suppressor gene, while mutation of a single copy of *RET* produces a gain of function, dominantly expressed effect.[11]

About 80% of patients with MEN1 will have a germline mutation identified in the *MEN1* gene coding for the menin protein.[1,3] *MEN1* germline mutations also have been identified in nearly two thirds of patients with sporadic MEN1.[10] There is presently no specific genotype-phenotype association of *MEN1* mutations to predict the clinical course or onset of this disease. More than 400 different mutations have been described in the MEN1 literature, and some of these have been reported in multiple, independent kindreds. Characterization of these same genetic variations in newly diagnosed individuals supports the role of the mutant allele in the disease etiology and is very helpful to the molecular pathology interpretation. As in many genetic syndromes, molecular identification of a pathogenic *MEN1* mutation in a family member with clinical features of MEN1 is most significant in the risk assessment of other first-degree relatives.

As reviewed by Guo and Sawicki, *MEN1* mutations are not limited to patients with MEN1.[10] A significant number of sporadic endocrine tumors also may harbor *MEN1* mutations but not with the same prevalence as the same types of tumors in the syndromic cases. *MEN1* mutations do not provide clinically useful information for sporadic tumor staging.[11] While one third of MEN1 patients will develop pituitary tumors and carry germline *MEN1* mutations, only about 1% of sporadic pituitary tumors carry *MEN1* mutations. The frequency of *MEN1* mutations also differs between different types of pancreatic endocrine tumors. In sporadic gastrinomas, the rate of *MEN1* mutations is 37% in contrast to 15% of sporadic nongastrinoma pancreatic tumors. Insulinomas, another pancreatic endocrine tumor of MEN1, have been characteristic of the MEN1 knockout mice,[18] but *MEN1* mutations have not been found in sporadic insulinomas.[19] Clearly the role of menin in familial and sporadic tumors is not the same. Sporadic parathyroid and pituitary adenomas with no detectable *MEN1* mutation may be characterized as phenocopies of MEN1, particularly when there is no family history of MEN1.[10]

Information about the penetrance and expressivity of the mutations in *MEN1* described to date suggests that there is nearly 100% risk for MEN1 in carriers of the same familial mutation by the age of 60 years.[3] Mutations leading to truncated protein expression create loss of functional menin protein. Missense mutations may disrupt menin interaction with cell growth regulatory molecules or its cellular localization. These mutations also may decrease the stability of menin protein or lead to its degradation. This is largely speculative at this time because the exact function of menin is unknown.

Clinical Significance

Nearly 50% of patients with MEN1 will die as a result of their disease; thus, genetic screening to identify family members with germline *MEN1* mutations may lead to closer monitoring, earlier identification of tumors, improved outcomes, and longer survival. Primary hyperparathyroidism is present in more than 90% of patients with MEN1, but parathyroid carcinomas are rarely seen. In addition, pancreatic tumors may be observed in about 60% of MEN1 patients. These most commonly are classified as gastrinomas or insulinomas. Pituitary adenomas also occur in about 30% of MEN1 patients, but pituitary carcinomas are not usually detected. Variable expressivity and reduced penetrance also have been reported in many MEN1 families, making distinction of a familial condition difficult to discern from sporadic tumor occurrence. Absence of a detectable mutation does not exclude the diagnosis of MEN1. Genetic variability in other genes with which the menin protein interacts also may contribute to variable phenotypes of family members with the same *MEN1* allele.

Identification of carriers of *MEN1* gene mutations has not been very successful in either predicting or preventing the course of MEN1. The utility of *MEN1* mutation detection in directing patient care is limited by the absence of very sensitive imaging methods for early detection and treatment of MEN1-associated tumors, coupled with the lack of adequate treatment options for the types of tumors observed in MEN1.[10] Knowledge of the carrier status is not an indicator for prophylactic surgery as in MEN2; thus, mutation screening in MEN1 patients does not avoid or cure the malignancy but may assist in clinical monitoring and lifestyle decisions. When a family-specific mutation can be identified, the clinical significance of a subsequent negative genetic analysis in an at-risk relative can aid the genetic counselor in reassuring that family member. Targeting therapy to the menin protein has not been attempted because its function and the protein domains required for tumorigenesis have not been identified.[11]

Quality Control and Laboratory Issues

Complete analysis of the *MEN1* gene sequence by PCR amplification of each exon and direct DNA sequencing fails to identify mutations in 20% to 25% of patients with MEN1. Some *MEN1* mutations have been identified repeatedly in apparently unrelated families, suggesting some tendency for mutational hot spots within the gene. Caution must be

used in the interpretation of the pathogenic consequences of DNA sequence variants. Predicting abnormal protein function from a DNA sequence change is imperfect. Several *MEN1* polymorphisms have been reported (Arg171Gln and Ala541Thr) that produce amino acid substitutions featuring different polar side chains but apparently are not associated with a disease phenotype.[10] Missense mutations in amino acids surrounding these same codons are consistent with MEN1 incidence in affected families.[3] Classification of polymorphisms as benign may change with long-term follow-up or ascertainment of additional MEN1 kindreds. Some frameshift or nonsense mutations do not completely abolish normal protein function. The existence of phenocopies (clinical presentations that may mimic familial disorders) also should be a consideration due to the relatively high frequency of sporadic parathyroid and pituitary adenomas,[10] particularly in the absence of a family history of MEN1. Genotype-phenotype prediction may be further complicated by variable expression among family members or different families with the same sequence change, or by slow growth of an MEN1 tumor even after a gene mutation has been characterized. Once a sequence change has been documented as present in individuals with clinical features of MEN1 and absent in unaffected individuals, one can be reasonably certain of the significance of the finding as a true disease-associated mutation.

Of mutations reported in the literature, approximately 70% are nonsense, frameshift, or splice-site mutations that predict the expression of a truncated menin protein product. We investigated the feasibility and utility of an *in vitro* protein truncation test (PTT) for diagnostic *MEN1* screening. This proved to be too problematic for routine clinical laboratory implementation, given the problems inherent in (*a*) the instability of abnormal mRNA in peripheral blood lymphocytes of patients with MEN1 and (*b*) the scarcity of methionine residues for ^{35}S-methionine substitution in labeling the amino terminus polypeptides of the menin protein. Newer approaches with nonisotopic PTT[14,15] may be more successful in applying this strategy to *MEN1* mutation detection. Identification of large deletions by Southern blot or fluorescence in situ hybridization (FISH) analysis may be useful for mutation screening because deletion of an entire exon or exons will result in amplification and sequencing of the remaining, normal allele and a false-negative test result. Genetic variants in the regulatory regions of *MEN1*, including promoter mutations or methylation defects, may indicate abnormal gene expression but have not yet been documented.

MEN2

Molecular Basis of Disease

In 1993, the *RET* protooncogene ("rearranged during transfection"), located on chromosome 10q11, was cloned and characterized as encoding a transmembrane tyrosine kinase with a cysteine-rich extracellular receptor domain (Figure 20-2).[20,21] *RET* spans 60 kb with 21 exons encoding a protein of approximately 1,100 amino acids.[6] Glial cell derived neurotrophic factor (GDNF) and neurturin are two ligands for the *RET* receptor domain.[22] The ligands initiate homodimerization of *RET* protein molecules resulting in phosphorylation and activation of the tyrosine kinase domain, leading to downstream signal transduction.[22] Missense mutations in the *RET* gene were characterized in individuals with the MEN2 syndromes[20,21,23,24] with a very restricted pattern of variations (Figure 20-2). In patients with MEN2A, nearly all the mutations in this gene are missense mutations localized to five conserved cysteine residues in *RET* exons 10 and 11 (Table 20-1) in a cysteine-rich extracellular domain adjacent to the transmembrane domain of the RET protein (Figure 20-2).[20,21,23] Most of these mutations produce constitutive dimerization and increased kinase activation of the aberrant RET protein.[25] Individuals with only FMTC have missense mutations in these same cysteine codons or in a few additional codons (768 and 804) located in exons 13 and 14 (Table 20-1), which specify intracellular portions of the RET protein next to the transmembrane domain (Figure 20-2).[26,27] Nearly 95% of all patients diagnosed with MEN2B have been characterized with a single missense mutation in codon 918 of *RET* exon 16, M918T, replacing a methionine residue with a threonine.[23,24] This variant, in the intracellular tyrosine kinase catalytic domain (Figure 20-2), alters substrate recognition, which leads to cellular transformation.[25] In each case, the mutation produces a gain-of-

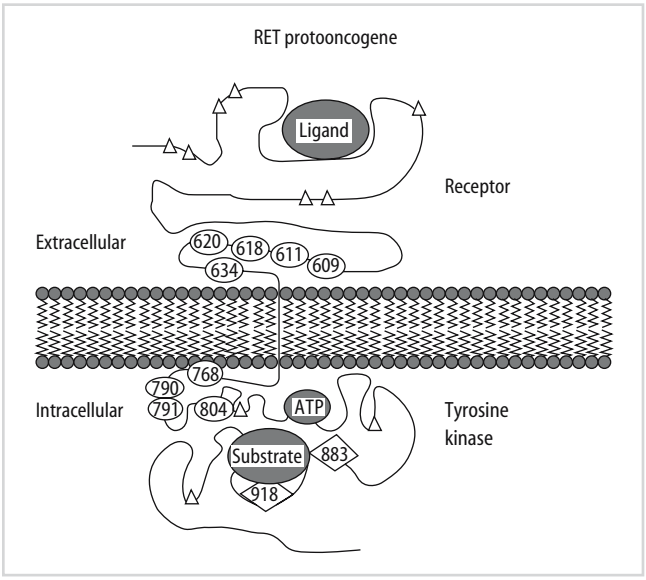


Figure 20-2. Diagram of *RET* gene product indicating the relative location of germline mutations characterized in patients with MEN2. Ovals are numbered to represent the codons that commonly contain germline missense mutations associated with MEN2A and FMTC. Diamonds are numbered to represent the codon location of germline mutations in MEN2B. Triangles indicate the positions of mutations in hereditary Hirschsprung disease. (From Phay JE, Moley JF, Lairmore TC. Multiple endocrine neoplasias. *Semin Surg Oncol* 2000;18:324–332, copyright © 2000. Reprinted by permission of Wiley-Liss, Inc, a subsidiary of John Wiley and Sons, Inc.)

Table 20-1. Common *RET* Gene Mutations in MEN2 and Associated Clinical Phenotypes

		Amino Acid Substitution†									
RET location	Codon*	Y TAC	F TTC	G GGC	R CGC	S TCC or AGC	W TGG	D GAC	M ATG	L TTG	T ACG
Exon 10	C609 TGC	2A, FMTC HSCR	2A	2A	2A	—‡	HSCR				
	C611 TGC	2A	2A	FMTC	2A	2A	2A				
	C618 TGC	FMTC	2A	2A, FMTC	2A, FMTC	2A, FMTC HSCR	—‡				
	C620 TGC	2A	FMTC	2A	2A, FMTC HSCR	FMTC	2A, HSCR				
Exon 11	C634 TGC	2A	2A	2A	2A, FMTC	2A, FMTC	2A				
Exon 13	E768 GAG							FMTC	—‡	—‡	—‡
Exon 14	V804 GTG							—‡	FMTC, 2A	FMTC	—‡
Exon 16	M918 ATG							—‡	—‡	—‡	2B

*The normal codon and sequence is listed in the Codon column.
†The Amino Acid Substitutions (single-letter abbreviations) and codon triplet sequences that result from the point mutations at each normal codon are listed across the top of the table.
‡Some substitutions have not been described at every codon and are indicated by a dash.
The clinical phenotypes that have been described for each substitution are displayed in the grid. 2A, MEN2A; FMTC, familial medullary thyroid carcinoma; HSCR, Hirschsprung disease; 2B, MEN2B. Example: C609Y, TGC → TAC has been described in cases of MEN2A, FMTC, and HSCR.
More extensive lists can be viewed at the Human Gene Mutation Database Cardiff Web site for RET (http://archive.uwcm.ac.uk/uwcm/mg/search/120346.html) and the RET Gene Card (Weizmann) (http://bioinfo.weizmann.ac.il/cards-bin/carddisp?RET).

function change for the aberrant RET protein. Online information about many other *RET* mutations is available from the Human Gene Mutation Database Cardiff (http://uwcmml1s.uwcm.ac.uk/uwcm/mg/search/120346.html) and the Weizmann Institute of Science GeneCards (http://www.genecards.org/cgi-bin/carddisp-pl?=RET).

Hirschsprung disease (HSCR) is also associated with somatic or germline *RET* mutations, either in codons 609, 618, or 620, or throughout other regions of the *RET* gene.[28,29] Some families have been described in which HSCR cosegregates with either MEN2A or FMTC.[30] *RET* mutations in patients with HSCR may produce either loss or gain of RET protein function and include frameshift and nonsense mutations. Papillary thyroid carcinoma is associated with somatic gene rearrangements of the *RET* gene.[31]

Available Assays

Screening for *RET* protooncogene mutations to confirm MEN2 diagnosis or to determine predisposition for MEN2 in asymptomatic relatives may be accomplished by a focused survey of a few exons of the *RET* gene. The majority of pathogenic mutations have been characterized in exons 10, 11, 13, 14, and 16, including codons 609, 611, 618, 620, 634, 768, 804, and 918, with few exceptions (Table 20-1). In addition, all mutations of these *RET* codons associated with MEN2 are missense substitutions.[20,21,23,24,26,27] Nonsense and frameshift mutations in *RET* have not been observed in MEN2, and small, in-frame deletions or insertions are rarely identified.[32] These cysteine codons are not mutated with equal frequency, with the majority of patients with MEN2A having alterations of codon C634 (Figure 20-3). Nearly every base position of these TGC cysteine codons has been a point of allelic variation, generating more than 20 distinct alleles; thus, mutation detection methods must be designed with the capability to interrogate each nucleotide position of the codon. In clinical laboratory practice, the preferred approach first is to characterize the causative (pathogenic) mutation in an affected individual using a comprehensive PCR plus direct DNA sequencing approach (Table 20-2) either of the commonly altered exons (Table 20-1) or of the entire *RET* gene. A variety of physical detection methods for rapid discovery of DNA sequence variation such as denaturing gradient gel

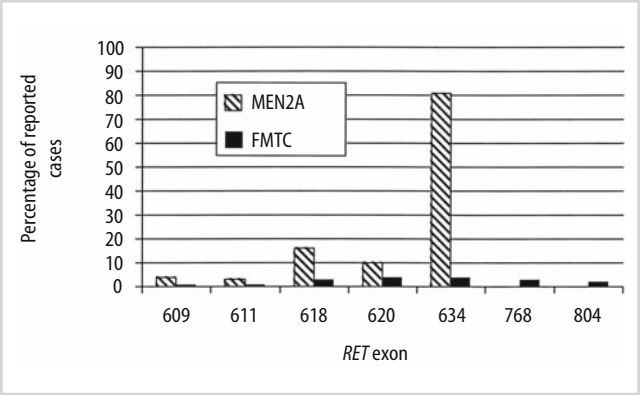


Figure 20-3. Distribution of *RET* mutations in MEN2A and FMTC. Summarized from the literature. Codon 634 in exon 11 represents a mutation hotspot in MEN2A.

electrophoresis (DGGE),[33] SSCP,[32,34,35] HA,[36] and denaturing high-performance liquid chromatography (DHPLC)[36] have been used extensively in research studies and are very useful for initial localization of a single-nucleotide sequence variation. Most of these methods identify a single-nucleotide sequence change by the altered mobility caused by the DNA conformational change using electrophoretic or melting profile shift analysis in comparison to the wild-type nucleotide sequence for the same DNA fragment. Sensitivity in the detection of the mobility shift is optimal in fragments of 200 to 400 base pairs (bp) for these approaches, and may include 50 to 600 bp targets or up to 1.5 kb with DHPLC. Determination of the exact position of the base change within these fragments (for example, TGC to TAC or TGC to GGC) can sometimes be deduced by comparison with an extensive panel of controls of defined mutant alleles previously confirmed by DNA sequencing.[33,36] The high GC content of some *RET* sequences additionally can confound physical detection methods by reducing the sensitivity of resolution between homoduplex and heteroduplex DNA molecules. Confirmation by direct DNA sequencing identifies both the exact position and nature of the base change localized by conformational screening methods. One major advantage of many of these conformational techniques is that they are frequently performed with automated instrumentation.

Following the identification of the specific pathogenic mutation in an affected relative, other family members may be tested for just this allele with a more targeted, follow-up test of a single codon within one exon of *RET*. The molecular methodology may not be limited to direct DNA sequencing.[37] For example, many missense mutations identified in the comprehensive screening also may be directly detected by PCR–restriction fragment length polymorphism (RFLP) using a variety of enzymes,[37] DGGE, DHPLC, SSCP, HA, or single nucleotide polymorphism (SNP) testing methods.[38] Site-specific, single-base extension assays, such as those used for single-nucleotide polymorphism detection, have been described[38] for detection of many of the *RET* mutations. Linkage analysis is feasible[39] when direct mutation analysis fails to identify one of the common mutations in a family with a clear pattern of heritable disease.

Interpretation

The sensitivity of *RET* mutation screening requires concise clinical description of the patient's phenotype. At least 95% of patients with MEN2 will have one of a limited number of well-characterized mutations in the *RET* gene leading to a gain of function. There is some genotype-phenotype

Table 20-2. Oligonucleotide Primer Sequences in Use for Clinical Diagnostic PCR and Direct DNA Sequencing of *RET* Exons 10,11,13,14, and 16

Primer Name	*RET* Exon	Forward or Reverse	Primer Sequence
10NF2	10	F	5′-TCAGGGGGCAGCATTGTTGGG-3′
10R2	10	R	5′-ATGGGGCCCACTCGCCTCCCAGC-3′
10F	10	F	5′-GCGCCCCAGGAGGCTGAGTG-3′
10R3	10	R	5′-CCTCCCTTGTTGGGACCTCAGATGT-3′
11F2	11	F	5′-AGAGCATACGCAGCCTGTACCC-3′
11NR	11	R	5′-GTCATCTCAGCTGAGGAG-3′
11F3	11	F	5′-ACACGGCAGGCTGGAGAGCCATGA-3′
11R	11	R	5′-CACCGGAAGAGGAGTAGCTG-3′
CRT4F	13	F	5′-GCAGGCCTCTCTGTCTGAACTT-3′
CRT4E	13	R	5′-GGAGAACAGGGCTGTATGGA-3′ (PCR only)
CRT4EN	13	Nested	5′-AGGCCCCATACAATTTGATG-3′ (for Seq. only)*
2		R	
14F	14	F	5′-TGGCTCCTAGAAGACCCA-3′
14R	14	R	5′-AGAGCCATATGCACGCAC-3′ (PCR only)
14RINT	14	Nested	5′-TCTCGGCCAGATACTGCATC-3′ (for Seq. only)*
		R	
16LC	16	F	5′-AGAGAGTTAGAGTAACTTCAATGTC-3′
16RC	16	R	5′-CTACATGTATAAGGGTGTTT-3′
16LH	16	F	5′-AGGGATAGGGCCTGGGCTTC-3′
16RH	16	R	5′-TAACCTCCACCCCAAGAGAG-3′

*Nested reverse primers (internal to the PCR reverse primer) are used for the cycle sequencing of exons 13 and 14 PCR products.

specificity for the separate MEN2 subtypes, as summarized in Table 20-1. There is significant overlap of clinical symptoms and missense variations at any of the five conserved cysteine codons (609, 611, 618, 620, 634) that may be diagnostic for either MEN2A (95%) or FMTC (80%) (Figure 20-3). Mutations at codon 768 or 804 in exons 13 and 14, respectively, are consistent with the diagnosis of FMTC alone.[26] The substitution of methionine at codon 918 with threonine, M918T, is diagnostic for the MEN2B variant and is observed in 95% to 98% of MEN2B patients. The specificity of *RET* mutation detection for these disorders is nearly 100%.[40]

Interpretation should be conducted within the context of family history and the clinical presentation of the types of endocrine tumors in an affected family member, when available. Specific classification of a patient's MEN2 subtype is essential for the most accurate prognosis and effective clinical management. Distinguishing FMTC or MEN2A may be difficult in small kindreds, with limited or relatively young affected family members, or when scant family medical history is documented because pheochromocytomas or parathyroid hyperplasia may primarily manifest in older relatives.

Identical point mutations have been observed in families with only FMTC and in those with MEN2A as well as in kindreds that have MEN2A plus HSCR. Thus, the genotype-phenotype associations of MEN2 expressivity are not clearly defined by *RET* mutations alone. Within particular kindreds, the clinical presentation is fairly consistent, considering the variable expression of tumor types in MEN2A. As a membrane tyrosine kinase molecule, the RET protein interacts with several coreceptor molecules and ligands including GDNF, neurturin, persephin, and artemin.[41] Base substitutions at the key conserved cysteine codons affect intermolecular associations, disrupt normal signaling, and change the activity or specificity of the kinase, or both.[25] It is hoped that clearer elucidation of concomitant variations in these other "modifier" genes will aid in the definition of more precise genotype-phenotype effects.

MEN2 syndromes account for approximately 25% of MTC cases, based on epidemiological studies of new index cases. The remainder of MTC cases are sporadic, unilateral, or manifest in older adults with negative genetic findings of germline *RET* mutations. A significant proportion of the sporadic tumors (30% to 67%) have somatic M918T mutations of *RET* in the absence of a constitutive germline mutation in leukocyte DNA of the same patient.[42,43] Sporadic MTC also may feature a much more diverse type and distribution of *RET* mutations than MEN2A and FMTC. Disease-causing germline *RET* mutations also have been reported in 1% to 24% of individuals with apparently sporadic MTC, but the lack of previous family history of MEN2 may be attributed to incomplete penetrance of some genetic variations. Somatic *RET* mutations have not been detected in sporadic pheochromocytoma,[43] but this tumor may also present in von Hippel Lindau (VHL) disease and neurofibromatosis type 1 (NF1). Neumann and colleagues showed the utility of additional mutation screening of the *VHL* gene associated with VHL disease in individuals with pheochromocytoma, but no other MEN2 symptoms and no identifiable *RET* gene mutation.[44] Screening for mutations in *VHL* and *NF1* genes should be considered in familial cases of pheochromocytoma with no detectable germline mutation in *RET*. HSCR is generally characterized by a broader spectrum of mutations throughout *RET* than those typical of MEN2, in both the familial and sporadic HSCR presentations.[29,30] The rare occurrence of an MEN2 family with no identifiable *RET* mutation presents the possibility of a clinical misdiagnosis.

Clinical Significance

The primary utility of genetic testing for *RET* mutations is to confirm the diagnosis of MEN2, to provide predictive risk-assessment testing for family members, and to facilitate prenatal diagnosis of MEN2. Early detection of a pathogenic *RET* mutation improves the prognosis for presymptomatic individuals by offering an opportunity for therapeutic intervention prior to advanced disease, metastasis, or both.[45–47] There is reduced morbidity and mortality achieved by increased clinical monitoring, prophylactic thyroidectomy (followed by thyroid hormone replacement therapy plus autotransplantation of the parathyroids), or both.[1,37] This can be a very effective treatment to prevent disease metastasis. Resected thyroid tissue from children and adults with positive genetic findings demonstrates C cell hyperplasia or microscopic foci of malignancy in the absence of biochemical screening abnormalities or clinical symptoms.[47] Thus, genetic testing for *RET* mutations is a more sensitive and specific screening tool than either physiological testing or pathology examination for assessing familial cancer risk for MEN2 in these families.[45–47]

The clinical significance of a positive finding of a *RET* mutation is considerable, given that there is virtually 100% penetrance of these mutations for MTC. But genetic findings cannot predict the age of disease onset; thus, continued surveillance for residual or recurrent MTC plus adrenal tumors is included in the follow-up care of MEN2 patients and asymptomatic carriers. About 50% of MEN2 patients (both subtypes A and B) will develop pheochromocytoma, while 20% to 30% of MEN2A patients may develop parathyroid hyperplasia.[1] Patients with FMTC typically develop only MTC and none of the other clinical manifestations of MEN2A or MEN2B. Each subtype appears to "breed true" within a family. Genetic screening and mutation identification are currently recommended by 5 years of age for children who are at risk in MEN2A families. Children should be screened even earlier if MEN2B has been diagnosed in close relatives, due to the earlier age of onset and aggressive clinical course of this variant.[45,46] Some of the other advantages of DNA testing in MEN2 are that the test is relatively noninvasive and low risk, it is usually better tolerated by the patient than biochemical screening by metabolic challenge,[37] genotype results are not subject to physiologic status, and serial genotype

testing is not necessary. Individuals in MEN2 families who are not carriers of the familial *RET* mutation have the lower, general population risk for sporadic incidence of these endocrine neoplasias and do not require the same frequent monitoring for abnormal thyroid or adrenal function in the absence of clinical symptoms. Several tyrosine kinase inhibitors that inhibit RET are being tested in clinical trials for treatment of MEN2 neoplasms.[48,49]

Quality Control and Laboratory Issues

Overall, germline DNA sequencing for *RET* mutations is favored in clinical molecular laboratories because the mutations are localized, there are limited regions to be sequenced, and the assay has a low false-negative rate (~5%). Some of the limitations of *RET* genetic testing approaches can be attributed to the inherent technical aspects of PCR, DNA sequencing, restriction endonuclease digestion, fragment comigration, and nonpathogenic sequence changes. PCR primer and target fragment design are critical for sensitive detection of mutations by DNA conformation methods. DNA fragment size and base composition directly influence the melting profile of the PCR product and alter the optimal analytic conditions for heteroduplex and homoduplex separation (temperature, buffer concentrations, resolution). DNA sequence polymorphisms within the PCR product at positions other than the ones being scored for pathogenic changes can alter the fragment mobility, prompting false-positive results. For example, a frequent silent polymorphism at *RET* codon 769 of exon 13 changes one leucine codon, CTT, to another leucine codon, CTG. This may produce altered mobility of the fragment, which could be attributed to a change at the adjacent codon 768, consistent with FMTC. Careful comparison with both wild-type and previously characterized mutant alleles as positive controls may assist with the appropriate interpretation of these variations, as will direct DNA sequence confirmation of the same template.

When reviewing DNA sequence electropherograms from automated instruments, one must critically examine both the base sequence indicated by the software and the nucleotide-specific peaks or signals in the trace, because true heterozygous base positions, such as the cysteine codon substitutions in exons 10 and 11 of *RET*, may be difficult to discern. The peaks should overlap with signals of nearly equal intensity, a condition that may be adversely affected by flanking sequence composition. For example, short repeats of G or C nucleotides may display decreasing peak heights for the more distal residues, thereby quenching adjacent signal intensity. Sequencing of both strands of the amplified DNA fragment should confirm the heterozygous base detection and is the standard approach in many laboratories using this technology for clinical molecular testing. The DNA oligonucleotide primers used for direct cycle sequencing reactions may be the same or nested within the PCR fragment but should be positioned 20 to 40 bases from the codons being scored in the mutation screening to optimize clear sequence interpretation. Some fluorescent dye chemistries used in automated DNA sequence techniques yield high background signals (particularly of thymidine nucleotides) resulting from incomplete removal of unincorporated dye terminator reagents prior to electrophoresis.

A few pitfalls specific to *RET* mutation screening have been encountered. We observed preferential (or biased) PCR amplification resulting from specific DNA sequence polymorphisms in primer binding regions. The polymorphisms caused poor primer annealing and failed amplification of the allele sequences adjacent to the polymorphism. We observed this variable performance with some *RET* PCR primer sequences for exon 11 in clinical genetic testing for direct detection of mutations in codon 634. Both affected and at-risk family members were screened by PCR-RFLP for the C634Y mutation (TGC→TAC) in codon 634 in *RET* exon 11 using the *Rsa* I enzyme. One individual had a DNA fragment pattern consistent with two copies of the mutant allele and no wild-type allele, which was unlikely in this autosomal dominant disorder. The patient's disease presentation and course were not dissimilar (neither more severe, nor more extensive, nor at an earlier age) from other affected relatives who were heterozygous for the same allele. Direct DNA sequencing of the PCR product did not resolve the problem. We next selected nonoverlapping PCR primers that flanked the same region of exon 11. These primers amplified both the mutant and wild-type alleles to demonstrate that the individual was heterozygous, and both PCR-RFLP and DNA sequencing detected both alleles. A downstream DNA polymorphism (codon 691) was identified in a *cis* configuration with the wild-type allele at codon 634. The original primer that overlapped this polymorphic region annealed poorly and did not amplify the adjacent wild-type C634 allele, resulting in efficient amplification of only the C634Y allele and a false homozygous mutant pattern (Figure 20-4). Another solution to restore the amplification of both alleles was to lower the annealing temperature slightly for the first primer pair, thus allowing the mismatch and facilitating the efficient PCR amplification of both alleles. Consequently, our standard procedure for *RET* screening employs two independent sets of PCR primers to amplify each exon, followed by enzymatic digestion of each product or DNA sequencing of both strands of the PCR products. Findings from both PCR products must be concordant before a valid clinical evaluation is reported. In our standard practice, positive (same mutant allele) and negative (no mutation) DNA controls are analyzed in parallel every time the assay is performed. PCR with no added nucleic acid is also included and must be free of amplified DNA fragments, indicative of contamination, in all reported assays. In some reactions, small primer-dimer products (<30 bp) may appear, but are distinct from diagnostic DNA patterns.

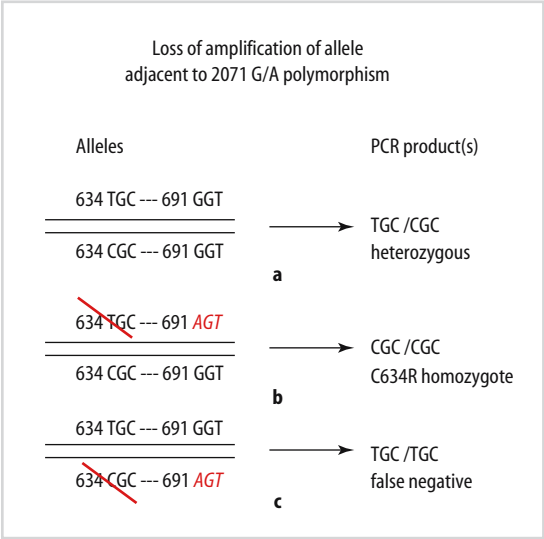


Figure 20-4. Schematic representation of the basis of unequal amplification of codon 634 in *RET*. When the GGT sequence is present at codon 691, there is complete homology with the PCR primer and efficient amplification of the upstream, adjacent 634 codon on both alleles (a). When AGT is present at codon 691 on one *RET* allele, the match with the PCR primer is imperfect and there is poor or absent amplification of the codon 634 region. In the heterozygous condition, this produces preferential amplification of only the 634 allele adjacent to the GGT 691 codon. This could produce either a false-positive (b; homozygous mutant, C634R for example) or a false-negative (c; wild type) genotype depending on whether the AGT variant is positioned in *trans* (b) or *cis* (c) configuration relative to the mutant C634 allele. Only the 634 alleles (mutant or wild type) linked to GGT at codon 691 will be amplified efficiently.

The clustering of these mutations in a limited region of the *RET* protooncogene allows the use of the Pyrosequencing technology as an alternative DNA-based diagnostic method of screening.[50] Reports of novel *RET* DNA alterations affecting noncysteine codons within exons 10 and 11[51] require extended family studies to determine the clinical significance and the *RET* activation status of these variants.

References

1. Phay JE, Moley JF, Lairmore TC. Multiple endocrine neoplasias. *Semin Surg Oncol.* 2000;18:324–332.
2. Eng C. Seminars in medicine of the Beth Israel Hospital, Boston. The RET proto-oncogene in multiple endocrine neoplasia type 2 and Hirschsprung's disease. *N Engl J Med.* 1996;335:943–951.
3. Wautot V, Vercherat C, Lespinasse J, et al. Germline mutation profile of MEN1 in multiple endocrine neoplasias type 1: search for correlations between phenotype and the functional domains of the MEN1 protein. *Hum Mutat.* 2002;20:35–47.
4. Chandrasekharappa SC, Guru SC, Manickam P, et al. Positional cloning of the gene for multiple endocrine neoplasia-type 1. *Science.* 1997;276:404–407.
5. Gagel RF, Marx SJ. In: Larsen PR, Kronenberg H, Melmed S, Polonsky K. eds. Multiple Endocrine Neoplasias. *Williams Textbook of Endocrinology.* 10th ed. Orlando, FL: WB Saunders; 2002:1717–1762.
6. Agarwal SK, Guru SC, Heppner C, et al. Menin interacts with the AP1 transcription factor JunD and represses JunD-activated transcription. *Cell.* 1999;96:143–152.
7. Bertolino P, Radovanovic I, Casse H, et al. Genetic ablation of the tumor suppressor menin causes lethality at mid-gestation with defects in multiple organs. *Mech Dev.* 2003;120:549–560.
8. Crabtre JS, Scacheri PC, Ward JM, et al. Of mice and MEN1: insulinomas in a conditional mouse knockout. *Mol Cell Biol.* 2003;23:6075–6085.
9. Libutti SK, Crabtree JS, Lorang D, et al. Parathyroid gland-specific deletion of the mouse *Men1* gene results in parathyroid neoplasia and hypercalcemic hyperparathyroidism. *Cancer Res.* 2003;63:8022–8028.
10. Guo SS, Sawicki MP. Molecular and genetic mechanisms of tumorigenesis in multiple endocrine neoplasia type 1. *Mol Endocrinol.* 2001;15:1653–1664.
11. Marx SJ. Molecular genetics of multiple endocrine neoplasia types 1 and 2. *Nat Rev Cancer.* 2005;5:367–375.
12. Mutch MG, Dilley WG, Sanjurjo F, et al. Germline mutations in the multiple endocrine neoplasias type 1 gene: evidence for frequent splicing defects. *Hum Mutat.* 1999;13:175–185.
13. Agarwal SK, Kester MB, Debelenko LV, et al. Germline mutations in the MEN1 gene in familial multiple endocrine neoplasias type 1 and related states. *Hum Mol Genet.* 1997;6:1169–1175.
14. Gite S, Lim M, Carlson R, et al. A high-throughput non-isotopic protein truncation test. *Nat Biotech.* 2003;21:194–197.
15. Traverso G, Shuber A, Levin B, et al. Detection of APC mutations in fecal DNA from patients with colorectal tumors. *N Engl J Med.* 2002;346:311–320.
16. Kishi M, Tsukada T, Shimizu S, et al. A large germline deletion of the MEN1 gene in a family with multiple endocrine neoplasia type 1. *Jpn J Cancer Res.* 1998;89:1–5.
17. Lairmore TC, Piersall LD, DeBenedetti MK, et al. Clinical genetic testing and early surgical intervention in patients with multiple endocrine neoplasias type1 (MEN1). *Ann Surg.* 2004;239:637–647.
18. Crabtree JS, Scacheri PC, Ward JM, et al. A mouse model of multiple endocrine neoplasia, type 1, develops multiple endocrine tumors. *Proc Natl Acad Sci U S A.* 2001;98:1118–1123.
19. Cupisti K, Hoppner W, Dotzenrath C, et al. Lack of MEN1 gene mutations in 27 sporadic insulinomas. *Eur J Clin Invest.* 2000;30:325–329.
20. Mulligan LM, Kwok JB, Healey CS, et al. Germ-line mutations of the *RET* proto-oncogene in multiple endocrine neoplasias type 2A. *Nature.* 1993;363:458–460.
21. Donis-Keller H, Dou S, Chi D, et al. Mutations in the RET proto-oncogene are associated with MEN2A and FMTC. *Hum Mol Genet.* 1993;2:851–856.
22. Pichel JG, Shen L, Sheng HZ, et al. Defects in enteric innervation and kidney development in mice lacking GDNF. *Nature.* 1996;382:73–76.
23. Hofstra RM, Landsvater RM, Ceccherini I, et al. A mutation in the *RET* proto-oncogene associated with multiple endocrine neoplasias type 2B and sporadic medullary thyroid carcinoma. *Nature.* 1994;367:375–376.
24. Carlson KM, Dou S, Chi D, et al. Single missense mutation in the tyrosine kinase catalytic domain of the RET proto-oncogene associated with multiple endocrine neoplasia type 2B. *Proc Natl Acad Sci U S A.* 1994;91:1579–1583.
25. Santoro M, Carlomagno F, Romano A, et al. Activation of *RET* as a dominantly transforming gene by germline mutations of MEN2A and MEN2B. *Science.* 1995;267:381–383.
26. Bolino A, Schuffernecker I, Luo Y, et al. *RET* mutations in exons 13 and 14 of FMTC patients. *Oncogene.* 1995;10:2415–2419.
27. Eng C, Smith DP, Mulligan LM, et al. A novel point mutation in the tyrosine kinase domain of the RET proto-oncogene in sporadic medullary thyroid carcinoma. *Oncogene.* 1995;10:509–513.
28. Edery P, Lyonnet S, Mulligan LM, et al. Mutations of the *RET* proto-oncogene in Hirschsprung's disease. *Nature.* 1994;367:378–380.
29. Romeo G, Ronchetto P, Luo Y, et al. Point mutations affecting the tyrosine kinase domain of the *RET* proto-oncogene in Hirschsprung's disease. *Nature.* 1994;367:377–378.

30. Borst MJ, VanCamp JM, Peacock ML, Decker RA. Mutational analysis of multiple endocrine neoplasias type 2A associated with Hirschsprung's disease. *Surgery.* 1995;117:386–391.
31. Grieco M, Santoro M, Berlingieri MT, et al. PTC is a novel rearranged form of the *RET* proto-oncogene and is frequently detected in vivo in human thyroid papillary carcinomas. *Cell.* 1990; 60:557–563.
32. Musholt PB, Musholt TJ, Goodfellow PJ, et al. Cold single-stranded conformational variants for mutation analysis of the *RET* proto-oncogene. *Surgery.* 1997;122:363–370.
33. Peacock ML, Borst MJ, Sweet JD, Decker RA. Detection of *RET* mutations in multiple endocrine neoplasias type 2A and familial medullary thyroid carcinoma by denaturing gradient gel electrophoresis. *Hum Mutat.* 1996;7:100–104.
34. Ceccherini I, Hofstra RMW, Luo Y. DNA polymorphisms and conditions for SSCP analysis of the 20 exons of the *RET* proto-oncogene. *Oncogene.* 1994;9:3025–3029.
35. Siegelman M, Mohabeer A, Fahey TJ III, et al. Rapid, nonradioactive screening for mutations in exons 10, 11, and 16 of the *RET* protooncogene associated with inherited medullary thyroid carcinoma. *Clin Chem.* 1997;43:453–457.
36. Liu W, Smith DI, Rechtzigel KJ, et al. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. *Nucleic Acids Res.* 1998;26:1396–1400.
37. Wells SA Jr, Chi DD, Toshima K, et al. Predictive DNA testing and prophylactic thyroidectomy in patients at risk for multiple endocrine neoplasia type 2A. *Ann Surg.* 1994;220:237–250.
38. Chen X, Zehnbauer B, Gnirke A, et al. Fluorescence energy transfer detection as a homogeneous DNA diagnostic method. *Proc Natl Acad Sci U S A.* 1997;94:10756–10761.
39. Howe JR, Lairmore TC, Mishra SK, et al. Improved predictive test for MEN2, using flanking dinucleotide repeats and RFLPs. *Am J Hum Genet.* 1992;51:1430–1442.
40. Noll WW. Utility of RET mutation analysis in multiple endocrine neoplasia type 2. *Arch Pathol Lab Med.* 1999;123:1047–1049.
41. Baloh RH, Tansey MG, Lampe PA, et al. Artemin, a novel member of the GDNF ligand family, supports peripheral and central neurons and signals through the *GFRa3-RET* receptor complex. *Neuron.* 1998;21:1291–1302.
42. Zedenius J, Larsson C, Bergholm U, et al. Mutations of codon 918 in the *RET* proto-oncogene correlate with poor prognosis in sporadic medullary thyroid carcinomas. *J Clin Endocrinol Metab.* 1995;80: 3088–3090.
43. Eng C, Mulligan LM. Mutations of the *RET* proto-oncogene in the multiple endocrine neoplasias type 2 syndromes, related sporadic tumors, and Hirschsprung disease. *Hum Mutat.* 1997;9:97–109.
44. Neumann HPH, Eng C, Mulligan LM, et al. Consequences of direct genetic testing for germline mutations in the clinical management of families with multiple endocrine neoplasia type II. *JAMA.* 1995;274:1149–1151.
45. Ledger GA, Khosla S, Lindor NM, et al. Genetic testing in the diagnosis and management of multiple endocrine neoplasia type II. *Ann Intern Med.* 1995;122:118–124.
46. Gagel RF, Core GJ, Bugalho MJMG, et al. Clinical use of molecular information in the management of multiple endocrine neoplasia type 2A. *J Intern Med.* 1995;238:331–341.
47. Lips CJM, Landsvater RM, Hoppener JWM, et al. Clinical screening as compared with DNA analysis in families with multiple endocrine neoplasia type 2A. *N Engl J Med.* 1994;331:828–835.
48. Carlomagno F, Vitagliano D, Guida T, et al. ZD6474, an orally available inhibitor of KDR tyrosine kinase activity, efficiently blocks oncogenic RET kinases. *Cancer Res.* 2002;62:7284–7290.
49. Moley JF. Medullary thyroid carcinoma. *Curr Treat Options Oncol.* 2003;4:339–347.
50. Kruckeberg KE, Thibodeau SN. Pyrosequencing technology as a method for the diagnosis of multiple endocrine neoplasia type 2. *Clin Chem.* 2004;50:522–529.
51. Ahmed SA, Snow-Bailey K, Highsmith WE, et al. Nine novel germline gene variants in the *RET* proto-oncogene identified in twelve unrelated cases. *J Mol Diagn.* 2005;7:283–288.

Chapter 21

The Neurofibromatoses

Karen Stephens

Neurofibromatosis type 1 and neurofibromatosis type 2 are two distinct genetic disorders that predispose to the development of tumors primarily of the nervous system (Table 21-1).[1] A recently recognized third form of neurofibromatosis, known as schwannomatosis,[2] is not included in this review, as molecular genetic testing is not available for this disorder.

NEUROFIBROMATOSIS TYPE 1

Molecular Basis of Disease

Neurofibromatosis type 1 (NF1) is an autosomal dominant progressive disorder with high penetrance but extremely variable expressivity (reviewed in References 1, 3, and 4). The cardinal features are café au lait macules, intertrigenous freckling, Lisch nodules, and multiple neurofibromas, although numerous other features and complications are common. Criteria for diagnosis of NF1 were established in 1987 by a National Institutes of Health Consensus Conference and are widely used.[5] Neurofibromas are benign nerve sheath tumors that arise on peripheral nerves. Cutaneous neurofibromas develop in virtually all cases of NF1, typically appear in the second decade of life, grow slowly, increase in number with age, and are considered at low risk for transformation to a malignant peripheral nerve sheath tumor (MPNST; previously known as neurofibrosarcoma). In contrast, diffuse plexiform neurofibromas and deep nodular plexiform neurofibromas are considered at increased risk for transformation to MPNST. Individuals affected with NF1 have a lifetime risk for MPNST of 8% to 13%.[6] Other neoplasms epidemiologically associated with NF1 include medulloblastoma, pheochromocytoma, astrocytoma, and adenocarcinoma of the ampulla of Vater. Children affected with NF1 are at increased risk for optic pathway and brainstem gliomas, rhabdomyosarcomas, and malignant myeloid leukemias. NF1 patients also are at increased risk for a second malignancy, some of which may be treatment related.

NF1 is caused by inactivating mutations in one copy of the *NF1* gene resulting in haploinsufficiency for the gene product neurofibromin (Table 21-1). About 85% to 90% of constitutional mutations are nonsense, splicing, and missense; they are distributed throughout the gene, although some exons appear to be mutation rich (Figure 21-1). An estimated 5% of mutations are large contiguous gene deletions typically of 1.4 megabases (Mb) that delete one entire *NF1* allele (reviewed in Reference 7). About one half of cases are familial (inherited from an affected parent) and one half are sporadic, resulting from a de novo *NF1* mutation. An unknown fraction of sporadic cases are due to postzygotic mutation of the *NF1* gene, which complicates mutation detection and genetic counseling. Neurofibromin functions as a negative regulator of the *RAS* oncogene by stimulating the conversion of active guanosine 5′-triphosphate (GTP)-bound RAS to the inactive guanosine 5′-diphosphate (GDP)-bound form by hydrolyzing GTP. Biochemical, cell culture, and genetic studies in both NF1 patients and mouse models are consistent with a model whereby a somatic mutation inactivates the remaining functional *NF1* gene (leading to increased activated RAS) in a progenitor Schwann cell as an early, probably initiating, event in the development of neurofibromas (reviewed in Reference 8). Biallelic inactivation of *NF1* also occurs in other types of progenitor cells that give rise to NF1-associated tumors such as gliomas and myeloid malignancies.

Clinical Utility of Testing

A diagnosis of NF1 can almost always be made based on clinical findings, particularly after 8 years of age. Clinical DNA-based testing is available from many licensed clinical laboratories (see GeneTests, http://www.genetests.org/). Testing is not typically used for diagnostic purposes, but can be useful for confirming a clinical diagnosis, reproductive counseling, and prenatal or preimplantation diagnosis. Blanket recommendations for diagnostic testing for NF1 cannot be made because the sensitivity of clinical

Table 21-1. Comparison of Features of the NF1 and NF2 Disorders

Feature	Neurofibromatosis 1	Neurofibromatosis 2
Alternate name	Peripheral neurofibromatosis; von Recklinghausen neurofibromatosis	Central neurofibromatosis; bilateral acoustic neuroma
OMIM accession number*	162200	101000
Mode of inheritance	Autosomal dominant	Autosomal dominant
Frequency of disorder	1/3000–1/4000	1/25,000
Fraction of sporadic cases	30–50%	~50%
Gene symbol	*NF1*	*NF2*
Chromosomal location	17q11.2	22q12.2
Gene size; transcript size	~350 kb; ~11–13 kb†	~110 kb; 2 kb†
GenBank accession no. (gene; cDNA)‡	NT_010799; NM_000267	Y18000; NM••000268
Number of exons	60	17
Tissue expression pattern	Widely expressed	Widely expressed
Protein product (size; no. of residues)	Neurofibromin (>220 kDa; 2818)	Merlin; also known as schwannomin (65 kDa; 595)
Normal functions of protein	Tumor suppressor; negative regulator of *RAS* oncogene	Tumor suppressor; associates with proteins of the cytoskeleton
Commonly associated tumors	Neurofibroma, MPNST, optic pathway and brainstem gliomas	Bilateral vestibular schwannomas, schwannomas of other central and peripheral nerves, meningiomas
Animal models	Mouse, fruit fly	Mouse, fruit fly

*Online Mendelian Inheritance in Man [database online] (http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?db=OMIM).
†Alternative splicing produces transcripts of varying lengths.
‡See Gene Lynx Human (http://www.genelynx.org/) for a compilation of and hyperlinks to gene, protein structure, and genomic resources.
MPNST, malignant peripheral nerve sheath tumor.

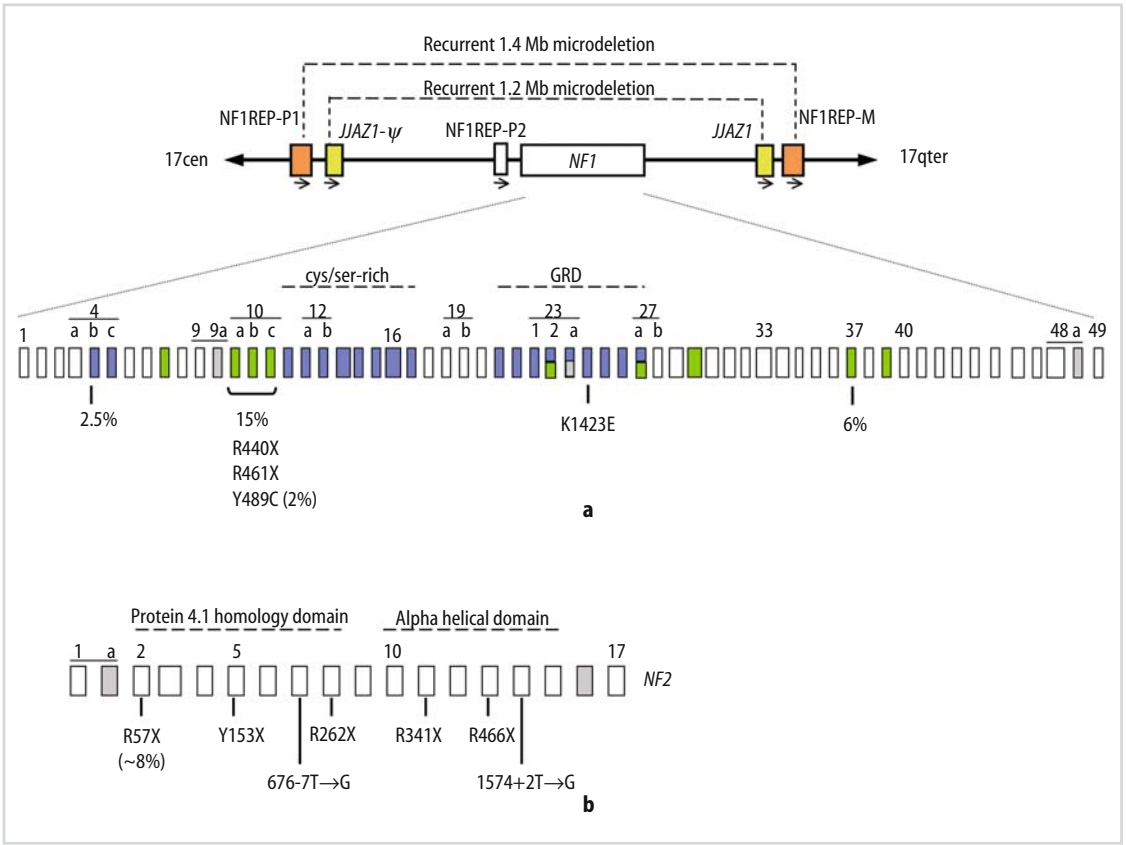


Figure 21-1. *NF1* and *NF2* genes: genomic structure and mutations. (a) At the top is a schematic of the *NF1* gene region at chromosome segment 17q11.2. The 350 kb *NF1* gene is flanked by two different sets of directly oriented paralogs. The 51 kb paralogs NF1REP-P1 (previously termed NF1REP-P) and NF1REP-M (orange boxes) and the 46 kb *JJAZ1* gene and pseudogene (−*ψ*) (yellow boxes) are shown. Homologous recombination between NF1REP elements results in a recurrent 1.4 megabase (Mb) microdeletion, while recombination between the *JJAZ1* paralogs results in a recurrent 1.2 Mb microdeletion.[11,16] NF1REP-P2 is a partial element with a limited role in mediating *NF1* microdeletions. The 60 exons of *NF1* are represented by boxes (not to scale), and exon numbering is sequential except as indicated. The GTPase-activating protein-related domain or GRD (exons 21-27a) and a cysteine/serine-rich domain with three cysteine pairs suggestive of ATP binding (exons 11-17) are indicated. Gray boxes indicate alternatively spliced exons that vary in abundance in different tissues. Mutations have been identified in virtually every exon. Exons where mutations are apparently in greater abundance than expected are indicated (green boxes).[21,22,25] In one study, exons 7, 10a, b, c, 23-2, 27a, 29, and 37 accounted for 30% of mutations, 15 of which were in exons 10a, b, and c, which harbor three recurrent mutations, including Y489C, which alone accounts for approximately 2% of mutations.[21] Blue boxes indicate exons that have clusters of missense or single base or codon deletion mutations, or both.[22] Some of the recurrent, although still infrequent, mutations are given below the exons. (b) The 17 exons of the *NF2* gene are represented by boxes (drawing not to scale), and the exon numbering is sequential. The protein 4.1-homology domain thought to mediate binding to cell surface glycoproteins (exons 2-8), the α-helical domain (exons 10-15), and the unique C-terminus (exons 16-17) are shown. Gray boxes indicate alternatively spliced exons; the inclusion of exon 16 creates an alternate termination codon resulting in a slightly truncated protein. Mutations have been identified involving each exon except for 16. Selected recurrent mutations found in a limited survey of references cited in the text are indicated. In several studies, R57X occurred in 8% of familial constitutional mutations.

diagnosis for NF1 is very high and the sensitivity of molecular testing is not 100%. Furthermore, the benefits of diagnostic testing are subjective and may differ from family to family. Early planning is necessary for couples considering prenatal diagnosis using amniocytes or chorionic villus sampling or preimplantation diagnosis. These tests are available only when the pathogenic familial germline mutation (or the predisposing haplotype in the case of linkage testing) has been identified previously in an affected parent, a process that can require weeks or months.

The primary genetic counseling issue related to molecular testing of *NF1* is the inability to predict the severity or course of the disorder in a patient or fetus. Even among family members who carry the same *NF1* mutation, there can be considerable variation in clinical manifestations and complications. For the majority of cases, there is no correlation between genotype and phenotype. For the approximately 5% of individuals who carry a constitutional *NF1* deletion (most commonly 1.4 Mb), there is a 2-fold increased lifetime risk of MPNST,[9] a predisposition to childhood overgrowth,[10] an early age of onset with excessive numbers of cutaneous neurofibromas, numerous internal neurofibromas, learning disabilities, vascular anomalies, and astrocytomas (References 7 and 11 to 14 and references therein). A recommendation for routine testing for *NF1* microdeletion has been proposed, with follow-up for increased suspicion for MPNST.[9]

NF1 testing may be useful to confirm a diagnosis in a patient with equivocal findings, such as a child who has a few café au lait macules and carries a presumptive diagnosis of NF1. In such cases, it is important to realize that the sensitivity and specificity of testing patients who do not fulfill the NIH diagnostic criteria for a diagnosis of NF1 is unknown but is likely to be quite low.

For unaffected parents of a child with sporadic NF1, recurrence risk is a concern. Although thought to be rare, germline mosaicism has been reported in an asymptomatic parent of a child with sporadic NF1.[15] Therefore, there is a small, but unknown, increased risk of recurrence even if the child's pathogenic mutation is not detected in the genomic DNA from parental leukocytes. Although the frequency is unknown, sporadic NF1 cases with postzygotic mutations resulting in somatic mosaicism may not be as rare as once thought. One study suggests that among NF1 microdeletion cases, the frequency of somatic deletion may be very high (~40%).[16] Assuming that 10% of NF1 cases have microdeletions, a frequency of 4% mosaicism is expected in the general NF1 population. This is certainly an underestimate, as it does not consider mosaicism for intragenic *NF1* mutations. Mosaic individuals carry the *NF1* mutation in only a fraction of their cells, depending on the developmental interval and the cell type in which the mutation occurred. The phenotype of mosaic individuals ranges from localized (segmental) disease to mild or severe generalized disease.[16,17] The sensitivity of mutation detection may be lower due to an increased signal to noise ratio, that is, a low level of a mutant allele in a background of two normal *NF1* alleles. Offspring that inherit an *NF1* mutation from a parent with mosaicism, however, will have a constitutional *NF1* mutation and may have more severe disease than their mosaic parent. Genetic counseling regarding the clinical and reproductive implications of *NF1* mosaicism is highly recommended.[17]

Available Assays

Mutation of the *NF1* gene is the only known cause of the disorder. Molecular tests for diagnostic, prenatal, and preimplantation diagnosis are available. The choice of assay and testing laboratory depends upon the reason for referral and mutation types and detection rates of their assay(s).

Fluorescent in situ hybridization (FISH; see chapter 2) with *NF1* probes of either metaphase or interphase white blood cells is the optimal test to rule out or confirm the approximately 5% of cases due to a submicroscopic NF1 microdeletion (Figure 21-1).[18] In the future, a first-tier test may employ an *NF1* deletion junction-specific polymerase chain reaction (PCR) assay.[19] The recent availability of high-resolution genomic microarrays of the NF1 deletion region will facilitate clinical testing by array–comparative genomic hybridization (CGH),[20] which may become clinically important in the future if deletions involving a subset of genes predispose to certain manifestations. The sensitivity of deletion-specific PCR and array-CGH assays to detect low-level *NF1* deletion mosaicism will need to be determined. Routine cytogenetic analysis is of limited clinical utility, as the *NF1* microdeletions are submicroscopic, and translocation and rearrangement involving *NF1* are extremely rare.

Linkage analysis is an indirect test that tracks the inheritance of the mutant *NF1* allele in members of a family. This may be the quickest, most economical NF1 test for at-risk individuals and fetuses of families that fulfill the testing criteria. The primary requirement is the availability and cooperation of multiple family members whose NF1 status is known by detailed clinical evaluation. Multiple *NF1* intragenic polymorphic markers are available that facilitate identification and tracking of the predisposing haplotype in a family and provide the specificity for linkage testing.

Efficient detection of subtle intragenic *NF1* gene mutations, for purposes of diagnostic testing or mutation typing for prenatal or preimplantation diagnosis, is complicated by the large number of exons and large size of the gene (Table 21-1), variation in type and distribution of mutations, and large fraction of private mutations. About 70% to 80% of mutations result in a premature translation termination codon, with nonsense and splicing defects being

the most common.[21] These mutations can be detected by the protein truncation test (PTT; see chapter 2), which detects truncated neurofibromin polypeptides synthesized by in vitro translation of multiple overlapping *NF1* complementary DNA (cDNA) segments. A detection rate of about 80% can be attained with an optimized PTT testing protocol (see Laboratory Issues below). The majority of such mutations are private to each individual or family, although there are recurrent mutations that may account for, at most, a few percent of cases (Figure 21-1a).

About 10% of *NF1* mutations are missense or in-frame insertions or deletions of a few nucleotides,[21,22] some of which show clustering (Figure 21-1a). Their identification requires direct sequence analysis of *NF1* exons and splice junctions in genomic DNA or cDNA segments. Prospective testing of NF1 subjects by direct genomic sequence analysis revealed a detection rate of 89%, which is more streamlined than PTT testing and allows for automation.[23] Various mutation scanning techniques of *NF1* genomic DNA or cDNA are employed by clinical laboratories, including denaturing high-performance liquid chromatography (DHPLC), temperature gradient gel electrophoresis (TGGE), single-strand conformation polymorphism (SSCP), and heteroduplex analysis (HA) (see chapter 2). Although high detection rates are reported in the literature using DHPLC (72–95%),[24,25] it is important to realize that the detection rates for mutation scanning protocols will be laboratory specific due to the degree of optimization of the specific technique. A survey of clinical laboratories is recommended prior to sample submission. DHLPC has the advantages of using genomic DNA and high-throughput capability compared to the cDNA/gel-based PTT; however, a recently reported high-throughput PTT may be available for clinical testing in the future.[26]

Interpretation of Test Results

The detection of a truncated neurofibromin polypeptide by PTT can result in false positives.[21] High specificity requires identifying the underlying mutation at the genomic DNA or cDNA level, or both, since false positives can arise during sample handling (see Laboratory Issues below). The interpretation of missense and subtle in-frame alterations as pathogenic mutations rather than neutral polymorphisms is complicated by the lack of a functional assay for neurofibromin. Apparent recurrence of a putative mutation requires careful study of the literature, since not all *NF1* mutation studies sequenced the entire gene. No comprehensive *NF1* mutation database is available; however, some mutations have been submitted to the Human Gene Mutation Database (http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html), and the largest NF1 database is actively managed and analyzed by Jan Friedman (http://www.medgen.ubc.ca/friedmanlab/).

Although most likely rare, affected family members with different independent *NF1* inactivating mutations have been reported,[27] presumably a reflection of the high mutation rate of the gene (~10^5/gamete/generation). The interpretation of FISH with *NF1* probes can be complicated by mosaicism for an *NF1* microdeletion; therefore, an appropriate number of cells must be analyzed.[16] The frequency of mosaicism for an *NF1* mutation is not known; however, this is likely the underlying mechanism for patients with segmental or localized signs of the disorder.

Laboratory Issues

Optimal detection of mutations that predict a truncated neurofibromin polypeptide occurs when the nonsense-mediated decay pathway is partially inhibited, thereby increasing the ratio of mutant transcripts with a premature termination codon to normal transcripts. A protein synthesis inhibitor, such as puromycin, in the culture medium is effective for Epstein-Barr virus (EBV)-transformed lymphoblasts or phytohemaglutinin-stimulated primary lymphocytes.[21,28] Furthermore, blood handling and shipping protocols must be used to reduce false positives in PTT resulting from environmental effects such as cold shock[29] or delay in messenger RNA isolation.[21,28]

There is no standardized proficiency program of interlaboratory comparison for *NF1* testing. Performance assessment must be conducted by participation in ungraded proficiency survey programs, split sample analysis with other laboratories, or other suitable and documented proficiency-testing methods. No *NF1* testing kits, probes, or controls are approved or cleared by the Food and Drug Administration. Intronic primers for amplification of *NF1* exons and associated splice junctions that apparently do not coamplify the *NF1* pseudogene fragments have been reported.[23,30] Two other factors require consideration during test development and interpretation. Reports in support of,[31] and in opposition to,[32] an apparent tandem duplication of the *NF1* gene region have been published. In addition, transcriptional activity from *NF1* pseudogenes or pseudogene fragments has been reported.[33] Some issues related to *NF1* testing have been reviewed recently.[34]

NEUROFIBROMATOSIS TYPE 2

Molecular Basis of Disease

The development of bilateral vestibular schwannomas is a hallmark of neurofibromatosis type 2 (NF2). Other commonly associated tumors include schwannomas of other central, spinal, and peripheral nerves and meningiomas (reviewed in References 1 and 35 to 38). This is a life-threatening disorder due to the location of the tumors, along with the propensity for development of multiple

tumors. Most patients become completely deaf and can have poor balance and vision, and weakness. The mean age of onset is 18 to 24 years and the mean age of death is 36 years. The age at onset of symptoms and age at diagnosis are predictors of vestibular schwannoma growth rates and risk of death (Reference 39 and references therein). Ependymomas and astrocytomas occur less frequently and are usually indolent central nervous system tumors. Patients affected with NF2 are at minimal increased risk for malignancy. Juvenile posterior subcapsular cataract is a common nontumor manifestation. The disorder may be underdiagnosed in children who present with ocular and skin manifestations. Early diagnosis improves management, which is primarily surgical and radiological. Modifications to the criteria for a diagnosis of NF2, initially established by the 1987/1991 National Institute of Health Consensus Conference, have been proposed to increase the specificity.[40] A consensus statement on management of the NF2 patient and family was recently published.[41]

NF2 is caused by haploinsufficiency for the tumor suppressor merlin (also known as schwannomin), the protein product of the *NF2* gene (Table 21-1). About one half of patients are the first case of NF2 in the family. These sporadic cases result from de novo mutation of the *NF2* gene, a significant fraction of which are postzygotic mutations that result in mosaicism. The majority of constitutional germline mutations are private, predict the truncation of merlin, and are distributed throughout the gene (see Clinical Utility of Testing below). In NF2 patients, a vestibular schwannoma develops from a progenitor Schwann cell that carries a somatic inactivating mutation in the single remaining *NF2* gene. Merlin is a protein of the cytoskeleton whose normal function remains to be determined, although it is known to associate with transmembrane proteins important in adhesion, proteins involved in signaling pathways, and cytoskeletal proteins (reviewed in Reference 42).

Clinical Utility of Testing

DNA-based clinical testing for NF2 is available (see GeneTests, http://www.genetests.org) and primarily is used for presymptomatic testing of at-risk individuals, typically young children of an affected parent. An early diagnosis of NF2 may improve outcome, and at-risk children who did not inherit the *NF2* mutation can be spared worry, costly brain imaging, and audiologic screening. Genetic counseling is recommended prior to testing presymptomatic at-risk children. Testing also is useful to confirm a clinical diagnosis, which may be most helpful in sporadic cases of NF2, particularly children who present with ocular or skin manifestations or adults with equivocal findings or mild disease. Some of these cases may be mosaic for an *NF2* mutation, as the estimated frequency of mosaicism is high (16.7% to 24.8% of sporadic cases).[43] Genetic counseling regarding the clinical and reproductive implications of *NF2* mosaicism is recommended.[17] Testing also is useful for reproductive counseling and prenatal or preimplantation diagnosis.[44] Prenatal diagnosis of NF2 using amniocytes or chorionic villus tissue is available only in cases where the pathogenic *NF2* germline familial mutation (or predisposing haplotype in the case of linkage testing) has been identified previously in an affected parent. For preimplantation genetic diagnosis, the specific parental *NF2* mutation must be known. Prepregnancy planning is important for couples considering prenatal diagnosis or preimplantation diagnosis. For sporadic NF2 patients undergoing surgery, it is advisable to freeze a portion of the tumor, which may be valuable at a later date for mutation identification if the patient is mosaic.

The primary genetic counseling issues regard predicting the course of the disorder and recurrence risks. There are genotype-phenotype correlations, but they cannot predict the age of onset or the course of disease for an individual patient. About 50% of *NF2* mutations are nonsense or frameshift, with about 24% splice site, 11% to 30% submicroscopic deletion, and 5% missense.[45] Typically, constitutional frameshift and nonsense mutations are associated with more severe NF2, defined by earlier age at onset and higher frequency and mean number of tumors.[46,47] Constitutional missense and small in-frame mutations are associated with mild disease,[47] and mutations in splice donor and acceptor sites result in variable clinical outcomes.[48] Interestingly, individuals with *NF2* splice-site mutations in exons 1 through 5 have an earlier age at onset and greater numbers of intracranial meningiomas compared to those with splice-site mutations in exons 11 through 15 (Figure 21-1b).[45] The type of constitutional germline *NF2* mutation also is correlated with the number of NF2-associated nonvestibular nervous system tumors including intracranial meningiomas, spinal tumors, and peripheral nerve tumors.[39] Individuals with constitutional nonsense or frameshift *NF2* mutations have significantly more of these tumors than individuals carrying missense, splice-site, or deletion mutations or somatic mosaicism.

Recurrence risks for asymptomatic parents of an affected child are unknown but are somewhat greater than the population risk, due to the possibility of germline mosaicism in a parent.[49] For mosaic patients, the risk of transmitting NF2 to offspring is ≤50%, depending on the proportion of gametes that carry the *NF2* mutation.[17] Offspring that do inherit the mutation, however, will have a constitutional *NF2* mutation and may have more severe disease than their mosaic parent. Testing asymptomatic parents of a child with NF2 has the potential to identify a mosaic mutation.

Available Assays

Mutation of the *NF2* gene is the only known cause of this disorder. Linkage analysis is clinically available for at-risk individuals and fetuses with multiple family members of

unambiguous clinical status regarding NF2 disease who are willing to participate in the testing process. The availability of highly informative intragenic *NF2* polymorphisms increases the specificity of this method. For certain families, linkage analysis will be the most cost- and time-effective test that gives a definitive diagnosis. It can sometimes be an option when mutation-scanning or sequencing test results are negative. See "Interpretation of Test Results," below, for cautions regarding linkage test interpretation.

Identifying an *NF2* mutation typically requires a multipronged testing protocol due to the high frequency of private constitutional mutations, the high frequency of postzygotic mutations, the different types of *NF2* mutations, and the distribution of mutations throughout the gene. Wallace et al.[50] describe a comprehensive testing service that includes four PCR reactions using a meta-PCR technique to link the amplicons into chimeric concatemers for direct sequencing, gene dosage PCR for deletions, loss of heterozygosity (LOH) studies, and subsequent sequencing of the gene in tumor tissue. In prospective studies, this approach yielded an 88% detection rate in familial NF2 cases and a 59% detection rate in sporadic NF2 cases. Direct sequencing or exon-scanning techniques (e.g., SSCP, TGGE, and HA; see chapter 2) of DNA from peripheral leukocytes, followed by direct sequencing to identify the underlying *NF2* mutation, generally have a lower detection rate.[46,47,51–53] The detection rate of either sequencing or exon scanning methods is significantly lower (34–51%) in sporadic cases in part due to the high frequency of postzygotic *NF2* mutations, which can be masked by the presence of normal alleles.[47,48,51,53] The mutation detection rate of mosaic cases can be increased significantly by analysis of tumor tissue.

Because schwannomas are clonal tumors with minimal cellular admixture, *NF2* mutations can be detected at high frequency in tumor tissue. Testing of tumor tissue is available clinically (see GeneTests, http://www.geneclinics.org/) and is most useful in cases where a mutation is not detected in primary lymphoblasts, where clinical manifestations are suggestive of somatic mosaicism, or where constitutional tissue is not available.[43,50,53] Moyhuddin et al.[53] nearly doubled the mutation detection rate among mosaic cases using vestibular schwannoma tissue rather than peripheral leukocytes. Mutations are likely to be germline (rather than somatic) if the identical mutation is detected in two or more pathologically or anatomically distinct tumors or if a tumor shows LOH for *NF2* intragenic or flanking loci, while constitutional tissue is heterozygous at these loci. Mutational analysis of tumor tissues is expected to have the greatest sensitivity for *NF2* somatic mosaic mutations[43] and sporadic cases with negative results from mutation scanning or sequencing tests.[54]

Efficient detection of the 11% to 30% of constitutional *NF2* deletions (typically multiexonic in nature) has been accomplished using numerous techniques, including FISH, various gene dosage PCR assays, multiplex ligation probe amplification (MLPA), and high-resolution genomic arrays.[53,55–57]

Note that for mutation scanning tests and deletion-detection assays, detection rates will be laboratory specific due to the varying degrees of optimization of the technique; therefore, a survey of testing laboratories is recommended prior to sample submission.

Interpretation of Test Results

Interpretation of the results of exon-scanning tests requires identifying the underlying *NF2* mutation at the genomic DNA or cDNA level, or both, to avoid false positives. Functional assays for merlin have been developed that can provide insight into the interpretation of missense and subtle in-frame alterations as pathogenic mutations rather than neutral polymorphisms;[58–60] however, such assays may not be part of a clinical testing protocol. An international *NF2* mutation database is available (Neurofibromatosis 2 Mutation Databases, http://uwcmml1s.uwcm.ac.uk/uwcm/mg/nf2/), and this site also recognizes the United Kingdom population-based registry. Some mutations are also detailed in the Human Gene Mutation Data-base Cardiff (http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html). Somatic mosaicism or *NF2* gene deletions must be considered in patients who have a negative mutation test using DNA from peripheral leukocytes, regardless of the severity of their manifestations. The risk of recurrence from mosaic parent to offspring is considered very low if an *NF2* mutation cannot be identified in the parent.[53] NF2 linkage tests should consider excluding the first affected member in a family; if they are mosaic, the linkage results will be misleading in the next generation.[61] For similar reasons, linkage analysis for presymptomatic testing of subjects in this "next" generation should be performed with caution.

Laboratory Issues

There is no standardized proficiency program of interlaboratory comparison for *NF2* testing; performance assessment must be conducted by participation in ungraded proficiency survey programs, split sample analysis with other laboratories, or other suitable and documented proficiency-testing methods. No *NF2* testing kits, probes, or controls are approved or cleared by the Food and Drug Administration. Direct gene sequencing may not be the optimal test to detect *NF2* mosaic mutations in lymphoblasts, since reliable detection of a low-level point mutation will be difficult. Exon-scanning techniques that are semiquantitative, such as TGGE, will detect relative intensity differences between heteroduplexes and homoduplexes that suggest possible mosaicism.[51] Depending on age, fixation, and storage conditions, some tumors may not yield nucleic acid of sufficient quality for mutational analysis.

References

1. Friedman JM, Butmann DH, MacCollin M, Riccardi VM, eds. *Neurofibromatosis: Phenotype, Natural History, and Pathogenesis.* 3rd ed. Baltimore: Johns Hopkins University Press; 1999.
2. MacCollin M, Chiocca EA, Evans DG, et al. Diagnostic criteria for schwannomatosis. *Neurology.* 2005;64:1838–1845.
3. Friedman JM. Neurofibromatosis 1: clinical manifestations and diagnostic criteria. *J Child Neurol.* 2002;17:548–554.
4. Friedman JM. Neurofibromatosis 1. In: GeneReviews at GeneTests: Medical Genetics Information Resource [database online]. Seattle: University of Washington; 1997–2005. Available at: http://www.genetests.org. Accessed October 20, 2005. Updated October 5, 2004.
5. DeBella K, Szudek J, Friedman JM. Use of the national institutes of health criteria for diagnosis of neurofibromatosis 1 in children. *Pediatrics.* 2000;105(pt 1):608–614.
6. Evans DG, Baser ME, McGaughran J, Sharif S, Howard E, Moran A. Malignant peripheral nerve sheath tumours in neurofibromatosis 1. *J Med Genet.* 2002;39:311–314.
7. Stephens K. Neurofibromatosis 1. In: Lupski JR, Stankiewicz P, eds. *Genomic Disorders: The Genomic Basis of Genetic Disease.* Trenton, NJ: Humana Press; 2006, pp. 207–219.
8. Dasgupta B, Gutmann DH. Neurofibromatosis 1: closing the GAP between mice and men. *Curr Opin Genet Dev.* 2003;13:20–27.
9. De Raedt T, Brems H, Wolkenstein P, et al. Elevated risk for MPNST in NF1 microdeletion patients. *Am J Hum Genet.* 2003;72:1288–1292.
10. Spiegel M, Oexle K, Horn D, et al. Childhood overgrowth in patients with common NF1 microdeletions. *Eur J Hum Genet.* 2005;13:883–888.
11. Dorschner MO, Sybert VP, Weaver M, Pletcher BA, Stephens K. NF1 microdeletion breakpoints are clustered at flanking repetitive sequences. *Hum Mol Genet.* 2000;9:35–46.
12. Gutmann DH, James CD, Poyhonen M, et al. Molecular analysis of astrocytomas presenting after age 10 in individuals with NF1. *Neurology.* 2003;61:1397–1400.
13. Venturin M, Guarnieri P, Natacci F, et al. Mental retardation and cardiovascular malformations in NF1 microdeleted patients point to candidate genes in 17q11.2. *J Med Genet.* 2004;41:35–41.
14. Kehrer-Sawatzki H, Kluwe L, Funsterer C, Mautner VF. Extensively high load of internal tumors determined by whole body MRI scanning in a patient with neurofibromatosis type 1 and a non-LCR-mediated 2-Mb deletion in 17q11.2. *Hum Genet.* 2005;116:466–475.
15. L'Azaro C, Gaona A, Lynch M, Kruyer H, Ravella A, Estivill X. Molecular characterization of the breakpoints of a 12-kb deletion in the NF1 gene in a family showing germ-line mosaicism. *Am J Hum Genet.* 1995;57:1044–1049.
16. Kehrer-Sawatzki H, Kluwe L, Sandig C, et al. High frequency of mosaicism among patients with neurofibromatosis type 1 (NF1) with microdeletions caused by somatic recombination of the JJAZ1 gene. *Am J Hum Genet.* 2004;75:410–423.
17. Ruggieri M, Huson SM. The clinical and diagnostic implications of mosaicism in the neurofibromatoses. *Neurology.* 2001;56:1433–1443.
18. Leppig KA, Viskochil D, Neil S, et al. The detection of contiguous gene deletions at the neurofibromatosis 1 locus with fluorescence in situ hybridization. *Cytogenet Cell Genet.* 1996;72:95–98.
19. Lopez-Correa C, Dorschner M, Brems H, et al. Recombination hotspot in NF1 microdeletion patients. *Hum Mol Genet.* 2001;10:1387–1392.
20. Mantripragada KK, Thuresson AC, Piotrowski A, et al. Identification of novel deletion breakpoints bordered by segmental duplications in the NF1 locus using high-resolution array-CGH. *J Med Genet.* 2006;43:28–38.
21. Messiaen LM, Callens T, Mortier G, et al. Exhaustive mutation analysis of the NF1 gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects. *Hum Mutat.* 2000;15:541–555.
22. Fahsold R, Hoffmeyer S, Mischung C, et al. Minor lesion mutational spectrum of the entire NF1 gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain. *Am J Hum Genet.* 2000;66:790–818.
23. Mattocks C, Baralle D, Tarpey P, Ffrench-Constant C, Bobrow M, Whittaker J. Automated comparative sequence analysis identifies mutations in 89% of NF1 patients and confirms a mutation cluster in exons 11–17 distinct from the GAP related domain. *J Med Genet.* 2004;41:e48.
24. Han SS, Cooper DN, Upadhyaya MN. Evaluation of denaturing high performance liquid chromatography (DHPLC) for the mutational analysis of the neurofibromatosis type 1 (NF1) gene. *Hum Genet.* 2001;109:487–497.
25. De Luca A, Buccino A, Gianni D, et al. NF1 gene analysis based on DHPLC. *Hum Mutat.* 2003;21:171–172.
26. Gite S, Lim M, Carlson R, Olejnik J, Zehnbauer B, Rothschild K. A high-throughput nonisotopic protein truncation test. *Nat Biotechnol.* 2003;21:194–197.
27. Klose A, Peters H, Hoffmeyer S, et al. Two independent mutations in a family with neurofibromatosis type 1 (NF1). *Am J Med Genet.* 1999;83:6–12.
28. Wimmer K, Eckart M, Rehder H, Fonatsch C. Illegitimate splicing of the NF1 gene in healthy individuals mimics mutation-induced splicing alterations in NF1 patients. *Hum Genet.* 2000;106:311–313.
29. Ars E, Serra E, de la Luna S, Estivill X, Lazaro C. Cold shock induces the insertion of a cryptic exon in the neurofibromatosis type 1 (NF1) mRNA. *Nucleic Acids Res.* 2000;28:1307–1312.
30. Li Y, O'Connell P, Breidenbach HH, et al. Genomic organization of the neurofibromatosis 1 gene (NF1). *Genomics.* 1995;25:9–18.
31. Gervasini C, Bentivegna A, Venturin M, Corrado L, Larizza L, Riva P. Tandem duplication of the NF1 gene detected by high-resolution FISH in the 17q11.2 region. *Hum Genet.* 2002;110:314–321.
32. Kehrer-Sawatzki H, Messiaen L. Interphase FISH, the structure of reciprocal translocation chromosomes and physical mapping studies rule out the duplication of the NF1 gene at 17q11.2. A reply. *Hum Genet.* 2003;113:188–190.
33. Yu H, Zhao X, Su B, et al. Expression of NF1 pseudogenes. *Hum Mutat.* 2005;26:487–488.
34. Thomson SA, Fishbein L, Wallace MR. NF1 mutations and molecular testing. *J Child Neurol.* 2002;17:555–561; discussion 71–72, 646–651.
35. Ruggieri M, Iannetti P, Polizzi A, et al. Earliest clinical manifestations and natural history of neurofibromatosis type 2 (NF2) in childhood: a study of 24 patients. *Neuropediatrics.* 2005;36:21–34.
36. Baser ME, Evans R, Gareth D, Gutmann DH. Neurofibromatosis 2. *Curr Opin Neurol.* 2003;16:27–33.
37. Evans DG, Sainio M, Baser ME. Neurofibromatosis type 2. *J Med Genet.* 2000;37:897–904.
38. Evans DG. Neurofibromatosis 2. In: GeneReviews at GeneTests: Medical Genetics Information Resource [database online]. Seattle: University of Washington, 1997–2005. Available at: http://www.genetests.org. Accessed October 20, 2005. Updated April 6, 2004.
39. Baser ME, Kuramoto L, Joe H, et al. Genotype-phenotype correlations for nervous system tumors in neurofibromatosis 2: a population-based study. *Am J Hum Genet.* 2004;75:231–239.
40. Baser ME, Friedman JM, Wallace AJ, Ramsden RT, Joe H, Evans DG. Evaluation of clinical diagnostic criteria for neurofibromatosis 2. *Neurology.* 2002;59:1759–1765.
41. Evans DG, Baser ME, O'Reilly B, et al. Management of the patient and family with neurofibromatosis 2: a consensus conference statement. *Br J Neurosurg.* 2005;19:5–12.
42. McClatchey AI, Giovannini M. Membrane organization and tumorigenesis—the NF2 tumor suppressor, merlin. *Genes Dev.* 2005;19:2265–2277.
43. Kluwe L, Mautner V, Heinrich B, et al. Molecular study of frequency of mosaicism in neurofibromatosis 2 patients with bilateral vestibular schwannomas. *J Med Genet.* 2003;40:109–114.
44. Abou-Sleiman PM, Apessos A, Harper JC, Serhal P, Winston RM, Delhanty JD. First application of preimplantation genetic diagnosis to neurofibromatosis type 2 (NF2). *Prenat Diagn.* 2002;22:519–524.

45. Baser ME, Kuramoto L, Woods R, et al. The location of constitutional neurofibromatosis 2 (NF2) splice site mutations is associated with the severity of NF2. *J Med Genet.* 2005;42:540–546.
46. Parry DM, MacCollin MM, Kaiser-Kupfer MI, et al. Germ-line mutations in the neurofibromatosis 2 gene: correlations with disease severity and retinal abnormalities. *Am J Hum Genet.* 1996;59:529–539.
47. Evans DG, Trueman L, Wallace A, Collins S, Strachan T. Genotype/phenotype correlations in type 2 neurofibromatosis (NF2): evidence for more severe disease associated with truncating mutations. *J Med Genet.* 1998;35:450–455.
48. Kluwe L, MacCollin M, Tatagiba M, et al. Phenotypic variability associated with 14 splice-site mutations in the NF2 gene. *Am J Med Genet.* 1998;77:228–233.
49. Sestini R, Vivarelli R, Balestri P, Ammannati F, Montali E, Papi L. Neurofibromatosis type 2 attributable to gonosomal mosaicism in a clinically normal mother, and identification of seven novel mutations in the NF2 gene. *Hum Genet.* 2000;107:366–371.
50. Wallace AJ, Watson CJ, Oward E, Evans DG, Elles RG. Mutation scanning of the NF2 gene: an improved service based on meta-PCR/sequencing, dosage analysis, and loss of heterozygosity analysis. *Genet Test.* 2004;8:368–380.
51. Kluwe L, Mautner VF. Mosaicism in sporadic neurofibromatosis 2 patients. *Hum Mol Genet.* 1998;7:2051–2055.
52. Evans DG, Wallace AJ, Wu CL, Trueman L, Ramsden RT, Strachan T. Somatic mosaicism: a common cause of classic disease in tumor-prone syndromes? Lessons from type 2 neurofibromatosis. *Am J Hum Genet.* 1998;63:727–736.
53. Moyhuddin A, Baser ME, Watson C, et al. Somatic mosaicism in neurofibromatosis 2: prevalence and risk of disease transmission to offspring. *J Med Genet.* 2003;40:459–463.
54. Kluwe L, Friedrich RE, Tatagiba M, Mautner VF. Presymptomatic diagnosis for children of sporadic neurofibromatosis 2 patients: a method based on tumor analysis. *Genet Med.* 2002;4:27–30.
55. Kluwe L, Nygren AO, Errami A, et al. Screening for large mutations of the NF2 gene. *Genes Chromosomes Cancer.* 2005;42:384–391.
56. Diebold R, Bartelt-Kirbach B, Evans DG, Kaufmann D, Hanemann CO. Sensitive detection of deletions of one or more exons in the neurofibromatosis type 2 (NF2) gene by multiplexed gene dosage polymerase chain reaction. *J Mol Diagn.* 2005;7:97–104.
57. Buckley PG, Mantripragada KK, Benetkiewicz M, et al. A full-coverage, high-resolution human chromosome 22 genomic microarray for clinical and research applications. *Hum Mol Genet.* 2002;11:3221–3229.
58. Stokowski RP, Cox DR. Functional analysis of the neurofibromatosis type 2 protein by means of disease-causing point mutations. *Am J Hum Genet.* 2000;66:873–891.
59. Gutmann DH, Hirbe AC, Haipek CA. Functional analysis of neurofibromatosis 2 (NF2) missense mutations. *Hum Mol Genet.* 2001;10:1519–1529.
60. Ryu CH, Kim SW, Lee KH, et al. The merlin tumor suppressor interacts with Ral guanine nucleotide dissociation stimulator and inhibits its activity. *Oncogene.* 2005;24:5355–5364.
61. Bijlsma EK, Wallace AJ, Evans DG. Misleading linkage results in an NF2 presymptomatic test owing to mosaicism. *J Med Genet.* 1997;34:934–936.

Chapter 22

TP53 Mutation and Li-Fraumeni Syndrome

Sydney D. Finkelstein and Frank S. Lieberman

Preeminent among human tumor suppressor genes in functionality and importance is *TP53*, reflected by its well-recognized characterization as "guardian of the genome."[1–5] TP53 is perhaps the most intensively studied human-cancer-associated oncoprotein, in keeping with its protean effects on critical cellular pathways, including transcriptional control of gene expression, cell cycle proliferation, DNA repair, apoptosis, and cellular maturation and differentiation.[6–8] Mutational damage to *TP53* is the single most common cancer DNA alteration, having been observed in more than 50% of all human cancers.[9] Detection of *TP53* mutation is now performed for several clinical indications including assessment of tumor biological aggressiveness, discrimination of tumor recurrence versus de novo cancer formation, determination of tumor anaplasia, and as part of a search for germline inherited mutational change associated with heightened cancer susceptibility.[10,11] The emerging diagnostic and prognostic role attributed to *TP53* mutation detection justifies testing in selected patients as part of the clinical molecular pathology workup of human cancer.

Soon after its initial molecular description, germline inheritance of *TP53* mutations was sought and found in families exhibiting a high susceptibility to cancer formation.[12–15] Known as the Li-Fraumeni syndrome (LFS; Online Mendelian Inheritance in Man [OMIM; database online] #151623], constitutional transmission of *TP53* mutations is a well-recognized, though relatively uncommon, form of inherited predisposition for cancer development.[12–15] While much remains to be learned regarding the frequency of germline passage of *TP53* mutations in the general population and the relationship between specific forms of mutational damage and phenotypic expression of cancer, a variety of methods exist for the detection of *TP53*-related genomic damage. These methods utilize different analytical techniques for molecular characterization and, when used in conjunction with one another, can provide accurate, detailed, and validated data to identify and follow patients at risk for subsequent cancer occurrence as well as to further characterize cancer growth and treatment responsiveness. This chapter briefly describes these analytical techniques and their current and potential future application. The discussion focuses mainly on practical testing for this gene and also comments on the LFS and its clinical detection.

Molecular Aspects of *TP53*

TP53 is a 393-amino-acid nuclear phosphoprotein encoded by a gene with 11 exons and measuring 53 kDa in size, hence the designation TP53.[16,17] The first exon of the *TP53* gene is noncoding and separated from the remaining ten coding exons by a large intron. The gene is situated on chromosome 17 at 17p13. The protein is highly conserved across vastly different species, in keeping with a fundamental role in cell growth and its regulation. The full genomic DNA sequence of *TP53* is readily available in a variety of data repositories. Numerous excellent reviews have been published discussing the structural, functional, and clinical aspects of TP53 and its role in human cancer.[1–5,18–24]

Transcriptional control of *TP53* is complex and includes two separate promoter regions situated in the 5′ nontranslated region as well as in the first intron of the gene.[12–15] TP53 is biologically highly pleiotropic with respect to its role in various aspects of cell homeostasis, including both transcriptional dependent and independent mechanisms for regulation of gene expression. Important downstream intermediates for TP53 action include GADD45, critical in DNA damage repair; $P21^{waf\text{-}1}$, which is a cell-cycle-dependent kinase inhibitor; MDM2, responsible for self regulation of TP53; and BAX protein, involved in cellular apoptosis.[1–8] This represents only a short list of critical growth regulatory and cell differentiation genes under the partial or total control of TP53. It is clear given TP53's ubiquitous interaction with many diverse cellular pathways that significant derangement in growth control and cancer development should follow mutational damage to this important regulatory gene.

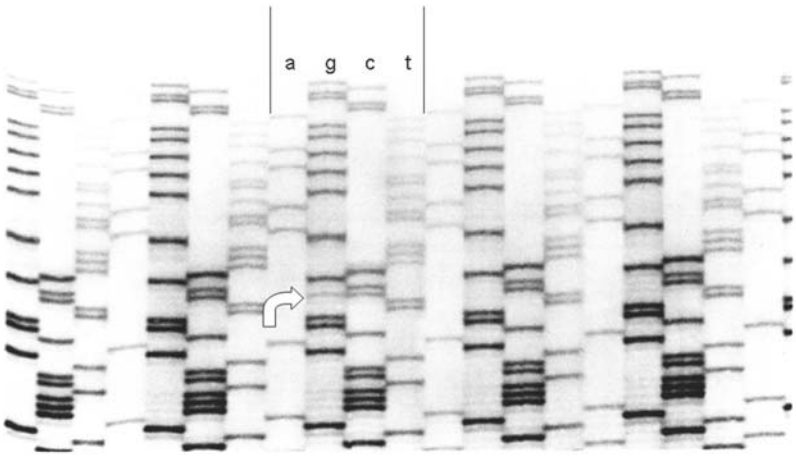


Figure 22-1. Missense point mutation in *TP53* exon 7. The arrow points to a G-to-A mutation in the DNA sequence, leading to insertion of an incorrect amino acid in the corresponding position in the protein. While only one copy of the gene is affected, biallelic inactivation is very likely to be present, leading to a loss of tumor suppressor function in the affected cell.

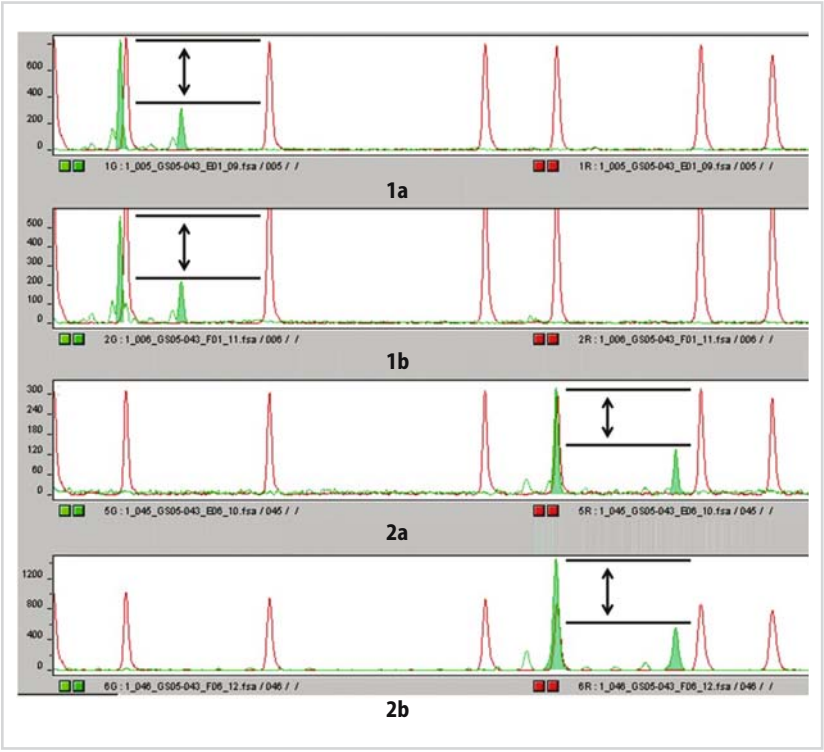


Figure 22-2. Allelic imbalance analysis is a highly accurate and sensitive method for determination of genomic deletion. Note the very close replication of samples performed in duplicate (1a and 1b; 2a and 2b). The degree of imbalance as determined by the ratio of peak heights is a measure of mutated tumor cell content in the sample.

Table 22-1. Microdissection Genotyping of High-Grade Glioma

Brain Neoplasm Genotyping	1p D1S 407	1p MYCL 5NT	1p D1S 1193	1p D1S 1172	3p26 D3S 1503	3p26 D3S 2303	5q23 D5S 592	5q23 D5S 615	9p21 D9S 254	9p21 D9S 251	10q23 D10S 520	10q23 D10S 1173	17p13 D17S 974	17p13 D17S 1289	17p13 D17S 1303	TP53 I1	19q D19S 400	19q D19S 559
Nonneoplastic brain tissue	I	I	NI	I	NI	I	I	NI	I	I	I	I	NI	I	NI	I	I	NI
High-grade glioma, area 1	LOH S	LOH S	NI	LOH L	NI	NO LOH	LOH S	NI	NO LOH	NO LOH	NO LOH	NO LOH	NI	LOH S	NI	LOH L	LOH S	NI
High-grade glioma, area 2	LOH S	LOH S	NI	LOH L	NI	NO LOH	LOH S	NI	LOH L	NO LOH	NO LOH	NO LOH	NI	LOH S	NI	LOH L	LOH S	NI
High-grade glioma, area 3	LOH S	LOH S	NI	LOH L	NI	NO LOH	LOH L	NI	NO LOH	LOH S	NO LOH	NO LOH	NI	LOH S	NI	LOH L	LOH S	NI
High-grade glioma, area 4	LOH S	LOH S	NI	LOH L	NI	NO LOH	LOH L	NI	NO LOH	NO LOH	NO LOH	NO LOH	NI	LOH S	NI	LOH L	LOH S	NI

I, informative microsatellite marker; NI, noninformative microsatellite marker; NO LOH, no significant allelic loss demonstrated; LOH S, significant allelic loss affecting the shorter polymorphic microsatellite allele; LOH L, significant allelic loss affecting the longer polymorphic microsatellite allele.

Mutational damage to *TP53* is the single most common human cancer alteration currently known.[9] Mutations of the gene can be detected in more than 50% of tumors of diverse histologic type, including carcinomas of various organs, sarcomas, cerebral gliomas, and hematopoietic malignancies (Figure 22-1).[9] Mutational damage itself can take many different forms but most often consists of missense point mutations leading to amino acid substitutions within highly conserved domains in the protein.[9] More than 95% of such point mutations are situated in the region spanning amino acids 91 through 309.[1–5] While missense mutations are typical of gene damage, alternative forms of DNA change also can be seen, including nonsense point mutations, deletions, insertions, and rearrangements.[1–5] Shown in Figure 22-1 is an example of a point mutation in a high-grade cerebral glioma. The frequency of different patterns of mutational damage varies according to the specific form of cancer being investigated. In certain instances there appears to be a very close relationship between the precise pattern of mutational damage and cancer etiology and histopathologic type. For example, missense mutations involving codon 249 leading to substitution of arginine by serine are closely tied to development of hepatocellular carcinoma in Asian and African patients. This form of damage is believed to be mediated by aflatoxin exposure.[25] Similarly, point mutations affecting codon 337 with substitution of arginine by histidine are seen in a high proportion of Brazilian patients with adrenocortical carcinoma. While an environmental toxin is suspected, its exact nature is unclear at this time.[26]

In contrast to this highly precise pattern of gene damage in specific forms of TP53-associated human cancer, most malignancies seen in the West do not show such a predictable relationship between *TP53* mutational change and cancer type. Mutational alterations can be encountered throughout exons 2 to 10, which necessitates a broad genotyping strategy for mutation detection. While mutation localization within the gene can be wide, certain hotspots for point mutation occurrence are well recognized and include codons 175, 245, 248, 252, 258, 273, and 282.[6–8] To assist in the collection and analysis of both common and uncommon forms of mutational damage, a variety of online databases can be easily accessed for reference.[27] The largest of these is maintained by the International Association for Research on Cancer, based in Lyon, France.[28] This database has recently been updated and includes more than 16,000 human cancers evaluated in detail for *TP53* mutational change. The database also provides valuable information on more than 200 germline *TP53* mutations (see discussion on LFS, below) as well as detailed information on DNA single-nucleotide and microsatellite tandem repeat polymorphisms.[27]

After initial mutation of the *TP53* gene, acquired in most instances as a somatic alteration and less frequently as an inherited constitutional alteration, additional DNA structural damage is acquired that eliminates all functional protein derived from the remaining unaffected allele (Figure 22-2).[1–5] This is the predictable two-step process initially described by Knudson[29–31] and shown to be operative for many tumor suppressor genes, including retinoblastoma and Wilms tumor genes.[29–31] The second event can take many different forms; however, most often it involves genomic deletion of the unaffected normal allele as an aberration during cell proliferation. The result of the two-step process is a complete lack of functional TP53 protein within the affected cell. This in turn is believed to confer significant growth advantage contributing to clonal expansion and further acquisition of somatic mutations targeting other growth regulatory genes (Figure 22-2). Allelic loss can be measured using simple techniques as shown in Figure 22-2 and described in Table 22-1.

The temporal relationship of the two steps can vary greatly in duration. In the case of inherited germline mutations of *TP53* (LFS), the interval can be 50 years or greater. In contrast, for most somatically acquired cancers, the two steps follow each other quickly in time. Critical events likely responsible for cancer phenotypic properties usually place *TP53* mutational damage as causally responsible for the transition from high-grade dysplasia (partial loss of growth control) to frank malignancy (total loss of growth control).[6–8] This has been well documented in colon cancer as well as many other forms of human malignancy.[6–8] From an analytical perspective, the presence of a second step assuming the form of a genomic deletion provides an alternative means to formulate testing designed to detect and characterize *TP53* mutational damage (Figure 22-2). This testing can be applied clinically to address specific clinical issues involved in the diagnosis of and treatment planning for cancer patients.

LI-FRAUMENI SYNDROME

Li-Fraumeni syndrome (LFS) is a rare, autosomal dominant, inherited cancer susceptibility disorder first reported in 1969.[32] While the disease has been attributed to germline mutations in *TP53*, not all affected subjects consistently manifest evidence for such mutational damage. When narrowly defined in strict terms as described below, more than 70% of individuals will reliably exhibit *TP53* point mutations that in turn can serve as the basis for testing of other family members for cancer susceptibility risk.[33] When the criteria for inclusion in this syndrome are relaxed, so called Li-Fraumeni-like syndromes (LFL), only 20% of affected individuals have evidence of *TP53* mutation.[33] Thus, *TP53* must be regarded as only one of an increasing number of known cancer-associated gene alterations responsible for increased cancer susceptibility in the human population.

LFS is noted for the diversity of cancer types within families and even within the same individual over time that

frequently cross histogenetic cell types. The most common forms of LFS malignancies include sarcomas, breast cancer, central nervous system gliomas, and hematopoietic neoplasms, most notably acute lymphocytic leukemia and different types of lymphomas.[33] Of particular note is the presence of adrenocortical carcinomas that, when present in a young person, should raise concern for the diagnosis of LFS. While the overall burden of multiple synchronous or metachronous tumors is not as great as with other forms of inherited cancer susceptibility (i.e., adenomatous polyposis coli, hereditary nonpolyposis colon cancer, BRCA1, and BRCA2), subjects can manifest multiple independent primary neoplasms. The early detection and characterization of independent cancer formation is an important element in the overall cancer surveillance and management of LFS patients and their family members.

Varying criteria have been used to define LFS. The strictest definition requires the occurrence of a sarcoma in a patient under 45 years of age in concert with a first-degree relative having either cancer of any type before age 45 or any form of sarcoma irrespective of age.[34,35] Using these clinical criteria, more than 70% of patients will have a germline mutation in *TP53*.[34,35] Family members with *TP53* mutation have a 50% risk for developing cancer by age 30 that then rises to 90% by age 70. When criteria for inclusion in LFS are relaxed, the incidence of demonstrable *TP53* mutations correspondingly decreases. This observation, together with a failure to document *TP53* mutations in a minority of otherwise classic LFS patients, suggests involvement of alternative genes operating in TP53 functional pathways as primary targets for germline mutational damage. It has become clear that inability to detect the presence of a *TP53* mutation does not exclude the diagnosis of LFS or LFL syndromes. The latter have been increasingly recognized as having a complex mutational basis targeting one or even several interactive cancer-associated genes. Further ongoing research is designed to improve our understanding of the genetic basis of this important cancer susceptibility syndrome.

Initial studies in LFS families suggested the presence of a narrow range of point mutational damage involving *TP53* exon 7.[36–38] As additional affected kindreds have been characterized, the range of specific germline mutations has expanded to include virtually the entire coding region of the gene. While certain hotspots associated with LFS are recognized, that is, codons 248, 252, and 258, testing for potential genomic involvement in an individual patient should not be limited to these codons. Any analysis for LFS or a related inherited cancer-susceptibility disorder should be done in concert with trained genetic counselors as part of a multidisciplinary team cognizant of the many challenges faced in the detection, evaluation, and management of affected individuals. Test results will significantly impact not only the affected patient but also many proximal and distal family members.

TP53 Mutation as a Predictor of Cancer Biological Aggressiveness

While approximately 50% of all human cancers manifest evidence of *TP53* mutations, the alterations tends to be closely associated with tumors manifesting histopathologic or clinical features, or both, of increased biological aggressiveness.[39–42] For example, well-differentiated forms of cerebral gliomas, that is, pilocytic astrocytomas, typically lack evidence of *TP53* mutations. In contrast, when *TP53* mutations are seen in gliomas, the tumors tend to be high grade in cellular appearance and classified as anaplastic astrocytomas or glioblastoma multiforme.[43–45] Thus, a search for *TP53* mutations in this tumor type can assist in the classification of human gliomas according to degree of expected biological aggressiveness. Discrimination between well-differentiated or low-grade forms of human gliomas and anaplastic or high-grade forms of gliomas is a most important and fundamental goal of surgical pathology evaluation of glial neoplasms. Methods to detect *TP53* mutations become an important parameter on which to base this distinction.

Approximately 50% of pulmonary epithelial cancers acquire *TP53* mutations during the course of tumor growth and progression. Several series have shown that disease-free and overall survival correlate inversely with the presence of *TP53* mutations.[46–48] Of note in this regard is the finding that small-cell neuroendocrine carcinoma of the lung, one of the most aggressive forms of human cancer, displays a very high rate of *TP53* mutation in excess of 90% of affected patients. Non-small-cell forms of lung cancer, which are intermediate in tumor biological aggressiveness, manifest a lower rate of *TP53* mutation, of approximately 50%.[46–48] Well-differentiated forms of lung cancer, for example, bronchoalveolar cell carcinoma, rarely show evidence of *TP53* mutation.[46–48] These findings, demonstrated in pulmonary neoplasia but proven for many other forms of human cancer, support the close relationship between *TP53* mutation acquisition and increased tumor biological aggressiveness. The data underscore the increasing role played in diagnostic pathology to characterize and classify neoplasms not solely on microscopic appearance but also according to critical forms of genetic damage held to be causally related to tumor biological aggressiveness.

Available Assays

Given the knowledge that *TP53* point mutations can occur throughout the coding region of the gene, a rigorous approach to mutation detection is recommended to effectively identify mutations in tumors of individual patients. In the past, the mutation testing was often limited to certain hotspots to simplify the analysis or minimize the cost of testing, or both. Given that a significant minority of patients with otherwise classic features for LFS will fail to

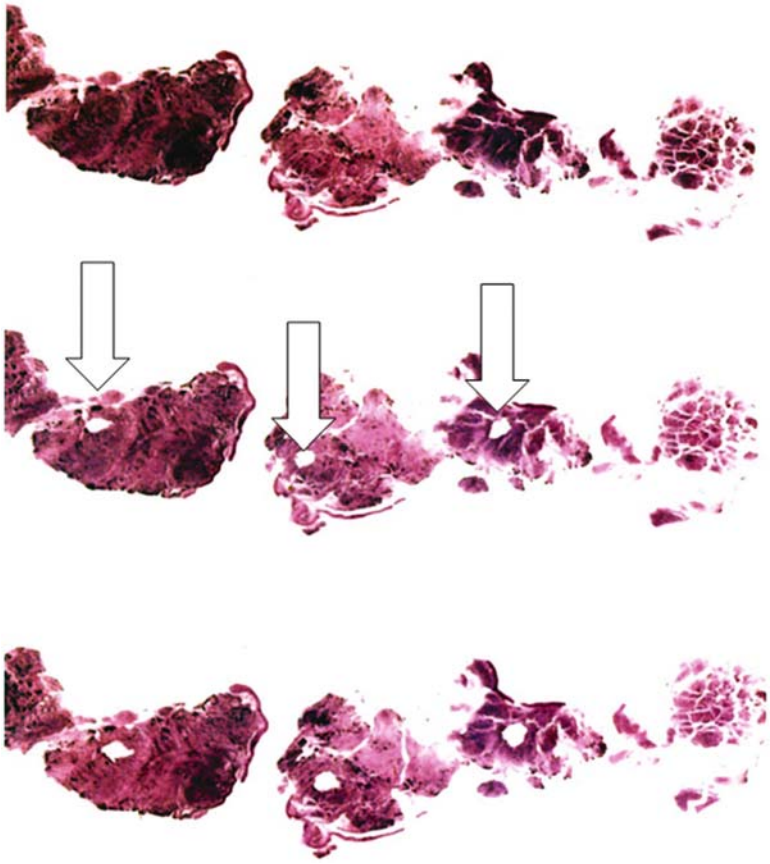


Figure 22-3. Tissue microdissection. Note that multiple sites have been carefully microdissected according to histopathologic features. Based on the cellular morphologic features, the very best samples are collected for mutation determination. The tumor has been microdissected at multiple locations to best understand intratumoral heterogeneity. This approach is capable of delineating the temporal sequence of mutation acquisition.

have identifiable *TP53* point mutations and that overlapping syndromes involving related genes may account for a subset of *TP53*-like effects, a thorough analysis of the gene for mutations becomes important in the diagnostic workup. Recently, a number of uncommon *TP53* mutations have been reported targeting the 5′ nontranslated region, intron 1 promoter regions, and other sites not directly responsible for coding of amino acids.[49–51] This has led to the use of ancillary tests in addition to DNA sequence analysis to better identify *TP53* mutations.

The gold standard for *TP53* mutation detection centers on determination of precise DNA sequence changes throughout the gene (Figure 22-1). The most commonly used method to achieve this goal is sequencing of PCR products from the coding portion of the gene. When applied to DNA extracted from fresh specimens in the form of fluid, blood, cells, or tissue samples, this can be carried out using a relatively small number of individual nucleic acid amplification reactions followed by sequencing using automated high-throughput equipment. The result is a timely, reliable, and detailed sequencing of the gene at relatively low cost and effort. Under these conditions, either DNA or RNA can be collected in abundant amounts to serve as starting material for DNA sequencing.[52,53] The availability of chip-based approaches using sequencing by hybridization also can rapidly deliver reliable results.[52,53,54] While the latter technique is highly specialized and not as widely available, submission of specimens for referral laboratory testing can be accomplished saving the user the effort needed to develop these assays.

As indicated above, a negative result for *TP53* mutations does not completely exclude the possibility of structural damage to the gene. Regions of the gene not analyzed by the specific testing method may be affected. Alternatively, associated proteins linked closely to TP53 function may be mutated. The significance of these changes will increase as our understanding of TP53 and its interaction with other cellular proteins is better understood.

The use of fixative-treated cancer tissue, rather than fresh or frozen tissue, for *TP53* mutation analysis is challenging (Figure 22-3). Such specimens often are the only available tissue on which a diagnosis of LFS can be confirmed. Alternatively, one may be searching for *TP53* point mutations on a fixative-treated cancer specimen for the purpose of characterizing tumor biological aggressiveness or establishing relatedness to other sites of cancer formation. Solid tumor tissue specimens typically are complex in topographic cellular heterogeneity. Careful, detailed histologic study is required as a preliminary step to identify important sites for mutational analysis (Figure 22-3). Careful attention must be given to the microscopic appearance of the specimen to focus the *TP53* mutation analysis on the most representative lesional sites. The specimen itself may be small in size, as in the form of a biopsy specimen (Figure 22-3). Thus, the desired cells for analysis will need to be carefully microdissected free of surrounding

normal supporting cells containing wild-type *TP53*. Inclusion of supporting cells will dilute the potentially mutated cell population and interfere with the ability to detect somatically acquired *TP53* mutations.

TP53 mutation detection on DNA from fixed tissues demands a different approach than that used when dealing with fresh, abundant cellular material. Tissues subjected to chemical fixation undergo degradation of RNA and chemical modification of DNA. Under these conditions, comprehensive sequence determination will entail greater effort, including multiple nucleic acid amplification reactions of individual exons. Chemical fixation limits the capacity to effectively generate amplicons that are more than 200 base pairs in length.[52,53] While several *TP53* exons are below 200 base pairs in length, some are in excess of this upper limit and require polymerase chain reaction (PCR) amplification in two or more individual reactions. A concern has been raised that chemical fixation can induce base modifications and lead to false-positive detection of point mutations.[55] This criticism has proven to be more theoretical than real and can be addressed by replicative analysis included as part of the internal quality control measures in the testing of fixed, microdissected tissue specimens.[55]

Alternative *TP53* mutation detection methods exist and have the potential to contribute meaningful information to validate and further define neoplastic progression. The simplest ancillary tool in this regard is TP53 immunohistochemical staining (Figure 22-4). Most missense point mutations interfere with the ability of TP53 protein to undergo effective DNA binding, thereby abrogating its capacity for transcriptional control. At the same time, the half-life of the mutant protein is significantly increased over that of wild-type protein, leading to TP53 accumulation in nuclei of affected cells (Figure 22-4). The presence of diffuse, strong positive staining of tumor cell nuclei using immunohistochemical staining for TP53 provides a simple and highly affordable means to infer the presence of *TP53* mutations. This approach must be used cautiously as both false-positive and false-negative staining can be encountered (see Quality Control and Laboratory Issues below).

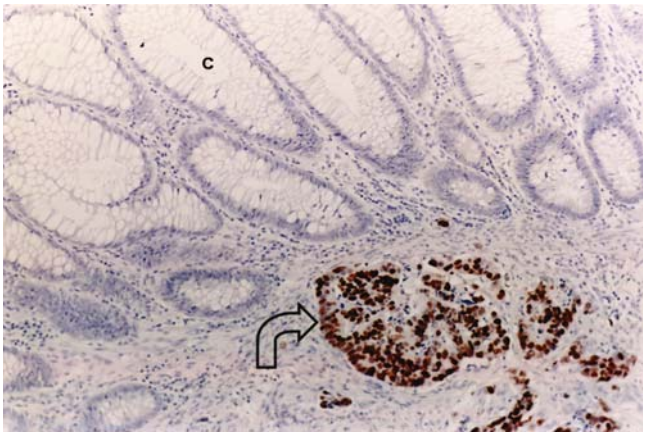


Figure 22-4. Immunohistochemistry for detection of TP53 protein. The longer stability and half-life of mutant TP53 protein lead to its collection in the cells and capacity for detection by immunohistochemistry (arrow). While immunohistochemical detection is sensitive and affords slide-based microscopic analysis, significant false-positive and false-negative detection is typically encountered, limiting the effectiveness of this approach. DNA analysis is used to validate immunohistochemical results.

As is true for many tumor suppressor genes, *TP53* undergoes a two-step process of mutation followed by loss of the residual unaffected allele, culminating in complete absence of effective wild-type protein in tumor cells.[1–5] The second step of allelic loss tends to encompass the entire normal allele and in so doing also results in the loss of genomic DNA adjacent to the gene itself.[6–8] Simple and sensitive methods, unrelated to direct sequence analysis, exist to detect this second event. These methods are based on DNA polymorphisms that are abundant throughout the genome in the form of single nucleotide polymorphisms (SNPs) or microsatellite tandem repeat polymorphisms (Figure 22-2). Given the interest in *TP53* as a preeminent tumor suppressor gene, a detailed understanding of polymorphisms in the region of *TP53*, both intragenic as well as outside the gene, exist that can be used in the design of allelic imbalance analysis. Most important, allelic loss infers the presence of point mutation damage irrespective of the precise location of the initial point mutation event. This is most advantageous since it provides an independent technique to assay for the presence of mutational change and is not limited by uncommon positioning of point mutations.

An example of *TP53* allelic loss from microdissected tissue targeting tetranucleotide repeat polymorphisms located in close proximity to *TP53* as well as other genes is shown in Figure 22-2 and Table 22-1.[56] This approach, employing microsatellites, is especially appealing because it requires only a small amount of DNA per reaction and uses a short amplicon length for quantitative assessment of the desired tandem repeat. As such, it is ideal for testing of microdissected, small, fixative-treated tissue samples (Figure 22-3).

Theoretically, microsatellite analysis of nonneoplastic tissue should yield polymorphic microsatellite amplicons that are equal to each other in amount. In practice, such a perfect relationship between amplified polymorphic alleles is not usually attainable. The reasons for lack of exactly balanced allelic content in nonneoplastic tissue are manifold but center on variations secondary to tissue-sectioning effects and slight differences in initial PCR amplifiability of individual alleles in any given tissue section, as well as differences in PCR efficiency in amplifying alleles of different size. This in turn entails defining thresholds for significant allelic imbalance beyond which definitive allelic loss is confidently predicted to exist (Figure 22-5). The threshold we have conservatively chosen utilizes an allele ratio of below 0.5 or above 2.00, which would correspond to the presence of a minimum of 50% mutated tumor cells in a given microdissected tissue sample (Figure 22-5).

The procedure for allelic imbalance analysis is described here. Multiple microsatellite loci are amplified from two

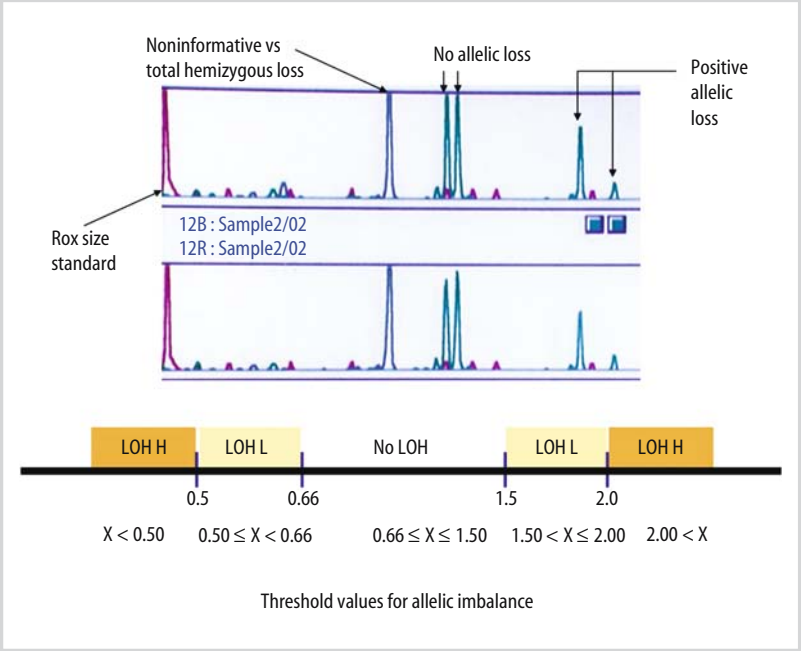


Figure 22-5. Allelic imbalance analysis for determination of loss of heterozygosity (LOH). The ratio of peak heights of polymorphic microsatellites is calculated and compared to a database of values representing the range of values for nonneoplastic samples in the general population. Peak height ratios falling outside 95% confidence limits represent true imbalance.

DNA samples, one from nonneoplastic and the other from neoplastic microdissected tissue samples. The reactions are electrophoresed, and the height of each amplicon peak from each allele is quantified. Nonneoplastic microdissected tissue samples are first analyzed to determine whether each specific locus is informative, that is, whether the locus is heterozygous (Figure 22-2), and therefore useful for analysis in the neoplastic tissue. When a particular microsatellite marker in a nonneoplastic tissue sample is homozygous and manifests only a single peak, the patient is designated as noninformative for that marker.

For informative heterozygous microsatellite markers, the ratio of the two alleles from the neoplastic tissue DNA is analyzed. The two alleles are in balance when the ratio of the individual allele peak heights falls within the range of 0.66 to 1.50. Values beyond this range are classified as demonstrating allelic imbalance, recognizing the following two categories. Low-level allelic imbalance is said to exist when the microsatellite allelic peak height ratios fall into the range 0.50 to 0.66 or 1.50 to 2.00. High-level allelic imbalance is said to be present when the allele ratios fall below 0.50 or above 2.00. To afford a conservative assessment for the presence of allelic loss, only high-level allelic loss is classified at this time as indicative of mutation. Using these criteria, it would be necessary for at least 50% of the cells in any given microdissection target to be mutated to detect allelic loss. These stringent conditions are justified given the ability of tissue microdissection to optimally sample tumor at a particular site coupled with the desire not to overcall mutational change that may, in fact, not be present (Figure 22-5).

In selected instances, most often due to small specimen size, nonneoplastic tissue may not be available. Recourse to nonneoplastic tissue can be sought in other tissue specimens from the same procedure, prior specimens, or residual blood samples from the patient. In a small number of cases, when even these measures do not provide nonneoplastic tissue, allelic imbalance can still proceed using tissue containing a significant admixture of nonneoplastic cellular elements apparent on histologic examination. Such samples reliably provide sufficient content of polymorphic alleles, albeit imbalanced, to ensure accurate determination of polymorphic marker informativeness. In the few instances when even these criteria cannot be satisfied, microsatellites in the form of single peaks are designated as being noninformative so that allelic loss will not be overcalled. These considerations allow even the smallest of fixed biopsy specimens to be suitable for *TP53* allelic loss testing.

Having established the presence of allelic imbalance based on polymorphic allele ratios as described above, the microdissected sample then is designated as positive for loss of heterozygosity (LOH) for the particular microsatellite marker (Table 22-1). In addition, the specific deleted allele is denoted by the designation as either "S" or "L," depending upon whether the shorter or the longer allele is relatively diminished in content in a particular patient. This is important because the presence of the identical deleted alleles in different microdissected DNA samples from the different tumor specimens supports the existence of the same deletion in all sites of tumor. Similarly, it is possible to identify two separate mutations of the same genomic region in different topographic tissue samples when deleted alleles are shown to be discordant (Table 22-1).

A similar approach can be applied to SNPs, which have the decided advantage of greater abundance throughout the genome. Many well-documented SNPs have been reported within the *TP53* gene.[27] There is, however, a significant disadvantage to the use of SNPs over microsatellites when

performing allelic imbalance determination of the type described above. Most microsatellites manifest a high heterozygosity rate within the human population, typically greater than 75%. Thus, more than 75% of patient analyses may be expected to yield cogent information on the presence or absence of *TP53* allelic loss and by consequence TP53 mutational change using microsatellite analysis. In contrast, SNPs show a far more limited degree of heterozygosity among humans, with the availability of only two alleles for a given polymorphic base. Polymorphic rates for SNPs are on the order of 20% to 30%, and thus most subjects will prove noninformative with respect to allelic imbalance analysis for a particular SNP. This can be compensated to a degree by extending the genotyping to additional markers; however, this may quickly exhaust the available microdissected tissue for a given case while at the same time increasing the overall effort and cost. The use of two or at most three microsatellite markers within or in close proximity to *TP53* is sufficient to confidently determine allelic imbalance, usually in a replicate fashion, and is therefore ideally suited for minute microdissected fixed-tissue samples. The same cannot be said for *TP53*-associated SNPs, for which the use of three separate SNP markers may not prove informative in all patient samples.

A useful potential application of allelic loss analysis as described above is for positive detection of LFS due to *TP53* mutational damage. As stated earlier, even classic forms of LFS will, in a small proportion of patients, fail to indicate the presence of a constitutional germline mutation. In these patients and in subjects with inadequate family histories or patients described as LFL syndromes, the question of *TP53* mutational change becomes very important. The mutational status of *TP53* becomes an essential question to address as completely as possible. If there truly is no mutation of *TP53*, efforts can be redirected to mutational assessment of other candidate genes. When no *TP53* mutation is identified, there remains a concern that the *TP53* mutation may have been missed for an individual patient.

Allelic imbalance determination can assist in this problematic area. If a germline allelic loss is present, then each individual neoplasm associated with that constitutional alteration will manifest the same pattern of allelic loss for *TP53*. The approach relies on the two-step process of tumor suppressor gene loss interrogating multiple independent neoplasms for identity, or concordance, in the pattern of allelic loss. For this testing to be operative, there must be multiple independent primary tumors that have proceeded through the two-step process with allelic loss. Multiple tumors may come from a single patient either as synchronous or metachronous neoplasms, or multiple tumors may be gathered from several affected family members. The greater the number of independent tumors showing the same pattern of allelic loss, the more confident is the conclusion that the patient or family has a germline *TP53* allelic loss. When allelic loss for microsatellite markers situated in proximity to *TP53* is found to be discordant—that is, different alleles are lost in different tumors—germline allelic loss is not favored. This approach based on searching for concordance versus discordance in the allelic loss pattern across different tumors provides indirect information concerning the presence of inherited germline alterations for tumor suppressor gene mutations such as *TP53* in LFS.

Quality Control and Laboratory Issues

DNA sequencing of *TP53* has been well studied for detection and characterization of germline and somatically acquired mutations. Standard quality measures, such as sequencing of internal negative and known positive controls, should be routinely included. Individual point mutations, when detected, can be confirmed in replicate analysis. Techniques such as the use of alternative sequencing primers or sequencing of the opposite strand provide strong assurance for analytical accuracy of mutation detection by sequencing. Advanced techniques such as DNA sequencing by chip-based hybridization or mass spectroscopy may similarly use replicate analysis to confirm the accuracy of mutation detection. DNA sequencing remains the gold standard and the procedure of choice for determination of germline alterations in blood, cellular, or fresh tissue samples.

For fixative-treated specimens we recommend the use of microdissection to secure multiple tissue targets for replicate mutational analysis. Sequencing of internally derived nonneoplastic tissue is essential to control for the effects of chemical fixation and to control for aberrations suspected to be related to nucleic acid amplification. We routinely sample fixed tumor specimens at multiple sites with the expectation that *TP53* point mutations will be the same in topographically distinct areas and confirm the *TP53* mutational status (Figure 22-3). The same mutation may not be present in multiple areas of the same tumor since somatic acquisition of *TP53* mutations can occur stochastically across the neoplasm in topographically distinct clonal populations of cells. When multiple mutations within the same tumor can be established, this provides strong support against the existence of a germline *TP53* mutation, that is, LFS, since a concordant pattern of allelic loss would be expected not only in all tumors present but also throughout the full extent of each neoplasm. This approach is especially useful in searching for evidence of germline mutation in tumor suppressor genes associated with cancer susceptibility. Sampling as many different neoplasms from the same patient or from related family members, or from both, and finding a concordant pattern of allelic loss support the presence of a germline mutation.

Immunohistochemical staining is a valuable ancillary tool to use in TP53 assessment and easily can be applied to fresh as well as fixed archival tissue specimens (Figure 22-4). A strong diffuse pattern of nuclear staining is strongly supportive of the presence of *TP53* mutation in a tumor

and as such should correlate well with alternative techniques such as DNA sequencing and allelic imbalance analysis. Careful attention must be given to variant patterns of staining that can lead to false-positive and false-negative correlation with *TP53* mutation status of the tumor. Most important is the absence of staining when the *TP53* mutation is a truncating mutation creating a stop codon or is a frameshift mutation caused by DNA deletions or insertions. These mutations typically are associated with absent immunostaining and can easily be misinterpreted as absence of DNA structural alterations. Truncation-type mutations are frequent, accounting for more than 30% of *TP53* mutations in certain forms of cancer such as ovarian epithelial malignancies.[57]

False-positive immunostaining can be seen secondary to TP53 hyperexpression, which may be seen in reactive states such as cellular regeneration or in neoplastic cell proliferation. TP53 hyperexpression in general produces a staining pattern that is not usually as strong as that resulting from missense point mutations. The morphologic distinction between *TP53* point mutation and reactive TP53 hyperexpression can be difficult and highly subjective. For these reasons, TP53 immunostaining alone is inadequate to properly characterize *TP53* mutation status. When exercised with caution, however, TP53 immunostaining can be a useful ancillary tool.

Probably the greatest limitation in TP53 immunostaining is the variability in extent and intensity of staining that often is encountered in neoplastic lesions associated with incomplete clonal expansion of mutated tumor cells (Figure 22-6). It is frequently the case that a given tumor will show an admixture of positive and negative staining cells with the former varying in the degree of intensity from intense to pale. The difference cannot be fully accounted for by variation in nuclear content consequent to histologic sectioning. Rather, the staining intensity pattern is likely reflective of an admixture of clonal populations of tumor cells with varying degrees of capacity for overgrowth of less aggressive precursor cells. In some cases a distinct minority cell population with *TP53* mutational change is present in the form of widely scattered strongly positive tumor cells admixed with other immunostain negative cells.[10,11] This staining pattern indicates somatic acquisition of *TP53* mutational damage that does not confer a strong growth selection advantage to the affected cell population. The correlation between immunostaining features and careful microdissection with DNA sequencing and allelic imbalance analysis will enable quality assessment of *TP53* mutational damage for individual cases.

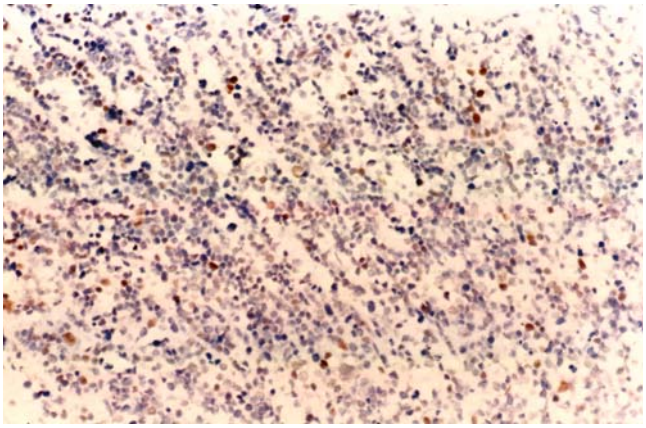


Figure 22-6. TP53 immunohistochemistry selectively stains only a subset of the nuclei in this section. This pattern is difficult to interpret and illustrates the need for more objective mutation testing by DNA analysis.

TP53 testing for LFS and LFL syndromes requires use of a comprehensive approach for DNA sequencing that is especially important when DNA sequencing alone fails to identify mutations. We recommend the use of allelic imbalance analysis using microsatellites or SNPs of multiple tumors when present in an individual patient or in several family members to search for concordant allelic loss for *TP53*. The finding of concordant allelic loss provides strong evidence for the presence of a germline mutation linking the various independent neoplasms from one or more related family members. Even in the face of an absence of documented *TP53* point mutations, the existence of high concordance rates among associated tumors should provide sufficient justification to expand the search for uncommon forms of *TP53* mutation.

In addition to its utility in suspected LFS, allelic loss concordance or discordance analysis is a valuable tool for discrimination between de novo primary tumor formation and tumor recurrence.[58,59] *TP53* is the most commonly mutated human cancer gene and as such qualifies for inclusion in genotyping surveys designed to distinguish cancer recurrence from new cancer formation. Discordance in the pattern of allelic loss between two neoplasms provides support that the two tumors may be independent primary neoplasms. It must, however, be kept in mind that this assessment must not be limited to only one mutational marker but rather extended to include a broad panel of polymorphic markers in the region of the *TP53* gene before a final conclusion is reached.

An example of this type of panel microsatellite marker analysis is shown in Figure 22-2 and Table 22-1 for a case of high-grade cerebral glioma. The glioma was microdissected at four locations designated as areas 1 through 4. Each microdissected tissue sample was analyzed at 18 different polymorphic microsatellite markers, four in close proximity to *TP53*, four on chromosome arm 1p, and two each from regions 3p26, 5q23, 9p21, 10q23, and 19q. Using nonneoplastic microdissected tissue as an internal source for nonmutated DNA (topmost lane in Figure 22-2), 12 of these microsatellite markers are established as being informative. Three of four *TP53* markers are informative. Six markers situated at 1p, 17p13, and 19q show allelic loss alterations and are concordant across all four microdissected sites. Single allelic loss alterations at 9p21 were present only in areas 2 and 3 consistent with somatic acquisition of this mutation focally within the glioma. The final discordant allelic loss alteration was identified at 5q23, with

the type of allelic loss alterations being discordant in areas 1 and 2 compared to areas 3 and 4 (Table 22-1). This discordant pattern of allelic loss suggests independent mutations affecting the same genomic region at 5q23. This pattern of allelic alteration with a predominance of concordant allelic loss (six of nine total allelic loss events are fully concordant across all the microdissected tumor areas) plus additional allelic changes in some tumor areas supports a single neoplasm with clonal evolution. Early acquisition of *TP53* mutation is supported by concordant evidence of allelic loss in all four areas of the glioma. Intratumoral heterogeneity is characterized by a mixture of concordant and discordant allelic loss (Table 22-1).

Use of *TP53* sequencing when allelic loss is demonstrated can provide compelling evidence to discriminate between de novo primary tumor formation and tumor recurrence. Given the great variety of individual forms of mutational damage, the finding of identical point mutations with concordant patterns of allelic loss establishes the relatedness of two sites of cancer formation, with one being the recurrence or metastasis of the other. On the other hand, discordant point mutations clearly indicate de novo primary cancer formation provided other markers of allelic loss also are discordant. The implications for distinguishing cancer recurrence from new cancer formation cannot be overemphasized since the two conditions are treated very differently, with identification of a new cancer providing the opportunity for curative measures while a recurrence is approached with an understanding of advanced spread of disease.

References

1. Lane DL. Cancer. p53, guardian of the genome. *Nature*. 1992;358: 15–16.
2. Finlay CA, Hinds PW, Levine AJ. The p53 proto-oncogene can act as a suppressor of transformation. *Cell.* 1989;57:1083–1093.
3. Kastan MB, Zhan Q, El-Deiry WS. A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia telangiectasia. *Cell.* 1992;71:587–597.
4. Kastan MB, Onyekwere O, Sidransky D, Vogelstein B, Craig RW. Participation of p53 in the cellular response to DNA damage. *Cancer Res.* 1991;51:6304–6311.
5. El-Deiry WS, Tokino T, Velculescu VE. WAF1, a potential mediator of p53 tumor suppression. *Cell.* 1993;75:817–825.
6. Diller L, Kassel J, Nelson CE. p53 functions as a cell cycle control protein in osteosarcomas. *Mol Cell Biol.* 1990;10:5772–5781.
7. Shaulsky G, Goldfinger N, Peled A, Rotter V. Involvement of wild-type p53 in pre-B cell differentiation in vitro. *Proc Natl Acad Sci U S A.* 1991;88:8982–8986.
8. Tarunina M, Jenkins JR. Human p53 binds DNA as a protein homodimer but monomeric variants retain full transcription transactivation activity. *Oncogene.* 1993;8:3165–3173.
9. Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 mutations in human cancers. *Science.* 1991;253:49–53.
10. Finkelstein SD, Przygodzki RM, Pricolo VE, etc. Prediction of biological aggressiveness in colorectal cancer by p53/K-ras-2 topographic genotyping. *Mol Diagn.* 1996;1:5–28.
11. Pollack IF, Hamilton RL, Finkelstein SD, et al. The relationship between mutations and overexpression of p53 and prognosis in malignant gliomas of childhood. *Cancer Res.* 1997;57:304–309.
12. Malkin D, Li FP, Strong LC. Germline p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science.* 1990;250:1233–1238.
13. Frebourg T, Friend SH. Cancer risks from germline p53 mutations. *J Clin Invest.* 1992;90:1637–1641.
14. Frebourg T, Kassel J, Lam KT. Germline mutations of the p53 tumor suppressor gene in patients with high risk for cancer inactivate the p53 protein. *Proc Natl Acad Sci U S A.* 1992; 89:6413–6417.
15. Varley JM, Evans DGR, Birch JM. Li-Fraumeni syndrome—a molecular and clinical review. *Br J Cancer.* 1997;1;1–14.
16. Lamb P, Crawford L. Characterization of the human p53 gene. *Mol Cell Biol.* 1986;6:1379–1385.
17. Pavletich NP, Chambers KA, Pabo CO. The DNA-binding domain of p53 contains the four conserved regions and the major mutation hot spots. *Genes Dev.* 1993;7:2556–2564.
18. Reisman D, Greenberg M, Rotter V. Human p53 oncogene contains one promoter upstream of exon 1 and a second stronger promoter within intron 1. *Proc Natl Acad Sci U S A.* 1988;85:5146–5150.
19. Holmquist GP, Gao S. Somatic mutation theory, DNA repair rates, and the molecular epidemiology of p53 mutations. *Mutat Res.* 1997;386:69–101.
20. Fields S, Jang SK. Presence of a potent transcription activating sequence in the p53 protein. *Science.* 1990;249:1046–1049.
21. Oren M, Prives C. p53: upstream, down-stream, and off stream. Review of the 8th p53 workshop. *Biochim Biophys Acta.* 1996;1288: R13–R19.
22. Momand J, Zambetti GP, Olson D, George D, Levine AJ. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell.* 1992;69:1237–1245.
23. Oltvai ZN, Milliman CL, Korsmeyer SJ. Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. *Cell.* 1993;74:609–619.
24. Bienz-Tadmor B, Zakut-Houri R, Libresco S, Givol D, Oren M. The 5′ region of the p53 gene: Evolutionary conservation and evidence for a negative regulatory element. *EMBO J.* 1985;4:3209–3213.
25. Hussain SP, Harris CC. Molecular epidemiology and carcinogenesis: endogenous and exogenous carcinogens. *Mut Res.* 2000;462:311–322.
26. DiGiammarino EL, Lee AS, Cadwell C, et al. A novel mechanism of tumorigenesis involving pH-dependent destabilization of a mutant p53 tetramer. *Nat Struct Biol.* 2002;9:12–16.
27. Beroud C, Soucci T. p53 gene mutation: software and database. *Nucleic Acids Res.* 1998;26:200–204.
28. Olivier M, Eeles R, Hollstein M, Khan MA, Harris CC, Hainaut P. The IARC TP53 database: new online mutation analysis and recommendations to users. *Hum Mutat.* 2002;19:607–614.
29. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A.* 1971;68:820–823.
30. Knudson AG. Genetics and etiology of human cancer. *Adv Hum Genet.* 1977, 8:1–66.
31. Knudson AG. Hereditary cancer: two hits revisited. *J Cancer Res Oncol.* 1996;122:135–140.
32. Li FP, Fraumeni F Jr. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann Intern Med.* 1969;71:747–752.
33. Malkin D, Li FP, Fraumeni JF Jr. Germ-line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science.* 1990;250:1233–1238.
34. Eng C, Hampel H, de la Chapelle A. Genetic testing for cancer predisposition. *Annu Rev Med.* 2001;52:371–400.
35. Eng C, Schneider K, Fraumeni JF, Li FP. Third international workshop on collaborative interdisciplinary studies of p53 and other predisposing genes in Li-Fraumeni syndrome. *Cancer Epidemiol Biomarkers Prev.* 1997;6:379–383.
36. Toguchida J, Yamaguchi T, Dayton SH. Prevalence and spectrum of germline mutations of the p53 among patients with sarcomas. *N Engl J Med.* 1992;326:1301–1308.
37. Birch JM, Hartley AL, Tricker KJ. Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer Res.* 1994,54:1298–1304.

38. Varley JM, McGown G, Thorncroft M. Germ-line mutations of TP53 in Li-Fraumeni families: an extended study of 39 families. *Cancer Res.* 1997;57:3245–3252.
39. Baker SJ, Fearon ER, Nigro JM. Chromosomal 17 deletions and p53 mutations in colorectal carcinomas. *Science.* 1989;244:217–221.
40. Righetti SC, Della Torre G, Pilotti S. A comparative study of p53 gene mutations, protein accumulation, and response to cisplatin-based chemotherapy in advanced ovarian carcinoma. *Cancer Res.* 1966; 56:689–693.
41. Kovach JS, Hartmann A, Blaszyk H, Cunningham J, Schaid D, Sommer SS. Mutation detection by highly sensitive methods indicates that p53 gene mutations in breast cancer can have important prognostic value. *Proc Natl Acad Sci U S A.* 1996;93:1093–1096.
42. Pricolo VE, Finkelstein SD, Hansen K, Cole BF, Bland KI. Mutated p53 gene is an independent adverse predictor of survival in colon carcinoma. *Arch Surg.* 1997;132:371–374.
43. Chozick BS, Weicker ME, Jackson CL, et al. Pattern of mutant P53 expression in human astrocytomas suggests the existence of alternative pathways of tumorigenesis. *Cancer.* 1994;73:406–415.
44. Mashiyama S, Murakami Y, Yoshimoto T, Sekiya T, Hayashi K. Detection of p53 gene mutations in human brain tumors by single strand conformation polymorphism analysis of polymerase chain reaction products. *Oncogene.* 1991;6:1313–1318.
45. Pollack IF, Finkelstein SD, Burnham J, et al. Age and TP53 mutation frequency in childhood malignant gliomas: results in a multi-institutional cohort. *Cancer Res.* 2001;61:7404–7407.
46. Takahashi T, Nan MM, Chiba I, Buchhagen DL, Mima JD. p53: a frequent target for genetic abnormalities in lung cancer. *Science.* 1989;246:491–494.
47. Harris CC. The p53 tumor suppressor gene: at the crossroads of molecular carcinogenesis, molecular epidemiology, and cancer risk assessment. *Science.* 1993;262:1980–1981.
48. Carbone DP, Misudomi T, Chiba I. p53 immunostaining positivity is associated with reduced survival and is imperfectly correlated with gene mutations in resected non-small cell lung cancer. A preliminary report of LCSG 871. *Chest.* 1994;106(suppl):377S-381S.
49. Eeles RA. Predictive testing for germline mutations in the p53 gene: are all the questions answered? *Eur J Cancer.* 1993;29A:1361–1315.
50. Jolly KW, Malkin D, Douglass EC, Brown TF, Sinclair AE, Look AT. Splice-specific mutation of the p53 gene in a family with hereditary breast cancer. *Oncogene.* 1994;9:97–102.
51. Felix CA, Strauss EA, D'Amico D, et al. A novel germline p53 splicing mutation in a pediatric patient with a second malignant neoplasm. *Oncogene.* 1993;8:1203–1210.
52. Dreesen JC, Bras M, Coonen E, Dumoulin JC, Evers JL, Geraedts JP. Allelic dropout caused by allele-specific amplification failure in single-cell PCR of the cystic fibrosis delta F508 deletion. *J Assist Reprod Genet.* 1996;13:112–114.
53. Finkelstein SD, Przygodzki R, Swalsky PA. Microdissection-based p53 genotyping: concepts for molecular testing. *Mol Diagn.* 1998;3: 179–191.
54. Wikman FP, Lu ML, Thykjaer T, et al. Evaluation of the performance of a p53 sequencing microarray chip using 140 previously sequenced bladder tumor samples. *Clin Chem.* 2000;46:1555–1561.
55. Safatle-Ribeiro AV, Ribeiro U Jr, Sakai P, et al. Integrated p53 histopathologic/genetic analysis of premalignant lesions of the esophagus. *Cancer Detect Prev.* 2000;24:13–23.
56. Hahn M, Fislage R, Pingoud A. Polymorphism of the pentanucleotide repeat d(AAAAT) within intron 1 of the human tumor suppressor gene p53 (17p13.1). *Hum Genet.* 1995;95:471–472.
57. Kanbour-Shakir A, Kounelis S, Papadaki H, et al. Relationship of p53 genotype to second look recurrence and survival in ovarian epithelial malignancy. *Mol Diagn.* 1996;1:121–129.
58. Ribiero U, Safatle-Ribiero A, Posner MC. Comparative P53 mutational analysis of multiple primary cancers of the upper aerodigestive tract. *Surgery.* 1996;120:45–53.
59. Rolston R, Sasatomi S, Hunt J, Swalsky P, Finkelstein SD. Distinguishing de novo second cancer formation from tumor recurrence: mutational fingerprinting by microdissection genotyping. *J Mol Diagn.* 2001;3:129–132.

Chapter 23

von Hippel-Lindau Disease

Catherine Stolle

Von Hippel-Lindau disease (VHLD) is an autosomal dominant cancer predisposition syndrome that gives rise to hemangioblastomas of the brain and spine, retinal angiomas, clear cell renal cell carcinoma, pheochromocytoma, endolymphatic sac tumors, tumors of the epididymis or broad ligament, and pancreatic tumors or cysts.[1] The incidence of VHLD is estimated to be about 1 in 40,000 live births in the white population. Onset is typically between the second and fourth decade of life, with penetrance for the disease nearly complete by the age of 65 years. In most cases, a family history of the disorder is apparent. In about 20% of cases, however, the proband appears to have acquired a new mutation.[2]

Molecular Basis of the Disease

The *VHL* tumor suppressor gene was isolated by positional cloning in 1993.[3] The gene, which consists of three exons spanning about 10 kilobases (kb) of genomic DNA, is highly conserved among worms, flies, rodents, and humans (reviewed in Reference 4). Two transcripts, 6.0 kb and 6.5 kb in size, are almost ubiquitously expressed and encode proteins of 213 and 159 amino acid residues. The latter isoform is the major product in most tissues and results from initiation of translation from an internal methionine codon at position 54. Both protein isoforms appear to be functional.

The VHL protein has been implicated in a variety of functions including transcriptional regulation, posttranscriptional gene expression, protein folding, extracellular matrix formation, and ubiquitylation (reviewed in Reference 4). In recent years, the role of VHL in the regulation of hypoxia-inducible genes through the targeted ubiquitylation and degradation of hypoxia-inducible factor 1 alpha subunit (HIF1A) has been elucidated, leading to a model of how disruption of the *VHL* gene results in the production of highly vascularized tumors.

Normal VHL binds to the protein elongin C, which forms a complex with elongin B and cullin-2 (CUL2). This complex resembles the SKP1-CUL1-F-box protein (SCF) ubiquitin ligase or E3 complex in yeast that catalyzes the polyubiquitylation of specific proteins and targets them for degradation by proteosomes. Under normoxic conditions, HIF1A is hydroxylated at a specific asparagine residue by a member of the egg-laying deficiency protein nine-like (EGLN) protein family of prolyl hydroxylase enzymes. VHL binds to hydroxylated HIF1A and targets it for degradation. Under hypoxic conditions, HIF1A is not hydroxylated, VHL does not bind, and HIF1A subunits accumulate. HIF1A forms heterodimers with HIF1B and activates transcription of a variety of hypoxia-inducible messenger RNAs (mRNAs) (i.e., *VEGF*, *EPO*, *TGFA*, *PDGFB*). Likewise, when VHL is absent or mutated, HIF1A subunits accumulate, resulting in cell proliferation and the neovascularization of tumors characteristic of VHLD.[4]

Predisposition to VHLD is inherited as an autosomal dominant trait. However, tumor formation requires loss of the second allele (loss of heterozygosity [LOH]), and so the disease is recessive at the level of the cell. Mutations known to result in predisposition to VHLD include partial or complete deletions of the gene and point mutations (missense, nonsense, frameshift, and splice site). Point mutations are predicted either to truncate the protein, to alter protein folding, or to interfere with the binding of VHL to elongin C, HIF1A, or other target proteins.[4,5] Although there are a handful of "common" mutations and one mutation hotspot in exon 3, point mutations are distributed over all three exons of the gene from codon 54 (internal initiator methionine) to the stop codon.

Four VHLD phenotypes have been described based on the likelihood of pheochromocytoma or renal cell carcinoma.[6] Type 1 is characterized by a low risk for pheochromocytoma. Truncating mutations or missense mutations that are predicted to grossly disrupt the folding of the VHL protein[5] are associated with VHLD type 1. VHLD type 2 is characterized by a high risk for pheochromocytoma. Patients with VHLD type 2 almost invariably have a missense mutation. VHLD type 2 is further subdivided into those with a low risk (type 2A) and those with a high risk

(type 2B) of renal cell carcinoma, as well as individuals at risk for pheochromocytoma only (type 2C). Some missense mutations correlate with a specific type 2 VHLD phenotype.[6]

A novel genotype-phenotype correlation has been reported.[7] Individuals with a complete deletion of the *VHL* gene are more likely to present with multiple hemangioblastomas of the brain or spine or both. At present, it is not clear why a complete deletion of the *VHL* gene would result in a phenotype distinctly different from that caused by a partial deletion or truncating mutation (i.e., VHL type 1).

Clinical Utility

Clinical molecular testing for VHLD has proven to be virtually 100% effective at detecting germline mutations in patients with pathology-proven disease.[8] For this reason, molecular testing may be used to confirm a clinical diagnosis in an affected patient, screen for a mutation in an unaffected individual with a family history of VHLD, or rule out the disease in individuals with one VHLD-like tumor but no family history. When the mutation in the family is known, molecular testing may be used for predictive testing of at-risk family members or for prenatal testing.

A clinical diagnosis of VHLD may be made in an individual with at least two typical VHLD tumors with or without a family history of VHLD, or in an individual with at least one typical tumor and a significant family history.[1] In such cases, a *VHL* gene mutation is almost invariably found. Identification of a mutation confirms the clinical diagnosis, establishes the need for periodic clinical screening, and facilitates predictive testing of at-risk relatives.

Predictive testing of at-risk family members for a known *VHL* gene mutation permits identification of presymptomatic mutation carriers and leads to early detection of tumors, timely intervention, and improved outcome. Identification of mutation-negative individuals eliminates the need for costly annual clinical screening. Genetic testing for a known *VHL* gene mutation is definitive.

Since the detection rate for germline mutations is so high, and since about 20% of patients have VHLD as the result of a sporadic mutation, testing is indicated in individuals with a single VHLD-type tumor and no family history of the disease. A negative-mutation screen greatly reduces the risk of VHLD in this circumstance. Since tumors have been reported in children as young as 4 years of age, and since the mutation status of at-risk individuals affects clinical care, predictive testing of asymptomatic children is appropriate. Prenatal testing for a *VHL* gene mutation is possible when the disease-causing mutation in an affected parent is known. However, prior to testing, consultation with a genetic counselor knowledgeable about the natural history of the disease and available treatment options is strongly recommended.

Available Assays

A 100% mutation detection rate has been reported in individuals with a germline mutation in the *VHL* gene using a combination of quantitative Southern blot analysis and DNA sequence analysis.[8] Quantitative Southern blot analysis is performed by digestion of high-molecular-weight genomic DNA with restriction enzymes such as *Eco* RI and *Ase* I that cut at the 5′ and 3′ boundaries of the gene to yield a fragment approximately 9.7 kb in size. Use of a double digest results in more efficient transfer of the gene fragment and avoids false negatives that occasionally result from digestion with *Eco* RI alone. Blots are hybridized with a *VHL* gene–specific probe (g7) to check for partial gene deletions and a probe for a two-copy gene (such as beta globin) to check for complete gene deletions. Partial gene deletions are detected as a decrease in intensity of the *VHL* gene band and the appearance of a band of altered mobility in the patient's sample (see Figure 23-1, lane 6). Complete gene deletions are detected as a decrease in intensity of the *VHL* gene band relative to the beta globin band in the patient compared to control samples (see Figure 23-1, lanes 3 and 5). Partial and complete gene deletions account

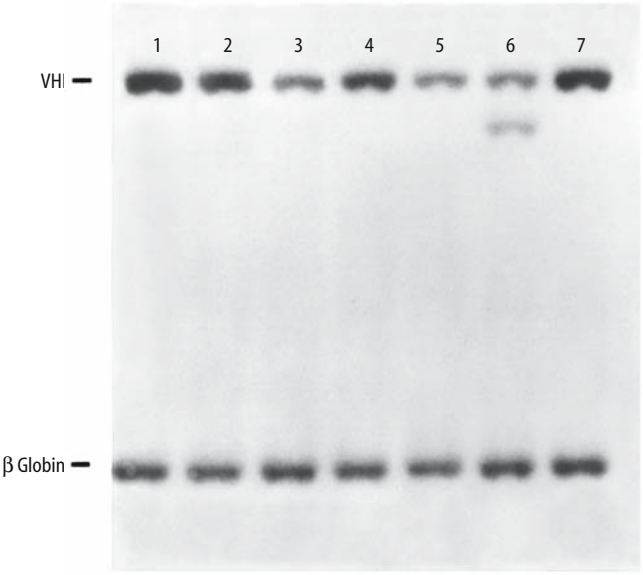


Figure 23-1. Quantitative Southern blot analysis for partial or complete deletions in the *VHL* gene. Genomic DNA (5 µg) is digested with *Eco* RI and *Ase* I. Digestion products are separated by agarose gel electrophoresis and transferred to a nitrocellulose membrane by capillary blotting. The filter is hybridized to random primer-labeled probes specific for the *VHL* gene (g7) and the beta globin gene (as an internal standard for DNA loading). Normal controls (lanes 1, 2, 4, and 7) exhibit a single intense *VHL* gene band. Patients with a partial deletion in the *VHL* gene exhibit a less-intense *VHL* band and a band of altered mobility (lane 6). Patients with a complete deletion of the *VHL* gene exhibit a band of decreased intensity (lanes 3 and 5) compared to controls when equal amounts of DNA are loaded on the gel. (Reprinted with permission from Chernoff A, Kasparcova V, Linehan WM, Stolle CA. Molecular analysis of the von Hippel-Lindau disease gene. In: *Methods in Molecular Medicine Series*, vol 53: Renal Cancer: Methods and Protocols. Mydlo JH, ed. Humana Press Inc., Totowa, NJ; 2001, p. 193–216.

for approximately 28% of all cases of VHLD, with complete deletions occurring in about 3% to 5% of patients. Recently, a real-time quantitative polymerase chain reaction (PCR) assay has been described for the rapid detection of complete and partial deletions in the *VHL* gene.[9]

PCR amplification and DNA sequence analysis of exons 1 to 3 of the *VHL* gene (including the adjacent splice donor and acceptor sequences) will detect all disease-causing point mutations. DNA sequence analysis is typically performed on an automated DNA sequencer using fluorescently labeled dideoxy terminator nucleotides. Both single base changes and small insertions or deletions may be detected as the appearance of two bases in any given position(s) of the DNA sequencing chromatogram (see Figure 23-2, panels b and c). A mutation-scanning method such as conformational sensitive gel electrophoresis (CSGE) may be used to identify mutation-containing exons prior to DNA sequence analysis. CSGE involves denaturing and reannealing PCR products to form heteroduplexes and homoduplexes that then are separated on a polyacrylamide gel. Exons exhibiting shifts (see Figure 23-3, lanes 4, 7, and 10) are sequenced to identify the specific mutation.

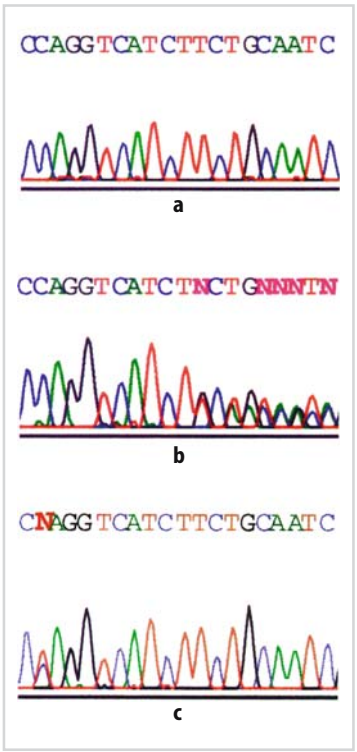


Figure 23-2. DNA sequence chromatogram from a portion of exon 1 of the *VHL* gene. The normal sequence is presented in panel a. A deletion or insertion of one or more bases (frameshift) results in a chromatogram with two bases in every position after the mutation (panel b). A single base change is recognized by the appearance of two bases in one position (panel c). Note that the height of the normal peak is generally reduced when a mutation is present in the same position. (Reprinted with permission from Chernoff et al., Molecular analysis of the von Hippel-Lindau disease gene. In: Mydlo JH, ed. *Renal Cancer: Methods and Protocols*. Humana Press Inc; 2001.)

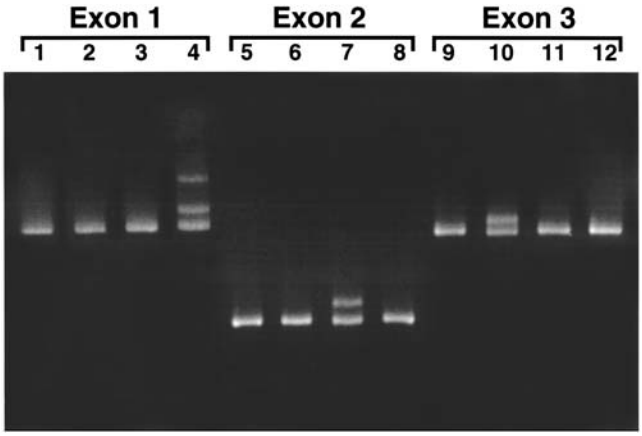


Figure 23-3. Conformational sensitive gel electrophoresis (CSGE) of the three *VHL* exons. PCR products from exons 1 to 3 of the *VHL* gene are denatured and reannealed. The mobility of heteroduplexes containing a normal and a mutant strand are "shifted" (lanes 4, 7, and 10), indicating the presence of a mismatched base (mutation) in the PCR product, compared to homoduplexes containing two normal strands after separation on a polyacrylamide gel. (Reprinted with permission from Chernoff et al., Molecular analysis of the von Hippel-Lindau disease gene. In: Mydlo JH, ed. *Renal Cancer: Methods and Protocols*. Humana Press Inc; 2001.)

Interpretation of Results

Using a combination of quantitative Southern blot analysis and DNA sequence analysis, germline mutations in the *VHL* gene have been identified in about 250 consecutive patients with well-documented (pathology-proven) VHL disease seen at the National Cancer Institute (NCI) since 1995, with no false negatives and no false positives. The sensitivity and specificity of the assays used to detect *VHL* gene mutations, therefore, are very high.

If a disease-causing mutation is identified in an affected or an at-risk individual, that individual will be predisposed to developing tumors characteristic of VHLD. Periodic screening of target organ systems is recommended for early detection and treatment of tumors. Molecular testing of offspring should be suggested, as they are at 50% risk of inheriting the disease-causing mutation. The parents of the proband should be offered molecular testing, as they may have unappreciated disease. The risk to siblings depends on the mutation status of the parents.

The interpretation of a negative mutation screen depends on the circumstances surrounding the testing and the strength of the clinical diagnosis in the affected family member. A negative test in an individual at risk for a known *VHL* gene mutation is definitive. The individual is not at risk for developing VHLD and need not undergo clinical screening for VHLD-associated tumors.

In screening of an unaffected at-risk individual for an unknown mutation in the *VHL* gene, however, a negative mutation screen indicates only that the individual does not have a germline *VHL* gene mutation. This may be because (1) the individual did not inherit the mutant allele from the affected relative (i.e., a parent), (2) the affected relative does not have VHLD (i.e., as a result of an incorrect or inconclusive diagnosis), or (3) if the individual is the

parent, he or she may be mosaic for a mutation detectable in affected offspring. Whenever possible, careful clinical evaluation or molecular analysis of an affected family member is suggested to improve the accuracy of the test interpretation.

For an individual with a clinical diagnosis of VHLD but no family history, a negative mutation screen indicates that either the patient is a phenocopy (coincidental occurrence of tumors typical of VHLD but without a gene mutation) or a mosaic (with a mutation in some, but not all, cells of the body). Mosaicism has been documented in the affected but mutation-negative parents of patients with a germline *VHL* gene mutation.[2] The frequency of mosaicism is not known but is believed to be low (i.e., <5%). If the diagnosis of VHLD is supported by pathologic findings (i.e., magnetic resonance imaging [MRI] of brain or spine, eye examination, abdominal ultrasound or computed tomography [CT] scan, urinary catecholamines, etc.), then the patient is more likely to be mosaic and should be periodically screened for additional tumors characteristic of VHLD. Offspring (if any) should be considered at risk and may wish to undergo periodic clinical screening for tumors or molecular analysis. The degree of risk (0–50%) depends on the extent of mosaicism and the potential for transmitting the mutant allele, neither of which can be determined.

For patients who do not fulfill the clinical criteria for VHLD diagnosis because they have a single type of tumor and no family history, a negative mutation screen would indicate that they are very unlikely to have VHLD. The possibility of mosaicism cannot be entirely ruled out, however. Although routine clinical surveillance for characteristic VHL tumors would not appear warranted, regular follow-up visits to a physician regarding the initial tumor would be expected. Likewise, routine screening of offspring would not be indicated unless the proband develops additional tumors (especially of another typical organ system) or if the offspring begins to develop symptoms of VHLD.

Occasionally, a base change in the *VHL* gene will be detected that is neither a known disease-causing mutation nor a polymorphism. These base changes of "unknown significance" with regard to VHLD are generally missense mutations or mutations in introns outside the consensus splice sequences. Novel mutations may be evaluated as disease-causing mutations or polymorphisms by criteria described by Cotton and Scriver.[10] In such cases, it is helpful to test other affected or unaffected family members, or both, to determine whether the mutation segregates with the disease.

Laboratory Issues

Mutation analysis of the *VHL* gene is performed as a laboratory-developed test. PCR primers and conditions have been published,[8] and probes for Southern blot analysis are available on request from the original researchers.[3] Control DNAs for Southern blot analysis are also available from the author on request. There is no organized proficiency-testing program for VHLD; therefore, proficiency testing is performed either by interlaboratory exchange or by repeat testing of samples.

Various sources of DNA are acceptable for testing. Ethylene diamine tetra-acetic acid (EDTA) anticoagulated peripheral blood is the most commonly submitted specimen; however, DNA, frozen tissue, or cultured cells also may be used. Paraffin-embedded tissue, cheek epithelial cells, unspun amniocytes, or direct chorionic villus sampling (CVS) specimens may be used for detecting point mutations by DNA sequence analysis; however, these samples do not provide sufficient quantities of high-molecular-weight genomic DNA for Southern blot analysis.

Mutation screening by Southern blot and DNA sequence analysis is completed in approximately 4 weeks. Interpretation of the assays is fairly straightforward (described above). Point mutations may be compared to an online database of *VHL* gene mutations in the Universal VHL-Mutation database (www.urnd.be/).[11]

The mutation detection rate quoted by laboratories in the United States varies from 95% to greater than 99%. Since all laboratories are using the same basic methodologies, this variability is likely due to uncertainty regarding the patient's diagnosis or inadequate clinical information rather than the inability to detect germline mutations in the *VHL* gene.

References

1. Glenn GM, Choyke PL, Zbar B, Linehan WM. von Hippel-Lindau disease: clinical review and molecular genetics. *Probl Urol.* 1990;4:312–330.
2. Sgambati MT, Stolle C, Choyke PL, et al. Mosaicism in von Hippel-Lindau disease: lessons from kindreds with germline mutations identified in offspring with mosaic parents. *Am J Hum Genet.* 2000;66:84–91.
3. Latif F, Tory K, Gnarra J, et al. Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science.* 1993;260:1317–1320.
4. Kaelin WG Jr. Molecular basis of the VHL hereditary cancer syndrome. *Nat Rev Cancer.* 2002;2:673–682.
5. Stebbins CE, Kaelin WG Jr, Pavletich NP. Structure of the VHL-elonginC-elonginB complex: implications for VHL tumor suppressor function. *Science.* 1999;284:455–461.
6. Zbar B, Kishida T, Chen F, et al. Germline mutations in the von Hippel-Lindau disease (VHL) gene in families from North America, Europe and Japan. *Hum Mutat.* 1996;8:348–357.
7. Hes F, Zewald R, Peeters T, et al. Genotype-phenotype correlations in families with deletions in the von Hippel-Lindau (VHL) gene. *Hum Genet.* 2000;106:425–431.
8. Stolle C, Glenn G, Zbar B, et al. Improved detection of germline mutations in the von Hippel-Lindau disease tumor suppressor gene. *Hum Mutat.* 1998;12:417–423.
9. Hoebeeck J, van der Luijt R, Poppe B, et al. Rapid detection of VHL exon deletions using real-time quantitative PCR. *Lab Invest.* 2005;85:24–33.
10. Cotton RGH, Scriver CR. Proof of "disease causing" mutation. *Hum Mutat.* 1998;12:1–3.
11. Beroud C, Joly D, Gallou C, Staroz F, Orfanelli MT, Junien C. Software and database for the analysis of mutations in the VHL gene. *Nucleic Acids Res.* 1998;26:256–258.

Section III

Solid Tumors

Section Editor: Karen L. Kaul

Chapter 24

Breast Cancer

Jeffrey S. Ross, Gerald P. Linette, James Stec, Edward Clark, Mark Ayers, Fraser Symmans, Gabriel N. Hortobagyi, and Lajos Pusztai

Based on current incidence rates, an American woman has a 1 in 9 chance of developing breast cancer at some time during her life.[1] According to the American Cancer Society, in 2001, there were 192,200 new cases of invasive breast cancer and 47,100 cases of in situ disease in the United States. That same year 40,200 American women died of the disease. This chapter considers the molecular pathology of breast cancer, focusing on the biomarker assays that are currently used in clinical management of the disease, excluding discussion of serum diagnostics, genetic predisposition testing, microarray-based RNA expression profiling, and micrometastasis detection, which are covered elsewhere in this book. The chapter concludes with a brief section on potential future assays, including the emerging field of pharmacogenomics.

Molecular Prognostic and Predictive Factors

In 1991, Hilsenbeck and collaborators described the requirements for adopting a new prognostic marker into clinical practice.[2] The overwhelming majority of proposed markers for breast cancer have failed to fulfill those requirements. Thus, although there are numerous biomarkers that have achieved promise on the basis of preliminary results, only HER2/NEU has been formally incorporated into standard practice over the past three decades.

EPIDERMAL GROWTH FACTOR RECEPTOR

Molecular Basis

The epidermal growth factor receptor (EGFR), also known as c-*erb*-B-1 and HER1, is a member of a family of transmembrane receptors that includes HER2, HER3, and HER4. HER1 has significant homology with the HER3/NEU protein, both of which have an intrinsic tyrosine kinase intracellular domain that is activated by ligand binding to the EGFR.

Available Assays and Interpretation

EGFR gene amplification has been detected by Southern blot, polymerase chain reaction (PCR), and fluorescence in situ hybridization (FISH). EGFR protein overexpression has been detected by Western blot and immunohistochemistry (IHC). EGFR status in breast cancer generally has been defined as overexpression when IHC tests show increased cellular staining or when FISH assays show significantly increased *EGFR* gene copy number.

Clinical Significance

EGFR is overexpressed in 14% to 90% of breast cancers, depending on the material tested and the method used for detection or quantitation of the receptor. EGFR overexpression has been linked to adverse prognosis in a variety of tumors, including breast cancer.[3,4] Conflicting studies of EGFR status in breast cancer have been reported, with some groups finding a correlation with prognosis[5] and others finding no correlation.[6] The frequent finding of gene amplification, gene mutation, or protein overexpression of EGFR, or some combination of these, in breast cancer, which has ranged from 67% to as high as 90% of breast cancers,[7] has prompted numerous clinical trials employing small molecule inhibitors[8] and antibodies[9] that target the EGFR pathway. To date, the clinical trials have yielded some evidence of efficacy, and one agent (ZD1839, Iressa) has been FDA approved for use in lung cancer. EGFR-targeted therapies have not moved forward for the treatment of breast cancer. In addition, levels of

EGFR expression have not correlated with therapeutic outcome.

Quality Control and Laboratory Issues

EGFR testing is not widely employed in clinical molecular laboratories and currently is considered to be for research use only. There is no current quality-assurance program for EGFR testing.

HER2/NEU

Molecular Basis

HER2/NEU (C-*erb*B-2) gene amplification, HER2 protein overexpression, or both, are identified in 10% to 34% of invasive breast cancers.[10] The ligand for the HER2/NEU protein receptor has not been identified, and its activation may occur through homo- and heterodimerization with other family members (EGFR, HER3, and HER4). Both morphologic and molecular techniques have been used to measure HER2/NEU status in breast cancer clinical samples (Table 24-1).[11] The vast majority of these studies have linked either gene amplification or protein overexpression of HER2/NEU with adverse prognosis in either node-negative or node-positive disease.[11] In general, when specimens have been carefully fixed, processed, and embedded, there has been excellent correlation between HER2/NEU gene copy status and protein expression levels.[11–14]

Table 24-1. Comparison of Methods for HER2/NEU Detection in Breast Cancer

Method	Target	FDA Status	Slide-Based
IHC	Protein	Approved*	Yes
FISH	DNA	Approved†	Yes
Chromogenic in situ hybridization (CISH)	DNA	Not approved	Yes
Southern blot	DNA	Not approved	No
Reverse transcription–polymerase chain reaction (RT-PCR)	mRNA	Not approved	No
Tumor enzyme-linked immunosorbent assay (ELISA)	Protein	Not approved	No
Serum ELISA	Protein	Approved‡	No

*Two IHC methods are FDA approved for Herceptin selection (Dako Herceptest and Ventana Pathway).
†One FISH method is approved for prognosis (Ventana Inform), and one FISH method is approved for prognosis and Herceptin selection (Abbott-Vysis Pathvysion).
‡Serum ELISA (Bayer Diagnostics, Tarrytown, NY) is FDA approved for monitoring.

Available Assays and Interpretation

IHC staining (Figure 24-1, lower panel), which has been the predominant method utilized for clinical testing of HER2/NEU, can be significantly affected by technical issues, especially for archival fixed paraffin-embedded tissues. Advantages of IHC testing include its wide availability, relatively low cost, easy preservation of stained slides, and use of routine microscopy. Difficulties with the technical aspects of IHC include the effects of preanalytic issues (including duration and type of fixation, and block storage conditions), intensity of antigen retrieval, type of antibody (polyclonal versus monoclonal), nature of control samples, and, most important, lack of an objective slide scoring system. Two commercially available HER2/NEU IHC kits, the Dako Herceptest and the Ventana Pathway, are FDA approved for determining the eligibility of patients to receive the anti-HER2/NEU therapeutic antibody trastuzumab (Herceptin). Problems with standardization in slide scoring recently have been highlighted and raise questions regarding the best method for determining HER2/NEU status for prediction of response to Herceptin.[15] Slide scoring can be improved by avoiding specimen edges, retraction artifacts, under- or overfixation, cases with significant staining of benign tissues, and tumor cells lacking a complete membranous staining pattern (the so-called chicken wire appearance). The use of a computerized image analysis system can reduce slide scoring variability among pathologists and improve the reproducibility of interpreting the IHC tests.[16]

The FISH technique (Figure 24-1, upper panel), which is morphology-driven and like IHC can be automated, has the advantages of an objective scoring system and the presence of a built-in internal control consisting of the two *HER2/NEU* gene signals present in all cells in the specimen. Disadvantages of FISH testing include the higher cost compared with IHC, longer time required for slide scoring, requirement of a fluorescence microscope, and inability to preserve the slide for storage and review. Two versions of the FISH assay are FDA approved: the Ventana Inform test, which measures only *HER2/NEU* gene copies, and the Abbott-Vysis Pathvysion test, which includes a chromosome 17 probe in a dual-color format. Published studies indicate that the two assays are highly correlative.[17] The chromogenic in situ hybridization method, or CISH technique (Figure 24-1, middle panel), features the advantages of both IHC (routine microscopy, lower cost, familiarity) and FISH (built-in internal control, objective scoring, more robust DNA target), but is not, to date, FDA approved for determining patient eligibility for Herceptin treatment.[18,19]

The reverse transcription–polymerase chain reaction (RT-PCR) technique[20,21] has predominantly been used to detect *HER2/NEU* messenger RNA (mRNA) from occult breast cancer cells found in peripheral blood and bone marrow samples. It has correlated more with gene amplification status than IHC levels,[22] and failed to predict

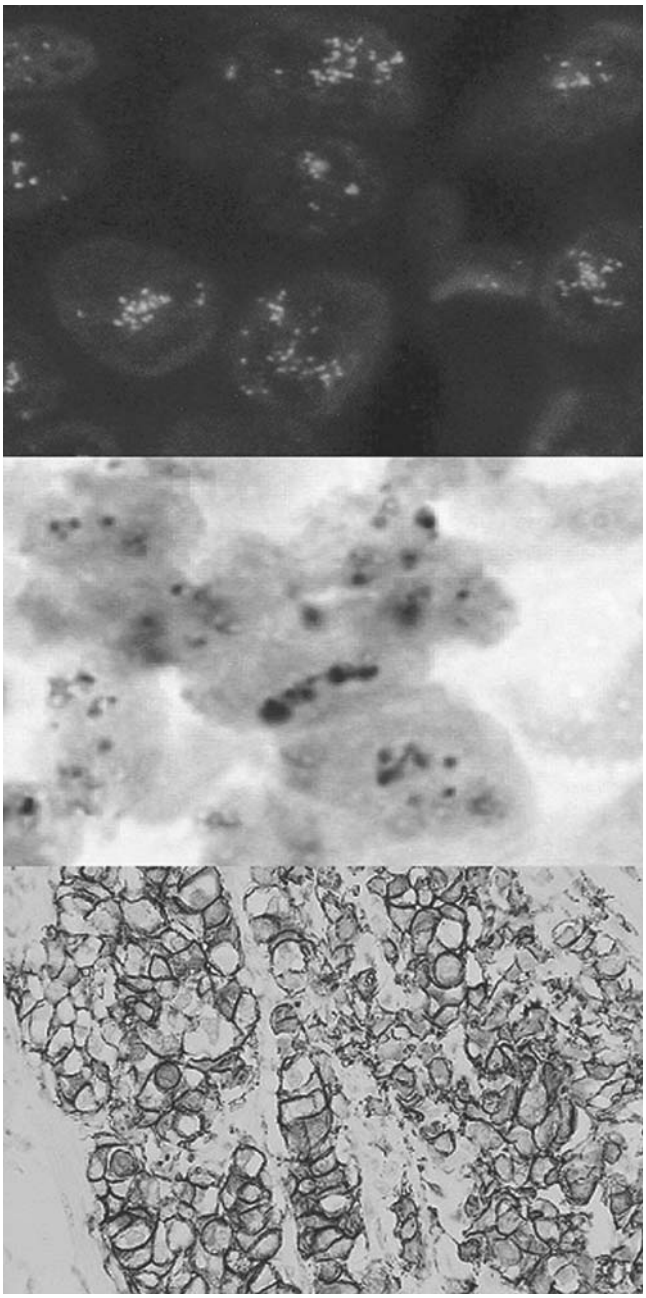


Figure 24-1. *HER2/NEU* testing in breast cancer. Upper panel: *HER2/NEU* gene amplification detected by FISH (Abbott-Vysis Pathvysion System). Middle panel: *HER2/NEU* gene amplification detected by CISH (Zymed System). Lower panel: IHC using Herceptest System (Dako Corp.) with continuous membranous 3+ positive immunostaining for HER2/NEU protein. (Figure prepared in collaboration with Dr. Ken Bloom of US Labs Inc, Irvine, CA.)

survival, but did correlate with estrogen receptor and progesterone receptor (ER/PR) status and tumor grade in one breast cancer outcome study of 365 patients.[23]

The enzyme-linked immunosorbent assay (ELISA) technique avoids the potential antigen damage associated with fixation, embedding, and uncontrolled storage, when performed on tumor cytosols made from fresh tissue samples. However, the small size of breast cancers associated with expanded screening programs in the United States generally precludes tumor tissue ELISA methods because insufficient tumor tissue is available to produce a cytosol.

Clinical Significance

Trastuzumab is a monoclonal IgG1-class humanized murine antibody developed by Genentech Corporation (South San Francisco, CA) using recombinant technologies. Clinical use of trastuzumab is targeted specifically for patients with advanced relapsed breast cancer that overexpresses the HER2 protein.[24] Since its launch in 1988, trastuzumab has become a major therapeutic option for patients with HER2-positive breast cancer and is being used not only for its approved indication as second-line treatment for advanced metastatic disease, but also for early-stage disease and in neoadjuvant treatment protocols.[25–28]

The best method to identify patients for trastuzumab therapy remains controversial. The original IHC technique used as the clinical trial assay was succeeded by the commercial Herceptest (Dako Corporation, Glostrup, Denmark). The Herceptest was originally criticized for yielding false-positive results,[29] although better performance ultimately was achieved when the manufacturer's instructions were precisely followed. Concern over IHC accuracy using standard formalin-fixed paraffin-embedded tissue sections[15] has encouraged the use of the FISH assay for its ability to predict trastuzumab response.[30] Well-documented lower response rates of 2+ IHC staining versus 3+ staining tumors[31] has resulted in an approach that uses either IHC as a primary screen followed by FISH testing for all 2+ cases or primary FISH testing.[32,33] In a recently published study using trastuzumab as a single agent, the response rate in 111 assessable patients with 3+ IHC staining was 35% and the response rate for 2+ cases was 0%; the response rates in patients with and without *HER2/NEU* gene amplification detected by FISH were 34% and 7%, respectively.[31] In a study of trastuzumab plus paclitaxel treatment of patients with HER2/NEU-overexpressing tumors, overall response rates ranged from 67% to 81% compared with 41% to 46% in patients with normal expression of HER2/NEU.[34] In summary, while the superiority of one method versus the other remains controversial,[35,36] most laboratories either screen all cases with IHC and triage the 2+ cases for FISH testing, or use FISH as the only method for *HER2/NEU* testing.

Prediction of Response of Breast Cancer to Other Therapies

The most frequent association of HER2/NEU status with a lack of therapeutic response to trastuzumab is the reported resistance of HER2/NEU-positive tumors to hormonal therapy alone.[37–41] Most recently, patients with estrogen (ER)-positive, HER2/NEU-positive tumors have an improved response to alternative hormonal therapy with letrozole, a third-generation aromatase inhibitor.[42] HER2/NEU protein overexpression also has been linked to

resistance of tumors in patients treated with cyclophosphamide, methotrexate, 5-fluorouracil (5-FU) cytoxin, methotrexater 5-fluorouracilt (CMF) adjuvant chemotherapy, as well as to taxane-based regimens.[43–45] HER2/NEU overexpression also has been associated with enhanced response rates to anthracycline-containing chemotherapeutic regimens.[46–50] HER2 and topoisomerease II are colocalized to the short arm of chromosome 17, and the 2 genes are usually coexpressed. These data have generated the reasonable hypothesis that HER2 overexpression identifies tumors that are more likely to respond to anthracycline therapy, especially at higher doses of the drug.

Serum HER2/NEU Antigen Levels as a Tumor Marker

Circulating levels of the cleaved extracellular domain of the HER2/NEU receptor protein have successfully predicted the presence and progression of HER2/NEU-positive breast cancer. Serum HER2/NEU level has been correlated with decreased survival and absence of clinical response to hormonal therapy in ER-positive tumors in some studies, but not in others.[10,51,52]

HER2/NEU Expression and Breast Pathology

HER2/NEU overexpression has been consistently associated with the more aggressive and extensive forms of ductal carcinoma in situ[53–55] and both mammary and extramammary Paget's disease.[56,57] The majority of studies that have compared the HER2/NEU status of paired primary and metastatic tumor tissues have found an overwhelming consistency of the HER2 status in both invasive and noninvasive tumors, regardless of the method of testing (IHC or FISH).[58–62] HER2/NEU amplification and overexpression has been associated with adverse outcome in some studies of male breast carcinoma[63–66] but not in others.[67,68] Finally, low-level HER2/NEU overexpression has been identified in biopsies of benign breast disease and associated with an increased risk of subsequent invasive breast cancer.[69]

Quality Control and Laboratory Issues

The advantages and disadvantages of the competing assays are discussed above. The widespread use of both the IHC and FISH techniques in clinical laboratories to assess HER2/NEU status in breast cancer has been the subject of many reviews as to performance and quality-control issues.[70,71] The College of American Pathologists has a proficiency-testing program that allows participating laboratories to assess their performance compared to a wide variety of private, public, and commercial laboratories.

TP53

Molecular Basis

TP53 is a tumor-suppressor gene localized to chromosome 17p that codes for a multifunctional DNA-binding protein involved in cell cycle arrest, DNA repair, differentiation, and apoptosis.[72] The *TP53* mutation rate is lower in breast cancer than in other epithelial cancers and has been associated with more aggressive disease and worse overall survival rate.[73]

Available Assays and Interpretation

The prognostic significance of TP53 status in breast cancer has been affected by the accuracy of IHC methods compared to molecular methods (single strand conformational polymorphism [SSCP], direct sequencing, microarray sequencing chip, and yeast colony functional assay) to identify *TP53* mutations.[72,73] IHC evaluations with and without image analysis–assisted slide scoring using either the DO-1 or the PAb1801 antibodies have yielded variable associations between TP53 stabilized mutant protein nuclear staining and outcome in breast cancer. In general, the DO-1 antibody has been favored by some investigators over the PAb1801;[74] however, the high number of false-positive and false-negative IHC results (compared with gene sequencing) precludes reliable use of IHC as an indicator for *TP53* gene mutation status in human breast cancer.[75]

Clinical Significance

In general, breast carcinomas with *TP53* mutations are consistently associated with high histologic grade, high mitotic index, high proliferation rate, aneuploid DNA content, negative assays for ER/PR,[76–78] and variable association with amplification of oncogenes such as *HER2, MYC*, *RAS,* and *INT2*.[79,80] *TP53* mutation has been correlated with resistance to hormonal, adjuvant, and neoadjuvant chemotherapy and combination chemotherapy for metastatic disease using a variety of agents, including anthracyclines and taxanes.[81–84]

Quality Control and Laboratory Issues

The best method for testing for *TP53* mutation status in clinical samples is controversial. While some groups believe that IHC testing, with its accuracy ranging from 80% to 90% for the prediction of *TP53* mutation is sufficient, others recommend that either SSCP or complete gene sequencing be performed to confirm *TP53* mutation status. In addition, since *TP53* expression can be silenced by epigenetic events, such as methylation of the promoter region of the gene, the relative impact of mutation detection versus expression silencing has only recently been

considered. *TP53* testing is generally performed for research purposes, and no clinical quality-assurance programs for laboratory accuracy are currently available.

ESTROGEN AND PROGESTERONE RECEPTOR PROTEINS

Molecular Basis

The estrogen (ER) and progesterone (PR) nuclear hormone receptors are activated and appear to play a role in the development of more than half of all breast cancers.[85–87] Positive ER and PR assays are associated with well-differentiated histology, negative lymph node status, diploid DNA content, low proliferation rate and a tendency for a relatively indolent clinical course.[88,89] ER/PR-negative tumors are often associated with aggressive disease, including amplification of the *HER2/NEU, MYC,* and *INT2* oncogenes, mutation of the *TP53* gene, and increased expression of invasion- and metastasis-associated growth factors, growth factor receptors, and proteases.[88,89]

Available Assays and Interpretation

The small size of newly diagnosed primary tumors has required a shift from the original methods of quantitative biochemical competitive binding assay with dextran-coated charcoal using fresh tumor protein extracts and cytosols to IHC methods.[90,91] Despite its limitations, IHC is currently the standard method for determining the ER and PR status of breast cancers. The ER and PR status of breast cancers remains a cornerstone of therapeutic planning and is likely to continue to have clinical significance for the foreseeable future. Microarray studies of gene expression patterns in breast cancers have shown that ER status correlates with differences in gene expression patterns in breast cancer (Figure 24-2). Microarray studies also have

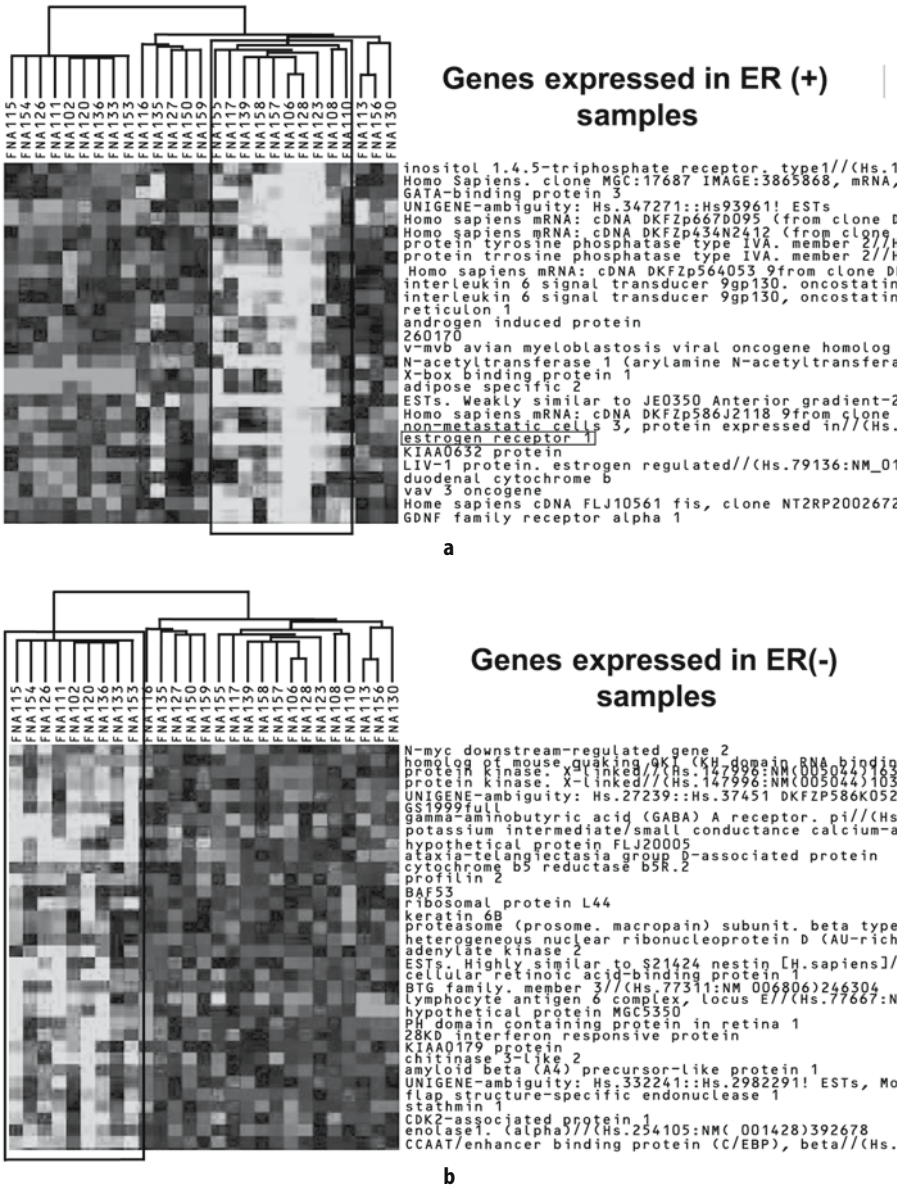


Figure 24-2. Breast cancer cDNA microarray results evaluated by hierarchical gene cluster analysis for defining specific gene expression signatures. Hierarchical clustering algorithm allows the clustering of individual tumor profiles on the basis of their similarities to their coexpression with the estrogen receptor alpha gene. Each row represents a single gene and each column a single tumor. Red indicates upregulation of gene expression in the tumor, green represents downregulation, and black indicates no change in relative gene expression. (a) Genes that are upregulated and downregulated in ER-positive tumors. (b) Overall gene expression pattern for ER-negative tumors.

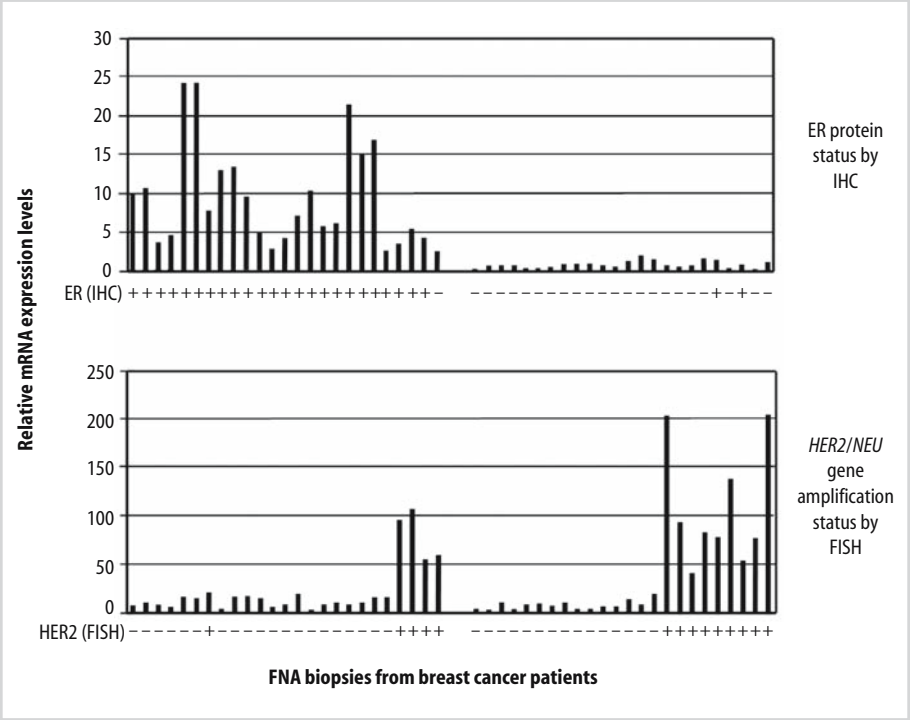


Figure 24-3. Comparison of *ER* and *HER2/NEU* mRNA expression detected by microarray profiling and corresponding ER protein expression measured by IHC and *HER2/NEU* gene copy number measured by FISH. The concordance between ER levels determined by IHC and *ER* levels determined by gene expression profiling was 95%. The concordance between *HER2/NEU* gene amplification status determined by FISH and *HER2/NEU* mRNA levels determined by gene expression profiling was 100%. Thus, a single gene expression profiling test combined with FNA biopsies from breast cancer patients may capture clinically relevant markers and guide therapy for the disease.

shown excellent concordance between protein expression of ER measured by IHC and *HER2/NEU* gene amplification measured by FISH and mRNA levels for both genes detected on the arrays (Figure 24-3).

Clinical Significance

The role of ER and PR receptor testing as markers of prognosis and predictors of response to antiestrogen therapy is established as a standard of care for patients with breast cancer.[88,89] The majority of breast cancers in the United States are evaluated for ER and PR status using IHC. Results of ER and PR IHC are used as the basis for selection of patients as candidates for hormonal therapy. The response to standard antiestrogen therapies such as Tamoxifen in first-line endocrine treatment of metastatic breast cancer averages 30% for either ER- or PR-positive tumors.[92] For patients with both ER- and PR-positive breast cancers, response rates have ranged from 50% to 70%.[92–94] For patients with ER- and PR-negative tumors, there is only a 5% to 10% chance that the patient will respond to hormonal therapy.[92–94] The introduction of specific estrogen response modulators[95] and aromatase inhibitors, such as anastrozole, letrozole, and extremestane,[96,97] have added new strategies of hormonal therapy for breast cancer patients.

Summary of Prognostic Factors in Breast Cancer

Table 24-2 summarizes the major prognostic factors for breast cancer. The ancillary biomarkers currently approved by the College of American Pathologists and the American Society of Clinical Oncology guidelines include ER/PR and HER2/NEU testing.

Pharmacogenetics and Pharmacogenomics

More than one million genetic markers known as single nucleotide polymorphisms (SNPs) have recently become available for genotyping and phenotyping studies.[98] SNP genotyping and gene sequencing have identified a variety of familial cancer predisposition syndromes based on single and multiple gene variants,[99] as well as genomic variations in drug-metabolizing enzymes such as the cytochrome system that are associated with variations in drug metabolism.[100] Pharmacogenetic strategies have been used to reduce the incidence of toxicity from such anticancer drugs as amonafide, 5-FU, 6-mercaptopurine (6-MP), irinotecan (camptosar, CPT-11), epirubicin, and flavopyridol.[101] The application of genotyping strategies to predict chemotherapeutic efficacy recently has emerged in a variety of clinical settings.[102,103] Relevant to breast cancer, recent publications have suggested that overexpression of thymidylate synthase is associated with resistance to 5-FU and related compounds.[104]

Two important challenges in the diagnosis and management of breast cancer are, first, how to identify patients at low risk for recurrence who can be spared from systemic adjuvant therapy, and, second, how to select the optimal systemic therapy for an individual who is at high risk for recurrence. Pharmacogenomics is an application of whole genome and protein expression data designed to predict the sensitivity or resistance of an individual's disease to a single

Table 24-2. Summary of Major Ancillary/Molecular Prognostic Factors in Breast Cancer

Bio marker	Assay	Target of Therapy	Therapeutic Agent	Current Status	Future Prospects
ER/PR	IHC binding assay	Yes	Tamoxifen SERMs Aromatase inhibitors	Standard of care	Improved IHC with antibodies that are negative when estrogen receptor is truncated to reduce false positives
HER2/NEU	IHC FISH	Yes	Herceptin Other antibodies Gene therapy	Standard of care	CISH assay may replace both IHC and FISH
DNA ploidy	Flow cytometry	No	—	Common use	Decreased use
S phase	Flow cytometry	No	—	Common use	Maintained use
Cell proliferation index	IHC	No	—	Common use	Increased use of Ki-67 IHC
Cyclin D	IHC	Possible	Flavopyridol Translocation targets	Clinical trials	May select new drug use such as proteasome inhibitors
Cyclin E	IHC Western	No	—	RUO	Prognostic significance must be validated
EGFR	IHC FISH	Yes	Iressa Tarceva Erbitux	Increasing use Clinical trials	Targeting the anti-EGFR drugs may likely be combined with pharmacogenomics
VEGF	IHC	Yes	Avastin Small molecules	Increasing use Clinical trials	Increasing use for prognosis; initial targeted therapy disappointing.
TP53	IHC SSCP Sequencing	Yes	Gene therapy	Increasing use Clinical trials	Targeted therapies disappointing to date
E-cadherin	IHC Methylation-PCR	Yes	5-azacytidine Demethylation	Increasing use Clinical trials	Diagnosis of pleomorphic lobular carcinoma
CD-44 v6	IHC	No	—	RUO	Predictive significance of v6 splice variant requires validation
Cathepsin D	Immunoassay	No	—	Common use in Europe	IHC studies disappointing; decreasing use
Upa/PAI1	Immunoassay	Yes	Small molecules	Common use in Europe	Targeted therapies in early stages; IHC assays not validated to date, restricting use in the United States
MMPs 2,9,11	IHC	Yes	Marmistat	Clinical trials RUO	Early results of targeted therapy disappointing
MDR	IHC	Yes	Small molecules	Clinical trials RUO	Continued use
BCL2	IHC	Yes	Genasense Proteasome inhibitors	Increasing use Clinical trials	Initial results of targeted therapies disappointing
Telomerase	TRAP assay IHC ISH	Yes	Small molecules	RUO	Increase use if slide-based assays are successful as prognostic factor
NFκB	IHC Western	Yes	Proteasome inhibitors	RUO	Will be used if targeted therapies are successful alone or in combination with cytotoxic drugs
Transcriptional profiling	cDNA array Oligonucleotide array	No	—	RUO	Continued expansion of use; predictive marker sets will require multiple validations; could become standard of care if initial results are confirmed

ER/PR, estrogen and progesterone receptor; IHC, immunohistochemistry; FISH, fluorescent in situ hybridization; RUO, research use only; EGFR, epidermal growth factor receptor; VEGF, vascular endothelial growth factor; SSCP, single strand conformational polymorphism; TRAP, telomeric repeat amplification protocol.

agent or group of drugs. Recently, powerful new technologies have emerged that can measure simultaneously the expression of several thousands of mRNA species in a biological specimen. With high-density DNA micro-arrays, technically it may be possible to monitor the expression of almost all human genes in a biological sample. The hierarchical clustering technique of data analysis from transcriptional profiling of clinical samples known to have responded or been resistant to a single agent or combination of anticancer drugs has recently been employed as a guide to anticancer drug therapy in cancers of the breast and other organs.[105,106] Microarray analysis of gene expression has achieved 81% accuracy for predicting a complete pathologic response after preoperative chemotherapy with sequential weekly paclitaxel and FAC in breast cancer.[107] More important, 75% of the patients who were predicted to have a complete pathologic response based on their gene expression profile indeed experienced a complete response. This compares very favorably with the 25% to 30% chance of complete response that unselected patients may expect with this treatment regimen.

The potential of pharmacogenomics as a novel tool that can truly identify clinically important subgroups of patients is enormous. However, the challenges that need to be solved before clinical application of this method also are significant. Despite the many technical and data interpretation difficulties associated with this approach and the complex nature of the profiling procedure itself, it is likely that DNA microarray technology will be applied in the clinic in the near future to direct therapy selection and predict clinical outcomes that cannot be predicted with current clinical testing methods.

References

1. Lacey JV Jr, Devesa SS, Brinton LA. Recent trends in breast cancer incidence and mortality. *Environ Mol Mutagen.* 2002;39:82–88.
2. Hilsenbeck SG, Clark GM, McGuire WL. Why do so many prognostic factors fail to pan out? *Breast Cancer Res Treat.* 1992;22:197–206.
3. Nicholoson S, Richard J, Sainsbury C, et al. Epidermal growth factor receptor (EGFR): results of a six-year follow-up study in operable breast cancer with emphasis on the node negative subgroup. *Br J Cancer.* 1991;63:146–150.
4. Castellani R, Visscher EW, Wykes S, et al. Interaction of transforming growth factor-alpha and epidermal growth factor receptor in breast carcinoma. *Cancer.* 1994;73:344–349.
5. Tsutsui S, Kataoka A, Ohno S, et al. Prognostic and predictive value of epidermal growth factor receptor in recurrent breast cancer. *Clin Cancer Res.* 2002;8:3454–3460.
6. Ferrero JM, Ramaioli A, Largillier R, et al. Epidermal growth factor receptor expression in 780 breast cancer patients: a reappraisal of the prognostic value based on an eight-year median follow-up. *Ann Oncol.* 2001;12:841–846.
7. Suo Z, Risberg B, Karlsson MG, et al. The expression of EGFR family ligands in breast carcinomas. *Int J Surg Pathol.* 2002;10:91–99.
8. Morris C. The role of EGFR-directed therapy in the treatment of breast cancer. *Breast Cancer Res Treat.* 2002;(suppl 1):S51–S55.
9. Solbach C, Roller M, Ahr A, et al. Anti-epidermal growth factor receptor-antibody therapy for treatment of breast cancer. *Int J Cancer.* 2002;101:390–394.
10. Volas GH, Leitzel K, Teramoto Y, et al. Serial serum C-erbB-2 levels in patients with breast carcinoma. *Cancer.* 1996;78:267–272.
11. Ross JS, Fletcher JA. HER-2/neu (c-erb-B2) gene and protein in breast cancer. *Am J Clin Pathol.* 1999;112(suppl 1):S53–S67.
12. Schnitt SJ, Jacobs TW. Current status of HER2 testing: caught between a rock and a hard place. *Am J Clin Pathol.* 2001;116:806–810.
13. Hayes DF, Thor AD. c-erbB-2 in breast cancer: development of a clinically useful marker. *Semin Oncol.* 2002;29:231–245.
14. Masood S, Bui MM. Prognostic and predictive value of HER2/neu oncogene in breast cancer. *Microsc Res Tech.* 2002;59:102–108.
15. Paik S, Bryant J, Tan-Chiu E, et al. Real-world performance of HER2 testing—National Surgical Adjuvant Breast and Bowel Project experience. *J Natl Cancer Inst.* 2002;94:852–854.
16. Wang S, Saboorian MH, Frenkel EP, et al. Assessment of HER-2/neu status in breast cancer. Automated Cellular Imaging System (ACIS)-assisted quantitation of immunohistochemical assay achieves high accuracy in comparison with fluorescence in situ hybridization assay as the standard. *Am J Clin Pathol.* 2001;116:495–503.
17. Wang S, Saboorian MH, Frenkel E, et al. Laboratory assessment of the status of Her-2/neu protein and oncogene in breast cancer specimens: comparison of immunohistochemistry assay with fluorescence in situ hybridisation assays. *J Clin Pathol.* 2000;53:374–381.
18. Tanner M, Gancberg D, Di Leo A, et al. Chromogenic in situ hybridization: a practical alternative for fluorescence in situ hybridization to detect HER-2/neu oncogene amplification in archival breast cancer samples. *Am J Pathol.* 2000;157:1467–1472.
19. Zhao J, Wu R, Au A, et al. Determination of HER2 gene amplification by chromogenic in situ hybridization (CISH) in archival breast carcinoma. *Mod Pathol.* 2002;15:657–665.
20. Pawlowski V, Revillion F, Hornez L, et al. A real-time one-step reverse transcriptase-polymerase chain reaction method to quantify c-erbB-2 expression in human breast cancer. *Cancer Detect Prev.* 2000;24:212–223.
21. Bieche I, Onody P, Laurendeau I, et al. Real-time reverse transcription-PCR assay for future management of ERBB2-based clinical applications. *Clin Chem.* 1999;45:1148–1156.
22. Tubbs RR, Pettay JD, Roche PC, et al. Discrepancies in clinical laboratory testing of eligibility for trastuzumab therapy: apparent immunohistochemical false-positives do not get the message. *J Clin Oncol.* 2001;19:2714–2721.
23. Pawlowski V, Revillion F, Hebbar M, et al. Prognostic value of the type I growth factor receptors in a large series of human primary breast cancers quantified with a real-time reverse transcription-polymerase chain reaction assay. *Clin Cancer Res.* 2000;6:4217–4225.
24. Huston JS, George AJ. Engineered antibodies take center stage. *Hum Antibodies.* 2001;10:127–142.
25. Hortobagyi GN. Overview of treatment results with trastuzumab (Herceptin) in metastatic breast cancer. *Semin Oncol.* 2001;28:43–47.
26. McKeage K, Perry CM. Trastuzumab: a review of its use in the treatment of metastatic breast cancer overexpressing HER2. *Drugs.* 2002;62:209–243.
27. Shawver LK, Slamon D, Ullrich A. Smart drugs: tyrosine kinase inhibitors in cancer therapy. *Cancer Cell.* 2002;1:117–123.
28. Ligibel JA, Winer EP. Trastuzumab/chemotherapy combinations in metastatic breast cancer. *Semin Oncol.* 2002;29:38–43.
29. Roche PC, Ingle JN. Increased HER2 with U.S. Food and Drug Administration-approved antibody. *J Clin Oncol.* 1999;17:434.
30. Mass RD, Press MF, Anderson S, et al. Evaluation of clinical outcomes according to HER2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab. *Clin Breast Cancer.* 2005;6:240–246.
31. Vogel CL, Cobleigh MA, Tripathy D, 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;20:719–726.

32. Nichols DW, Wolff DJ, Self S, et al. A testing algorithm for determination of HER2 status in patients with breast cancer. *Ann Clin Lab Sci.* 2002;32:3–11.
33. Kobayashi M, Ooi A, Oda Y, et al. Protein overexpression and gene amplification of c-erbB-2 in breast carcinomas: a comparative study of immunohistochemistry and fluorescence in situ hybridization of formalin-fixed, paraffin-embedded tissues. *Hum Pathol.* 2002;33:21–28.
34. Seidman AD, Fornier MN, Esteva FJ, et al. Weekly trastuzumab and paclitaxel therapy for metastatic breast cancer with analysis of efficacy by HER2 immunophenotype and gene amplification. *J Clin Oncol.* 2001;19:2587–2595.
35. Schnitt SJ, Jacobs TW. Current status of HER2 testing: caught between a rock and a hard place. *Am J Clin Pathol.* 2001;116: 806–810.
36. Press MF, Slamon DJ, Flom KJ, et al. 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;20:3095–3105.
37. Piccart M, Lohrisch C, Di Leo A, et al. The predictive value of HER2 in breast cancer. *Oncology.* 2001;(suppl 2):73–82.
38. Dowsett M. Overexpression of HER-2 as a resistance mechanism to hormonal therapy for breast cancer. *Endocr Relat Cancer.* 2001;8: 191–195.
39. Muss HB. Role of adjuvant endocrine therapy in early-stage breast cancer. *Semin Oncol.* 2001;28:313–321.
40. Schmid P, Wischnewsky MB, Sezer O, et al. Prediction of response to hormonal treatment in metastatic breast cancer. *Oncology.* 2002;63:309–316.
41. Nunes RA, Harris LN. The HER2 extracellular domain as a prognostic and predictive factor in breast cancer. *Clin Breast Cancer.* 2002;3:125–135.
42. Ellis MJ, Coop A, Singh B, et al. Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1-and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial. *J Clin Oncol.* 2001;12: 3808–3816.
43. Berns EM, Foekens JA, VanStaveren IL, et al. Oncogene amplification and prognosis in breast cancer: relationship with systemic treatment. *Gene.* 1995;159:11–18.
44. Sparano JA. Taxanes for breast cancer: an evidence-based review of randomized phase II and phase III trials. *Clin Breast Cancer.* 2000;1:32–40.
45. Yu D. Mechanisms of ErbB2-mediated paclitaxel resistance and trastuzumab-mediated paclitaxel sensitization in ErbB2-overexpressing breast cancers. *Semin Oncol.* 2001;28(suppl 16):12–17.
46. Kim R, Tanabe K, Uchida Y, et al. The role of HER-2 oncoprotein in drug-sensitivity in breast cancer. *Oncol Rep.* 2002;9:3–9.
47. Hamilton A, Larsimont D, Paridaens R, et al. A study of the value of p53, HER2, and Bcl-2 in the prediction of response to doxorubicin and paclitaxel as single agents in metastatic breast cancer: a companion study to EORTC 10923. *Clin Breast Cancer.* 2000;1:233–240.
48. Di Leo A, Larsimont D, Gancberg D, et al. HER-2 and topoisomerase llalpha as predictive markers in a population of node-positive breast cancer patients randomly treated with adjuvant CMF or epirubicin plus cyclophosphamide. *Ann Oncol.* 2001;12:1081–1089.
49. Petit T, Borel C, Ghnassia JP, et al. Chemotherapy response of breast cancer depends on HER-2 status and anthracycline dose intensity in the neoadjuvant setting. *Clin Cancer Res.* 2001;7:1577–1581.
50. Harris LN, Yang L, Liotcheva V, et al. Induction of topoisomerase II activity after ErbB2 activation is associated with a differential response to breast cancer chemotherapy. *Clin Cancer Res.* 2001;7: 1497–1504.
51. Lipton A, Ali SM, Leitzel K, et al. Elevated serum Her-2/neu level predicts decreased response to hormone therapy in metastatic breast cancer. *J Clin Oncol.* 2002;20:1467–1472.
52. Voss A, Neumann R. Serum HER-2/neu: a new predictive marker. *Int J Clin Pharmacol Ther.* 2002;40:584–585.
53. Bose S, Lesser ML, Norton L, et al. Immunophenotype of intraductal carcinoma. *Arch Pathol Lab Med.* 1996;120:81–85.
54. Moreno A, Lioveras B, Figueras A, et al. Ductal carcinoma in-situ of the breast: correlation between histologic classification and biologic markers. *Mod Pathol.* 1997;10:1088–1092.
55. Mack L, Kerkzelit N, Doig G, et al. Relationship of a new histological categorization of ductal carcinoma in-situ of the breast with size and the immunohistochemical expression of p53, C-*erb* B2, bcl2 and ki-67. *Hum Pathol.* 1997;28:974–979.
56. Wolber RA, DuPuis BA, Wick MR. Expression of *cerb* B2 oncoprotein in mammary and extramammary Paget's disease. *AM J Clin Pathol.* 1991;96:243–247.
57. Fu W, Lobocki CA, Silberberg BK, et al. Molecular markers in Paget disease of the breast. *J Surg Oncol.* 2001;77:171–178.
58. Masood S, Bui MM. Assessment of Her-2/neu overexpression in primary breast cancers and their metastatic lesions: an immunohistochemical study. *Ann Clin Lab Sci.* 2000;30:259–265.
59. Dittadi R, Zancan M, Perasole A, Gion M. Evaluation of HER-2/neu in serum and tissue of primary and metastatic breast cancer patients using an automated enzyme immunoassay. *Int J Biol Markers.* 2001;16:255–261.
60. Simon R, Nocito A, Hubscher T, et al. Patterns of her-2/neu amplification and overexpression in primary and metastatic breast cancer. *J Natl Cancer Inst.* 2001;93:1141–1146.
61. Vincent-Salomon A, Jouve M, Genin P, et al. HER2 Status in patients with breast carcinoma is not modified selectively by preoperative chemotherapy and is stable during the metastatic process. *Cancer.* 2002;94:2169–2173.
62. Xu R, Perle MA, Inghirami G, et al. Amplification of Her-2/neu gene in Her-2/neu-overexpressing and -nonexpressing breast carcinomas and their synchronous benign, premalignant, and metastatic lesions detected by FISH in archival material. *Mod Pathol.* 2002;15: 116–124.
63. Joshi MG, Lee AK, Loda M, et al. Male breast carcinoma: an evaluation of prognostic factors contributing to a poorer outcome. *Cancer.* 1996;77:490–498.
64. Pich A, Margaria E, Chiusa L, Oncogenes and male breast carcinoma: c-erbB-2 and p53 coexpression predicts a poor survival. *J Clin Oncol.* 2000;18:2948–2956.
65. Wang-Rodriguez J, Cross K, Gallagher S, et al. Male breast carcinoma: correlation of ER, PR, Ki-67, Her2-Neu, and p53 with treatment and survival, a study of 65 cases. *Mod Pathol.* 2002;15: 853–861.
66. Rayson D, Erlichman C, Suman VJ, et al. Molecular markers in male breast carcinoma. *Cancer.* 1998;83:1947–1955.
67. Shpitz B, Bomstein Y, Sternberg A, et al. Angiogenesis, p53, and c-erbB-2 immunoreactivity and clinincopathological features in male breast cancer. *J Surg Oncol.* 2000;75:252–257.
68. Bloom KJ, Govil H, Gattuso P, et al. Status of HER-2 in male and female breast carcinoma. *Am J Surg.* 2001;182:389–392.
69. Stark A, Hulka BS, Joens S, et al. HER-2/neu amplification in benign breast disease and the risk of subsequent breast cancer. *J Clin Oncol.* 2000;18:267–274.
70. Jacobs TW, Gown AM, Yaziji H, et al. HER-2/neu protein expression in breast cancer evaluated by immunohistochemistry. A study of interlaboratory agreement. *Am J Clin Pathol.* 2000;113:251–258.
71. Tubbs RR, Stoler MH. The quality of Her-2/Neu predictive immunohistochemistry: something FISHy? *Mod Pathol.* 2000;13:1–3.
72. Liu MC, Gelmann EP. P53 gene mutations: case study of a clinical marker for solid tumors. *Semin Oncol.* 2002;29:246–257.
73. Gasco M, Shami S, Crook T. The p53 pathway in breast cancer. *Breast Cancer Res.* 2002;4:70–76.
74. Cunningham JM, Ingle JN, Jung SH, et al. p53 gene expression in node positive breast cancer: relationship to DNA ploidy and prognosis. *J Nat Cancer Inst.* 1994;86:1871–1873.
75. Lohmann D, Ruhri C, Schmitt M, et al. Accumulation of p53 protein as an indicator for p53 gene mutation in breast cancer. *Diagn Mol Pathol.* 1993;2:36–41.

76. Bhargava V, Thor A, Deng G, et al. The association of p53 immunopositivity with tumor proliferation and other prognostic indicators in breast cancer. *Modern Pathol.* 1994;7:361–368.
77. Rosanelli GP, Steindorfer P, Wirnsberger GH, et al. Mutant p53 expression and DNA analysis in human breast cancer. Comparison with conventional clinicopathological parameters. *Anticancer Res.* 1995;15:581–586.
78. Pelosi G, Bresaola E, Rodella S, et al. Expression of proliferating cell nuclear antigen, Ki-67 antigen, estrogen receptor protein, and tumor suppressor p53 gene in cytologic samples of breast cancer: an immunochemical study with clinical, pathobiological, and histologic correlations. *Diagn Cytopathol.* 1994;11:131–140.
79. Beck T, Weller EE, Weikel W, et al. Usefulness of immunohistochemical staining for p53 in the prognosis of breast carcinomas: correlation with established prognosis parameters and with the proliferation marker, MIB-1. *Gynecol Oncol.* 1995;57:96–104.
80. Peyrat J-P, Bonneterre J, Lubin R, et al. Prognostic significance of circulating P53 antibodies in patients undergoing surgery for local regional breast cancer. *Lancet.* 1995;345:621–622.
81. Daidone MG, Veneroni S, Benini E, et al. Biological markers as indicators of response to primary and adjuvant chemotherapy in breast cancer. *Int J Cancer.* 1999;84:580–586.
82. Kandioler-Eckersberger D, Ludwig C, et al. TP53 mutation and p53 overexpression for prediction of response to neoadjuvant treatment in breast cancer patients. *Clin Cancer Res.* 2000;6:50–56.
83. Bertheau P, Plassa F, Espie M, et al. Effect of mutated TP53 on response of advanced breast cancers to high-dose chemotherapy. *Lancet.* 2002;360:852–854.
84. Sjostrom J, Blomqvist C, Heikkila P, et al. Predictive value of p53, mdm-2, p21, and mib-1 for chemotherapy response in advanced breast cancer. *Clin Cancer Res.* 2000;6:3103–3110.
85. Palmieri C, Cheng GJ, Saji S, et al. Estrogen receptor beta in breast cancer. *Endocr Relat Cancer.* 2003;9:1–13.
86. Nilsson S, Gustafsson JA. Biological role of estrogen and estrogen receptors. *Crit Rev Biochem Mol Biol.* 2002;37:1–28.
87. Sommer S, Fuqua SA. Estrogen receptor and breast cancer. *Semin Cancer Biol.* 2001;11:339–352.
88. Osbome CK. Steroid hormone receptors in breast cancer management. *Breast Cancer Res Treat.* 1998;51:227–238.
89. Locker GY. Hormonal therapy of breast cancer. *Cancer Treat Rev.* 1998;24:221–240.
90. Wilbur DC, Willis J, Mooney RA, et al. Estrogen and progesterone detection in archival formalin-fixed paraffin embedded tissue from breast carcinoma: a comparison of immunocytochemistry with dextran coated charcoal assay. *Modern Pathol.* 1992;5:79–84.
91. Esteban JM, Battifora H, Warsi Z, et al. Quantitation of estrogen receptors on paraffin-embedded tumors by image analysis. *Modern Pathol.* 1991;4:53–57.
92. Jordan VC. Tamoxifen: a most unlikely pioneering medicine. *Nat Rev Drug Discov.* 2003;2:205–213.
93. Schmid P, Wischnewsky MB, Sezer O, et al. Prediction of response to hormonal treatment in metastatic breast cancer. *Oncology.* 2002;63:309–316.
94. Clemons M, Danson S, Howell A. Tamoxifen ("Nolvadex"): a review. *Cancer Treat Rev.* 2002;28:165–180.
95. Ibrahim NK, Hortobagyi GN. The evolving role of specific estrogen receptor modulators (SERMs) *Surg Oncol.* 1999;8:103–123.
96. Miller WR, Anderson TJ, Dixon JM. Anti-tumor effects of letrozole. *Cancer Invest.* 2002;20:15–21.
97. Buzdar AU, Robertson JF, Eiermann W, et al. An overview of the pharmacology and pharmacokinetics of the newer generation aromatase inhibitors anastrozole, letrozole, and exemestane. *Cancer.* 2002;95:2006–2016.
98. Taylor JG, et al. Using genetic variation to study human disease. *Trends Mol Med.* 2001;7:507–512.
99. Weber W, Estoppey J, Stoll H. Familial cancer diagnosis. *Anticancer Res.* 2001;21:3631–3635.
100. Ingelman-Sundberg M. Genetic susceptibility to adverse effects of drugs and environmental toxicants. The role of the CYP family of enzymes. *Mutat Res.* 2001;482:11–19.
101. Innocenti F, Ratain MJ. Update on pharmacogenetics in cancer chemotherapy. *Eur J Cancer.* 2002;38:639–644.
102. Roses AD. Pharmacogenetics. *Hum Mol Genet.* 2001;10:2261–2267.
103. Relling MV, Dervieux T. Pharmacogenetics and cancer therapy. *Nat Rev Cancer.* 2001;1:99–108.
104. Nishimura R, Nagao K, Miyayama H, et al. Thymidylate synthase levels as a therapeutic and prognostic predictor in breast cancer. *Anticancer Res.* 1999;19:5621–5626.
105. Weinstein JN. Pharmacogenomics—teaching old drugs new tricks. *N Engl J Med.* 2000;343:1408–1409.
106. Slonim DK. Transcriptional profiling in cancer: the path to clinical pharmacogenomics. *Pharmacogenomics.* 2001;2:123–136.
107. Ayers M, Symmans WF, Stec J, et al. Gene expression profiling of fine needle aspirations of breast cancer identifies genes associated with complete pathologic response to neoadjuvant taxol/FAC chemotherapy [abstract 220]. *Breast Cancer Res Treat.* 2002;76: S64.

Chapter 25

Urothelial Carcinoma

Kevin C. Halling

Molecular Basis of Disease

The two main types of urothelial carcinoma (UC) are papillary UC (pTa) and "flat" UC (pTis), also known as noninvasive carcinoma in situ. Approximately 75% to 80% of UC are papillary and approximately 20% to 25% are CIS. Papillary tumors tend to recur but not progress to invasive cancer. CIS is aggressive and tends to progress to muscle-invasive cancer.

UC tumorigenesis is a multistep process. Papillary UC are thought to arise from areas of urothelial hyperplasia or from urothelial papillomas. Although most papillary tumors are low-grade tumors that have little tendency to progress to invasive tumors, a small proportion are high-grade tumors that have significant potential to progress to invasive UC. Most invasive UC tend to arise through the following sequence of events: normal urothelium to dysplasia to CIS to invasive cancer.

At the chromosomal level, DNA ploidy studies demonstrate that the majority of low-grade papillary tumors are diploid or near-diploid tumors, while the majority of high-grade papillary UC, CIS, and invasive UC (pTa tumors) are aneuploid tumors.[1] By comparative genomic hybridization (CGH) studies, noninvasive papillary UC have relatively few chromosomal abnormalities except for loss of all or part of chromosome 9, while tumors with lamina propria invasion (pT1 tumors) have a high number of chromosomal gains and losses.[2] The pT1 tumors also have loss of all or part of chromosome 9 but have numerous additional chromosomal abnormalities, which include whole or partial chromosomal losses and gains. Fluorescence in situ hybridization (FISH) studies provide further evidence that low-grade pTa UC tumors have relatively few chromosomal abnormalities, while CIS, high-grade pTa, and invasive UC have a high frequency of chromosomal abnormalities.[3]

Microsatellite analysis (MA) has revealed that frequent sites of allelic imbalance (AI) in UC include 3p, 4p, 8p, 9p, 9q, 11p, 13q, 17p, and 18q.[4–6] Regions with high rates of AI are the sites of known or putative tumor suppressor genes. Many of the regions that show high rates of AI correspond to the areas of chromosomal gains and losses detected by CGH.

Two important molecular genetic alterations that contribute to UC tumorigenesis are mutational and epigenetic alterations that inactivate the *P16* gene, the *TP53* gene, or both. *P16* loss is one of the earliest events in the development of both papillary and flat/invasive UC.[2,6–8] Mutations that inactivate the *TP53* gene appear to be found primarily in CIS and invasive UC and not low-grade papillary tumors and in part may be responsible for the aggressive behavior of these tumors.[9,10] Other oncogenes and tumor suppressor genes that have been found to be mutated in a subset of UC include *MYC*, *HRAS*, and *PTEN*.

Defective DNA mismatch repair (MMR) is manifested as microsatellite instability (MSI) at >30% of microsatellite markers examined and in most cases is associated with a loss of expression of one of the DNA MMR proteins, hMSH2, hMLH1, hMSH6, or hPMS2. MMR is rarely observed in UC of the bladder but is found in approximately 20% to 30% of upper tract UC.[11,12] The finding of defective MMR in an upper tract UC should prompt an investigation into the possibility that the patient may have hereditary nonpolyposis colorectal cancer (HNPCC) and a germline mutation of one of the DNA MMR genes.

Chromosomal instability (CIN) is present in invasive UC and CIS. It is likely that genes that maintain genomic stability are inactivated early during invasive UC tumorigenesis. CIN drives tumorigenesis and tumor progression by accelerating the mutation rate in tumor cells.[13] The genes responsible for CIN in invasive UC are not known, and the role of *TP53* inactivation in CIN has been a matter of debate. pTa tumors show no evidence of CIN but, as noted above, tend to be diploid or near-diploid tumors with relatively few chromosomal alterations. Until recently, the genetic alterations known to contribute to the formation of low-grade pTa tumors have been restricted to chromosome 9 and *P16* alterations. Most low-grade papillary UC and urothelial papillomas have missense mutations of the

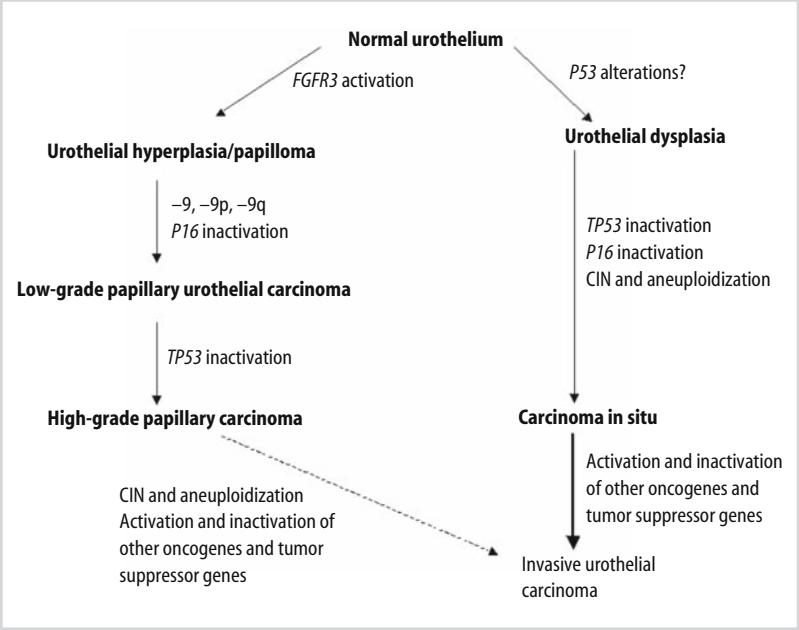


Figure 25-1. Genetic pathways for urothelial carcinoma tumorigenesis. Noninvasive papillary tumors are characterized by early activating mutations of the *FGFR3* gene, inactivating mutations or epigenetic alterations of the *P16* gene, and a diploid or near-diploid DNA content. CIS and invasive tumors are characterized by early inactivating mutations within the *TP53* and *P16* genes, chromosomal instability (CIN), and an aneuploid DNA content. A small proportion of papillary tumors may acquire *TP53* alterations or alterations of other as-yet unknown genes that cause invasive potential of these tumors.

fibroblast growth factor 3 (*FGFR3*) gene, while mutations of this gene are less common in invasive UC and CIS.[14]

Taken together, the various studies suggest that there are two genetic pathways leading to the development of UC.[2,7] One pathway leads to the formation of noninvasive papillary UC and the other to the development of CIS/invasive UC (Figure 25-1). The pathway for noninvasive papillary UC is characterized by the presence of *FGFR3* mutations and/or chromosome 9 alterations and *P16* inactivation. The pathway for invasive UC is characterized by early alterations in the *TP53* and *P16* genes, late alterations of other tumor suppressor genes and oncogenes, chromosomal instability, and aneuploidy. The genetic differences between noninvasive papillary and CIS/invasive tumors likely explain the markedly different behavior and prognosis of these tumors.[15]

Available Assays

DNA Ploidy Analysis by Flow Cytometry or Digital Image Analysis

Flow cytometry and digital image analysis (DIA) can be used to detect aneuploid cells in the urine that are consistent with a diagnosis of UC.[16,17] The sensitivity of flow cytometry is limited by the fact that it will fail to detect UC cells if the aneuploid cells represent only a small proportion of all the cells in the urine. Rare UC cells cannot be distinguished from normal cells in S phase in the "aneuploid" regions of the DNA ploidy histogram. DIA is a technique in which the cells on the slide are stained with a Feulgen stain, which stains the DNA in a stoichiometric fashion, followed by image analysis in an image cytometer to assess for aneuploidy. The advantage that DIA has over flow cytometry is that one can select the cells that are assessed for ploidy status. Cells that are clearly nonneoplastic can be ignored. This makes it possible to identify smaller populations of aneuploid cells in a background of diploid cells. Several studies suggest that the sensitivity and specificity of DIA ploidy analysis for the detection of UC are relatively high.[16,17] For example, Cajulis et al. showed that the sensitivity of DIA, flow cytometry, and conventional cytology for UC were 91%, 72%, and 61%, respectively, while the specificity of DIA, flow cytometry, and conventional cytology were 83%, 80%, and 100%, respectively. The main disadvantage of DIA is that it is relatively labor-intensive.

Fluorescence In Situ Hybridization

Fluorescence in situ hybridization is a technique that utilizes fluorescently labeled DNA probes to detect chromosomal abnormalities. There are two general types of FISH probes: chromosome enumeration probes (CEP) and locus-specific indicator (LSI) probes. CEP provide useful information on whether the cells exhibit significant aneusomy and CIN. Aneusomy is defined as an abnormal copy number for a specific chromosome. LSI probes hybridize to specific loci and can be used to detect alterations of the *TP53*, *HER2/NEU*, *MYC*, and other cancer-related genes.

Most UC are characterized by numerical and structural chromosomal abnormalities and a marked degree of CIN with variation in the chromosomal abnormalities found from cell to cell. The finding of aneusomy and CIN in a population of cells by FISH is virtually pathognomonic of malignancy.

A FISH assay for the detection of UC in urine has been developed. This assay utilizes a multiprobe mixture that contains CEP3, CEP7, CEP17, and LSI 9p21 probes labeled with red, green, aqua, and yellow fluorophores, respectively.[3,18] This probe set, now known as UroVysion (Abbott

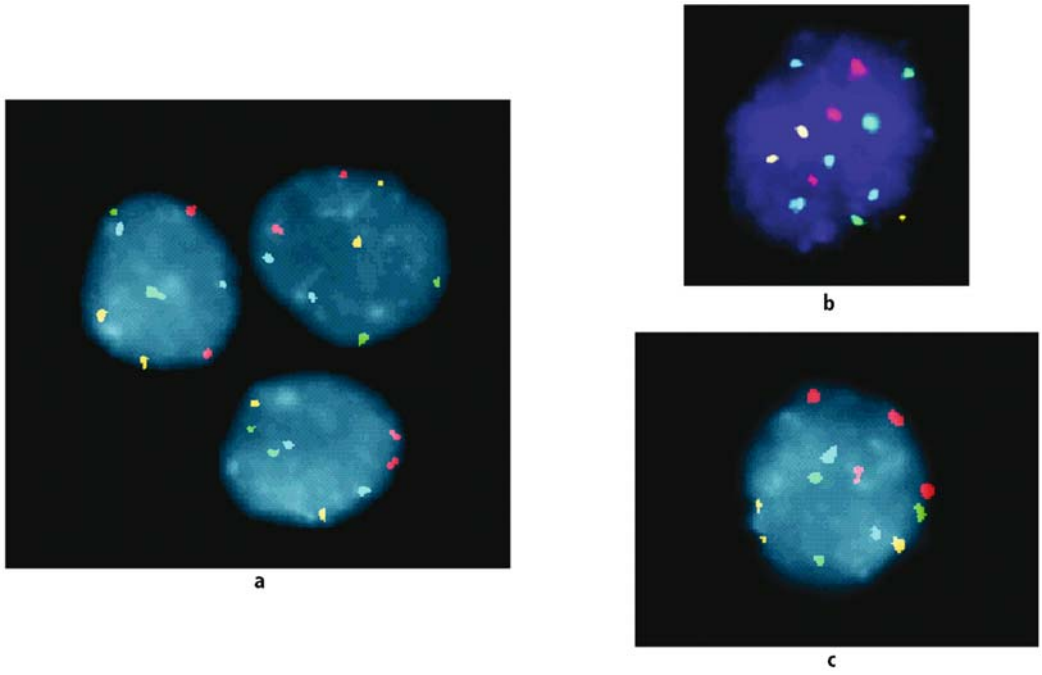


Figure 25-2. Representative examples of nonneoplastic urothelial cells (panel a) and UC cells (panels b and c) with FISH. Nonneoplastic cells generally have two signals for each of the four probes, though occasional nonneoplastic cells show only one signal for one or more of the probes due to random overlap of signals or imperfect hybridization efficiency. UC cells generally exhibit gains for two or more of the probes (i.e., polysomy) of the UroVysion probe set. The finding of just a few cells with polysomy is virtually pathognomic of malignancy. Red signals, CEP3; green signals, CEP7; aqua signals, CEP17; yellow signals, LSI 9p21.

Laboratories, Waukegan, IL), received FDA approval in August 2001 for monitoring UC patients for tumor recurrence. Representative examples of patients with FISH-positive and -negative findings are shown in Figure 25-2.

FISH using the UroVysion probe set is significantly more sensitive than urine cytology for the detection of recurrent UC.[3,16,19–21] (Note that the author receives industry funding from Abbott Laboratories and royalties from the sale of the UroVysion probe set). The sensitivity of UroVysion for the detection of CIS, invasive UC, and high-grade papillary tumors is greater than 95%.[3] The sensitivity of UroVysion is lower for low-grade papillary tumors than other UC but is still significantly better than cytology for low-grade tumors. Though more studies need to be performed, it is possible that the low-grade tumors that FISH fails to detect are less dangerous and the intervals between cystoscopy could be extended.

UroVysion not only has good sensitivity for UC in patients with biopsy-proven UC but also can detect recurrent UC before it is clinically evident by cystoscopy.[3,19–21] In the trial that led to FDA approval,[21] Sarosdy et al. reported that there were 36 patients with a negative cystoscopic examination but a positive FISH result. With continued follow-up, 15 (41.7%) of these cases were found to have biopsy-proven tumor recurrence with time to tumor diagnosis of 3 to 16 months (mean 6.0 months). Conversely, among 68 patients who had a negative cystoscopy and a negative FISH result, only 13 (19.1%) had a biopsy-proven recurrence at 3 to 19 months (mean 11.2 months). The time to recurrence was significantly less ($p = 0.014$) for the patients with a positive FISH result but a negative cystoscopy than for patients with a negative FISH result and a negative cystoscopy.[21]

One of the greatest advantages of urine cytology over other assays for UC is its high specificity. New assays that improve on the sensitivity of UC detection but have worse specificity than urine cytology are not particularly useful since the urologist has to perform cystoscopy and perhaps other studies to determine whether the result is a true- or false-positive result. A number of the antigen-based assays such as the FDA-approved BTA stat test have been shown to have higher sensitivity than cytology but significantly worse specificity. For example, the specificity of the BTA stat test is typically about 75%.[22] By contrast, FISH with UroVysion, like cytology, has high specificity for UC detection, typically exceeding 95%.[3,16,19–21]

The primary disadvantage of the FISH assay is that it requires more effort than conventional cytology or point-of-care assays such as the BTA stat test. Typical turnaround times for the FISH assay are 1 to 2 days, though the test can be performed in a single day. Automated "dot counters" are under development, and these may increase the ease of FISH test performance, reduce the cost of testing, and increase the throughput and sophistication of the data that can be obtained. Another shortcoming of the FISH test is its inability to detect some low-grade papillary tumors. An assay for UC cells that harbor *FGFR3* mutations (see *FGFR3* section below) may complement FISH and allow for the detection of virtually all UC.

Microsatellite Analysis

Sidransky and coworkers have described an assay that utilizes MA to detect recurrent UC.[23,24] The principles underlying this test are that UC exhibit frequent AI of certain loci (e.g., 9p21 and 17p13) and the detection of AI in urinary cells provides evidence that the patient has UC. The test also is able to identify tumor if there is evidence of MSI in the DNA extracted from the cells of the urine. In practice,

however, the majority of cases found to be positive by MA are positive because there is evidence of AI.[12]

A reference normal specimen, most commonly peripheral blood or buccal mouthwash, is required to perform this test. DNA is isolated from urinary cells and the normal tissue specimen from the same patient. Polymerase chain reaction (PCR) of 20 or more polymorphic microsatellite markers is performed separately for both the normal and tumor tissue, and the PCR products are analyzed by polyacrylamide gel or capillary electrophoresis. The markers used include loci that are frequently altered in UC, such as loci on 3p, 8p, 9p, 11p, 17p, and 18q. Differences in the pattern of PCR products are used to identify LOH and/or MSI in the tumor cells. Representative examples of AI and MSI can be seen in Figure 23-1 in chapter 23.

An initial study, in a relatively small population of patients (20 patients with biopsy-proven bladder cancer), revealed that the sensitivity of MA was 95% compared to 50% for urine cytology.[23] Most cases were detected due to 9p21 loss, with only a few cases being detected due to the presence of MSI.

Several studies have demonstrated the utility of MA for the detection of UC in urine specimens.[25–28] In a phase II trial, MA was compared to the BTA stat test and cytology.[25] The sensitivity of MA, the BTA stat test, and cytology were 74%, 56%, and 22%, respectively, while the specificity of MA, the BTA stat test, and cytology were 82%, 79%, and 95%, respectively.

The MA method is promising but currently has several limitations, which include labor-intensiveness, poorly defined criteria for a positive result, and poor analytical sensitivity. The criteria used to decide whether a case is positive or negative for LOH, MSI, or both have not been well defined. Studies are needed to define these criteria and determine how varying these criteria affects the sensitivity and specificity of the test. As for other tests, less-stringent criteria for a positive case would increase sensitivity and decrease specificity, while more-stringent criteria would decrease sensitivity but increase specificity. Most studies suggest that MA has high sensitivity and specificity. However, a study by Christensen et al. suggests that MA analysis lacks specificity.[29] They found that patients with benign prostatic hyperplasia and cystitis showed evidence of MSI and LOH.

The analytical sensitivity (and consequently diagnostic sensitivity) of MA is limited by the fact that as the percentage of neoplastic cells in the urine becomes smaller, detection of LOH or MSI becomes more difficult. The presence of urinary leukocytes compromises the diagnostic sensitivity of MA.[30] For this reason, MA would likely not be useful for detecting recurrent tumor in superficial UC patients receiving Bacillus Calmette-Guerin (BCG) therapy, since the these patients frequently have numerous leukocytes in their urine. The sensitivity of MA also would be expected to be low for urine samples with high percentages of benign squamous cells, a common finding in nonclean catch urine specimens.

Telomerase

Telomerase is an enzyme that normally maintains the length of the telomeres in stem cells. Interestingly, telomerase is upregulated in most malignant cells and is thought to be necessary for the maintenance of the "immortal" phenotype that typifies malignant cells. Because telomerase is overexpressed in almost all malignancies but not in most normal tissues, it has been suggested that telomerase may be a good marker for the presence of malignant cells in various cytologic specimens, including urine and oral washings.

The utility of telomerase for detecting UC in urine specimens has been evaluated.[31–35] Most of the early studies assessed for the presence of telomerase activity with the telomere repeat amplification protocol (TRAP) assay.[36] Representative examples of patients with TRAP-positive and -negative results are shown in Figure 25-3. The sensitivity of the TRAP assay for UC has varied widely. Three studies have shown that the sensitivity of the TRAP assay for UC is higher for bladder wash specimens than for voided urine specimens.[33,37] Most of the studies demonstrate that the TRAP assay has high specificity for UC. One possible reason for the relatively poor sensitivity of the TRAP assay observed in some studies is that the telomerase enzyme is labile. For example, an initial study by Ramakumar et al. had a much higher sensitivity of telomerase than a follow-up study by the same group.[31,38]

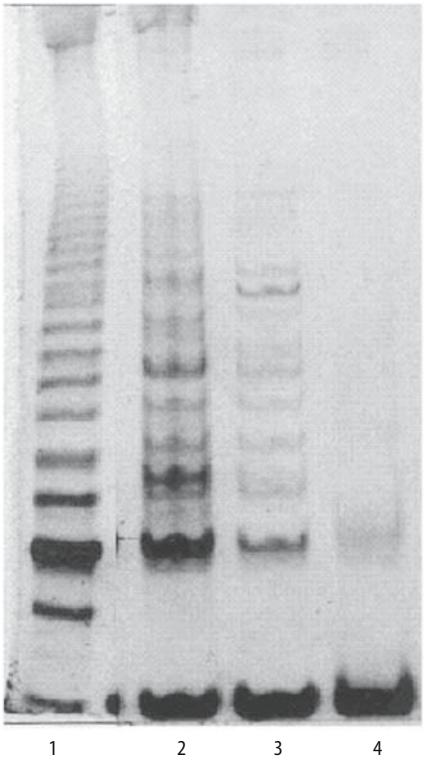


Figure 25-3. Representative examples of telomerase positive and negative cases. Lane 1, thyroid cancer cell line (positive control); lanes 2 and 3, positive telomerase results for urinary sediments from patients with UC; lane 4, negative telomerase result for urinary sediment from patient without UC. The positive cases demonstrate the ladder typically seen when telomerase activity is present in the specimen.

The most likely explanation is that in the original study, the urine specimens were placed on ice and processed quickly for telomerase activity, while in the follow-up study, the urine specimens were collected at ambient temperature and processed several hours after collection.

Nonfastidious assays for telomerase are needed for practical reasons. Several investigators have attempted to overcome telomerase lability by assaying for the RNA component of telomerase using a reverse transcription–polymerase chain reaction (RT-PCR) assay rather than for telomerase activity.[34,35] These studies suggest that the RT-PCR assay is more sensitive but less specific than the TRAP assay. For example, Neves et al. reported a sensitivity and specificity of 75% and 69%, respectively, for RT-PCR detection of telomerase.[32] It should be possible to define cutoffs for the RT-PCR assay that improve the specificity of the assay.

FGFR3

The detection of cells in the urine that harbor *FGFR3* mutations is a promising way to detect the low-grade papillary tumors that are not discovered with assays such as FISH or MA.[14,30,39] FGFR3 is a tyrosine kinase receptor. Germline point mutations in various domains of *FGFR3* are associated with human skeletal disorders such as hypochondroplasia and achondroplasia, and somatic mutations of *FGFR3* have been identified in bladder cancer and myeloma. Interestingly, two groups have demonstrated a high frequency of somatic *FGFR3* point mutations in low-grade papillary UC and urothelial papilloma but not in high-grade papillary UC, CIS, or invasive UC.[14,39] Billerey et al. found that the frequency of *FGFR3* mutations by stage was pTa 74%, pTis 0%, pT1 21%, and pT2 to pT4 16%. Grade 1 showed 84%, grade 2 showed 55%, and grade 3 showed 7%. The most commonly identified *FGFR3* mutation was an S249C mutation (33 of 48 tumors; 69%), but R248C, G372C, Y375C, and K652E mutations also were identified. The difference in the frequency of *FGFR3* mutations between low-grade and high-grade tumors was highly significant ($P < 0.0001$) and is consistent with the current model of bladder tumor progression in which the most common precursor of invasive UC is CIS (Figure 25-1).

Other Molecular Diagnostic Assays for UC

TP53 mutations are common in UC, especially in high-grade UC.[40] Assays that assess for *TP53* status could potentially be used to assess prognosis and detect tumor recurrence. Some studies have shown that TP53 overexpression detected by immunohistochemical analysis of paraffin-embedded tumors is associated with worse prognosis and higher risk of muscle invasion,[41,42] while others have not.[43] Immunohistochemical analysis of bladder tumors for TP53 expression has not been widely utilized by urologists or pathologists. An assay that detects urinary cells with *TP53* mutations could be used to detect recurrent tumors.[44] However, the heterogeneity of *TP53* mutations present in different bladder cancers makes it difficult to design an assay for clinical use.

A few studies have shown that the antiapoptotic protein survivin may be a sensitive and specific marker for the detection of recurrent UC,[45] but blinded prospective studies are needed to further evaluate the clinical utility of this assay. Alterations in certain genes such as glutathione S-transferase M1 and N-acetyltransferase, that encode proteins that metabolize carcinogens, may increase an individual's risk of developing bladder cancer especially among smokers, but assays for these alterations also have not been used clinically.[46]

Clinical Significance

Urine cytology has been the primary laboratory method for diagnosing and monitoring UC for the past 50 years. Urine cytology has excellent specificity but poor sensitivity for the detection of UC.[3] A survey of 20 published studies has shown that the sensitivity of urine cytology for grade 1, 2, and 3 tumors over the past 20 years has been 21%, 53%, and 78%, respectively.[3] The problem with false-negative urine cytology test results is that the patient's tumor, if not detected by cystoscopy either, will have 3 or more months to progress to a higher, potentially incurable state before it is detected. This is of particular concern for grade 3 UC, which inexorably progress if not removed or treated.

The suboptimal sensitivity of urine cytology has prompted investigators to develop new tests with improved sensitivity for UC detection. For the most part, assays for the detection of UC fall into one of three categories: morphologic tests, antigen-detection tests, and molecular tests. The only assay that uses cellular morphologic changes alone to determine whether cancer is present is urine cytology. Other assays such as Immunocyt and FISH use morphology as part of the assay but also utilize the detection of tumor-associated antigens or genetic alterations, respectively, as part of the detection method. Antigen-based assays such as the BTA stat test and NMP22 rely on the detection of antigens that are present at higher concentrations in the urine of patients with UC. Genetic assays such as MA, FISH, and DNA ploidy analysis by image or flow cytometry are based on the detection of genetic alterations in cells that strongly suggest the presence of UC or other tumors. Antigen-based tests often have statistically better sensitivity but worse specificity (approximately 70%) than urine cytology. False-positive results can occur with inflammatory states that lead to an increase in the concentration of the antigens that are being measured. The high false-positive rate for many of the antigen-based tests leads to the need for confirmatory cystoscopy to determine whether UC truly is present.

Molecular tests generally have high specificity and sensitivity but are more laborious than antigen-based assays. Molecular tests tend to have high specificity because the genetic alterations are generally specific for cancer and not found in normal or reactive cells. Reactive cells are non-neoplastic epithelial cells that exhibit atypical cytologic features due to the presence of inflammation, chemotherapy, or radiation.

Currently, there are no therapies that are directed to specific molecular targets in UC. However, therapeutic agents similar to Gleevec and Herceptin, which are used for patients with specific genetic alterations in chronic myelogenous leukemia (*BCR*/*ABL* translocation) and breast cancer (*HER2*/*NEU* amplification), respectively, will presumably become available for treatment of UC patients. When such agents become available, it will likely be important to evaluate the patient's tumor for the presence or absence of the genetic alteration that the drug targets.

Quality Control and Laboratory Issues

While a number of molecular tests have been developed for the detection of recurrent UC in patients, additional carefully designed and controlled studies are needed to further evaluate the performance characteristics of most of these assays before they can be widely introduced into clinical practice. Nonetheless, it is likely that one or more of these tests will be widely used by urologists to improve the management of patients who have been diagnosed with or are being evaluated for UC. Test validation studies should evaluate the diagnostic sensitivity and specificity of these assays relative to standard tests (such as cytology) in well-characterized patient populations. When applicable, receiver-operator curves should be evaluated to determine how different cutoff values affect the diagnostic sensitivity and specificity of the test. Other factors that need to be assessed as a part of test validation include the analytical sensitivity and reproducibility of the test, as well as specimen stability issues. For clinical implementation of a test, appropriate controls should be included with each run. For PCR-based tests, this would include positive, negative, and "no DNA" controls, and possibly analytical sensitivity and precision controls when a quantitative result is produced.

References

1. Bittard H, Lamy B, Billery C. Clinical evaluation of cell deoxyribonucleic acid content measured by flow cytometry in bladder cancer. *J Urol.* 1996;155:1887–1891.
2. Richter J, Jiang F, Gorog JP, et al. Marked genetic differences between stage pTa and stage pT1 papillary bladder cancer detected by comparative genomic hybridization. *Cancer Res.* 1997;57:2860–2864.
3. Halling KC, King W, Sokolova IA, et al. A comparison of cytology and fluorescence in situ hybridization for the detection of bladder cancer. *J Urol.* 2000;164:1768–1775.
4. Dalbagni G, Presti J, Reuter V, Fair WR, Cordon-Cardo C. Genetic alterations in bladder cancer. *Lancet.* 1993;342:469–471.
5. Knowles MA, Elder PA, Williamson M, Cairns JP, Shaw ME, Law MG. Allelotype of human bladder cancer. *Cancer Res.* 1994;54:531–538.
6. Rosin MP, Cairns P, Epstein JI, Schoenberg MP, Sidransky D. Partial allelotype of carcinoma in situ of the human bladder. *Cancer Res.* 1995;55:5213–5216.
7. Cairns P, Shaw ME, Knowles MA. Initiation of bladder cancer may involve deletion of a tumour-suppressor gene on chromosome 9. *Oncogene.* 1993;8:1083–1085.
8. Ruppert JM, Tokino K, Sidransky D. Evidence for two bladder cancer suppressor loci on human chromosome 9. *Cancer Res.* 1993;53:5093–5095.
9. Esrig D, Spruck CH, Nichols PW, et al. p53 nuclear protein accumulation correlates with mutations in the p53 gene, tumor grade, and stage in bladder cancer. *Am J Pathol.* 1993;143:1389–1397.
10. Tsutsumi M, Sugano K, Yamaguchi K, Kakizoe T, Akaza H. Correlation of allelic loss of the P53 gene and tumor grade, stage, and malignant progression in bladder cancer. *Int J Urol.* 1997;4:74–78.
11. Gonzalez-Zulueta M, Ruppert JM, Tokino K, et al. Microsatellite instability in bladder cancer. *Cancer Res.* 1993;53:5620–5623.
12. Hartmann A, Zanardo L, Bocker-Edmonston T, et al. Frequent microsatellite instability in sporadic tumors of the upper urinary tract. *Cancer Res.* 2002;62:6796–6802.
13. Jallepalli PV, Lengauer C. Chromosome segregation and cancer: cutting through the mystery. *Nat Rev Cancer.* 2001;1:109–117.
14. Billerey C, Chopin D, Aubriot-Lorton MH, et al. Frequent FGFR3 mutations in papillary non-invasive bladder (pTa) tumors. *Am J Pathol.* 2001;158:1955–1959.
15. Sauter G, Mihatsch MJ. Pussycats and baby tigers: non-invasive (pTa) and minimally invasive (pT1) bladder carcinomas are not the same. *J Pathol.* 1998;185:339–341.
16. Dalquen P, Kleiber B, Grilli B, Herzog M, Bubendorf L, Oberholzer M. DNA image cytometry and fluorescence in situ hybridization for noninvasive detection of urothelial tumors in voided urine. *Cancer Cytopathol.* 2002;96:374–379.
17. Cajulis RS, Haines GK, Frias-Hidvegi D, McVary K, Bacus JW. Cytology, flow cytometry, image analysis, and interphase cytogenetics by fluorescence in situ hybridization in the diagnosis of transitional cell carcinoma in bladder washes: a comparative study. *Diagn Cytopathol.* 1995;13:214–223.
18. Sokolova IA, Halling KC, Jenkins RB. The development of a multitarget, multicolor fluorescence in situ hybridization assay for the detection of urothelial carcinoma in urine. *J Mol Diagn.* 2000;2:116–123.
19. Bubendorf L, Grilli B, Sauter G, Mihatsch MJ, Gasser TC, Dalquen P. Multiprobe FISH for enhanced detection of bladder cancer in voided urine specimens and bladder washings. *Am J Clin Pathol.* 2001;116:79–86.
20. Skacel M, Fahmy M, Brainard JA, et al. Multitarget fluorescence in situ hybridization assay detects transitional cell carcinoma in the majority of patients with bladder cancer and atypical or negative urine cytology. *J Urol.* 2003;169:2101–2105.
21. Sarosdy MF, Schellhammer P, Bokinsky G, et al. Clinical evaluation of a multi-target fluorescent in situ hybridization assay for detection of bladder cancer. *J Urol.* 2002;168:1950–1954.
22. Glas AS, Roos D, Deutekom M, Zwinderman AH, Bossuyt PM, Kurth KH. Tumor markers in the diagnosis of primary bladder cancer. A systematic review. *J Urol.* 2003;169:1975–1982.
23. Mao L, Schoenberg MP, Scicchitano M, et al. Molecular detection of primary bladder cancer by microsatellite analysis. *Science.* 1996;271:659–662.
24. Steiner G, Schoenberg MP, Linn JF, Mao L, Sidransky D. Detection of bladder cancer recurrence by microsatellite analysis of urine. *Nat Med.* 1997;3:621–624.
25. van Rhijn BW, Lurkin I, Kirkels WJ, van der Kwast TH, Zwarthoff EC. Microsatellite analysis—DNA test in urine competes with cys-

toscopy in follow-up of superficial bladder carcinoma: a phase II trial. *Cancer.* 2001;92:768–775.
26. Amira N, Mourah S, Rozet F, et al. Non-invasive molecular detection of bladder cancer recurrence. *Int J Cancer.* 2002;101:293–297.
27. Utting M, Werner W, Dahse R, Schubert J, Junker K. Microsatellite analysis of free tumor DNA in urine, serum, and plasma of patients: a minimally invasive method for the detection of bladder cancer. *Clin Cancer Res.* 2002;8:35–40.
28. Seripa D, Parrella P, Gallucci M, et al. Sensitive detection of transitional cell carcinoma of the bladder by microsatellite analysis of cells exfoliated in urine. *Int J Cancer.* 2001;95:364–369.
29. Christensen M, Wolf H, Orntoft TF. Microsatellite alterations in urinary sediments from patients with cystitis and bladder cancer. *Int J Cancer.* 2000;85:614–617.
30. van Rhijn BW, Lurkin I, Chopin DK, et al. Combined microsatellite and FGFR3 mutation analysis enables a highly sensitive detection of urothelial cell carcinoma in voided urine. *Clin Cancer Res.* 2003;9: 257–263.
31. Halling KC, King W, Sokolova IA, et al. A comparison of BTA stat, hemoglobin dipstick, telomerase and Vysis UroVysion assays for the detection of urothelial carcinoma in urine. *J Urol.* 2002;167:2001–2006.
32. Neves M, Ciofu C, Larousserie F, et al. Prospective evaluation of genetic abnormalities and telomerase expression in exfoliated urinary cells for bladder cancer detection. *J Urol.* 2002;167:1276–1281.
33. Kinoshita H, Ogawa O, Kakehi Y, et al. Detection of telomerase activity in exfoliated cells in urine from patients with bladder cancer. *J Natl Cancer Inst.* 1997;89:724–730.
34. Ito H, Kyo S, Kanaya T. Detection of human telomerase reverse transcriptase messenger RNA in voided urine samples as a useful diagnostic tool for bladder cancer. *Clin Cancer Res.* 1998;4:2807–2810.
35. Muller M, Krause H, Heicappell R, Tischendorf J, Shay JW, Miller K. Comparison of human telomerase RNA and telomerase activity in urine for diagnosis of bladder cancer. *Clin Cancer Res.* 1998;4:1949–1954.
36. Kim NW, Piatyszek MA, Prowse KR, et al. Specific association of human telomerase activity with immortal cells and cancer. *Science.* 1994;266:2011–2015.
37. Heine B, Hummel M, Muller M, Heicappell R, Miller K, Stein H. Non-radioactive measurement of telomerase activity in human bladder cancer, bladder washings, and in urine. *J Pathol.* 1998;184: 71–76.
38. Ramakumar S, Bhuiyan J, Besse JA, et al. Comparison of screening methods in the detection of bladder cancer. *J Urol.* 1999;161:388–394.
39. van Rhijn BW, Montironi R, Zwarthoff EC, Jobsis AC, van der Kwast TH. Frequent FGFR3 mutations in urothelial papilloma. *J Pathol.* 2002;198:245–251.
40. Uchida T, Wada C, Ishida H, et al. p53 mutations and prognosis in bladder tumors. *J Urol.* 1995;153:1097–1104.
41. Esrig D, Elmajian D, Groshen S, et al. Accumulation of nuclear p53 and tumor progression in bladder cancer. *N Engl J Med.* 1994;331: 1259–1264.
42. Sarkis AS, Dalbagni G, Cordon-Cardo C, et al. Nuclear overexpression of p53 protein in transitional cell bladder carcinoma: a marker for disease progression. *J Natl Cancer Inst.* 1993;85:53–59.
43. Shariat SF, Weizer AZ, Green A, et al. Prognostic value of P53 nuclear accumulation and histopathologic features in T1 transitional cell carcinoma of the urinary bladder. *Urology.* 2000;56:735–740.
44. Sidransky D, Von Eschenbach A, Tsai YC, et al. Identification of p53 gene mutations in bladder cancers and urine samples. *Science.* 1991;252:706–709.
45. Smith SD, Wheeler MA, Plescia J, Colberg JW, Weiss RM, Altieri DC. Urine detection of survivin and diagnosis of bladder cancer. *JAMA.* 2001;285:324–328.
46. Brockmoller J, Cascorbi I, Kerb R, Roots I. Combined analysis of inherited polymorphisms in arylamine N-acetyltransferase 2, glutathione S-transferases M1 and T1, microsomal epoxide hydrolase, and cytochrome P450 enzymes as modulators of bladder cancer risk. *Cancer Res.* 1996;56:3915–3925.

Chapter 26

Sarcomas

Giovanni Tallini and Pei Hui

Introduction

Oncologic molecular pathology focuses on identifying and understanding molecular and genetic alterations underlying the development and progression of neoplastic processes. Mesenchymal malignancies may be classified into two pathogenetic types: sarcomas with complex genetic alterations and sarcomas with specific recurrent chromosomal translocations. The first type includes the majority of high-grade, pleomorphic mesenchymal malignancies that are characterized by complex chromosomal abnormalities, for example, malignant fibrous histiocytoma (MFH), osteogenic sarcoma, and embryonal rhabdomyosarcoma. In the second type, the sarcomas are translocation specific, that is, harboring a recurrent chromosomal translocation leading to an in-frame fusion of coding sequences from each of the two rearranged genes. The translocation results in the production of a chimeric transcript encoding a fusion protein with oncogenic activity. Histologically, the translocation-specific sarcomas are generally a monomorphic proliferation of neoplastic cells. This pathogenetic classification appears biologically relevant and is best illustrated by the sarcoma types observed in Li-Fraumeni syndrome and therapy-related malignancies. Li-Fraumeni patients, with *TP53* germline mutation, are prone to sarcomas with complex karyotypes, such as osteogenic sarcoma and MFH, which constitute major mesenchymal cancers. On the other hand, translocation-specific sarcomas virtually never occur in Li-Fraumeni patients.[1]

During the past 10 years, tremendous progress has been made in characterizing many translocation-specific sarcomas. These chromosomal translocations and associated fusion genes are highly specific to given sarcomas and occur in such a high prevalence that they are essentially used to define these neoplasms at both the pathobiological and clinical diagnostic levels.[2,3] These sarcomas are the focus of this chapter.

Molecular Basis of Disease

Chromosomal Translocations and Gene Rearrangements in Sarcomas

Chromosomal abnormalities (translocations, inversions, deletions, and insertions) are associated with DNA recombination and structural rearrangement of the genes located at the DNA breakpoints. If the breakpoint occurs within the involved gene, an altered gene structure arises. If the breakpoint occurs outside the gene, it may involve control elements critical for gene expression. In either case, the result can be a dramatic change in the gene structure, expression levels, or both.

Molecular analysis has identified two general mechanisms through which chromosomal translocations result in altered gene function. The first mechanism is gene fusion, in which chimeric or fusion genes are the result of joining of two parent genes (one upstream, or 5′, the other downstream, or 3′, to the breakpoint). Both genes are truncated by the translocation involving the coding portions of the parent genes. In general, translocation breakpoints are located in noncoding introns and the normal splicing mechanism removes the chimeric intron sequence. The exons are spliced "in frame" for the translational reading frame and can be translated into a novel fusion protein. In rare instances, the breakpoints are located in the exons of the parent genes. This may result in a novel chimeric product if the translational reading frame is maintained, or it may produce a truncated protein (encoded by the 5′ gene sequence) if the reading frame is lost. Transcription of chimeric genes is usually under the control of the upstream parent gene promoter but may be influenced by DNA sequences in or close to the downstream gene.

The second mechanism through which chromosomal translocations result in altered gene function is promoter swapping or exchange. The breakpoint occurs at the 5′ end of the coding region of the involved gene. This results in

the replacement of its promoter region with enhancer elements or the promoter from the translocation partner. Promoter swapping leads to transcriptional activation, but the protein is wild type.

Numerous fusion genes have been identified in malignant tumors of the soft tissues and in mesenchymal tumors in general (see Table 26-1). The majority of sarcoma translocations result in in-frame fusion genes, resulting

Table 26-1. Chromosomal Alterations and Aberrant Gene Products in Sarcomas

Tumor Type	Chromosomal Changes	Fusion Gene	Prevalence (%)	Transcript Variability (number of fusion transcript variants)	Aberrant Function	Reference
Ewing's sarcoma (ES)/peripheral neuroectodermal tumor (PNET)	t(11;22)(q24;q12)	*EWS-FLI1*	95	Large (18)	Transcription factor	Delattre. *Nature.* 1992;359:162.
	t(21;22)(q22;q12)	*EWS-ERG*	5	Large (4)	Transcription factor	Sorensen. *Nat Genet.* 1994;6:146. Zucman. *EMBO J.* 1993;12:4481.
	t(7;22)(p22;q12)	*EWS-ETV1*	<1	Unknown	Transcription factor	Jeon. *Oncogene.* 1995;10:1229.
	t(17;22)(q12;q12)	*EWS-E1AF*	<1		Transcription factor	Urano. *Biochem Biophys Res Commun.* 1996; 219:608.
	t(2;22)(q33;q12)	*EWS-FEV*	<1	Unknown	Transcription factor	Peter. *Oncogene*. 1997; 14:1159.
Desmoplastic small round cell tumor (DSRCT)	t(11;22)(p13;q12)	*EWS-WT1*	100	Small (3)	Transcription factor	Ladanyi. *Cancer Res.* 1994;54:2837.
Alveolar rhabdomyosarcoma (ARMS)	t(2;13)(q35;p14)	*PAX3-FKHR*	95	None (1)	Transcription factor	Galili. *Nat Genet.* 1993; 5:230. Shapiro. *Cancer Res.* 1993;53:5108.
	t(1;13)(p36;p14)	*PAX7-FKHR*	10	None (1)	Transcription factor	Davis. *Cancer Res.* 1994;54:2869.
Alveolar soft part sarcoma	der (17) t(X;17)(p11;q25)	*ASPL-TFE3*	100	Small (2)	Transcription factor	Ladanyi. *Oncogene.* 2001;20:48–57.
DFSP/giant cell fibroblastoma	t(17;22)(q22;q13)	*COL1A1-PDGFB*	100	Large (>8)	Autocrine growth factor	Simon. *Nat Genet.* 1997; 15:95.
Synovial sarcoma	t(X;18)(p11;q11)	*SYT-SSX1*	65	None (1)	Transcription factor	Clark. *Nat Genet.* 1994; 7:502.
	t(X;18)(p11;q11)	*SYT-SSX2*	35	None (1)	Transcription factor	Crew. *EMBO J.* 1995;14: 2333. De Leeuw. *Hum Mol Genet.* 1995;4:1097.
	t(X;18)(p11;q11)	*SYT-SSX4*	<1	None (1)	Transcription factor	Skytting. *J Natl Cancer Inst.* 1999;91:974.
Clear cell sarcoma (malignant melanoma of soft parts)	t(12;22)(q13;q12)	*EWS-ATF1*	90	Small (2)	Transcription factor	Zucman. *Nat Genet.* 1993;4:341.
Congenital fibrosarcoma	t(12;15)(p13;q25)	*ETV6-NTRK3*	100	None (1)	Receptor tyrosine kinase	Knezevich. *Nat Genet.* 1998;18:184.
Inflammatory myofibroblastic tumor*	t(1;2)(q25;p23)	*TPM3-ALK*	20	None (1)	Receptor tyrosine kinase	Griffin. *Cancer Res.* 1999;59:2776.
	t(2;19)(p23;p13)	*TPM4-ALK*	10	None (1)	Receptor tyrosine kinase	Lawrence. *Am J Pathol.* 2000;157:377.
	t(2;17)(p23;23)	*CLTC-ALK*	Uncertain	None (1)	Receptor tyrosine kinase	Bridge, *Am J Pathol.* 2001;159:41.
	12q15 rearrangement	*HMGIA2* rearrangement	Uncertain	Unknown	Transcription cofactor	Kazmierczac. *Cancer Genet Cytogenet.* 1999; 112:156.

Table 26-1. Chromosomal Alterations and Aberrant Gene Products in Sarcomas (*Continued*)

Tumor Type	Chromosomal Changes	Fusion Gene	Prevalence (%)	Transcript Variability (number of fusion transcript variants)	Aberrant Function	Reference
Extraskeletal myxoid chondrosarcoma	t(9;22)(q22;q12)	*EWS-CHN*	75	Small (2)	Transcription factor	Labelle. *Hum Mol Genet.* 1995;4:2219. Clark. *Oncogene.* 1996;12:229.
	t(9;17)(q22;q11)	*TAF2N-CHN*	25	Unknown	Transcription factor	Sjogren. *Cancer Res.* 1999;59:5064. Panagopoulos. *Oncogene.* 1999;18:7594. Attwooll. *Oncogene.* 1999;18:7599.
	t(9;15)(q22;q21)	*TCS12-CHN*	Uncertain	Unknown	Transcription factor	Sjogren. *Cancer Res.* 2000;60:6832.
Lipoma	12q15 rearrangement	*HMGIA2-LPP* Other *HMGIC* rearrangement	50†	Unknown	Transcription cofactor	Ashar. *Cell.* 1995;82:57. Schoenmakers. *Nat Genet.* 1995;10:436. Petit, *Genomics.* 1996;36:118.
	6p21	*HMGA1* rearrangement	10†	Unknown	Transcription cofactor	Tkachenko. *Cancer Res.* 1997;57:2276.
Lipoblastoma	8q11–13 rearrangement	*HAS2-PLAG1*	>90	None (1)	Transcription factor	Hibbard. *Cancer Res.* 2000;60:4869.
	t(7;8)(q22;q12)	*COL1A2-PLAG1*	Uncertain	Unknown	Transcription factor	Hibbard. *Cancer Res.* 2000;60:4869.
Myxoid/round cell liposarcoma	t(12;16)(q13;p11)	*TLS-CHOP*	>95	Small (3)	Transcription factor	Crozat. *Nature.* 1993;363:640. Rabbitts. *Nat Genet.* 1993;4:175.
	t(12;22)(q13;q12)	*EWS-CHOP*	2–5	Unknown	Transcription factor	Panagopoulos. *Oncogene.* 1996;12:489.
Well differentiated liposarcoma	12q13–15 amplification	*HMGIC*, *MDM2*, *SAS*, *CDK4*	>60	Unknown	Gene amplification resulting in cell growth dysregulation	Meza-Zepeda. *Cancer.* 2001;31:264. Pedeutour. *Genes Chromosomes Cancer.* 1994;10:85.
Endometrial stromal tumors	t(7;17) (p15;q21)	*JAZF1-JJAZ1*	40–100	None (1)	Transcription factor	Koontz. *Proc Natl Acad Sci* USA. 2001;98:6348.
Aggressive angiomyxoma	12q13–15 rearr	*HMGIC* rearrangement	Uncertain	Unknown	Transcription cofactor	Nucci. *Genes Chromosomes Cancer.* 2001;32:172.
Angiomatoid fibrous histiocytoma	t(12;16)(q13;p11)	*TLS-ATF1*	Uncertain	Unknown	Transcription factor	Waters. *Cancer Genet Cytogenet.* 2000;121:109.
Extrarenal rhabdoid tumor	del22q11.2 or 22q11.2 rearrangement	*INI1* deletion‡	>90	Unknown	Loss of tumor suppressor gene	Newsham. *Genomics.* 1994;19:433.
Low grade fibromyxoid sarcoma	t(7;16)(q33;p11)	*FUS-CREB3L2* (*BBF2H7*)	>95 %	Large	Transcription factor	Mertens. *Lab Invest.* 2005;85,408. Reid, *Am J Surg Pathol.* 2003;27, 1229.
	t(11;16)(q11;p11)	*FUS-CREB3L1*	<5 %	Unknown	Transcription factor	Mertens. *Lab Invest.* 2005;85,408.

*The proportion of inflammatory myofibroblastic tumors with ALK gene rearrangement is 35%, overall.
† Percentages indicate the proportion of tumors with abnormal karyotypes which have 12q13–15 or 6p21 chromosomal alterations. (*Sources:* Heim S, Mitelman F. *Cancer Cytogenetics,* 2nd ed. New York: Wiley-Liss, 1995; Fletcher CD, Akerman M, Dal Cin P, et al. Correlation between clinicopathological features and karyotype in lipomatous tumors. A report of 178 cases from the Chromosomes and Morphology (CHAMP) Collaborative Study Group. *Am J Pathol*. 1996;148:623–630.
‡ No fusion gene is present.

in abnormal chimeric transcription factors.[3] In a few cases, the gene fusion results in an aberrant tyrosine kinase or an autocrine growth factor.[4–6] A recent investigation identified the der(17) associated with the nonreciprocal t(X;17)(p11.2;q25) of human alveolar soft part sarcoma.[7] The translocation produces a chimeric transcript between the *TFE3* transcription factor gene and *ASPL*, a novel gene at 17q25. *ASPL* encodes a UBX-like domain at the C-terminus of the encoded protein. In alveolar soft part sarcoma, the 5′ end of *ASPL* is fused to exon 3 or 4 of *TFE3*, resulting in a fusion protein retaining the C-terminal TFE3 DNA-binding domain, a possible aberrant transcriptional regulator.

A recurrent t(7;17)(p15;q21) has been identified in endometrial stromal tumors.[8] Two new zinc finger genes are fused as a result of the translocation: *JAZF1* and *JJAZ1*. Protein products of the zinc finger genes usually function as transcriptional regulators via specific DNA binding through the zinc finger motif. The chimeric protein in endometrial stromal tumors has a tumor-specific mRNA transcript containing 5′ *JAZF1* and 3′ *JJAZ1* sequences including the zinc finger encoding regions from both parent genes. Since gene expression of wild-type *JAZF1* is present in normal endometrial stromal cells, the *JAZF1-JJAZ1* fusion gene present in endometrial stromal tumors likely results in aberrant transcriptional regulation in a lineage-specific manner.

Specificity and Oncogenic Nature of Sarcoma Fusion Transcripts

Unlike many epithelial neoplasms, where diverse genetic alterations usually underlie the stepwise progression of precursor lesions leading ultimately to the emergence of malignant clones, soft tissue malignancies have no identifiable precursor lesions, and clear, stepwise progressive mutational events have not been described. Rather, more often than not, a single genetic alteration is found in a particular type of sarcoma. In addition, chromosomal fusions in soft tissue sarcomas do not seem to represent a form of generalized genomic instability, as occurs with germline *TP53* mutations[1] or with microsatellite instability associated with colon carcinoma.[9] Benign tumor counterparts of soft tissue sarcomas usually carry quite different genetic or chromosomal abnormalities, or both. For example, the specific sets of chromosomal alterations found in soft tissue lipomas are not among those consistently observed in liposarcoma.[10,11] Similar to leukemogenesis and lymphomagenesis, the fusion gene in a given sarcoma is speculated to be oncogenic only in a specific cell type at a specific differentiation stage.[12]

In general, the genes involved in sarcoma translocation are transcription factors or cofactors. Many of the chimeric proteins include a strong transcriptional activator N-terminal domain encoded by one partner gene fused with a DNA-binding domain encoded by the other partner gene. In fact, fusion of domains capable of activating transcription with other domains featuring specific DNA-binding function appears to be a common theme shared among neoplasms of mesenchymal derivation, such as soft tissue tumors and leukemia. EWS (and its homologous TLS/FUS) is a powerful transcription activator[13] and provides a paradigm for this type of oncogenic mechanism, as also indicated by its "promiscuity" as a fusion partner (Table 26-1).

Available data indicate that the fusion genes produced in translocation-specific sarcomas are the initiating events that either are necessary or sufficient for the genesis of these malignant soft tissue tumors. In vitro and in vivo experiments have shown evidence of tumorigenesis with the expression of these fusion genes. The most common chromosomal translocation in myxoid liposarcomas, t(12;16)(q13;p11), creates a *TLS/FUS-CHOP* fusion gene. Transgenic mice expressing the altered form of FUS-CHOP created by an in-frame fusion of the FUS domain to the carboxy end of CHOP develop liposarcomas. No tumors of other tissues were found in these transgenic mice despite widespread activity of the transgene. The results provided evidence that the FUS domain of FUS-CHOP plays a specific and critical role in the pathogenesis of liposarcoma.[14] Alveolar rhabdomyosarcoma (ARMS) is consistently associated with the characteristic translocation t(2;13)(q35;q14) or t(1;13)(p36;q14), which encode the PAX3-FKHR or PAX7-FKHR fusion oncoproteins, respectively. PAX3-FKHR fusion protein contributes to oncogenesis through abnormal control of growth, apoptosis, differentiation, or motility.[15]

Available Assays and Interpretation

A significant number of sarcomas have consistent chromosomal abnormalities that are detectable by standard cytogenetics or molecular genetic approaches (Figure 26-1).[2,3,16]

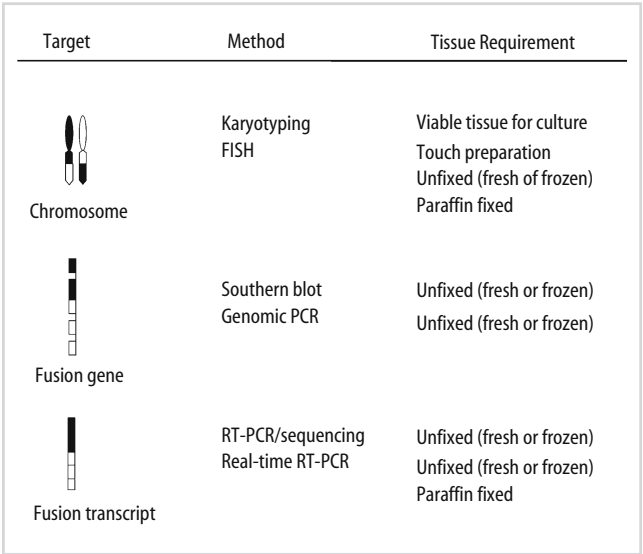


Figure 26-1. Molecular and cytogenetic methods for sarcoma diagnosis.

Sarcomas in this group should be defined by their specific molecular and cytogenetic alterations, although it will continue to be important to determine the sensitivity and specificity of these translocations for specific sarcoma types and the relative roles of molecular and histological classification of each sarcoma.

Karyotyping

Karyotyping is the classic cytogenetic approach for identifying chromosomal alterations including translocations in sarcomas.[17] Optimally, the procedure requires a substantial volume of viable, sterile tumor tissue, usually 1 cm^3 to 2 cm^3. The specimen should be harvested, placed in culture medium as soon as possible, and transported to the laboratory. Specimens also can be sent over a long distance at either room temperature or refrigerated for up to 48 hours. Small samples (limited incisional or needle biopsies) can be successfully cultured and karyotyped, although they may require a longer incubation time (1 to 2 weeks) to obtain enough dividing tumor cells for analysis. Characteristic and diagnostic chromosomal alterations seen in human soft tissue tumors are listed in Table 26-1.

An obvious drawback of conventional cytogenetic analysis is the requirement of adequate tumor cell growth to obtain metaphase spreads. If tumor cells do not grow, a false-negative result of a normal karyotype may occur due to the presence of normal fibroblasts in the specimen. Another pitfall of karyotyping is its limited resolution for identification of cryptic alterations, which may occur in sarcomas. Furthermore, translocations involving specific chromosomal regions may not necessarily represent the characteristic gene fusion for a specific sarcoma. Confirmation by fluorescence in situ hybridization (FISH) or molecular testing, therefore, is recommended for questionable cases or to reach a definitive diagnosis.

Fluorescence In Situ Hybridization

FISH provides a powerful diagnostic modality to demonstrate a specific gene fusion or chromosomal alteration. Metaphase chromosomal preparation from the tumor can be used to demonstrate translocations by chromosomal painting using chromosome-specific probes or gene-specific probes. However, this is not always possible since cell culture of sarcoma tissue to obtain metaphase chromosomes may be unsuccessful. In these cases, interphase FISH provides an excellent alternative using touch preparations from fresh or frozen tumor specimens without the requirement for tissue culture. One major advantage of interphase FISH over traditional karyotyping is the ability to detect cryptic gene rearrangements. In fact, the chromosomes of interphase nuclei are much more extended than metaphase or prometaphase chromosomes. As a result, FISH analysis performed on interphase nuclei permits higher resolution and can help to determine the physical mapping order of large DNA probes. Dual-colored DNA probes from the two rearranged genes involved in the translocation are most commonly used for the diagnosis of sarcomas with a specific translocation (Figure 26-2). Chromosomal centromeric probes and DNA probes spanning the translocation breakpoint also can be used. The interpretation of the FISH result depends on the types of hybridization probes.

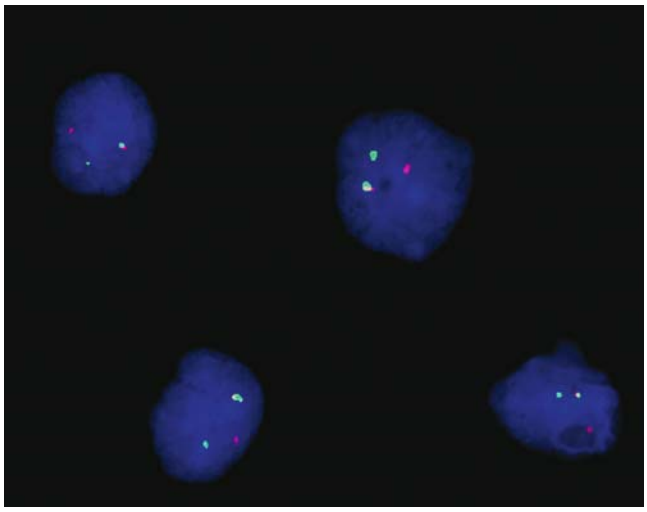


Figure 26-2. Dual-color interphase FISH is used to detect gene fusions in soft tissue sarcomas. Breakpoint flanking probes (one red and one green) are hybridized to nuclei preparation from paraffin-embedded tumor samples. The juxtaposed red and green signals, which can appear yellow in some nuclei, represent the chromosomal translocation (present in all four nuclei shown).

The possibility of performing interphase FISH on fixed paraffin-embedded tissue, either on histology sections or on nuclei preparations from paraffin blocks, has recently increased its potential diagnostic applications.[18,19] For interphase FISH of paraffin-embedded specimens, analysis of a nuclei preparation offers several advantages over the analysis of histological sections. The nuclei are more individualized on the slide, offering easy access to probe hybridization and interpretation. The complete genome of each nucleus is available for FISH, because there is no trimming effect as seen with histologic sections, in which some nuclei are only partially present on a slide due to the sectioning of the specimen (4–5 µm thick). On the other hand, direct examiniation of histological sections allows for correlation of FISH results with the architectural features of the tumor.

Southern Blot

Southern blot analysis detects sarcoma-specific translocations using labeled DNA probes specific to the fusion genes. The procedure is highly specific and particularly useful in detecting translocations with frequently variable translocation breakpoints, as for Ewing's sarcoma (ES). The main limitations of the Southern blot method are that it is labor intensive and requires fresh or frozen tissue

to obtain high-molecular-weight genomic DNA. False-negative results may arise when DNA fragments surpass the upper limits of DNA length (>15–20 kilobases [kb]), or with fusion genes encompassing large intron(s). Increasing the number of restriction enzymes usually resolves this problem. False-positive results may be seen due to incomplete restriction enzyme digestion, contamination by a cloning vector or bacteria, structural polymorphism of the genes involved in the translocation, and restriction recognition site variation.

Genomic Polymerase Chain Reaction

Polymerase chain reaction (PCR) amplification of the region encompassing the breakpoint of a fusion gene is often problematic. First, the breakpoint positions within the introns of the two rearranged genes are often variable, requiring selection of PCR primers within the adjacent exon and resulting in a variable size of the PCR product from case to case. Also, the introns can be very large, resulting in very large PCR products that may not be completely amplified and reducing the sensitivity of the test. Therefore, genomic PCR of sarcoma-specific translocations is not used for clinical testing.

Reverse Transcription–PCR

Chromosomal translocations in sarcomas give rise to aberrant fusion transcripts that are highly specific to a given tumor type (Table 26-1). More important, the structures of these fusion transcripts are highly consistent. Although the translocation breakpoint may involve various nucleotide positions of an intron at the DNA level, the resulting fusion transcript structure is the same due to RNA splicing. In a given sarcoma type, the gene fusion point at the messenger RNA (mRNA) level is highly precise to the single ribonucleotide. It is the tumor specificity and structural consistency of the fusion transcripts that make reverse transcription–PCR (RT-PCR) the preferred method for molecular detection of specific fusion transcripts.[20]

After reverse transcription of tumor mRNA into complementary DNA (cDNA), PCR primers complementary to the exons that flank the translocation breakpoint are used to amplify fusion transcript-specific RT-PCR products. Variations of RT-PCR methods have been used for the detection of sarcoma-specific fusion transcripts. Since the fusion joining point may be far away from the poly(A) tail of the aberrant transcript, random hexamers or gene-specific downstream primers should be used instead of oligo d(T) for the reverse transcription of RNA into cDNA. Instead of the traditional two-step RT-PCR method, reverse transcription and PCR can be performed in a single reaction, a procedure called one-step RT-PCR. The reverse transcriptase works first at low temperature to convert mRNA into cDNA, while the Taq DNA polymerase is inactive. The temperature then is raised to inactivate the reverse transcriptase, to activate the Taq DNA polymerase, and to initiate the amplification reaction at the same time. Reactions with multiple primer sets are difficult to optimize but are very convenient for molecular testing of sarcomas with numerous fusion gene variants such as ES/peripheral neuroectodermal tumor (PNET).[21] Nested PCR using an additional pair of primers to further amplify the first-round RT-PCR product can greatly increase the sensitivity, although it also increases the likelihood of cross contamination.

Identification of a positive RT-PCR result usually relies on detection of the expected-sized product on a common DNA-separating gel. When fresh or frozen tissue is the starting material, the target size of the RT-PCR product is designed carefully to be less than 500 base pairs (bp) and preferably around 200 bp. For fusion genes that have little size variations, such as *SYT-SSX* fusion of synovial sarcoma, the presence of a single RT-PCR product of the expected size generally is considered adequate for positive identification, although post-PCR confirmatory analysis of the product sequence can be performed. For fusion genes with molecular size variability, such as *TLS-CHOP* in myxoid liposarcoma, additional analysis of the PCR products is mandatory to confirm the specificity of the PCR products. Confirmatory methods include blot transfer with subsequent hybridization with a fusion gene-specific probe, DNA sequencing of the PCR product, restriction endonuclease digestion analysis, or an additional nested PCR step. Inclusion of positive and negative control reactions is necessary because the high sensitivity of the amplification process and cross-contaminations are major problems with RT-PCR detection of fusion transcripts for clinical diagnosis.

Paraffin-embedded blocks of soft tissue tumors may be used as the source of RNA,[22] although with some caution.[10] As a result of the fixation and embedding processes, the RNA from tissue blocks is substantially degraded. RNA degradation requires that the size of the amplification product be designed to be considerably smaller (around 100 bp) compared to assays designed for use of fresh or frozen tissue. RT-PCR from tissue blocks usually requires more amplification cycles, and a nested approach may be necessary for the detection of the aberrant transcript. These assay-altered conditions account for an increased risk of false-positive results due to carryover contamination. Very stringent conditions and additional control reactions are necessary to ensure test specificity.

Consideration must be given to the variability of fusion gene transcripts for each type of sarcoma (see Table 26-1) when interpreting RT-PCR test results. For instance, in ES/PNET, there are considerable variations among the different exons involved in the gene fusion.[21] Although the tumors may be cytogenetically indistinguishable from one another, they differ in the exon composition of the final fusion transcript. Variant chromosomal translocations involving different chromosomal partners also give rise to

variability in approximately 5% of ES/PNET cases. In fact, the t(21;22), instead of the far more common t(11;22), results in the fusion of *EWS* with *ERG*.[23] Very rarely, *EWS* is fused with other loci including *ETV1* or *E1AF* in ES/PNET.[21] Awareness of these variants is very important when a negative RT-PCR for *EWS-FLI1* occurs.

Real-Time RT-PCR

Real-time RT-PCR is a recent advancement in analysis of mRNA.[24] Although designed for quantitative measurement, real-time RT-PCR has become a reliable detection method for fusion gene transcripts.[25,26] It has three major advantages over the conventional endpoint RT-PCR. First, real-time RT-PCR greatly enhances the specificity of detection by incorporating a specific probe, complementary to the region of the gene internal to the PCR primers. Second, real-time PCR measures the geometric (exponential) phase of a PCR reaction (cycles 20–25), offering an accurate quantification of the fusion transcript. Third, the PCR product accumulation is measured in real time by fluorescence detection through the semitransparent plastic cap of the reaction tube and, therefore, post-PCR manipulation and possible PCR contamination are reduced. Since its introduction for molecular analysis, real-time RT-PCR has shown conclusively high specificity and sensitivity, which are important when analyzing very small samples or target genes with very low expression levels.[25] Simultaneous amplification of multiple targets by multiplex real-time RT-PCR is possible by using target-specific probes labeled with different fluorescent labels.[27]

With careful optimization, real-time RT-PCR can be highly reliable for detection of fusion transcripts using all types of tissue sources.[26] The early geometric phase detection with enhanced specificity and elimination of post-PCR manipulation makes real-time RT-PCR ideal for detection of sarcoma fusion transcripts, even if the RNA source is paraffin-embedded tissue.[28] The specific and quantitative detection power of real-time RT-PCR using paraffin-embedded tissue is illustrated in Figure 26-3, in which amplification of the *SYT-SSX* fusion transcript of synovial sarcoma is shown. Positive identification of fusion transcripts can be achieved using as little as 20 ng of RNA extracted from one paraffin block, making small specimens, such as core needle biopsies, suitable for definitive molecular diagnosis. Overall, real-time RT-PCR offers significant advantages over conventional RT-PCR. It is technically easier, more rapid, and more sensitive. Because a specific probe is included in the reaction, secondary confirmation is unnecessary, and the problem of carry-over cross-contaminations is minimized. The operating-cost profile and time to reporting also are in favor of real-time RT-PCR over conventional RT-PCR.[25] The routine clinical applications of real-time RT-PCR for sarcoma diagnosis and minimal disease detection are highly promising.

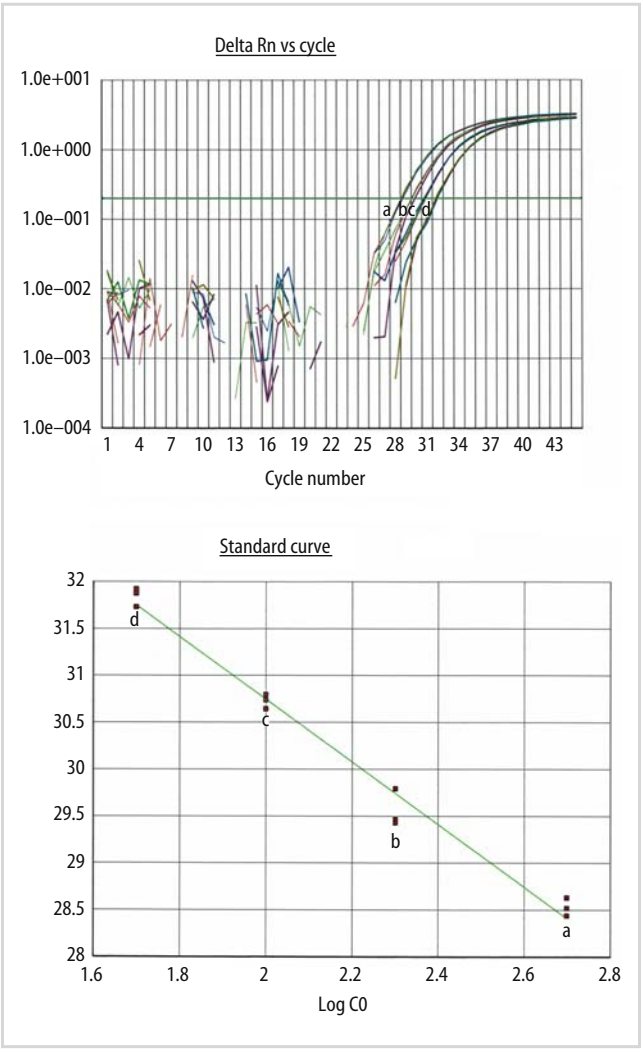


Figure 26-3. Real-time RT-PCR detection of gene fusion in a sarcoma. Top panel: Real-time RT-PCR plot of PCR cycle number versus fluorescence signal for *SYT-SSX* fusion transcript amplification in a synovial sarcoma with (a) 50 ng, (b) 100 ng, (c) 200 ng, and (d) 500 ng of total input RNA, each amplified in triplicate. The RNA was extracted from formalin-fixed, paraffin-embedded tissue. Bottom panel: Log gradient plot of the data from the top panel, demonstrating linear quantitative amplification consistent with the amount of input RNA template. The y-axis represents the threshold cycle number, and the x-axis is the log of the RNA concentration. The quantity of the specific RNA measured by real-time RT-PCR is inversely proportional to the threshold cycle number.

Clinical Utility of Testing

The molecular characterization of human sarcomas is resulting in rapid progress in sarcoma classification, diagnosis, disease monitoring, prediction of outcome, and design of novel therapeutic strategies.

Establishing Definitive Diagnosis and Redefining Tumor Entity

Soft tissue tumors comprise a vast and heterogeneous group of neoplasms. Because different tumors often have different biological behaviors and respond differently to

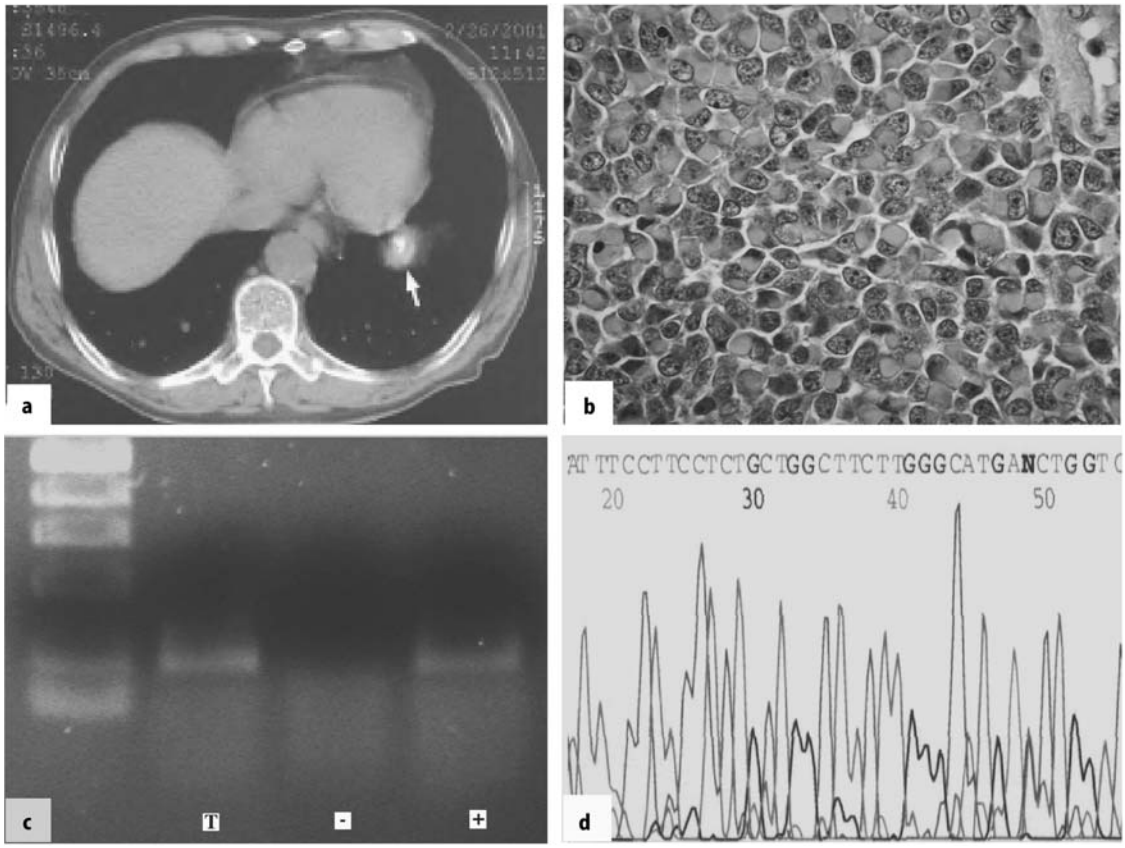


Figure 26-4. RT-PCR detection of *SYT-SSX2* fusion transcript in a primary poorly differentiated synovial sarcoma of the lung with rhabdoid features. (Courtesy of Dr. Thomas Ciesielski.) (a) CT scan of a 72-year-old with a subpleural 2 cm nodule. An extensive physical and radiographic examination of the patient failed to reveal a separate primary lesion. (b) Proliferating tumor cells with rhabdoid cytoplasm. This histological variant of synovial sarcoma may be difficult to diagnose without molecular testing. (c) RT-PCR of formalin-fixed tumor tissue shows the presence of a 97 bp RT-PCR product from the *SYT-SSX2* fusion transcript in lane 2 (T, tumor). Lanes 3 and 4 are the negative (–) and positive (+) controls, respectively. (d) Direct DNA sequencing of the RT-PCR product from panel c shows a *SYT-SSX2*-type fusion transcript.

therapeutic strategies, precise classification is of clinical importance. The majority of soft tissue tumors were first delineated on the basis of morphologic and clinical findings into homogeneous groups. However, it has increasingly been appreciated that many tumors with similar histological and pathological characteristics actually are heterogeneous groups of tumors that differ in their clinical behavior and underlying pathogenesis. Correlation of cytogenetic and molecular data with pathologic findings is essential for understanding both the biological significance and the clinical value of specific molecular changes.

A surprising correlation has emerged among chromosomal alterations, gene rearrangements, and distinct histopathologic entities. Morphologic and cytogenetic/molecular observations in fact have been validating each other in the classification of soft tissue tumors, and a combined approach has resulted in a more rational classification of soft tissue tumors. This approach has been particularly helpful in understanding the so-called small round blue cell tumors. It is now accepted that ES and PNET are essentially a single tumor type defined by the characteristic translocation between chromosomes 11 and 22, resulting in an *EWS* gene rearrangement. Despite the morphologic similarities, typical olfactory neuroblastoma lacks *EWS* rearrangement that argues against the previously proposed inclusion of this tumor in the ES/PNET group.[29] The finding of *PAX-FKHR* chimeric products is of considerable value in defining ARMS. In fact, the small proportion of primitive embryonal-like RMS containing a *PAX-FKHR* gene fusion likely represents unrecognized solid ARMS.[30] Molecular diagnosis also allows the certain identification of previously unrecognized variants of known tumors. One such example, illustrated in Figure 26-4, is a poorly differentiated synovial sarcoma with rhabdoid features.

Another area that has significantly benefited by this combined approach is that of adipose tissue neoplasms. The finding of ring or giant marker chromosomes cytogenetically defines the group of atypical lipomatous tumors (well-differentiated liposarcoma) and justifies the distinction of atypical lipomatous tumors from spindle cell and pleomorphic lipomas.[11] The finding of *TLS-CHOP* fusion transcripts in both myxoid liposarcoma and round cell liposarcoma demonstrates that they represent a continuum, an observation nicely correlating with the not infrequent observation of cases with mixed histology.[31] The t(17;22), which produces the *COL1A1-PDGFB* fusion, is present in both dermatofibrosarcoma protuberans (DFSP) and giant cell fibroblastoma, indicating that these are adult and pediatric presentations of a single tumor entity.[5,32]

Providing Prognostic Information

Synovial sarcomas are histologically classified into biphasic and monophasic types according to the presence or absence of glandular epithelial differentiation. Synovial sarcomas are characterized by a t(X;18)(p11;q11). The *SYT* gene at 18q11 is fused with either *SSX1* or *SSX2* at Xp11. Although there is only a 13-amino-acid difference between the two types of fusion genes, the *SYT-SSX* fusion type correlates with overall survival and histologic features of the tumor.[33,34] In patients with localized disease at presentation, *SYT-SSX* fusion type appears to be the only significant prognostic factor in a multivariate analysis. Of 202 patients with localized synovial sarcoma at diagnosis, the median and 5-year survival for the *SYT-SSX2* group were 13.7 years and 77%, respectively, compared to 9.2 years and 61%, respectively, for the *SYT-SSX1* group.[34]

Studies in ES and ARMS have also established that alternative forms of the specific aberrant fusion transcripts may have an important impact on one or more clinical features or prognosis. The *PAX3-FKHR* fusion is an adverse prognostic factor in ARMS.[35] Therefore, not only is *PAX-FKHR* fusion transcript the defining feature of ARMS, but also the specific type of gene fusion present in the tumor identifies, among patients with metastatic disease, a high-risk subgroup (with *PAX3-FKHR*) and a favorable-outcome subgroup (with *PAX7-FKHR*), as shown by a recent study of the Children's Oncology Group.[36]

Minimal Disease Detection

Molecular detection of occult tumor cells can be a prognostic indicator important for clinical staging and selection of therapeutic strategies. Detection of low-level expression of *EWS-FLI* fusion transcripts in bone marrow and blood has been reported in sarcomas.[37,38] Micrometastases were detected in 31% of patients with localized disease and 50% of patients with clinically metastatic tumors. RT-PCR detection of specific fusion transcripts has been used to study the marrow and body fluids of patients with ES, ARMS, and desmoplastic small round cell tumors (DSRCT). Molecular evidence of metastasis was detected in all patients with clinically evident metastatic tumors but not in the patients with localized disease.[39] Such applications are just beginning to be explored. Although correlative studies are needed to validate the prognostic significance of occult tumor cells, it is clear that molecular staging is an emerging field that may change the approach to the clinical management of patients with soft tissue sarcomas.

Developing Tumor-Specific Therapy

The therapeutic implications of sarcoma-specific gene fusion events are tremendous, since each fusion is a potential target for tumor-specific therapies. A tumor-specific cytotoxic T-cell response can be induced by peptides derived from the fusion point of the SYT-SSX chimeric protein in synovial sarcoma that may function in vitro as a neoantigen.[40] Other forms of tumor-specific immunotherapies are under development. STI571 (Gleevec) is a potent, relatively selective inhibitor for the BCR-ABL tyrosine kinase of chronic myelogenous leukemia. STI571 is being successfully used to treat gastrointestinal stromal tumors with KIT mutational activation,[41] and it has been shown to reduce the growth of *COL1A1-PDGFB*-transformed animal cells as well as primary cultures of DFSP tumor cells in vitro and in vivo.[6,42] These successes have opened the door to the design of tumor-specific drugs targeting the fusion proteins present in soft tissue sarcomas. Furthermore, specific fusion transcripts also may be targeted through the RNA interference mechanism[43] or by other gene therapy approaches.[44,45]

Quality Control and Laboratory Issues

In the clinical laboratory, the availability of two or more unrelated test methods for the diagnosis of each sarcoma type can facilitate the confirmation of unexpected results or the resolution of discrepant results. One key issue regarding the accurate molecular diagnosis of sarcomas is the selection of the appropriate procedure. This is largely determined by the type of sarcoma under consideration, the tissue specimen type, and the specimen quality. For fusion genes with limited structural variations, such as synovial sarcoma, ARMS, clear cell sarcoma of soft parts, extraskeletal myxoid chondrosarcoma, DSRCT, alveolar soft part sarcoma, and inflammatory myofibroblastic tumor, RT-PCR and real-time RT-PCR represent the first choice method. Interphase FISH is the best choice for sarcomas having fusion genes with chromosomal and intramolecular variations, such as myxoid/round cell liposarcoma, DFSP, and ES/PNET. Ideally, all specimens with a working diagnosis of sarcoma should have viable tissue sent for karyotyping. A portion of the diagnostic specimen can be temporarily stored in culture medium at room temperature or 4°C until a provisional histologic diagnosis is made, and then sent for karyotyping if appropriate. Whenever possible, a portion of the tumor specimen should be frozen and stored, to provide the best source of nucleic acid suitable for virtually all types of molecular tests (Southern blot, RT-PCR, real-time RT-PCR, and genomic PCR) used for molecular testing. Unfixed, fresh or frozen tissue also is suitable for protein analysis, although this is rarely used clinically. Touch preparations for FISH should be made and stored for possible use. When only paraffin-embedded archival material is available, interphase FISH and real-time RT-PCR are the best testing choices.

Turnaround time varies significantly among the different testing methods. Cytogenetic karyotyping requires

short-term culture (3 to 10 days). The length of time for culturing reflects the quality as well as the quantity of the specimen source, as well as the tumor proliferation rate in vitro. Turnaround time for interphase FISH also depends on specimen type. Touch preparation and frozen sections allow for rapid FISH analysis (1 to 3 days). For archival tissue sources, an additional 2 or 3 days are required to obtain the nuclear suspension or the tissue sections. RT-PCR has a 2- or 3-day turnaround time for common fusion gene detection but requires a longer time for more complex fusion variant detection. Real-time RT-PCR offers a shorter turnaround time (1 or 2 days) because it does not require post-PCR manipulation steps or confirmatory testing. Longer turn-around times may be expected if the laboratory batches specimens and only runs the test once per week, for example.

Conclusion

Cytogenetics and molecular cytogenetics have contributed significantly to the study of the pathobiology of soft tissue tumors. Identification of recurrent chromosomal changes has resulted in the definition of an increasing number of tumor-related molecular changes and is fostering a continual refinement of the classification of these lesions. The recognition of distinctive rearrangements in tumor subsets is providing powerful tools to complement the histological diagnosis in a field in which pathological diagnosis and clinical management are often difficult, due to both the variety of tumor phenotypes and their relative rarity. In this respect, novel technical approaches such as real-time RT-PCR, and multiplex RT-PCR reactions for the detection of fusion products specific for different tumor types offer practical and cost-effective methods for complementing histological diagnosis.

Acknowledgment

The authors thank Chaline Johnson for her outstanding technical assistance.

References

1. Kleihues P, Schauble B, zur Hausen A, et al. Tumors associated with p53 germline mutations: a synopsis of 91 families. *Am J Pathol.* 1997;150:1–13.
2. Ladanyi M, Bridge JA. Contribution of molecular genetic data to the classification of sarcomas. *Hum Pathol.* 2000;31:532–538.
3. Bennicelli JL, Barr FG. Chromosomal translocations and sarcomas. *Curr Opin Oncol.* 2002;14:412–419.
4. Coffin CM, Patel A, Perkins S, et al. ALK1 and p80 expression and chromosomal rearrangements involving 2p23 in inflammatory myofibroblastic tumor. *Mod Pathol.* 2001;14:569–576.
5. Simon MP, Pedeutour F, Sirbent N, et al. Deregulation of the platelet-derived growth factor beta-chain via fusion with collagen gene COL1A1 in dermatofibrosarcoma protuberans and giant cell fibroblastoma. *Nat Genet.* 1997;15:95–98.
6. Greco A, Roccato E, Miranda C, et al. Growth-inhibitory effect of STI571 on cells transformed by the COL1A1-PDGFB rearrangement. *Int J Cancer.* 2001;92:354–360.
7. Ladanyi M, Lui MY, Antonescu CR, et al. The der(17)t(X;17)(p11;q25) of human alveolar soft part sarcoma fused the TFE3 transcription factor gene to ASPL, a novel gene at 17q25. *Oncogene.* 2001;20:48–57.
8. Koontz JI, Soreng AL, Nucci M, et al. Frequent fusion of the JAZF1 and JJAZ1 genes in endometrial stromal tumors. *Proc Natl Acad Sci U S A.* 2001;98:6348–6352.
9. Chung DC, Rustgi AK. The hereditary nonpolyposis colorectal cancer syndrome: genetics and clinical implications. *Ann Intern Med.* 2003;138:560–570.
10. Rosai J, Akerman M, Dal Cin P, et al. Combined morphologic and karyotypic study of 59 atypical lipomatous tumors: evaluation of their relationship and differential diagnosis with other adipose tissue tumors (a report of the CHAMP study group). *Am J Surg Pathol.* 1996;20:1182–1189.
11. Fletcher CDM, Fletcher JA, Cin PD, et al. Diagnostic gold standard for soft tissue tumors: morphology or molecular genetics? *Histopathology.* 2001;39:100–103.
12. Ladanyi M. Diagnosis and classification of small round-cell tumors of childhood. *Am J Pathol.* 1999;155:2181–2182.
13. Rossow KL, Janknecht R. The Ewing's sarcoma gene product functions as a transcriptional activator. *Cancer Res.* 2001;61:2690–2695.
14. Perez-Losada J, Sanchez-Martin M, Rodriguez-Garcia MA, et al. Liposarcoma initiated by FUS/TLS-CHOP: the FUS/TLS domain plays a critical role in the pathogenesis of liposarcoma. *Oncogene.* 2000;19:6015–6022.
15. Barr FG. Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma. *Oncogene.* 2001;20:5736–5746.
16. Barr FG, Chatten J, D'Cruz CL, et al. Molecular assays for chromosomal translocations in the diagnosis of pediatric soft tissue sarcomas. *JAMA.* 1995;273:553–557.
17. Bridge JA, Sandberg AA. Cytogenetic and molecular genetic techniques as adjunctive approaches in the diagnosis of bone and soft tissue tumors. *Skeletal Radiol.* 2000;29:249–258.
18. Adem C, Gisselsson D, Dal Cin P, et al. EVT6 rearrangements in patients with infantile fibrosarcomas and congenital mesoblastic nephromas by fluorescence in situ hybridization. *Mod Pathol.* 2001;14:1246–1251.
19. Hui P, Howe G, Crouch J, et al. Real-time RT-PCR of cyclin D1 mRNA in mantle cell lymphoma: comparison with FISH and immunohistochemistry. *Leuk Lymphoma.* 2003;44:1385–1394.
20. Athale UH, Shurtleff SA, Jenkins JJ, et al. Use of reverse transcriptase polymerase chain reaction for diagnosis and staging of alveolar rhabdomyosarcoma, Ewing sarcoma family of tumors, and desmoplastic small round cell tumor; *Am J Pediatr Hematol Oncol.* 2001;23:99–104.
21. Dagher R, Pham RA, Sorbara L, et al. Molecular confirmation of Ewing's sarcoma. *Am J Pediatr Hematol Oncol.* 2001;23:221–224.
22. Guillou L, Coindre J, Gallaghr G, et al. Detection of the synovial sarcoma translocation t(X;18)(SYT;SSX) in paraffin-embedded tissues using reverse transcriptase-polymerase chain reaction: a reliable and powerful diagnostic tool for pathologists: a molecular analysis of 221 mesenchymal tumors fixed in different fixatives. *Hum Pathol.* 2001;32:105–112.
23. Sorensen PH, Lessnick SL, Lopez-Terrada D, et al. A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat Genet.* 1994;6:146–151.
24. Giulietti A, Overbergh L, Valckx D, et al. An overview of real-time quantitative PCR: applications to quantify cytokine gene expression. *Methods.* 2001;25:386–401.
25. Cummings TJ, Brown NM, Stenzel TT. TaqMan junction probes and the reverse transcriptase polymerase chain reaction: detection of alveolar rhabdomyosarcoma, synovial sarcoma, and desmoplastic small round cell tumor. *Ann Clin Lab Sci.* 2002;32:219–224.
26. Hill DA, Riedley SE, Patel AR, Shurtleff SA, et al. Real-time polymerase chain reaction as an aid for the detection of SYT-SSX1

and SYT-SSX2 transcripts in fresh and archival pediatric synovial sarcoma specimens: report of 25 cases from St. Jude Children's Research Hospital. *Pediatr Dev Pathol.* 2003;6:24–34.
27. Peter M, Gilbert E, Delattre O. A multiplex real-time PCR assay for the detection of gene fusions observed in solid tumors. *Lab Invest.* 2001;81:905–912.
28. Hoestein I, Menard A, Bui BN, et al. Molecular detection of the synovial sarcoma translocation t(X;18) by real-time polymerase chain reaction in paraffin-embedded material. *Diagn Mol Pathol.* 2002;11:16–21.
29. Kumar S, Perlman E, Pack S, et al. Absence of EWS/FLI1 fusion in olfactory neuroblastomas indicates these tumors do not belong to the Ewing's sarcoma family. *Hum Pathol.* 1999;30:1356–1360.
30. Parham DM, Shapiro DN, Downing JR, et al. Solid alveolar rhabdomyosarcomas with the t(2;13). Report of two cases with diagnostic implications. *Am J Surg Pathol.* 1994;18:474–478.
31. Knight JC, Renwick PJ, Cin PD, et al. Translocation t(12;16)(q13;p11) in myxoid liposarcoma and round cell liposarcoma: molecular and cytogenetic analysis. *Cancer Res.* 1995;55:24–27.
32. Sandberg AA, Bridge JA. Updates on the cytogenetics and molecular genetics of bone and soft tissue tumors: dermatofibrosarcoma protuberans and giant cell fibroblastoma. *Cancer Genet Cytogenet.* 2003;140:1–12.
33. Kawai A, Woodruff J, Healey JH, et al. SYT-SSX gene fusion as a determinant of morphology and prognosis in synovial sarcoma. *N Engl J Med.* 1998;338:153–160.
34. Ladanyi M, Antonescu CR, Leung DH, et al. Impact of SYT-SSX fusion type on the clinical behavior of synovial sarcoma: a multi-institutional retrospective study of 243 patients. *Cancer Res.* 2002;62:135–140.
35. Anderson J, Gordon T, McManus A, et al. Detection of the PAX3-FKHR fusion gene in paediatric rhabdomyosarcoma: a reproducible predictor of outcome? *Br J Cancer.* 2001;85:831–835.
36. Sorensen PH, Lynch JC, Qualman SJ, et al. PAX3-FKHR and PAX7-FKHR gene fusions are prognostic indicators in alveolar rhabdomyosarcoma: a report from the Children's Oncology Group. *J Clin Oncol.* 2002;20:2672–2679.
37. Fidelia-Lambert MN, Zhuang Z, Tsokos M. Sensitive detection of rare Ewing's sarcoma cells in peripheral blood by reverse transcriptase polymerase chain reaction. *Hum Pathol.* 1999;30:78–80.
38. Sumerauer D. Vicha A, Kucerova H, et al. Detection of minimal bone marrow infiltration in patients with localized and metastatic Ewing sarcoma using RT-PCR. *Folia Biol.* 2001;47:206–210.
39. Schleiermacher G, Delattre O. Detection of micrometastases and circulating tumor cells using molecular biology techniques in solid tumors. *Bull Cancer.* 2001;88:561–570.
40. Worley BS, Van den Broekes LT, Goletz TJ, et al. Antigenicity of fusion proteins from sarcoma-associated chromosomal translocations. *Cancer Res.* 2001;61:6868–6875.
41. Van Oosterom AT, Judson I, Verweij J, et al. Safety and efficacy of imatinib (STI571) in metastatic gastrointestinal stromal tumours. A phase I study. *Lancet.* 2001;358:1421–1423.
42. Sjoblom T, Shimizu A, O'Brien KP, et al. Growth inhibition of dermatofibrosarcoma protuberant tumors by the platelet-derived growth factor antagonist STI571 though induction of apoptosis. *Cancer Res.* 2001;61:5778–5783.
43. Simeoni F, Morris MC, Heitz F, et al. Into the mechanism of the peptide-based gene delivery system MPG: implications for delivery of siRNA into mammalian cells. *Nucleic Acids Res.* 2003;31:2717–2724.
44. Tanaka K, Iwakuma T, Harimaya K, et al. EWS-FLI-1 antisense oligodeoxynucleotide inhibits proliferation of human Ewing's sarcoma and primitive neuroectodermal tumor cells. *J Clin Invest.* 1997;99:239–247.
45. Uchida A, Seto M, Hashimoto N, et al. Molecular diagnosis and gene therapy in musculoskeletal tumors. *J Orthop Sci.* 2000;5:418–423.

Chapter 27

Gliomas

Robert Jenkins

Introduction

More than 18,000 people were diagnosed with and more than 13,000 died from primary brain tumors in 2003.[1] The most common of the primary brain tumors is gliomas. Depending on the grade and morphologic type of glioma, newly diagnosed patients receive watchful waiting, surgical resection, radiotherapy, or chemotherapy, or some combination of these therapies. Regardless of therapy, most patients will progress and have a high risk of mortality and reduced quality of life. Thus, there has been intense interest in understanding the biology and genetics of gliomas, to provide better diagnostic tools and new therapeutic approaches. Molecular pathology markers are being identified that have been or will soon prove to be clinically useful in the practice of clinical neurooncology (see Table 27-1).

Pathology Considerations

Gliomas are classified into tumors of astrocytic, oligodendroglial, ependymal, and mixed lineage. Diffuse astrocytic tumors are divided into three different grades by the World Health Organization (WHO) system:[2] grade 2 (or astrocytomas [A]), grade 3 (or anaplastic astrocytomas [AA]), and grade 4 (or glioblastoma multiforme [GBM]). These tumors occur predominantly in adults. Diffuse astrocytic tumors can be subclassified into several histological types, including fibrillary, gemistocytic, small cell, and giant cell subtypes. Several of these histological types can be found within a single diffuse astrocytic tumor. Pilocytic astrocytomas are nondiffuse and localized. These tumors are often grade 1 and predominantly occur in children and young adults. GBM can be subclassified into two groups based on the duration of symptoms.[3] Primary GBM presents de novo with a short duration of symptoms (often less than 3 months). Secondary GBM presents with a longer duration of symptoms or with a preceding grade 2 or 3 glioma. Primary and secondary GBM usually present in patients greater than 60 and less than 40 years of age, respectively.

Oligodendrogliomas are divided into two grades: grade 2 (oligodendroglioma) and grade 3 (anaplastic oligodendroglioma). Some neuropathologists have described a grade 4 oligodendroglioma.[4] Ependymomas are also classified as low and high grade.

One of the problems with the neuropathologic diagnosis of gliomas is their histologic heterogeneity. AA exhibits increased cellularity, nuclear atypia, and mitotic activity. Endothelial proliferation and necrosis are additional features of GBM. Depending on the extent of surgical sampling, some of the histologic elements that distinguish AA and GBM may not be present in the specimen submitted for pathological diagnosis, and some gliomas may be undergraded. Thus, molecular markers that assist in tumor grading are clinically useful.

In addition to histologic heterogeneity, diffuse gliomas often exhibit cell type heterogeneity. Both astrocytic and oligodendroglial elements are present in most diffuse gliomas, especially those of grades 2 and 3. When both of these elements comprise a significant proportion of the tumor, the glioma may be classified as a mixed oligoastrocytoma (MOA). However, the assessment of the appropriate proportions is subjective, with a continuum from "pure" astrocytomas, through mixed astrocytomas, to "pure" oligodendrogliomas. In a comparison of corresponding grade tumors, patients with oligodendrogliomas have a better survival than patients with astrocytic tumors (the survival of patients with MOA is intermediate). Thus, molecular tools that assist with morphological classification are clinically useful.

Molecular Pathology Tools for Glioma Grading

Gliomas have been extensively characterized by cytogenetic, molecular cytogenetic, and molecular genetic methods (reviewed in References 3, 5, and 6). Figure 27-1

Table 27-1. Molecular Markers of Gliomas

Marker	Indication
Markers Currently Entering Clinical Practice	
1p/19q deletion testing	Oligodendroglioma diagnosis and prognosis Prediction of oligodendroglioma therapeutic response
Markers with Clinical Potential	
EGFR amplification (overexpression)	Small cell GBM diagnosis and differentiation from oligodendroglioma Prediction of high-grade glioma therapeutic response to novel EGFR pathway inhibitors
Chromosome 10 loss/ *PTEN* mutation	Prognosis of anaplastic astrocytomas Subclassification of mixed oligoastrocytomas
TP53 mutation	Prognosis of anaplastic astrocytomas
1p/19q deletion testing	Subclassification of mixed oligoastrocytomas

summarizes the major genetic alterations found in diffuse gliomas. Some of these alterations are highly correlated with the grade of gliomas, especially for astrocytic tumors.

Several genetic alterations have been associated with GBM, especially primary GBM. These alterations include loss of chromosome 10, mutation of the *PTEN* gene, and amplification of the *EGFR* gene with resulting over expression of the EGFR protein. While the clinical use of these markers has not been completely validated, there is growing evidence that they may prove useful as a molecular adjunct for glioma grading. The presence of *EGFR* amplification, chromosome 10 loss, or *PTEN* mutation, or some combination of the three, in a glioma indicates that the tumor is likely a GBM. Importantly, a small but significant proportion of tumors with the histologic features of AA have one or more of these alterations.[7–11] It is reasonable to hypothesize that since patients with GBM have a poor prognosis, then patients whose AA shares molecular characteristics with GBM also may have a poor survival.

A recent Mayo Clinic/North Central Cancer Treatment Group (NCCTG) study indicated that AA with chromosome 10 loss or *PTEN* mutation behaves like GBM.[8] The median survival of patients with AA and chromosome 10 loss or *PTEN* mutation was approximately 4 months, a survival worse than that of patients with GBM (who had a median survival of approximately 12 months). The median survival of patients whose AA lacked these alterations was 34 months. This difference in survival was statistically significant even after adjustment for patient age, performance score, and extent of resection, all of which are important prognostic variables for patients with gliomas. Importantly, the patient survival for GBM with and without chromosome 10 loss or *PTEN* mutations was very similar.[8] It is possible that sampling bias (in terms of the tissue available for histologic and molecular analysis) may have accounted for the differences. However, sampling bias is a difficult problem to overcome in routine clinical practice, and molecular testing for chromosome 10 loss or *PTEN* mutations in a sample of a tumor with the features of an AA indicates that the tumor may be a GBM.

Interestingly, the Mayo Clinic/NCCTG study did not show that AA with *EGFR* amplification behaved like GBM.[8] However, careful stratification indicated that AA and GBM with *EGFR* amplification in young patients (i.e., less than 40 years old) progress much faster than AA and GBM

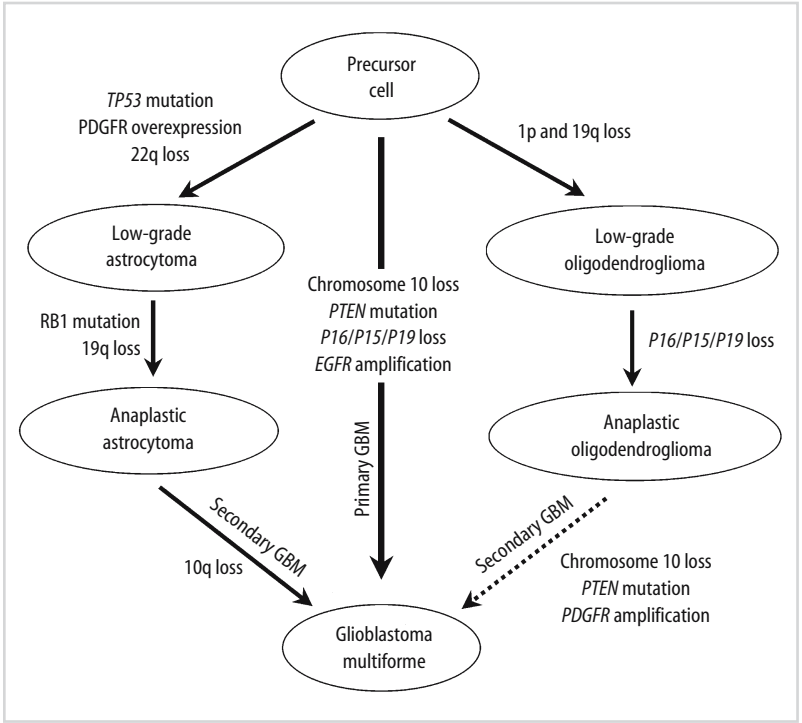


Figure 27-1. Possible genetic pathways of diffuse glioma formation and progression (modified from References 5 and 6). Histologic classification is based on current WHO guidelines.[2] Although the molecular genetic evidence is increasing, it is not widely accepted that anaplastic oligodendroglioma progresses to GBM (dashed arrow).

without *EGFR* amplification in this age group.[8] Conversely, AA and GBM with *EGFR* amplification in older patients (i.e., greater than 60 years old) had a better prognosis than those without this alteration. Importantly, this result has been confirmed by another group.[12] These differences in age-related prognosis may reflect the different biology of primary and secondary high-grade astrocytic tumors. *EGFR* amplification is accompanied by sequence alterations in the *EGFR* gene (reviewed in Reference 13). The most common alteration is deletion of exons 2 to 7 (the EGFRvIII variant), which generates an EGFR protein that is active independently of ligand binding. However, there are several other amplified and mutated EGFR proteins (for example, the C958 variant is truncated after amino acid 958), some of which remain ligand dependent. Current translational studies are investigating the prevalence of these different amplified and mutated EGFR variants in primary and secondary GBM, and in GBM from patients of different ages.[14]

TP53 mutations are found in about half of astrocytomas and AA. About one quarter of GBM have been observed to have *TP53* mutations, and these tumors often have the clinical presentation of a secondary GBM. Since *TP53* alterations usually are associated with low-grade tumors, it has been hypothesized that high-grade astrocytic tumors with *TP53* alterations may have a better prognosis. Except for the rare giant cell and gemistocytic variants (in which *TP53* mutations are especially prevalent[15,16]), *TP53* alterations do not seem to predict survival for GBM.[17–20] However, in the Mayo Clinic/NCCTG series, the median survival of patients with AA and *TP53* mutations was 5 years.[8] This survival was significantly longer than those without *TP53* mutations.[8]

Although *PTEN*, chromosome 10, *TP53*, and *EGFR* alterations have been found to be correlated with clinical grade and thus with patient prognosis, they have yet to be routinely used in clinical practice for two reasons. First, the studies require broader validation. Second, and more importantly, the presence or absence of such alterations does not currently change the clinical management of these patients. High-grade glioma trials designed to specifically evaluate the clinical usefulness of these markers are needed to move these molecular markers into clinical practice.

Molecular Pathology Markers of Morphologic Glioma Type

Oligodendrogliomas have a better prognosis, grade for grade, compared with MOA and astrocytic gliomas. In addition, Cairncross et al. demonstrated in 1994 that some anaplastic oligodendrogliomas respond to combination chemotherapy and radiation therapy.[21] Simultaneously, several investigators observed that approximately 70% of oligodendrogliomas have deletions of the chromosome 19 q arm, 70% have deletions of the chromosome 1p arm, and 60% have deletions of both arms (reviewed in Reference 5). Usually, the entire chromosomal arms are deleted.[22] In contrast, about 40%, 20%, and 10% of astrocytomas have deletion of 19q, 1p, and both 1p and 19q, respectively.[22] The proportion of MOA with 1p and 19q deletions is intermediate, suggesting that some of these tumors can be grouped with the oligodendrogliomas.[22,23] A significant proportion of MOA also have *P53* mutations and other genetic anomalies, suggesting that these tumors can be grouped with the astrocytomas.[22,23]

Two retrospective studies have provided evidence that the oligodendrogliomas with 1p deletions, 19q deletions, or both have a better prognosis and respond better to chemotherapy and radiation therapy.[24,25] While these results need to be confirmed by prospective trials, the data strongly suggest that 1p/19q deletions are associated with tumors of oligodendroglial lineage, with oligodendrogliomas that have a better prognosis, and with oligodendrogliomas that respond to chemotherapy and radiation therapy.

Small cell GBM can be confused with high-grade oligodendrogliomas. *EGFR* amplification is highly prevalent in small cell GBM, occurring in 22 of 30 tumors in one study.[26] Thus, the absence of 1p/19q deletions and the presence of *EGFR* amplification increase the likelihood that gliomas with large numbers of small, relatively regularly shaped cells are likely to be small cell GBM.

Other alterations are thought to be prevalent in oligodendrogliomas that lack 1p and 19q deletions. *CDKN2A/CDKN2B/ARF* (*P16*, *P15*, *P19*), *PTEN*, and chromosome 10 are frequently deleted in high-grade oligodendrogliomas. The expression of *CDKN2A* and *CDKN2B* often is inactivated by promoter methylation. While the presence of these alterations is associated with a poorer prognosis,[24,27,28] their detection has not become a part of standard clinical neurooncology.

Available Assays

Despite intensive effort, the target genes on 1p and 19q have not been identified. However, molecular testing for 1p/19q deletion can be performed by loss of heterozygosity (LOH) analysis (currently by polymerase chain reaction [PCR] analysis of microsatellite loci or by high-throughput analysis of single nucleotide polymorphisms) and by fluorescence in situ hybridization (FISH) with locus-specific bacterial artificial chromosome (BAC) probes. FISH can be performed on two paraffin-embedded sections and does not require microdissection, which is an advantage over the other non–in situ methods. An advantage of FISH is that 1p/19q deletion can be focal,[29] and normal cells are frequently present in gliomas. FISH can directly evaluate lesions and tumor cells of interest. An advantage of LOH analysis is that it can detect mitotic recombination (which is not detected by FISH). However, this LOH mechanism has been shown to be rare in gliomas.[22]

FISH detection of 1p and 19q deletions has been described.[22,25] Figure 27-2 illustrates typical 1p/19q FISH results for two anaplastic oligodendrogliomas, one with and one without 1p/19q deletion. Approximately 60% of the nuclei in the tumor with 1p and 19q deletion contain one 1p36 or 19q13.3 probe signal and two 1q24 or 19p13 control probe signals (Figure 27-2a). The remaining cells with two signals for each probe are likely contaminating normal glial, neuronal, or endothelial cells or tumor cells that lack deletion. The mean 1p/1q and 19q/19p signal ratios per nucleus are 0.62 and 0.71, respectively. Normal value studies and an evaluation of a large series of tumors by multiple methods indicate that ratios less than 0.80 are associated with 1p or 19q deletion. By comparison, approximately 60% of the nuclei in the tumor without 1p and 19q deletion contain two signals for all four probes (Figure 27-2b). The mean 1p/1q and 19q/19p signal ratios per nucleus are 1.04 and 1.01, respectively. The nuclei with one signal for each probe are likely a result of truncation of signals by paraffin sectioning. The nuclei with three or more signals are likely a result of nuclear overlap.

When used with appropriate control probes (e.g., the BAC probes on 1q42 and 19p13), FISH is able to detect 1p and 19q deletions in polyploid or aneuploid tumors. In addition, trisomy 19 is detected by this FISH method, an alteration long known to be associated with high-grade astrocytic tumors.[30] Based on normal value studies, we do not report chromosome 1 or 19 aneusomy until the proportion of nuclei with three or more signals for each probe exceeds 30% of the nuclei evaluated.

Quantitative analysis of microsatellite alleles (QuMA) also has been shown to detect 1p and 19q deletions (and trisomy 19).[31] Both QuMA and current LOH analysis methods have the capacity for automation and rapid throughput. However, both require microdissection for accurate deletion detection, and this process increases turnaround time.

For similar reasons, FISH is the optimal method for the detection of chromosome 10 loss and *EGFR* amplification. In a careful method-comparison study, semiquantitative PCR analysis and Southern blot analysis did not detect small foci of *EGFR* amplification compared to FISH.[8] The most accurate means to detect the EGFRvIII variant in gliomas is by immunohistochemistry with specific antibodies.[32] The other *EGFR* variants can be detected by DNA sequencing of tumor specimens.

TP53 and *PTEN* mutation detection is best performed by high-throughput DNA sequencing/alteration methods (see chapter 2). The Mayo Clinic/NCCTG study included full sequencing of the *PTEN* gene and exons 5 to 8 of the *TP53* gene.[8]

Future Directions

Array-based gene expression and gene dosage analyses have the potential to generate new molecular markers for many cancers, including gliomas. Several groups are currently evaluating these modalities in selected gliomas. Using an 1,100-gene expression array, Watson et al. have identified 196 transcripts that are associated with oligodendrogliomas of different grades.[33] Similarly, two groups have used 12,000-gene arrays to discover differential expression of a limited number of transcripts that can be used to classify high-grade gliomas.[34,35] Array-based comparative genomic hybridization (CGHa) likely will lead to the discovery of novel gene dosage alterations that can be used to classify gliomas.[36,37] Importantly, the markers (and marker panels) ascertained by these methods will have to be carefully validated. However, there is the exciting possibility that a limited number of protein or genetic markers could be developed to enhance the molecular identification of gliomas, and potentially to standardize neuropathologic diagnosis.

Much of the current translational research in gliomas is focused on developing therapeutic approaches that target specific protein or genetic alterations, or both. The expression array studies described above have, as one goal, the identification of new drug targets, but current research already has identified promising targets for novel therapeutic approaches. For example, several components of the receptor tyrosine kinase pathway are altered in gliomas: EGFR is overexpressed, *PTEN* is mutated, and AKT2 is activated. Small molecule inhibitors, immunotherapeutic modalities, and viral therapeutic agents are currently being developed that target this pathway (for examples, see References 38 and 39) and are currently being tested in glioma patients. Other pathways also are being targeted for novel therapeutic approaches.

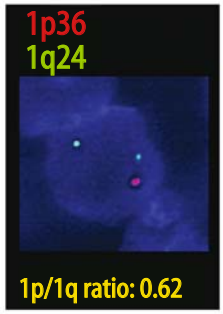
The design of current and future trials will need to develop the concept that therapeutic approaches will be customized to the molecular alterations found in specific tumors. Clinical molecular pathology will play an important role in this customization of clinical neurooncology practice.

Summary

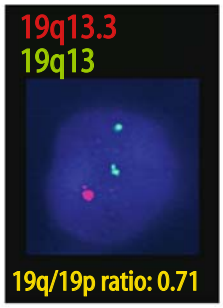
Many advances have been made in evaluating the molecular genetics of gliomas. For example, the presence of 1p and 19q deletions is associated with gliomas of oligodendroglial lineage, with oligodendrogliomas that have a better survival and with oligodendrogliomas that respond to chemotherapy and radiation therapy. While prospective data are needed to confirm these associations, 1p and 19q deletion testing has become a part of routine neurooncology practice. Alterations of chromosome 10, *EGFR*, *PTEN*, and *TP53* are associated with different grades of astrocytic gliomas and with patient survival. While their clinical utility requires further investigation, these markers (and others) will be increasingly used to stratify patients to current and new therapeutic approaches. In the future, it is likely that the survival rate of primary gliomas will be lengthened using biologically based treatment options linked with specific diagnostic biomarker profiles.

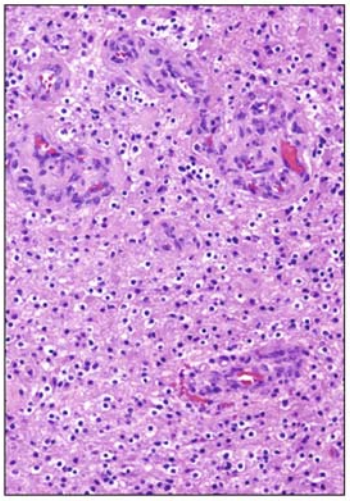
Anaplastic oligodendroglioma with loss of 1p and 19q

1q24 copy number	1p36 copy number 0	1	2	3	4	Total
0	0	0	0	0	0	0
1	0	17	1	0	0	18
2	0	70	12	0	0	82
3	0	0	0	0	0	0
4	0	0	0	0	0	0
Total	0	87	13	0	0	100



19p13 copy number	19q13.3 copy number 0	1	2	3	4	Total
0	0	0	0	0	0	0
1	0	24	0	0	0	24
2	0	51	23	0	0	74
3	0	0	1	1	0	2
4	0	0	0	0	0	0
Total	0	75	24	1	0	100

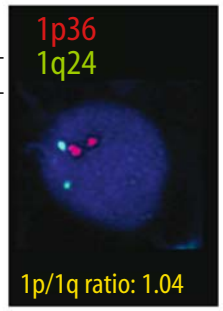




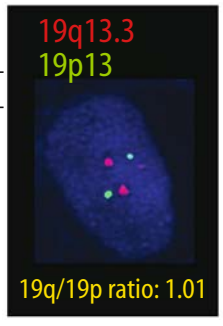
a

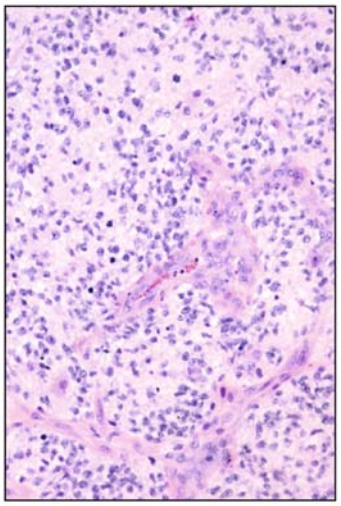
Anaplastic oligodendroglioma without loss of 1p and 19q

1q24 copy number	1p36 copy number 0	1	2	3	4	Total
0	0	0	0	0	0	0
1	0	15	11	0	0	26
2	0	6	57	4	0	67
3	0	0	1	5	0	6
4	0	0	0	0	1	1
Total	0	21	69	9	1	100



19p13 copy number	19q13.3 copy number 0	1	2	3	4	Total
0	0	0	0	0	0	0
1	0	13	8	0	0	21
2	0	6	66	0	0	72
3	0	0	0	6	0	6
4	0	0	0	0	1	1
Total	0	19	74	6	1	100





b

Figure 27-2. Representative histology and FISH results for two primary grade 3 anaplastic oligodendrogliomas. (a) An anaplastic oligodendroglioma with 1p and 19q deletion. (b) An anaplastic oligodendroglioma without 1p and 19q deletion. The values in the tables summarize the percent nuclei with the indicated number of 1p36 and 1q24 or 19p13 and 19q13.3 signals.

References

1. Jemal A, Murray T, Samuels A, Ghafoor A, Ward E, Thun MJ. Cancer statistics, 2003. *CA Cancer J Clin.* 2003;53:5–26.
2. Kleihues P, Louis DN, Scheithauer BW, et al. The WHO classification of tumors of the nervous system. *J Neuropathol Exp Neurol.* 2002;61:215–225.
3. Kleihues P, Ohgaki H. Primary and secondary glioblastomas: from concept to clinical diagnosis. *Neuro-oncol.* 1999;1:44–51.
4. Smith JS, Wang XY, Qian J, et al. Amplification of the platelet-derived growth factor receptor-A (PDGFRA) gene occurs in oligodendrogliomas with grade IV anaplastic features. *J Neuropathol Exp Neurol.* 2000;59:495–503.
5. Smith JS, Jenkins RB. Genetic alterations in adult diffuse glioma: occurrence, significance, and prognostic implications. *Front Biosci.* 2000;5:D213–D231.
6. Kitange GJ, Smith JS, Jenkins RB. Genetic alterations and chemotherapeutic response in human diffuse gliomas. *Expert Rev Anticancer Ther.* 2001;1:595–605.
7. Lin H, Bondy ML, Langford LA, et al. Allelic deletion analyses of MMAC/PTEN and DMBT1 loci in gliomas: relationship to prognostic significance. *Clin Cancer Res.* 1998;4:2447–2454.
8. Smith JS, Tachibana I, Passe SM, et al. PTEN mutation, EGFR amplification, and outcome in patients with anaplastic astrocytoma and glioblastoma multiforme. *J Natl Cancer Inst.* 2001;93:1246–1256.
9. Tada K, Shiraishi S, Kamiryo T, et al. Analysis of loss of heterozygosity on chromosome 10 in patient with malignant astrocytic tumors: correlation with patient age and survival. *J Neurosurg.* 2001;95:651–659.
10. Rasheed A, Herndon JE, Stenzel TT, et al. Molecular markers of prognosis in astrocytic tumors. *Cancer.* 2002;94:2688–2697.
11. Terada K, Tamiya T, Daido S, et al. Prognostic value of loss of heterozygosity around three candidate tumor suppressor genes on chromosome 10q in astrocytomas. *J Neurooncol.* 2002;58:107–114.
12. Simmons ML, Lamborn KR, Takahashi M, et al. Analysis of complex relationships between age, p53, epidermal growth factor receptor, and survival in glioblastoma patients. *Cancer Res.* 2001;61:1122–1128.
13. Frederick L, Wang XY, Eley G, James CD. Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas. *Cancer Res.* 2000; 60:1383–1387.
14. Feldkamp MM, Roncari L, Lau N, Salhia B, Guha A. Expression of EGFRvIII is a better prognostic marker for GBM than either EGFR expression or the Ki-67 labeling index. *Neuro-oncol.* 2002;4:343.
15. Peraud A, Watanabe K, Schwechheimer K, Yonekawa Y, Kleihues P, Ohgaki H. Genetic profile of the giant cell glioblastoma. *Lab Invest.* 1999;79:123–129.
16. Kosel S, Scheithauer BW, Graeber MB. Genotype-phenotype correlation in gemistocytic astrocytomas. *Neurosurgery.* 2001;48:187–193.
17. Nieder C, Petersen S, Petersen C, Thames HD. The challenge of p53 as prognostic and predictive factor in gliomas. *Cancer Treat Rev.* 2000;26:67–73.
18. Kraus JA, Glesmann N, Beck M, et al. Molecular analysis of the PTEN, TP53 and CDKN2A tumor suppressor genes in long-term survivors of glioblastoma multiforme. *J Neurooncol.* 2000;48:89–94.
19. Kraus JA, Wenghoefer M, Glesmann N, et al. TP53 gene mutations, nuclear p53 accumulation, expression of Waf/p21, Bcl-2, and CD95 (APO-1/Fas) proteins are not prognostic factors in de novo glioblastoma multiforme. *J Neurooncol.* 2001;52:263–272.
20. Shiraishi S, Tada K, Nakamura H, et al. Influence of p53 mutations on prognosis of patients with glioblastoma. *Cancer.* 2002;95:249–257.
21. Cairncross G, Macdonald D, Ludwin S, et al. Chemotherapy for anaplastic oligodendroglioma. National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol.* 1994;12:2013–2021.
22. Smith JS, Alderete B, Minn Y, et al. Localization of common deletion regions on 1p and 19q in human gliomas and their association with histological subtype. *Oncogene.* 1999;18:4144–4152.
23. Mueller W, Hartmann C, Hoffmann A, et al. Genetic signature of oligoastrocytomas correlates with tumor location and denotes distinct molecular subsets. *Am J Pathol.* 2002;161:313–319.
24. Cairncross JG, Ueki K, Zlatescu MC, et al. Specific genetic predictors of chemotherapeutic response and survival in patients with anaplastic oligodendrogliomas. *J Natl Cancer Inst.* 1998;90:1473–1479.
25. Smith JS, Perry A, Borell TJ, et al. Alterations of chromosome arms 1p and 19q as predictors of survival in oligodendrogliomas, astrocytomas, and mixed oligoastrocytomas. *J Clin Oncol.* 2000;18:636–645.
26. Burger PC, Pearl DK, Aldape K, et al. Small cell architecture—a histological equivalent of EGFR amplification in glioblastoma multiforme? *J Neuropathol Exp Neurol.* 2001;60:1099–1104.
27. Sasaki H, Zlatescu MC, Betensky RA, Ino Y, Cairncross JG, Louis DN. PTEN is a target of chromosome 10q loss in anaplastic oligodendrogliomas and PTEN alterations are associated with poor prognosis. *Am J Pathol.* 2001;159:359–367.
28. Sanson M, Leuraud P, Aguirre-Cruz L, et al. Analysis of loss of chromosome 10q, DMBT1 homozygous deletions, and PTEN mutations in oligodendrogliomas. *J Neurosurg.* 2002;97:1397–1401.
29. Randazzo L, Scarpace L, Lemke N, et al. Tumoral and chromosomal heterogeneity in oligodendrogliomas detected with medium-throughput 1p/19q LOH analysis. *Neuro-oncol.* 2002;4:346.
30. Bigner SH, Mark J, Burger PC, et al. Specific chromosomal abnormalities in malignant human gliomas. *Cancer Res.* 1988;48:405–411.
31. Nigro JM, Takahashi MA, Ginzinger DG, et al. Detection of 1p and 19q loss in oligodendroglioma by quantitative microsatellite analysis, a real-time quantitative polymerase chain reaction assay. *Am J Pathol.* 2001;158:1253–1262.
32. Ohman L, Gedda L, Hesselager G, et al. A new antibody recognizing the vIII mutation of human epidermal growth factor receptor. *Tumour Biol.* 2002;23:61–69.
33. Watson MA, Perry A, Budhjara V, Hicks C, Shannon WD, Rich KM. Gene expression profiling with oligonucleotide microarrays distinguishes World Health Organization grade of oligodendrogliomas. *Cancer Res.* 2001;61:1825–1829.
34. Mischel PS, Shai R, Choe GY, et al. Identification of distinctive subtypes of primary glioblastoma multiforme by gene expression profiling. *Neuro-oncol.* 2002;4:330.
35. Nutt CL, Mani DR, Betensky RA, et al. An accurate microarray class prediction model for high grade gliomas. *Neuro-oncol.* 2002;4:345.
36. Misra A, Pellarin M, Moore D, et al. A high-resolution genome wide map of genetic aberrations in human astrocytoma. *Neuro-oncol.* 2002;4:330.
37. Brennan C, Kim H, Cauwels C, et al. Identification of candidate genetic alterations from array-CGH in human glioblastomas. *Neuro-oncol.* 2002;4:327.
38. Lal A, Glazer CA, Martinson HM, et al. Mutant epidermal growth factor receptor up-regulates molecular effectors of tumor invasion. *Cancer Res.* 2002;62:3335–3339.
39. Sausville EA, Elsayed Y, Monga M, Kim G. Signal transduction-directed cancer treatments. *Annu Rev Pharmacol Toxicol.* 2003;43:199–231.

Chapter 28

Molecular Detection of Occult Tumor Cells

Karen L. Kaul

One of the key roles performed by pathologists is determination of the presence or absence of tumor cells in clinical samples. This is the basis for most approaches to staging, monitoring response to treatment, and detecting relapse of neoplasia and, as such, is a critical step in determining the course of patient management. Pathologists have utilized a variety of approaches, continually seeking to improve performance and thus patient outcome. The literature reflects this quest, including reports assessing the increased sensitivity afforded by step sectioning, immunohistochemistry, flow cytometry, and, more recently, molecular approaches for the detection of tumor cells in blood, bone marrow, and lymph node samples. The goal is, of course, the more accurate detection of disease spread and, ultimately, better patient care.

This chapter addresses some of the recent work in this area, focusing on molecular and, to some degree, immunohistochemical approaches for the detection of rare tumor cells in clinical samples. A synopsis of the hundreds of articles published to date is beyond the scope of this chapter; instead, more general issues and findings are addressed, along with presentation of selected work. Several reviews are available for more detailed reading.[1–6]

Molecular Basis of Disease

Cancer metastasis occurs when tumor cells acquire the ability to escape their local environment, enter the circulation to reach distant sites, attach at the distant site, and proliferate to form a metastatic lesion. Depending on the type of tumor, cells enter either the venous or lymphatic circulation (or both) and thus are spread to distant tissues (such as lung, liver, or bone marrow), or local lymph nodes, respectively, prior to the development of clinically detectable metastatic lesions. The fact that significant proportions of patients with organ-confined tumors who undergo theoretically curative surgery later suffer recurrence of their disease argues that current approaches to cancer staging are, to some degree, inadequate. Sensitive detection of tumor cells by immunohistochemical or molecular methods could lead to improved staging and monitoring of cancer patients. Such techniques also have been applied to the study of stem cell harvests, surgical resection margins, and various body fluids. Molecular techniques also have been proposed as tools for early detection of certain cancers using samples such as sputum, stool, and urine.

Since most solid tumors lack suitable genomic alterations useful for detection, gene expression markers are frequently employed. Thus, the basis for molecular detection is expression of a messenger RNA (mRNA) or protein marker unique to the tumor cells that can be detected via reverse transcription–polymerase chain reaction (RT-PCR) or immunohistochemistry (IHC), respectively. Unfortunately, truly specific tumor cell mRNA or protein markers are rare, and for carcinoma (derived from epithelial cells) markers are almost nonexistent. Most studies have therefore utilized markers reflecting the general type of cell or organ from which the tumor originates, and thus detect the presence of an expression-positive cell amidst the background of theoretically nonexpressing cells, such as leukocytes. Such markers cannot distinguish tumor from normal cells of the same cell type. Prostate-specific antigen (PSA), cytokeratins, carcinoembryonic antigen (CEA), and tyrosinase have been used as tumor markers, detected by either RT-PCR or IHC assays. Novel molecular and protein markers are being derived from new analytical approaches, such as expression microarray analysis.

Available Assays

IHC methods for tumor cell detection have been applied widely to preparations of cells from bone marrow, lymph node aspirates, and peripheral blood mononuclear cells. Cells can be smeared or centrifuged by cytospin onto slides, or sections taken from frozen or formalin-fixed, paraffin-embedded tissue. IHC is performed using standardized methods. Slides can be scanned by eye, though

to achieve a high level of sensitivity, thousands to millions of cells must be screened, which is tedious and time-consuming. Thus, image analysis has become a popular approach for screening.[7] Alternatively, flow cytometry has been utilized to achieve high number and rapid cell analysis.[8]

Tumor cells also can be assessed using DNA or RNA targets. Genomic targets can be utilized, if available, for the detection of a particular type of tumor. DNA targets have the advantage of greater stability compared to mRNA transcripts and, unlike protein targets, can be amplified to obtain greater analytic sensitivity. Some types of solid tumors do possess chromosomal translocations or point mutations that can be utilized for early diagnosis or detection of residual disease. Examples include certain sarcomas characterized by specific translocations, or *KRAS* mutations common in pancreatic and colorectal carcinoma, which have been used for detection of cells in pancreatic fluid and stool, respectively.[9–12] Microsatellite markers have been employed to assess cancer in certain body fluids.[13,14] However, for most cancers, such tumor-specific targets are lacking, and thus mRNA expression unique to the cell of origin for the tumor must be utilized.

For this purpose, RT-PCR has been widely used for detection of occult tumor cells in a variety of sample types. This approach permits analysis of a large number of cells simultaneously, and with great analytical sensitivity owing to the use of nucleic acid amplification. Key to the success of this approach is the careful choice of the mRNA target and meticulous design of the assay. Since few, if any, tumor-cell-specific mRNA targets exist, target transcripts that are unique to the tissue, such as cytokeratin for epithelial cells or PSA for prostate cancers, generally are used. In some cases, enrichment of tumor cells via immunoselection can enhance both the sensitivity and specificity of the RT-PCR assay.[15,16] Amplification approaches other than PCR also can be utilized effectively.[17]

Some investigators have targeted free nucleic acids in serum or plasma, rather than in the cellular fraction. Both DNA and RNA exist free in the circulation and have been studied in patients with breast cancer, melanoma, and colon cancer.[14,18–20]

Interpretation

In general, the interpretation of occult disease detection assays consists of either a positive or negative result, given that the assay controls are appropriate. For IHC markers, nonspecific or aberrant expression of protein markers requires that careful evaluation of the cytologic characteristics of the positive cells be performed. The ability to visualize cell morphology with IHC can reduce false-positive results because the interpretation of a positive result can be limited to IHC positive cells with tumor cell morphology. In the case of genomic or mRNA targets, cytologic evaluation is not possible, making target choice, and assay design and validation even more critical. Carefully designed and validated molecular assays will include positive and negative controls, as well as internal controls to demonstrate the presence of intact and amplifiable nucleic acid isolated from the patient sample. In particular, the internal control should be present at a constant level, and its expression should be low enough to accurately reflect RNA degradation in the sample.

Issues with sensitivity and specificity exist for all occult disease detection assays. The analytical sensitivity of each assay should be determined using serial dilutions of a positive cell line or cell type into a negative cell population. However, actual tumor cells may vary widely in the expression of the target mRNA or protein, and thus actual sensitivity may be significantly less than the sensitivity using the control. Additionally, some markers may fail to be expressed by tumor cells, or may be expressed inappropriately by other cell types.[21,22] Some groups have utilized multiplex assays in an effort to detect an increased range of tumor cells.[23–25] Others have used quantitative assays to define a threshold value above which a positive signal indicates the presence of tumor cells, and below which nonspecific or inappropriate expression of the marker is assumed.[26,27] However, because tumor cells may vary widely in their expression of a particular marker, it is difficult to define a clinically meaningful positive threshold above background. These and many issues must be addressed for assays to become more standardized, and thus for meaningful for clinical use.

Clinical Significance

Perhaps the greatest challenge for the development and validation of molecular assays for the detection of occult tumor cells is the demonstration of clinical utility. Carefully designed assays can successfully demonstrate the presence of tumor cells in clinical samples. Determining the clinical significance of these cells is considerably more difficult. Significant questions remain for most tumors regarding the issue of whether the presence of a small number of occult tumor cells in any sample type heralds the progression or recurrence of disease. While genomic abnormalities may indicate the presence of clearly malignant cells, protein or mRNA expression indicates cells of a particular organ or cell type, which may or may not be malignant. If malignant, these assays presently do not reflect the capacity for metastasis or progression. Additionally, the wide variation in assay protocols makes comparison of results between studies difficult.

Hundreds of studies using IHC and molecular methods to detect tumor cells have been performed. The sections below address some of the specific issues for more commonly examined sample types, and Table 28-1 summarizes a few studies, particularly those showing correlation to clinical parameters, organized by tumor type.

Table 28-1. Selected Molecular and Immunohistochemical Studies on Detection of Occult Tumor Cells

Sample Type	Finding	Reference
Breast		
Blood	In patients with metastases progression-free and overall survival correlated with circulating tumor cells	28
Nodes	All histologically positive nodes and 36% of negative nodes yielded positive CK 19 RT-PCR	29
Marrow	IHC detection of CK ⊕ cells correlated with progression	30, 31
Prostate		
Blood	PSA RT-PCR correlation with ⊕ margins	32
	RT-PCR not useful in staging, not prognostically relevant in clinically localized prostate cancer	33, 34
	PSA RT-PCR correlation relapse in African American men	35
	RT-PCR positives reduced in response to androgen deprivation	36
Nodes	IHC and RT-PCR correlated with relapse	37, 38
Marrow	RT-PCR results correlated with disease-free survival	39
Colorectal		
Blood	RT-PCR results correlated with stage	40
	CK 20 RT-PCR results correlated with prognosis	4
Nodes	CEA RT-PCR in stage 2 patients correlated with decreased survival	41, 42
	CK RT-PCR positives correlated with survival	
Marrow	CK 20 RT PCR correlated with outcome	4, 40
Melanoma		
Blood	RT-PCR predicted disease-free and overall survival	1
	RT-PCR used to monitor interferon treatment	43
Nodes	RT-PCR upstaged 25% of patients; associated with recurrence	44
Marrow	RT-PCR an independent predictor of disease-free survival	1
Pancreas		
Blood	CEA RT-PCR correlated with survival	45
Lung		
Blood	RT-PCR correlated with survival, positives reduced in patients on chemotherapy	46

CK, cytokeratin; PSA, prostate-specific antigen; CEA, chorioembryonic antigen.

Lymph Nodes

The proportion of patients who relapse from tumors of all types suggests that many patients are understaged and that better methods are needed to detect nodal metastasis. Tumor cells in lymph nodes can be detected by RT-PCR, and cases of histologically negative nodes that are molecularly positive are not uncommon. There is little question that molecular testing can increase the sensitivity of occult tumor cell detection. However, the clinical significance of detecting occult tumor cells has not been fully clarified. Several studies demonstrating a relationship between a positive molecular result for lymph nodes and patient outcome have been reported,[29,41,47,48] while others have shown no relationship. Many factors contribute to these confusing results, including variations in the assays (differing targets, variable sensitivity and specificity) and the amount of lymph node sampled (small metastatic foci may be missed). Additionally, the clinical relevance of tiny micrometastatic nodal metastases, even those evident by histology, remains in question.[49,50]

With the availability of rapid-cycle PCR technology, the application of molecular methods to the intraoperative assessment of patient samples has been proposed. In patients undergoing esophageal resection, RT-PCR has been used to detect *CEA* transcripts in lymph nodes during surgery and has revealed molecularly positive nodes that are histologically negative. In this small series, the positive patient suffered an earlier recurrence of cancer than those who were RT-PCR negative.[51]

Blood

Many studies evaluating circulating tumor cells have been published during the past decade. In a variety of tumor types, a correlation between the incidence of blood-borne cells and tumor stage has been demonstrated.[4,32,40,52,53] Unfortunately, the data do not support the reverse approach. At present, it does not appear possible to utilize molecular means to more accurately stage a patient.[33] Ultimately, the presence of circulating tumor cells may provide valuable information on the systemic spread of tumor in a manner different from conventional staging approaches. In prostate cancer, small but aggressive tumors may shed cells into the circulation while a more indolent but larger tumor may not; such a situation may appear to correlate poorly with tumor stage but may more accurately reflect tumor

aggressiveness. The detection of circulating breast cancer cells has been correlated with vascular invasion in the primary tumor. Larger studies using appropriate markers are needed to determine how to integrate these molecular test results into cancer staging strategies.

While a relationship between circulating tumor cells and tumor stage is evident, few studies have demonstrated a correlation between circulating cells and cancer recurrence or progression. Such studies require lengthy patient follow-up. Recently, however, a correlation between blood-borne cells and disease-free survival following radical prostatectomy has been reported.[35]

Marrow

In general, better correlation has been observed between patient outcome and molecular or IHC detection of tumor cells in the bone marrow compared to blood.[4,30,39,40] It is possible that circulating cells have the ability to get into the bloodstream but lack the ability to survive at a metastatic site, while cells in the marrow are at an "advanced" stage in the metastatic continuum, having acquired the capacity to survive in the circulation, attach, and grow in a remote environment. This, however, is speculation until we better understand the molecular events underlying the metastatic process.

While few studies using single markers have shown results that correlate with clinical or pathologic parameters, correlation with survival/outcome was observed using IHC to detect cytokeratin-positive cells in bone marrow samples from breast cancer patients.[30] Similar findings have been reported for prostate[39] and colorectal cancer.[41] In the United States, such studies can be hampered by clinical practices that do not include bone marrow sampling as part of routine staging for many tumors.

Margins

Technical advances in molecular methods are achieving rapid testing capabilities that may facilitate the use of molecular testing of surgical margins and lymph nodes in the intraoperative setting. While not practical at present, RT-PCR has been used for assessment of biopsies of the prostatic fossa in men undergoing radical prostatectomy.[54,55] In both studies, positive RT-PCR results correlated with clinical and histopathologic data.

Body Fluids

Cytologic assessment of peritoneal fluid is important in assessing the spread of tumor during surgical resection for ovarian and colorectal cancer. Studies report that positive results for IHC or RT-PCR testing of peritoneal fluid are an independent prognostic factor, correlating with disease progression.[4,56] Assessment of other body fluids by molecular methods may become useful in assessing the spread of tumors, as has been reported for pleural effusions from patients with lung cancer.[57]

Some types of fluid samples can be obtained noninvasively and thus present attractive approaches to screening and early diagnosis of certain cancers. Exfoliated cells are present in the sputum of lung cancer patients.[58] Other fluids, such as pancreatic fluid and breast ductal lavage samples may become valuable diagnostic samples.[12] However, RNA stability and the presence of adequate numbers of cells in the sample will be considerations for test development and validation. DNA alterations, if present, may be a more reliable marker for tumor cell detection.

Stem Cell Harvests

Whether the presence of tumor cell contamination in peripheral stem cell collections correlates with earlier relapse of breast cancer, as well as whether different approaches to stem cell mobilization may lead to differing levels of tumor cell contamination, has been debated for several years. The literature reflects this confusion, with numerous conflicting studies.[4,59,60] Larger studies with standardized methods are needed to determine the clinical utility of this type of testing.

Monitoring Treatment

Application of molecular methods to monitoring treatment response has become routine in the management of patients with acute leukemia. While such approaches are enticing for solid tumors, little data are yet available. In prostate cancer patients receiving androgen-deprivation treatment, a decreased incidence of circulating cells has been reported.[36] In a small study of breast cancer patients, those patients on chemotherapy or hormonal therapy with clinically evident metastasis showed a significantly lower incidence of circulating cells than patients not on treatment.[15] Similar results were observed for lung cancer.[46] RT-PCR has been used to monitor interferon response in melanoma patients.[43] Further studies are warranted.

Quality Control and Laboratory Issues

As with all nucleic acid–based tests, a variety of controls are required to ensure that reliable results are produced. A positive control must be included, and may be diluted to allow assessment of the assay sensitivity for each run. Negative controls also are critical and should provide controls for amplicon carryover during sample preparation and amplification. Perhaps the most critical control for RT-PCR assays is an internal control to demonstrate the presence of amplifiable mRNA in each sample. The lability of mRNA, and the challenge of procuring clinical samples for

processing in a timely fashion, makes this an important measure of the suitability of samples for analysis.

Target choice is critical and must be able to distinguish between the tumor cells and other cells in the sample. Assessment of assay specificity is a critical task in development of an RT-PCR assay, and must include analysis of a large number of clinical samples taken from patients without disease as well as with benign disease of the organ from which the tumor has developed. Numerous commonly used mRNA targets are known to potentially lack specificity, such as cytokeratin 19, which is reported to be expressed in normal leukocytes.[21] As has been mentioned, immunoselection of the tumor cells is one approach to avoid false-positive results in this setting. Pseudogenes for cytokeratin 19 and other genes can give rise to amplicons that appear by size to be derived from mRNA, but in fact reflect contaminating DNA present in the processed sample. Amplification reactions must be designed to detect gene transcripts rather than the gene itself, by using PCR primers that span intron/exon boundaries.

Conclusions and Future Directions

Since 1991, when the first report of detection of circulating cells in melanoma was published,[61] investigation of the clinical relevance of occult tumor cells in blood, bone marrow, and other sites has eluded many teams of investigators. The challenges in development and validation of RT-PCR assays aimed at expression of a normal gene (rather than a mutated, translocated, or chimeric gene) are significant and have undoubtedly led to much confusion within this field. Little standardization of these assays has yet occurred, making comparison of results of the various studies difficult. Carefully designed and controlled multicenter clinical studies are needed to further clarify the utility of this approach for any tumor or sample type. The International Union Against Cancer (UICC) staging protocol was recently revised to include molecular results in some of the categories, at this point largely to facilitate data gathering.[62]

Finally, the optimal marker, or combination of markers, for maximal sensitivity and specificity has not been determined. Recently, investigators have utilized gene expression approaches to identify panels of markers for use in the detection of occult tumor cells. For example, Martin and coworkers selected a set of 170 candidate genes using differential display analysis comparing gene expression in blood samples from normal and breast cancer patients.[64] Further analysis used gene expression microarray methods to refine this set to 12 genes that were upregulated in breast cancer patients, though a significant number of false positives were observed.[63] Other laboratories[64] have used the serial analysis of gene expression (SAGE) method to identify genes with abundant expression in breast cancer. Markers are needed that indicate not only the presence of tumor cells, but also the biologic relevance of the detected cells. Uciechowshi and colleagues, for example, have demonstrated that increasing genomic alterations in circulating breast cancer cells are associated with an increased risk of recurrence.[65] The proliferative ability of these cells may be another indicator of biologic potential.[66]

Thus, while promising, the IHC and molecular detection of occult tumor cells remains for the most part a research tool rather than a clinical assay. Significant efforts are needed before we know how to appropriately apply results to patient management. In many respects, this field parallels that of *HER2/NEU* in the early days of investigation, when experimental variation clouded clear conclusions. Hopefully, in time, molecular detection of occult tumor cells will have clear clinical utility and result in improved patient outcomes.

References

1. Ghossein RA, Bhattacharya S, Coit DG. Reverse transcriptase polymerase chain reaction (RT-PCR) detection of melanoma-related transcripts in the peripheral blood and bone marrow of patients with malignant melanoma. What have we learned? *Cancer Res.* 2001;158:63–77.
2. Pantel K, Cote RJ, Fodstad O. Detection and clinical importance of micrometastatic disease. *J Natl Cancer Inst.* 1999;91:1113–1124.
3. Raj GV, Moreno JG, Gomela LG. Utilization of polymerase chain reaction technology in the detection of solid tumors. *Cancer.* 1998;82:1419–1442.
4. Vogel I, Kalthoff H. Disseminated tumour cells. Their detection and significance for prognosis of gastrointestinal and pancreatic carcinomas. *Virchows Arch.* 2001;439:109–117.
5. Calaluce R, Miedema BW, Yohannes WY. Micrometastases in colorectal carcinoma: a review. *J Surg Oncol.* 1998;67:194–202.
6. 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;906:110–123.
7. Bauer KD, de la Torre-Bueno J, Diel IJ, et al. Reliable and sensitive analysis of occult bone marrow metastases using automated cellular imaging. *Clin Cancer Res.* 2000;6:3552–3559.
8. Tvasellas G, Huang A, McCullough T, Patel H, Araia R, Allen-March TG. Flow cytometry correlates with RT-PCR for detection of spiked but not circulating colorectal cancer cells. *Clin Exp Metastasis.* 2002;19:495–502.
9. Barr FG, Ladanyi M. Sarcomas. In: Leonard DGB, ed. *Diagnostic Molecular Pathology: Major Problems in Pathology.* Vol 41. Philadelphia: WB Saunders Co; 2003:53–76.
10. Kimura W, Zhao B, Futakawa N, et al. Significance of Ki-ras codon 12 point mutation in pancreatic juice in diagnosis of carcinoma of the pancreas. *Hepatogastroenterology.* 1999;46:532–539.
11. Wenger FA, Zieren J, Peter FJ, et al. K-ras mutations in tissue and stool samples from patients with pancreatic cancer and chronic pancreatitis. *Langenbecks Arch Surg.* 1999;384:181–186.
12. Uehara H, Nakaizumi A, Baba M, et al. Diagnosis of pancreatic cancer by K-ras point mutation and cytology of pancreatic juice. *Am J Gastroenterol.* 1996;91:1616–1621.
13. Mao L, Schoenberg MP, Scicchitano M, et al. Molecular detection of primary bladder cancer by microsatellite analysis. *Science.* 1996;271:659–662.
14. Nawroz H, Koch W, Anker P, Stroun M, Sidransky D. Microsatellite alterations in serum DNA of head and neck cancer patients. *Nat Med.* 1996;2:1035–1037.
15. Luke S, Kaul K. Detection of breast cancer cells in blood using immunomagnetic bead selection and reverse transcription-polymerase chain reaction. *Mol Diagn.* 1998;3:149–155.

16. Naume B, Borgen E, Nesland JM, et al. Increased sensitivity for detection of micrometastases in bone marrow/peripheral blood stem-cell products from breast-cancer patients by negative immunomagnetic separation. *Int J Cancer.* 1998;78:556–560.
17. Burchill SA, Perebolte L, Johnston C, Top B, Selby P. Comparison of the RNA-amplification based methods RT-PCR and NASBA for the detection of circulating tumour cells. *Br J Cancer.* 2002;86:102–109.
18. Silva JM, Garcia JM, Dominguez G, et al. Persistence of tumor DNA in plasma of breast cancer patients after mastectomy. *Ann Surg Oncol.* 2002;9:71–76.
19. Silva JM, Rodriguez R, Garcia JM, et al. Detection of epithelial tumour RNA in the plasma of colon cancer patients is associated with advanced stages and circulating tumour cells. *Gut.* 2002;50:530–534.
20. Gocke CD, Kopreski MS, Evans D. Telomerase mRNA is detectable in the plasma of pancreatic cancer patients [abstract]. *Clin Chem.* 2001;47:369.
21. Traweek ST, Liu J, Battifora H. Keratin gene expression in non-epithelial tissues. Detection with polymerase chain reaction. *Am J Pathol.* 1993;142:1111–1118.
22. Lambrechts AC, Bosma AJ, Klaver SG, et al. Comparison of immunohistochemistry, reverse transcriptase polymerase chain reaction, and nucleic acid sequence-bases amplification for the detection of circulating breast cancer cells. *Breast Cancer Res Treat.* 1999;56:219–231.
23. Mitas M, Mikhitarian K, Walters C, et al. Quantitative real-time RT-PCR detection of breast cancer micrometastasis using a multigene marker panel. *Int J Cancer.* 2001;93:162–171.
24. Taback B, Chan AD, Kuo CT, et al. Detection of occult metastatic breast cancer cells in blood by a multimolecular marker assay: correlation with clinical stage of disease. *Cancer Res.* 2001;61:8845–8850.
25. Hu X-C, Chow LWC. Detection of circulating breast cancer cells with multiple-marker RT-PCR assay. *Anticancer Res.* 2001;21:421–424.
26. Slade MJ, Smith BM, Sinnett HD, Cross NC, Coombes RC. Quantitative polymerase chain reaction for the detection of micrometastases in patients with breast cancer. *J Clin Oncol.* 1999;17:780–789.
27. Bostick P, Chatergee S, Chi D, et al. Limitations of RT-PCR in the detection of micrometastases in the lymph nodes of breast cancer patients. *J Clin Oncol.* 1998;16:2632–2640.
28. Christofanilli M, Hayes DF, Budd GT, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med.* 2004;351:781–791.
29. Schoenfeld A, Luqmani Y, Smith D, et al. Detection of breast cancer micrometastases in axillary lymph nodes by using polymerase chain reaction. *Cancer Res.* 1994;54:2986–2990.
30. Braun S, Pantel K, Müller P, et al. Cytokeratin-positive cells in the bone marrow and survival or patients with stage I, II, or III breast cancer. *N Engl J Med.* 2000;342:525–533.
31. Schoenfeld A, Kruger KH, Gomm J, et al. The detection of micrometastases in the peripheral blood and bone marrow if patients with breast cancer using immunohistochemistry and reverse polymerase chain reaction for keratin 19. *Eur J Cancer.* 1997;33:845–861.
32. Katz AE, de Vries GM, Begg MD, et al. Enhanced reverse transcriptase-polymerase chain reaction for prostate specific antigen as an indicator of true pathologic stage in patients with prostate cancer. *Cancer.* 1995;75:1642–1648.
33. Ignatoff JM, Oefelein MG, Watkin W, Chmiel JS, Kaul KL. Prostate specific antigen (PSA) reverse transcriptase polymerase chain reaction (RT-PCR) assay in preoperative staging of prostate cancer. *J Urol.* 1997;158:1870–1875.
34. Shariat SF, Gottenger E, Nguyen C, et al. Preoperative blood reverse transcriptase-PCR assays for prostate-specific antigen and human glandular kallikrein for prediction of prostate cancer progression after radical prostatectomy. *Cancer Res.* 2002;62:5974–5979.
35. Bianco FJ Jr, Powell IJ, Cher ML, Wood DP Jr. Presence of circulating prostate cancer cells in African American males adversely affects survival. *Urol Oncol.* 2002;7:147–152.
36. Su SL, Heston WD, Perrotti M, et al. Evaluating neoadjuvant therapy effectiveness on systemic disease: use of a prostatic-specific membrane reverse transcription polymerase chain reaction. *Urology.* 1997;49:95–101.
37. Shariat SF, Roudier MP, Wilcox GE, et al. Comparison of immunohistochemistry with reverse transcription-PCR for the detection of micrometastatic prostate cancer in lymph nodes. *Cancer Res.* 2003;63:4662–4670.
38. Shariat SF, Kattan MW, Erdamar S, et al. Detection of clinically significant, occult prostate cancer metastases in lymph nodes using a splice variant-specific rt-PCR assay for human glandular kallikrein. *J Clin Oncol.* 2003;21:1223–1231.
39. Wood DP Jr, Banerjee M. Presence of circulating prostate cells in the bone marrow of patients undergoing radical prostatectomy is predictive of disease-free survival. *J Clin Oncol.* 1997;15:3451–3457.
40. Soeth E, Roder C, Juhl H, Kruger U, Kremer B, Kalthoff H. The detection of disseminated tumor cells in bone marrow from colorectal cancer patients by a cytokeratin 20 specific nested reverse transcriptase—polymerase chain reaction is related to the stage of disease. *Int J Cancer.* 1996;69:278–282.
41. Liefers GJ, Cleton-Jansen AM, van de Velde CJH, Hermans J, van Krieken JH, Cornelisse CJ. Micrometastases and survival in stage II colorectal cancer. *N Engl J Med.* 1998;339:223–228.
42. Rosenberg R, Hoos A, Mueller J, et al. Prognostic significance of cytokeratin-20 reverse transcriptase polymerase chain reaction in lymph nodes of node-negative colorectal cancer patients. *J Clin Oncol.* 2002;20:1049–1055.
43. Mellado B, Del Carmen Vela M, Colomer D, et al. Tyrosinase mRNA in blood of patients with melanoma treated with adjuvant interferon. *Clin Oncol.* 2002;20:4032–4039.
44. Kuo CT, Hoon DS, Takeuchi H, et al. Prediction of disease outcome in melanoma patients by molecular analysis of paraffin-embedded sentinel lymph nodes. *J Clin Oncol.* 2003;21:3566–3572.
45. Uchikura K, Takao S, Nakajo A, et al. Intraoperative molecular detection of circulating tumor cells by reverse transcription-polymerase chain reaction in patients with biliary-pancreatic cancer is associated with hematogenous metastasis. *Ann Surg Oncol.* 2002;9:364–370.
46. Bessho A, Tabata M, Kiura K, et al. Detection of occult tumor cells in peripheral blood from patients with small cell lung cancer by reverse transcriptase-polymerase chain reaction. *Anticancer Res.* 2000;20:1149–1154.
47. Brown HM, Ahrendt SA, Komorowski RA, Doffek KM, Wilson SD, Demeure MJ. Immunohistochemistry and molecular detection of nodal micrometastases in pancreatic cancer. *J Surg Res.* 2001;95:141–146.
48. Godfrey TE, Kim S-H, Chavira M, et al. Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5′ nuclease quantitative reverse transcription-polymerase chain reaction. *J Mol Diagn.* 2000;2:84–91.
49. International (Ludwig) Breast Cancer Study Group. Prognostic importance of occult axillary lymph node micrometastases from breast cancer. *Lancet.* 1990;335:1565–1568.
50. Keene SA, Demeure MJ. The clinical significance of micrometastases and molecular metastases. *Surgery.* 2001;129:1–5.
51. Raja S, Luketich JD, Kelly LA, Gooding WE, Finkelstein SD, Godfrey TE. Rapid, quantitative reverse transcriptase-polymerase chain reaction: application to intraoperative molecular detection of occult metastases in esophageal cancer. *J Thorac Cardiovasc Surg.* 2002;123:475–483.
52. Olsson CA, de Vries GM, Benson MC, et al. The use of RT-PCR for prostate-specific antigen assay to predict potential surgical failures before radical prostatectomy: molecular staging of prostate cancer. *Br J Urol.* 1996;77:411–417.
53. Hoon DS, Wang Y, Dale PS, et al. Detection of occult melanoma cells in blood with a multiple marker polymerase chain reaction assay. *J Clin Oncol.* 1995;13:2109–2116.

54. Theodorescu D, Frierson HF, Sikes RA. Molecular determination of surgical margins using fossa biopsies at radical prostatectomy. *J Urol.* 1999;161:1442–1448.
55. Straub B, Müller M, Krause H, et al. Reverse transcriptase-polymerase chain reaction for prostate-specific antigen in the molecular staging of pelvic surgical margins after radical prostatectomy. *Urology.* 2001;57:1006–1011.
56. Guller U, Zajac P, Schnider A, et al. Disseminated single tumor cells as detected by real-time quantitative polymerase chain reaction represent a prognostic factor in patients undergoing surgery for colorectal cancer. *Ann Surg.* 2002;236:768–776.
57. Saito T, Kobayashi M, Harada R, Uemura Y, Taguchi H. Sensitive detection of small cell lung carcinoma cells by reverse transcriptase polymerase chain reaction for prepro-gastrin-releasing peptide mRNA. *Cancer.* 2003;97:2504–2511.
58. Lacroix J, Becker HD, Woerner SM, Rittgen W, Drings P, von Knebel Doeberitz M. Sensitive detection of rare cancer cells in sputum and peripheral blood samples of patients with lung cancer by prepro-GRP-specific RT-PCR. *Int J Cancer.* 2001;92:1–8.
59. Schulze R, Schulze M, Wischnik A, et al. Tumor cell contamination of peripheral blood stem cell transplants and bone marrow in high-risk breast cancer patients. *Bone Marrow Transplant.* 1997;19:1223–1228.
60. Cooper BW, Moss TJ, Ross AA, Ybanez J, Lazarus M. Occult tumor contamination of hematopoietic stem cell products does not affect clinical outcome of autologous transplantation in patients with breast cancer. *J Clin Oncol.* 1998;16:3509–3517.
61. Smith B, Selby P, Southgate J, Pittman K, Bradley C, Blair GE. Detection of melanoma cells in peripheral blood by means of reverse transcriptase and polymerase chain reaction. *Lancet.* 1991;338:1227–1229.
62. Hermanek P, Sobin LH, eds. *UICC TNM Classification of Malignant Tumors*. 5th ed. Berlin: Springer; 1998.
63. Martin KJ, Graner E, Li Y, et al. High-sensitivity array analysis of gene expression for the early detection of disseminated breast tumor cells in peripheral blood. *Proc Natl Acad Sci U S A.* 2001;98:2646–2651.
64. Bosma AJ, Weigelt B, Lambrechts AC, et al. Detection of circulating breast tumor cells by differential expression of marker genes. *Clin Cancer Res.* 2002;8:1871–1877.
65. Uciechowski P, Eder C, Böckmann B, et al. Prognostic value of genomic alterations in minimal residual cancer cells purified from the blood of breast cancer patients. *Br J Cancer.* 2000;83:1664–1673.
66. Bianco FJ Jr, Wood DP Jr, Gomes de Oliveira J, Nemeth JA, Beaman AA, Cher ML. Proliferation of prostate cancer cells in the bone marrow predicts recurrence in patients with localized prostate cancer. *Prostate.* 201;49:235–242.

Chapter 29

Microarray Studies (Beyond Histology)

Cyrus V. Hedvat

Molecular Basis of Disease

In 1999, the National Cancer Institute (NCI) began a funding initiative called the Director's Challenge: Toward a Molecular Classification of Tumors, with the goal of defining comprehensive profiles of molecular alterations in tumors that could be used to identify subsets of patients. These molecular profiles would provide the basis for future studies to validate the clinical utility of molecular-based classification schemes. A further goal of this initiative was the development and implementation of a plan for the timely release of the extensive data sets expected to result from these projects. At that time, it was becoming apparent that the knowledge of the entire human genome combined with rapidly improving technology for comprehensive analysis of the 20,000 to 25,000 genes and encoded proteins would provide an opportunity for a deeper understanding of the molecular basis of disease processes that would guide the evolution of improved therapies based on new diagnostic schemes. Specific molecular profiles correlate with important clinical parameters, which should allow physicians to base management decisions on the molecular characteristics of an individual patient's tumor, and improve a physician's ability to determine the primary tumor site for tumors of unknown origin at the time of their detection.

With the development and application of DNA microarrays, the expression of almost all human genes now can be systematically examined in human malignancies. This has led to the identification of candidate molecular targets for therapeutic intervention and biomarkers for early detection of these diseases. Various studies have demonstrated that patterns of gene expression can distinguish between tumors of different anatomic origin and define new subgroups of cancers with similar histologic appearances but distinct molecular profiles. Some of these new molecular subclasses of tumor appear to correlate with clinical behavior. Tests for characteristic gene expression patterns have not yet been translated into routine use in the clinical molecular laboratory and are unlikely to replace conventional microscopy in the near future. It is clear, however, that the practice of surgical pathology will significantly change as these findings are able to be incorporated into daily use. Instead of translation of findings from "bench to bedside," as in clinical oncology, discoveries will be translated from "bench to microscope."

Available Assays

DNA Microarray Technology

Comprehensive reviews of microarray techniques are available.[1–3] DNA microarrays are orderly, high-density arrangements of nucleic acid gene segments (either spotted complementary DNAs [cDNAs] or oligonucleotides), which enable broad-scale gene expression profiling. Labeled probes are generated from RNA by either reverse transcription or an in vitro transcription reaction using cDNA as a template and incorporating a specific label (Figure 29-1). The labeled probes are hybridized to the gene array, and relative signals detected, quantified, and compared.

cDNA Arrays

cDNA arrays are produced by spotting individual cDNA probes (usually purified PCR products) onto a solid support (nylon membrane or glass slide) using a robotic device, such that the entire array of spots represents a cDNA library. The cDNA array is used to compare the gene expression patterns of two conditions, such as normal compared to diseased tissue, benign compared to malignant tissue, or treated compared to untreated states. The RNA populations from the two comparative conditions are labeled with two different fluorescently labeled Cy-3 or Cy-5 tagged nucleotides. The two RNA-derived labeled probe sets are hybridized to separate arrays, or pooled in equal

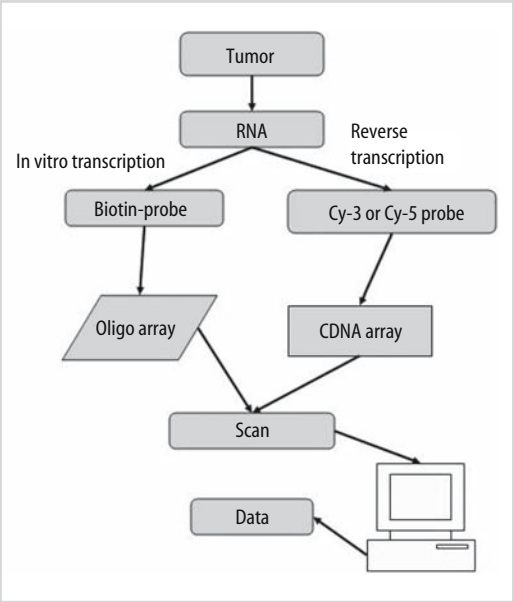


Figure 29-1. Gene expression analysis of tumors. RNA is extracted from an unfixed tumor sample and a labeled probe is generated from the RNA. The labeled probe is hybridized to the microarray, which then is scanned. The expression values are derived from the resulting hybridized probe signals, providing raw data for analysis.

amounts and hybridized to the same array. The fluorescent signals are measured and displayed as a ratio of the two fluorescent signals.

Oligonucleotide Arrays

Oligonucleotide arrays use small DNA sequences (20 to 70 nucleotides in length), either synthesized chemically and spotted onto a solid support or synthesized directly on the solid support or chip, such that the entire array of spots represents a set of genes. The most common version of this type of microarray is produced commercially by Affymetrix (Santa Clara, CA). In this version, the short oligonucleotides (length varies depending on the version of the chip) are synthesized in situ on the chip, with each gene represented by several oligonucleotides distributed along the length of the gene plus separate oligonucleotides with a single base mismatch with the gene sequence and used as a negative background control. Unlike spotted cDNA arrays, the two RNA populations to be compared are fluorescently labeled and hybridized to separate chips, and the fluorescent signal is detected. Control complementary RNAs (cRNAs) are spiked into each RNA population probe and used as an internal hybridization control. The supplied software converts the fluorescent intensity into a signal value for each gene, which is directly related to the amount of target mRNA in the probe. The relative signals for the same gene in the two RNA populations are then compared.

Proteomics

As in other high-throughput techniques, proteomics is in the rapid phase of development, with perhaps greater current attention to the technological advances than in the field of DNA microarrays. Proteomics is devoted to large-scale analysis of the proteome, which is the protein counterpart of the transcriptome. The proteome represents the result of gene expression, translation and protein degradation, and governs the phenotype of the cell, perhaps more than the transcriptome. Like the transcriptome, the proteome shows qualitative and quantitative changes during pathologic processes.

Proteomic analysis started with conventional technologies such as two-dimensional gel electrophoresis coupled with mass spectrometry.[4] Two classes of proteome array-based methodologies are distinguished by the nature of the arrayed molecules: protein or protein-binding particles (DNA, RNA, antibody, or other ligand). In one type of protein array, a large amount of protein is spotted on a solid support at a defined location and tested to characterize either a biochemical activity or a molecular interaction. Protein microarrays may be comprised of antibodies, whole cell or microdissected cellular lysates, recombinant proteins, small molecule drugs, or phage or antibody-like molecules. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), often referred to as gel-based proteomics, multidimensional chromatography, and protein biochips, in combination with mass spectrometry, are among the proteomic tools that are currently available for analysis of the proteome.

Proteomics encompasses many platform technologies for protein separation and identification, for determining biomolecular interactions, function, and regulation, and for annotating, storing, and distributing protein information. Today, the application of novel technologies from proteomics to the study of cancer is slowly shifting to the analysis of clinically relevant samples such as biopsy specimens and fluids. In the future, serum proteomic pattern analysis might be applied to medical screening as a supplement to diagnostic testing.

Tissue Microarrays

DNA microarrays provide invaluable tools for defining the global expression profile of a particular tissue; however, these studies are limited in that they are relatively expensive and time-consuming and rely on fresh frozen tissue that often is not readily available. Tissue microarrays (TMAs) allow the high-throughput analysis of paraffin-embedded tissues. This technique, originally described by Kononen et al.,[5] allows hundreds of tissue cores to be placed in a single paraffin block known as a TMA. Two instruments for the production of TMA blocks are available from Beecher Instruments (Sun Prairie, WI) and from Chemicon International (Advanced Tissue Arrayer,

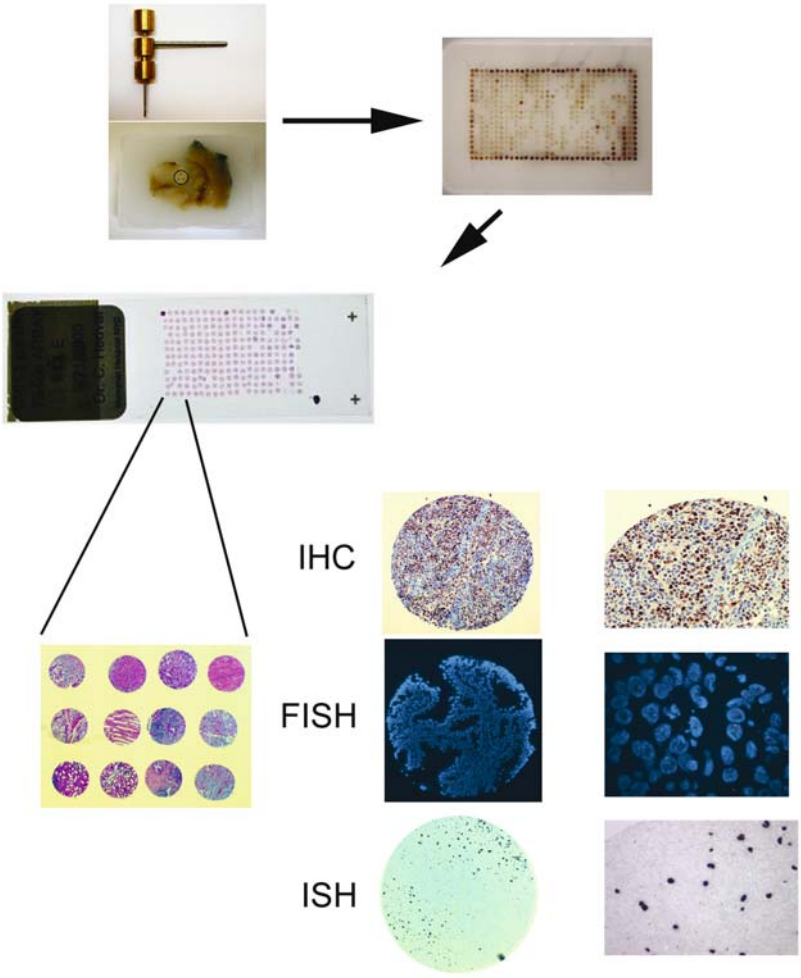


Figure 29-2. Tissue microarray construction and analysis. Tissue specimens embedded in paraffin blocks are sampled with a metallic punch, and the tissue cores are transferred to an empty paraffin block to produce a tissue microarray. Sections taken from the tissue microarray block will contain sections of all the sampled tissues, which can be studied simultaneously by immunohistochemistry (IHC), in situ hybridization (ISH), or fluorescent in situ hybridization (FISH).

Temecula, CA). The construction of a TMA (Figure 29-2) involves assembling a paraffin block containing hundreds of tissue cores (0.6 mm in diameter) derived from different "donor" blocks. Since the cores are very small, minimal damage is done to the original block. The technology is based on precision micrometers, which move a needle (0.6 mm to 2 mm in diameter) in the x- and y-axis. Micrometer drives are used to position the punch assembly with respect to the donor and recipient blocks. Once the TMA block is created, at least 50 histologic sections can be produced from the TMA block for either immunohistochemisty or in situ hybridization (fluorescent, radioactive, or nonradioactive probes).

TMAs can be created from single or multiple tumor types, different grades or stages of tumors, using human, mouse, or other animal tissues, or even cultured cell lines. They do present a significant challenge of data management, which is being addressed with new technologies. Technologies also are available to produce TMA blocks from frozen tissues.[6,7] Several scanners based on digital imaging and automated interpretation of immunohistochemical results have been commercially developed, including the BLISS imaging system (Bacus Laboratories, Lombard, IL), the ScanScope slide scanner (Aperio Technologies, Vista, CA), the ACIS II (Clarient, San Juan Capistrano, CA), and the Ariol system (Applied Imaging, San Jose, CA). The results from TMAs have been validated in several studies, which have shown that given adequate redundancy, immunostained TMAs are equivalent to standard whole sections.[8–10]

Interpretation of Results

Expression Data Analysis

The use of DNA microarrays generates a large number of individual data points, which must then be analyzed and archived. Optimal analysis requires expertise in statistics and bioinformatics, and the time and effort required to progress from the initial data acquisition to the extraction of relevant biological information is substantial. Some of the key aspects involved include image processing, data normalization, differential expression analysis, and database management. Each of the areas is complex, and a comprehensive discussion is beyond the scope of this chapter but can be explored further in several references.[11,12]

The process begins with image acquisition and analysis, which is largely dependent on the technology platform used. Data normalization issues are specific to the

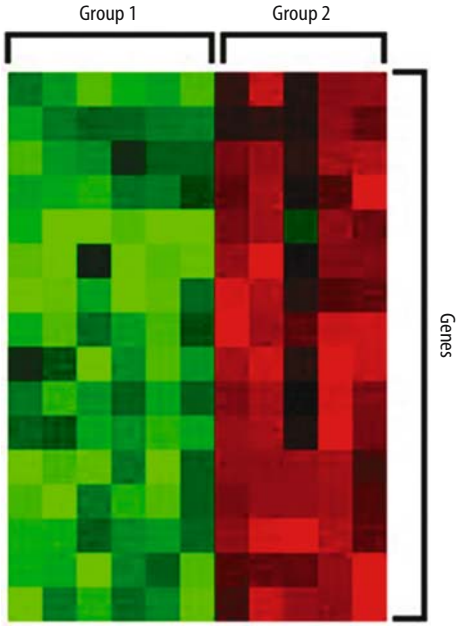


Figure 29-3. Hierarchical clustering analysis. Tumors can be divided based on similarities in gene expression patterns. Cells with log ratios of 0 are colored black, increasingly positive log ratios with reds of increasing intensity, and increasingly negative log ratios with greens of increasing intensity. Each gene is represented by a single row and each tumor sample by a single column.

microarray platform used for analysis, whether cDNA or oligonucleotide.

Investigators have adopted two general strategies for interpreting microarray data, commonly referred to as "unsupervised" and "supervised," to establish the degree of similarity between the gene expression levels in multiple samples. Unsupervised analysis does not take advantage of previous knowledge about samples in order to avoid any a priori relational assumption about function, structure, or any other experimental variable. The purpose is to identify relationships between samples based on similarities in gene expression. The most widely used of these techniques is hierarchical clustering, the object of which is to group together genes or samples with similar properties. The results of this analysis are typically displayed as a dendrogram representing the relatedness of the genes expressed within samples or of the samples themselves (Figure 29-3). The genes are commonly represented by colors, the intensity of which is proportional to the level of gene expression. Eisen et al.[13] developed a widely used hierarchical clustering and visualization software package called Cluster and TreeView, which can be used for analysis and display of microarray data.

Supervised methods, developed in statistics, are designed to classify samples according to key properties and can be used to test the strength of newly discovered sample relationships. This method also is used to look for statistically significant differences between two or more groups of samples. These methods rely on prior assumptions about the samples, such as the grade or stage of the tumors.

A typical microarray experiment yields expression data for thousands of genes from a relatively small number of samples, and gene-class correlations, therefore, can be revealed by chance alone. This issue can be addressed by collecting more samples from each class studied, but this is often difficult with clinical cancer specimens. Another approach is to perform exploratory data analysis on an initial data set and apply the findings to an independent test set. Findings confirmed in this fashion are less likely to be a result of chance alone. Permutation testing, which involves randomly permuting class labels and determining gene-class correlation, also has been used to determine statistical significance. Observed gene-class correlations that are stronger than those seen in permuted data are considered statistically significant.

Procedures to assess and maintain sample quality and calibrate scanning instrumentation to assure integrity of the data sets have not been standardized. Microarray technology is evolving rapidly, and end users are only beginning to learn how to interpret changes in largely unfamiliar study endpoints. Therefore, proscriptive guidelines for data interpretation also are not standardized. Establishment of such guidelines may be detrimental because it may limit exploratory research and the advancement of the science. Once we have a better understanding of study design, an accurate picture of the limitations of the technology, and improved understanding of data interpretation, such common guidelines may be possible.

Clinical Significance

The use of expression profiling for cancer diagnosis has been demonstrated for the classification of several cancers, including breast cancer, lung cancer, and others (Table 29-1). The potential applications of this new technology for elucidating disease mechanisms are vast. The possible genes and proteins for therapeutic drug and gene targeting may be identified with significant diagnostic, therapeutic, and disease-monitoring implications. The ability to work with limited amounts of starting RNA has improved such that sufficient quantities of RNA for microarray analysis can be extracted from formalin-fixed, paraffin-embedded tissues. This advance should increase

Table 29-1. DNA Microarrays Used to Classify Solid Tumors

Tumor	Technology	Reference
Lung	Oligonucleotide array	14
		15
Prostate	Oligonucleotide array	16
Breast	cDNA array	17
	cDNA array	18
	Spotted oligonucleotide array	19
Small round blue cell tumors	Oligonucleotide array	20
Melanoma	cDNA array	21

the potential clinical applications for archival pathology tissues. Neoplastic cells purified by microdissection or flow cytometry sorting can be isolated from small clinical biopsies to increase the number of samples for a given study or for clinical testing in the future.

Solid Tumors

Tumors have been classified historically based on their morphologic resemblance to normal tissue types. Microarray analysis has been applied to many types of solid tumors, including several studies involving breast cancer specimens. Perou et al.[17] explored expression-based breast cancer classification by studying 65 breast adenocarcinomas from 42 patients. Hierarchical cluster analysis defined three separate subtypes in this highly heterogeneous tumor class, based on patterns of gene expression. One subtype was known (*HER2/NEU/ERBB2* positive), and two others were previously unknown (luminal-like and basal-like). The clinical significance of the two novel cancer subsets has not been established. In a subsequent study of 78 independent tumors, clustering with the gene list from the first study identified similar patient subclasses, which then were found to be statistically significantly associated with differences in overall patient survival and relapse-free survival.[18]

Metastatic Tumors

A study comparing the gene expression profile of localized and metastatic prostate cancer demonstrated the power of microarrays in the development of predictive markers.[22] Varambally and colleagues[23] found overexpression of the polycomb group protein enhancer of zeste homolog 2 (*EZH2*) in hormone-refractory metastatic prostate cancer. The overexpression of EZH2 was validated using tissue arrays to show that the protein is overexpressed in the metastatic tumors, and application of an emerging powerful technology called RNA interference to disrupt expression of *EZH2* resulted in growth inhibition of a cell line derived from human metastatic prostate cancer.

Two microarray studies have challenged the hypothesis that tumors develop a metastatic potential as they grow. The studies tested the hypothesis that primary tumors harbor the necessary information to determine which tumors are more likely to metastasize. By comparing a series of breast tumors from node-negative and node-positive patients, van de Vijver et al.[19] found an expression profile that correlated with a poor prognosis and higher risk of metastasis. A second study suggests a broader applicability of these findings. Ramaswamy et al. discovered a genetic signature predictive of outcome by analyzing a variety of primary and metastatic adenocarcinomas.[24] A set of 17 differentially expressed genes present in the metastases was found in certain primary tumors without known metastases, and was associated with a poor prognosis. By using these specific gene expression signatures, the metastatic potential of even early tumors could be established, thereby identifying patients that should be treated more aggressively.

Quality Control and Laboratory Issues

Microarrays can be used to identify markers that can be validated using traditional diagnostic applications such as immunohistochemistry and adapted for routine clinical use. However, immunohistochemistry is generally non-quantitative, standardization of antibodies can be laborious, and multiplexing is difficult. More sophisticated and high-throughput validation methods are required for translation of microarray findings into clinical tests. An alternative approach would be to use microarrays in the clinical laboratory without the need for translation to another testing platform. This would require either custom arrays for different indications or whole genome analysis of every sample coupled with an analysis of relevant genes. As commercially available, low-cost, technically simple arrays and easy-to-use analytic software become available, routine clinical use of microarrays can be explored. In addition, the resulting data from clinical microarray testing could be used to generate large expression databases that could serve as an increasing database of centralized and standardized references to which new cancer samples could be compared. The routine clinical use of microarrays, however, has yet to be established. As molecular signatures of disease are identified, clinical testing for specific markers will become essential for the use of tailored therapies for individual patients.

The success of fully exploiting these powerful approaches depends on several factors: development of appropriate specimen handling techniques to assure target integrity; the accurate selection, amplification, and location of probe molecules; accurate reference sequence information; accurate distinction among multiple products of a single gene; precision image scanning; and reproducible and accurate transformation of image files to numerical data. For DNA and protein microarrays to be reliable tools, they must possess probe sequences that hybridize with high sensitivity and specificity, thereby allowing precise detection of their intended targets. Results must be highly reproducible, and quality control and quality assurance methods must be established. Determining the appropriate level of analytical and biological validation needed for each medical application of microarrays and their supporting computer-based bioinformatics systems raises new challenges for scientists in industry, academia, and regulatory agencies.

Measurements of genomic and proteomic alterations may be used to aid in risk assessment of patient subpopulations, to establish more specific diagnoses, to select optimal therapies, and to monitor patients' responses to therapies, for a broad variety of diseases, most notably

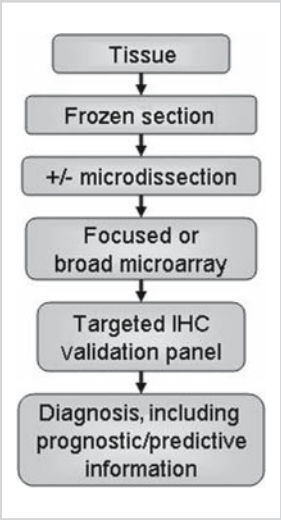


Figure 29-4. Potential future pathology specimen workflow.

cancer. Clinical trials of new drugs and biologics present unique opportunities for concomitant studies of diagnostics, including ones developed using microarray technologies, in a manner that meets both scientific and regulatory needs. Failure to address diagnostic issues when designing studies that promise new therapies can delay the scientific development and regulatory approval of the new diagnostics, impede full characterization of the new diagnostics due to a loss of controlled patient samples, and compromise evaluation of the therapy itself.

Once the strength of the linkage of genomic and proteomic measurements to associated biological outcomes is established, microarray testing methods must be available with sufficient sensitivity, specificity, reproducibility, robustness, reliability, accuracy, and precision for clinical use. Regardless of the evolution of these high-throughput techniques, the practice of surgical pathology will change. One possible view of the future workflow for patient samples is shown in Figure 29-4. As the power of microarrays and proteomics continues to develop, clinicians will increasingly ask for the availability of clinical microarray and proteomic testing of patient samples at the time of diagnosis. The pathologist will have to be actively involved in the implementation of such studies to ensure the integrity of the studies and the resulting clinical tests.

References

1. Maughan NJ, Lewis FA, Smith V. An introduction to arrays. *J Pathol.* 2001;195:3–6.
2. Holloway AJ, van Laar RK, Tothill RW, Bowtell DD. Options available—from start to finish—for obtaining data from DNA microarrays II. *Nat Genet.* 2002;32(suppl):481–489.
3. Churchill GA. Fundamentals of experimental design for cDNA microarrays. *Nat Genet.* 2002;32(suppl):490–495.
4. MacBeath G. Protein microarrays and proteomics. *Nat Genet.* 2002;(suppl):526–532.
5. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med.* 1998;4:844–847.
6. Hoos A, Cordon-Cardo C. Tissue microarray profiling of cancer specimens and cell lines: opportunities and limitations. *Lab Invest.* 2001;81:1331–1338.
7. Fejzo MS, Slamon DJ. Frozen tumor tissue microarray technology for analysis of tumor RNA, DNA, and proteins. *Am J Pathol.* 2001;159:1645–1650.
8. Hoos A, Urist MJ, Stojadinovic A, et al. Validation of tissue microarrays for immunohistochemical profiling of cancer specimens using the example of human fibroblastic tumors. *Am J Pathol.* 2001;158:1245–1251.
9. Hedvat CV, Hegde A, Chaganti RS, et al. Application of tissue microarray technology to the study of non-Hodgkin's and Hodgkin's lymphoma. *Hum Pathol.* 2002;33:968–974.
10. Rubin MA, Dunn R, Strawderman M, Pienta KJ. Tissue microarray sampling strategy for prostate cancer biomarker analysis. *Am J Surg Pathol.* 2002;26:312–319.
11. Brazma A, Vilo J. Gene expression data analysis. *Microbes Infect.* 2001;3:823–829.
12. Quackenbush J. Computational analysis of microarray data. *Nat Rev Genet.* 2001;2:418–427.
13. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A.* 1998;95:14863–14868.
14. Bhattacharjee A, Richards WG, Staunton J, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A.* 2001;98:13790–13795.
15. Garber ME, Troyanskaya OG, Schluens K, et al. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A.* 2001;98:13784–13789.
16. Singh D, Febbo PG, Ross K, et al. Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell.* 2002;1:203–209.
17. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature.* 2000;406:747–752.
18. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A.* 2001;98:10869–10874.
19. 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.
20. Pomeroy SL, Tamayo P, Gaasenbeek M, et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature.* 2002;415:436–442.
21. Bittner M, Meltzer P, Chen Y, et al. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature.* 2000;406:536–540.
22. Dhanasekaran SM, Barrette TR, Ghosh D, et al. Delineation of prognostic biomarkers in prostate cancer. *Nature.* 2001;412:822–826.
23. Varambally S, Dhanasekaran SM, Zhou M, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature.* 2002;419:624–629.
24. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. *Nat Genet.* 2003;33:49–54.

Section IV

Neoplastic Hematopathology

Section Editor: Adam Bagg

Chapter 30

Acute Myeloid Leukemia

David Grimwade

The diagnostic entity of acute myeloid leukemia (AML) encompasses a heterogeneous group of diseases whose prognosis differs substantially according to the nature of the underlying molecular lesion and the age of the patient. AML is predominantly a disease of the elderly with a dramatic increase in incidence in individuals over 60 years of age. Traditionally, cases of AML have been classified as primary (de novo) or secondary depending on the absence or presence of recognized predisposing factors (see Table 30-1).

Molecular Basis of Disease

The molecular basis of AML is to some extent dependent on the age of the population under consideration. In children and younger adults, balanced chromosomal translocations are relatively common, while in older patients AML is characterized by a more common picture of chromosomal losses and gains, often in the context of a complex karyotype. Evidence to date suggests that karyotype identifies biologically distinct subsets of disease and is highly predictive of response to therapy and overall survival (reviewed in Reference 1).

Genetic Consequences of Chromosomal Translocations Associated with AML

To date, more than 80 recurring balanced chromosomal rearrangements associated with AML have been characterized at the molecular level (see Table 30-2). Such rearrangements typically lead to the formation of chimeric fusion proteins, which are believed to play a key role in mediating the leukemic phenotype. A clear pattern is emerging whereby chromosomal rearrangements disrupt genes encoding transcription factors that under normal circumstances are directly involved in the regulation of hematopoiesis (reviewed in References 2 and 3). Notable examples include disruption of genes encoding the heterodimeric core binding factor (CBF) complex, which is comprised of α and β subunits by the t(8;21)(q22;q22) and inv(16)(p13q22)/t(16;16)(p13;q22) rearrangements, respectively, which lead to formation of AML1(CBFα2/RUNX1)-ETO and CBFβ-MYH11 fusions, respectively. A further example is provided by the involvement of RARα in fusion proteins resulting from rearrangements of 17q21 in acute promyelocytic leukemia (APL). Retention of dimerization, DNA-binding domains, or both within the resultant chimeric fusion proteins provides the potential to alter the pattern and repertoire of regulation of downstream target genes. A growing family of oncogenic fusion proteins, which include AML1-ETO and those involving TEL and RARα, have been shown to recruit nuclear corepressor complexes including histone deacetylase (HDAC), leading to transcriptional repression of genes implicated in normal hematopoietic differentiation.

Involvement of genes that encode proteins with a direct influence on chromatin remodeling or with transcriptional regulatory properties provides another recurring theme in the development of AML. Amongst this group are translocations involving the *MLL* gene at 11q23, as well as rearrangements disrupting genes encoding MOZ, TIF2, P300, and CBP. Deregulation of class I homeobox (*HOX*) genes, which encode transcription factors that play a key role in pattern formation and organogenesis during embryonic development, as well as contributing to the organization and regulation of hematopoiesis, provides a further common mechanism in the development of AML. A rarer target for AML-associated translocations involves genes encoding components of the nuclear pore complex, which plays a role in transport between nucleus and cytoplasm, for example, *NUP98* and *CAN (NUP214)* disrupted by rearrangements of 11p15 and 9q34 (most commonly t(6;9)(p23;q34), leading to DEK-CAN fusion), respectively.

The generation of chimeric fusion proteins, through chromosomal translocations, which lead to a block in differentiation and contribute to the biological characteristics of different subsets of leukemia, is generally believed to be a primary event in the pathogenesis of AML. The

Table 30-1. Factors Predisposing to the Development of Secondary AML

Genetic Predisposition
Down syndrome
Fanconi anemia
Other inherited bone marrow failure syndrome:
Shwachman-Diamond
Diamond-Blackfan
Dyskeratosis congenita
Kostmann's
Familial platelet disorder
DNA repair defects, e.g., Bloom's syndrome
Other tumor predisposition syndromes, e.g., Li-Fraumeni
Prior Hematologic Disorder
Chronic myeloid leukemia
Other myeloproliferative disorders
Myelodysplastic syndrome
Paroxysmal nocturnal hemoglobinuria
Exposure to Environmental or Therapeutic Agents
Chronic exposure to benzene and derivatives
Ionizing radiation
Chemotherapeutic agents
Alkylating agents
Topoisomerase II–targeting drugs

functional properties of the majority of AML-associated fusion proteins are quite distinct from those generated by chromosomal translocations associated with myeloproliferative disorders, which typically involve genes encoding components of signal transduction pathways and lead to formation of chimeric proteins with aberrant kinase activity that provide bone marrow progenitors with a proliferative advantage (see chapter 35). Such fusion proteins are rare in AML, with BCR-ABL occurring in only approximately 1% of cases. Nevertheless, over the last few years it has become apparent that mutations in genes encoding components of the signal transduction pathways are frequent in AML; these include activating mutations of FLT3, RAS, KIT, and SHP2. This latter class of mutations has been proposed to cooperate with the chimeric oncoproteins generated by AML-associated translocations, thereby endowing hematopoietic progenitors with the combined effects of a block in differentiation coupled with a proliferative advantage.[2] These are critical events in the development of the AML phenotype.

Other Mutations Implicated in the Pathogenesis of AML

Apart from the fusion genes resulting from chromosomal translocations identified in AML, it has become apparent over the last few years that other mutations play a role in the pathogenesis of the disease.[2,3] These fall into three categories: (1) mutations disrupting the function of transcription factors that are implicated in hematopoiesis, (2) mutations involving receptor tyrosine kinases and other components of the signaling cascades, and (3) mutations in the *nucleophosmin* (*NPM1*) gene encoding nucleophosmin.

Mutations in Genes Encoding Transcription Factors

The function of a number of key transcription factors that regulate normal hematopoietic development is disrupted in AML due to the presence of inactivating or dominant negative mutations (see Table 30-3). The first such example to be identified involved the *AML1* gene, which was found to be mutated in familial platelet disorder (FPD; reviewed in Reference 2). This rare autosomal dominant condition is characterized by progressive pancytopenia and dysplasia and ultimately progresses to AML in a high percentage of cases. *AML1* is mutated in approximately 3% to 5% of sporadic cases of AML, particularly those with acquired trisomy 21 and cases with minimal differentiation (French-American-British [FAB] type M0; details of FAB classification of AML are provided in Reference 7). Indeed, *AML1* mutations are detected in up to 25% of M0 cases and are frequently biallelic.

The CCAAT/enhancer binding protein-α (CEBPA), which plays an important role in granulopoiesis, also is a relatively common target in AML, being potentially deregulated by the AML1-ETO and PML-RARα oncoproteins. Furthermore, mutations involving *CEBPA* are present in approximately 10% of cases of AML.[3,8]

GATA1 is a key transcription factor involved in megakaryocytic differentiation and is mutated in cases of transient abnormal myelopoiesis (TAM) and acute megakaryoblastic leukemia (AMKL, FAB type M7) arising in infants with Down syndrome.[9] The role of PU.1, which has been implicated in normal differentiation of multiple hematopoietic lineages, in the pathogenesis of AML is somewhat controversial and awaits further investigation. A number of studies have detected partial tandem duplications of the *MLL* gene in AML, particularly in cases with trisomy 11 (reviewed in Reference 1).

Mutations in Genes Encoding Components of Signal Transduction Pathways

Activating mutations in hematopoietic tyrosine kinases are frequent events in AML.[2] Such mutations are likely to confer a proliferative advantage, although evidence to date suggests that they are unlikely to be the primary transforming event. Mutation of the *FLT3* gene encoding Fms-like tyrosine kinase 3 is one of the most common genetic lesions identified in AML thus far, being detected in almost one third of cases (reviewed in Reference 5). The majority are length mutations (internal tandem duplications, ITDs) involving the juxtamembrane region of the receptor, whereas approximately 4% to 7% of AML cases have mutations involving the activation loop. Nevertheless, both

Table 30-2. Frequencies of Specific Cytogenetic Abnormalities Determined in a Series of 1584 Patients Derived from the MRC AML10 Trial

Chromosomal Aberrations Associated with AML (grouped according to underlying molecular abnormality)	Molecular Consequence	Frequency in Children and Younger Adults (0–55 years)	Associated Clinical Features
Core binding factor gene rearrangement: ***AML1 (CBFA2)*** **gene at 21q22 or** ***CBFB*** **gene at 16q22**		**12%**	
t(3;21)(q26;q22)	*EAP/MDS1/EVI1-AML1* fusion	Rare	
t(8;21)(q22;q22)	*AML1-ETO* fusion	8%	Often AML M2, associated with chloromas
inv(16)(p13q22)/t(16;16)(p13;q22)	*CBFB-MYH11* fusion	4%	M4 with abnormal eosinophils (M4Eo)
t(16;21)(q24;q22)	*MTG16-AML1* fusion	Rare	Therapy-related AML
RARA **gene rearrangement (17q12~21):**		**12%**	5 fusion partners of *RARA* identified to date;* associated with APL
t(5;17)(q35;q12~21)	*NPM-RARA* fusion	Rare	AML M3, ATRA sensitive
t(11;17)(q23;q21)	*PLZF-RARA* fusion	Rare	AML M2/M3, ATRA and arsenic resistant
t(15;17)(q22;q12~21)	*PML-RARA* fusion	12%	AML M3, ATRA and arsenic sensitive
MLL **gene rearrangement (11q23):**		**5%**	More than 30 fusion partners identified to date; therefore, only the commonest abnormalities detected in AML are listed; associated with M4 and more commonly M5. Some cases are therapy related, secondary to topoisomerase II–targeted drugs, particularly etoposide
t(6;11)(q27;q23)	*AF6-MLL* fusion	<0.5%	
t(9;11)(p21~22;q23)	*AF9-MLL* fusion	0.8%	
t(10;11)(p11~13;q23)	*AF10-MLL* fusion	0.8%	
t(11;19)(q23;p13.3)	*MLL-ENL* fusion†	<0.5%	
Other		2.7%	
Involvement of zinc finger encoding genes by rearrangements of 3q: ***EVI1*** **gene at 3q26 or** ***MDS/EVI1*** **homologue** ***MEL1*** **at 1p36**			
inv(3)(q21q26)/t(3;3)(q21;q26)	EVI1 overexpression	2%	Trilineage dysplasia, dysmegakaryopoiesis
t(3;12)(q26;p13)	*MDS1/EVI1-TEL* fusion	<0.5%	
t(1;3)(p36;q21)	MEL1 overexpression	<0.5%	
Involvement of genes encoding factors involved in signal transduction: ***ABL*** **at 9q34**			
t(9;22)(q34;q11)	*BCR-ABL* fusion	0.8%	Typically associated with CML and ALL Detected in AML M0/1 and 2
Nuclear pore component genes: ***CAN (NUP214)*** **at 9q34** ***NUP98*** **at 11p15**			
t(6;9)(p23;q34)	*DEK-CAN* fusion	0.8%	AML M2 with basophilia
t(11)(p15)	*NUP98* fusions	Rare	*NUP98* rearrangements commonly associated with previous exposure to topoisomerase II–targeted drugs
Other genes encoding transcriptional regulators/chromatin modulators, including: ***MOZ*** **gene rearrangements at 8p11** ***CBP*** **gene rearrangements at 16p13**			
t(1;22)(p13;q13)	*OTT-MAL fusion*	<0.5%	Associated with infant AML M7
inv(8)(p11q13)	*MOZ-TIF2* fusion	Rare	AML M4/5 with prominent erythrophagocytosis
t(8;16)(p11;p13)	*MOZ-CBP* fusion	<0.5%	
t(8;22)(p11;q13)	*MOZ-P300* fusion	Rare	
t(10;16)(q22;p13)	*MORF-CBP* fusion	Rare	Childhood M5a
Ets family transcriptional factor genes: ***TEL (ETV6)*** **gene rearrangements at 12p13** ***ERG*** **gene rearrangement at 21q22**			
t(12)(p13)	*TEL (ETV6)* fusions	Rare	Associated with wide range of FAB types
t(16;21)(p11;q22)	*FUS-ERG* fusion	Rare	Reported in AML M1,2,4,5,7
Other recurrent translocations in AML: ***NPM*** **on 5q34~35**			
t(3;5)(q25;q34)	*NPM-MLF1* fusion	<0.5%	*NPM* is also disrupted in t(5;17) in APL and t(2;5) in anaplastic lymphoma
AF10 **on 10p12~p13**			
t(10;11)(p12~p13;q14~q21)	*CALM-AF10* fusion	<0.5%	Associated with poorly differentiated AML; also detected in T-ALL

* *NuMA-RARA* and *STAT5b-RARA* associated with t(11;17)(q13;q21) and der(17), respectively, have been identified in only single patients.
† Three different *MLL* fusion partners have been identified in cases with t(11;19), with *ENL* being the commonest in AML.
ATRA, all-*trans* retinoic acid.

Table 30-3 Genes Subject to Mutation in AML

Gene	Frequency of Mutation in AML	Associated Clinical Features
Transcription Factors		
AML1	3% (excluding M0)	High frequency in AML M0 (~21%) and AML with acquired trisomy 21
CEPBA	10%	Associated with AML M1, M2, and standard-risk karyotype; mutations not detected in AML1-ETO-positive AML
GATA1	Rare	Mutations detected in transient abnormal myelopoiesis and AML M7 in Down syndrome
PU1	Not established	Reported in cases of M0, M4, and M5*
MLL		
Cleavage	3%	DNA cleavage at TOPOII site
ITD	3%	More common in cases with trisomy 11
Tyrosine Kinases and Other Signal Transduction Molecules		
FLT3		
ITD	20–25%	More common in AML with PML-RARα (particularly M3v), DEK-CAN/t(6;9), and cases with normal karyotype; associated with higher presenting blast counts; ITD relatively rare in cases with adverse karyotypic features
Activation loop mutations†	4–7%	
KIT	3%	More common in core binding factor leukemias; potential cooperating event in leukemogenesis; more common in AML with trisomy 4
NRAS/KRAS	20%	Does not appear to be an independent prognostic indicator
PTPN11	4%	*PTPN11* encodes SHP-2; mutations associated with AML M5 and monosomy 7
Chromatin Modulators		
NPM1	~35%	Mutated in >50% of normal karyotype AML; associated with FLT3-ITD, M5 morphology, CD34 negativity, and elevated leukocyte count

*Based on report by Mueller et al.,[4] but findings remain to be confirmed by other studies.
†Point mutations (e.g., D835/I836) and length mutations (FLT3-840GS).[5,6]

types of mutation lead to constitutive activation of the receptor. *FLT3* ITDs are associated with higher presenting peripheral blast counts and are commonly detected in AML with normal karyotype and in *PML-RARA*-associated APL, particularly the hypogranular variant form (M3v).

Mutations in *NRAS* or *KRAS* also are relatively common in AML, occurring in approximately 20% of cases. Interestingly, these rarely coexist with mutations of *FLT3*. Indeed, *RAS* mutations are quite common in CBF leukemias, while *FLT3* ITDs are rare, the converse of what occurs in APL. The RAS pathway also is activated in the presence of mutations of the *PTPN11* gene, which encodes the SHP2 tyrosine phosphatase. *PTPN11* is commonly mutated in juvenile myelomonocytic leukemia (JMML) but also is involved in approximately 4% of AML, being associated with M5 morphology, elevated leukocyte count, and monosomy 7.[10,11] Mutations affecting KIT occur at a similar overall frequency in AML but are particularly common in CBF leukemias, raising the possibility that in this context they could provide a second hit, contributing to leukemogenesis.

Nucleophosmin *Mutations*

Recently, mutation of NPM1, which is a multifactorial nucleocytoplasmic shuttling protein implicated in regulating gene expression and the ARF-TP53 tumor-suppressor pathway, has been identified as a common mechanism in the pathogenesis of AML.[12] Mutations disrupt C-terminal tryptophan residues that are required for nuclear localization, leading to a characteristic cytoplasmic staining pattern by immunohistochemistry using antibodies against the NPM1 protein (that are in common use in diagnosis of lymphoma with NPM-ALK fusion).[12] Overall, approximately 35% of AML have *NPM1* mutations; however, mutation is particularly prevalent among cases with a normal karyotype (48–62%).[12–14] *NPM1* mutations are associated with *FLT3* ITD, CD34 negativity, presentation with elevated leukocyte count, and M4/M5 FAB type.[12–14]

Classification of AML

With the advent of routine cytogenetic characterization of AML, coupled with improved understanding of the molecular basis of the disease achieved through cloning of chromosomal breakpoint regions and identification of mutations in a wide range of genes, it has become clear that many primary AML-associated lesions that define biological entities are heterogeneous when classified according to FAB type.[7] This underlies the finding that the diagnostic karyotype is more highly predictive of initial response to

therapy and risk of relapse than is the FAB type, which reflects to a limited degree the molecular and clinical diversity of the disease. In the latter respect it has become apparent that cases of AML arising following previous chemotherapy or radiotherapy or on the background of a prior hematologic disorder such as myelodysplasia (MDS) are associated with a generally poorer prognosis than cases arising de novo. Moreover, there is increasing evidence, for example, karyotypic similarities and high rates of P-glycoprotein (PGP) overexpression, to suggest that a significant proportion of AML cases arising in older patients are biologically akin to secondary leukemias. These additional facets, that is, genetic and clinical features that do not form a part of the FAB classification but are highly pertinent to determining treatment approach for patients with AML and predicting the likelihood of cure, have been partly taken into account in the recent World Health Organization (WHO) classification (reviewed in Reference 15; see Table 30-4). The WHO system still, however, retains a morphological classification for all cases that do not fall into its "specific" clinical, morphological, or cytogenetic groups. These morphological groups are based on the FAB categories M0, M1, M2 M4, M5, M6, and M7, with the addition of three new categories, namely, acute basophilic leukemia, acute panmyelosis with marrow fibrosis, and myeloid sarcoma. A major difference between the FAB and WHO classifications lies in the minimum bone marrow blast percentage as a cutoff separating MDS from AML, with the WHO system dropping the FAB threshold of 30% blasts down to 20%. Moreover, cases with clonal abnormalities in the form of t(15;17)(q22;q12~21), inv(16), or t(8;21) are deemed to have AML irrespective of blast percentage, according to the WHO system.

Table 30-4. Classification of AML According to the World Health Organization

AML with Recurrent Genetic Abnormalities
t(8;21)/*AML1-ETO*
AML with abnormal eosinophils and inv(16)(p13q22)/t(16;16)(p13;q22)/*CBFB-MYH11*
APL with t(15;17)(q22;q12~21)/*PML-RARA* and variants
AML with 11q23/*MLL* abnormalities
AML with Multilineage Dysplasia
Following MDS or myeloproliferative disorder (MPD)
No previous history of MDS or MPD, but with dysplasia in ≥50% of cells in two or more myeloid lineages
AML and Myelodysplastic Syndromes, Therapy Related
Alkylating agent/radiation-related
Topoisomerase II inhibitor-related
Others
AML, Not Otherwise Categorized, Classified As
AML, minimally differentiated
AML without maturation
AML with maturation
Acute myelomonocytic leukemia
Acute monoblastic/acute monocytic leukemia
Acute erythroid leukemia (erythroid/myeloid and pure erythroleukemia)
Acute megakaryoblastic
Acute basophilic
Acute panmyelosis with myelofibrosis
Myeloid sarcoma

Table 30-5. Prognostic Factors for Subsequent Relapse in Patients with Newly Diagnosed AML

Pretreatment Predictors	
Major	**Other**
Karyotype	Elevated WBC
Molecular	Elevated LDH
FLT3 ITD	Secondary disease*
NPM1 mutation	Dysplastic features*
Older Age	BCL2/BAX ratio
	Resistance protein expression e.g. PGP*
	TP53 mutation*
	Autonomous growth of AML blasts in culture*
	Immunophenotype e.g. CD34 expression*
	Expression of chemotherapy metabolizing enzymes
	Expression of *BAALC, EVI1*, and *ERG*
Posttreatment Predictors	
Initial response (post-induction blast count)	
MRD assessment	

Many factors (indicated by *) previously considered to confer adverse risk have been shown to be closely related to cytogenetic risk group.
WBC, presenting white blood cell count; LDH, lactate dehydrogenase; PGP, P-glycoprotein; MRD, minimal residual disease.

Indications for Testing

Thorough laboratory investigation of AML is fundamental to the successful clinical management of the disease, fulfilling diagnostic and prognostic roles. Prognostic factors in AML can be subdivided into pretreatment and posttreatment factors (see Table 30-5). With the exception of age and type of AML (de novo vs secondary), pretreatment determination of the likely outcome to therapy is dependent on laboratory investigation, with karyotype and *NPM1* and *FLT3* mutation status being the most important factors identified to date.[5,13,14] Predicting likely outcome following initiation of therapy falls entirely within the realm of the laboratory and, for patients in morphological complete remission (CR) after initiation of treatment, is dependent on the detection of minimal residual disease (MRD), for which a number of technologies have been developed (see Rationale for MRD Monitoring below).

Establishing a Diagnosis of AML

Morphological analysis of peripheral blood and bone marrow smears stained with Wright Giemsa or May-

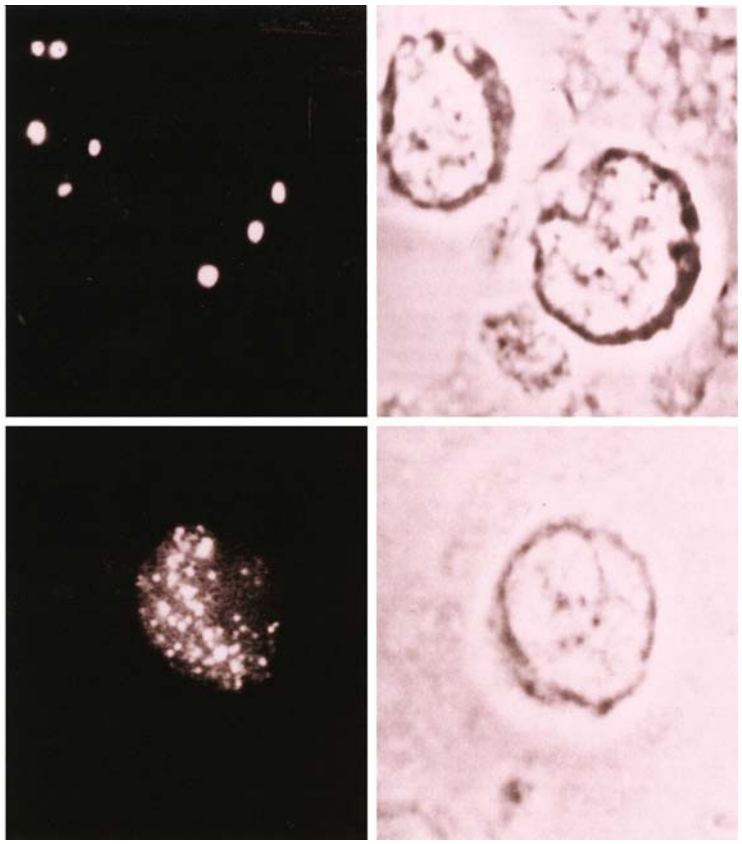


Figure 30-1. Rapid diagnosis of acute promyelocytic leukemia with the PML-RARA fusion by PML immunostaining. In leukemic blasts from cases of PML-RARA-negative AML and in normal cells, PML protein is localized within nuclear body structures (PML nuclear bodies). In such cells PML antisera detect a wild-type staining pattern comprising less than 30 (typically 5–20) discrete nuclear dots, e.g., a case of AML M1 (upper left panel). Whereas in APL cases with the PML-RARA fusion, PML nuclear bodies are disrupted, leading to a characteristic microspeckled/microparticulate nuclear staining pattern (>30 nuclear dots) with PML antisera, which detects PML-RARA and wild-type PML proteins (lower left panel). Nuclear integrity is confirmed by a nuclear stain or in this case by phase contrast microscopy (right-hand panels). These studies were performed with a PML polyclonal antibody, with FITC conjugated secondary antibody.

Grünwald Giemsa is clearly the first step in establishing a diagnosis for patients with suspected leukemia. After confirmation that a patient has acute leukemia on the basis of morphology, cytochemistry, and immunophenotyping, patients with AML may be further classified according to blast characteristics and degree of differentiation in the marrow.[7] This initial workup may reveal features that are helpful in predicting the presence of particular molecular lesions (see Table 30-2), and may be especially valuable in situations in which the expected cytogenetic abnormality is lacking, yet the predicted fusion gene is formed as a result of a cryptic rearrangement, for example, insertion event, which may be detected by fluorescence in situ hybridization (FISH) or reverse transcription polymerase chain reaction (RT-PCR).[16,17]

It is critical to identify patients who could possibly have APL, since these patients are at high risk of sudden induction death due to hemorrhage if not treated promptly and appropriately. Furthermore, this group requires early introduction of molecularly targeted therapy in the form of all-trans retinoic acid (ATRA) in order to achieve optimal treatment results. For the rapid diagnosis of APL, immunofluorescent methods using antibodies directed against the PML protein are of value, with the conversion from a normal pattern of 5 to 20 discrete dots to a microparticulate nuclear staining pattern (>30 microspeckles; see Figure 30-1) being indicative of the presence of the *PML-RARA* fusion gene. Alternatively, RT-PCR for detection of the *PML-RARA* fusion transcript may be used to confirm the diagnosis of APL.

Immunophenotyping is a valuable component of the routine diagnostic workup for patients with AML for a number of reasons, including (1) establishing, confirming, and further refining a diagnosis of AML, (2) identification of potential prognostic markers, and (3) establishing expression profiles suitable for subsequent MRD monitoring. Although immunophenotypic studies are assuming a more important role in the investigation of AML (reviewed in Reference 18), they are not discussed in any great detail in this chapter, which is restricted to the role of molecular genetic analyses.

Role of Karyotype Assessment

Karyotype analysis should be considered a mandatory component of the workup for all patients with AML, since it provides important diagnostic and prognostic information. Approximately 60% of cases of AML are characterized by a clonal abnormality, and indeed, according to the new WHO classification, detection of t(15;17), t(8;21), inv(16)/t(16;16), and *MLL* gene rearrangement can form the basis of a diagnosis of AML in cases found to have less than 20% blasts in the bone marrow.[15] Cytogenetic analysis is also of value in identifying patients who are candidates for molecularly targeted treatment approaches. In

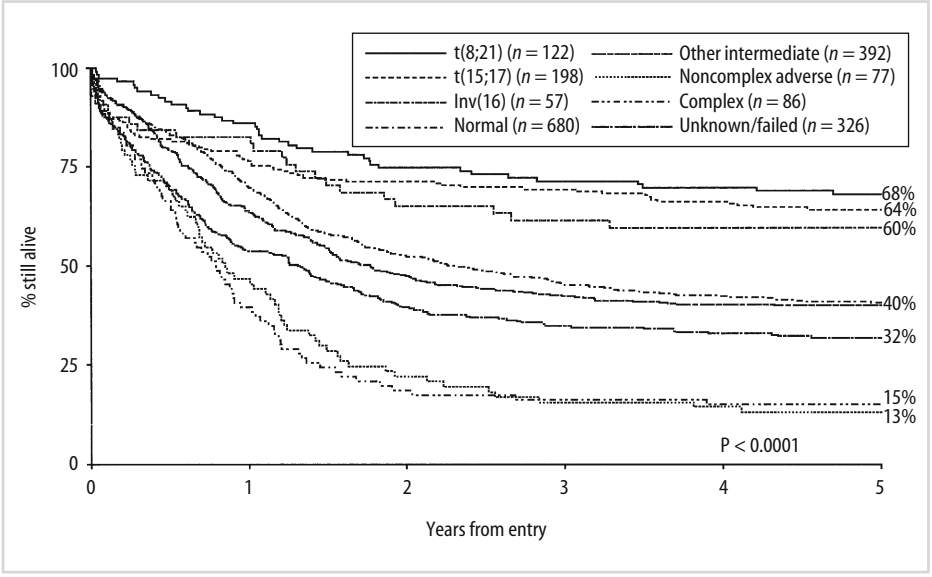


Figure 30-2. Prognostic impact of diagnostic karyotype in children and younger adults entered into UK Medical Research Council AML10 trial. Patients with t(8;21), t(15;17), and inv(16) were found to have a superior overall survival and are assigned to the favorable risk group in the MRC classification system. Patients (lacking one of the above abnormalities) with complex karyotype (defined as five or more unrelated abnormalities) or presence of -5/del(5q), -7, or 3q abnormalities in the context of a noncomplex karyotype were found to have a very poor outcome and comprise the MRC adverse risk group. Patients with normal karyotype or other structural or numerical abnormalities (other intermediate) comprise the MRC standard risk group. (Figure prepared by Georgina Harrison, Clinical Trials Service Unit, Oxford, UK. Reprinted from Grimwade D. The clinical significance of cytogenetic abnormalities in acute myeloid leukaemia. *Baillieres Best Pract Res Clin Haematol*. 2001;14:497–529; copyright 2001, with permission from Elsevier.)

particular, presence of the t(15;17), which generates the *PML-RARA* fusion in APL, predicts a favorable response to ATRA and arsenic trioxide, while detection of the rare, in AML, t(9;22)(q34;q11), which leads to the BCR-ABL fusion, identifies a group of patients with dismal prognosis when treated with conventional chemotherapy, who could potentially derive some benefit from imatinib.

Karyotype is the most important independent predictor of outcome identified to date for patients with AML. Analysis of large cohorts of patients receiving comparable treatment approaches has permitted the determination of the prognostic impact of the most common recurring cytogenetic abnormalities in AML (see Figure 30-2). Patients with t(15;17), t(8;21), and inv(16)/t(16;16), which lead to *PML-RARA*, *AML1-ETO,* and *CBFB-MYH11* gene fusions, respectively, have a relatively favorable outcome; whereas monosomies of chromosome 5 or 7, del(5q), 3q abnormalities, and t(9;22) are associated with an adverse prognosis. Consideration of the outcome of cases with more than one cytogenetic abnormality, in which the karyotype includes features that in their own right would confer favorable or adverse risk, respectively (see Figure 30-3), has led to the definition of hierarchical risk groups (see Table 30-6). The distribution of the cytogenetic risk groups varies substantially according to age. A greater proportion of younger patients have favorable risk abnormalities, while older patients are more likely to have AML with chromosome 5 and 7 abnormalities, often in the context of a complex karyotype, accounting for the expansion of the adverse risk group in the age group of individuals 55 years and older (see Figure 30-4). Cytogenetic classification of AML has proved to be of immense clinical value and is used to determine treatment approach in a number of clinical trials. It is now apparent that routine use of autologous or

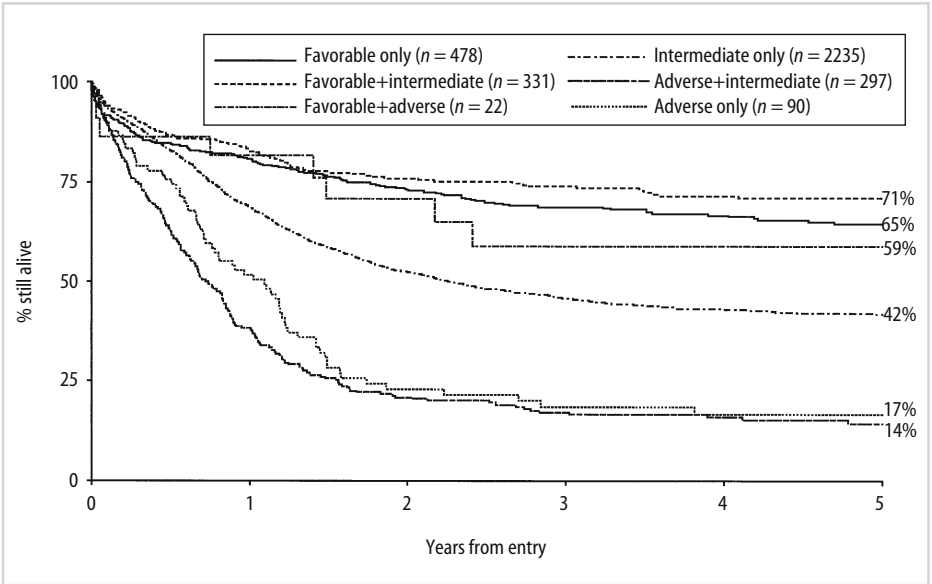


Figure 30-3. Prognostic impact of additional cytogenetic abnormalities in children and younger adults entered into UK Medical Research Council AML10 and AML12 trials. Additional cytogenetic abnormalities from either intermediate or adverse risk categories did not confer a deleterious effect on outcome among patients with favorable karyotypic features. In the remaining patients, presence of adverse karyotypic features was associated with a poorer prognosis in comparison to patients with intermediate risk abnormalities alone (provided that these did not number more than four unrelated abnormalities, i.e., a very complex karyotype, which is associated with a poor prognosis). Favorable, intermediate, and adverse risk groups are as defined in Table 30-6. (Figure prepared by Georgina Harrison, Clinical Trials Service Unit, Oxford, UK. Reprinted from Grimwade D. The clinical significance of cytogenetic abnormalities in acute myeloid leukaemia. *Baillieres Best Pract Res Clin Haematol*. 2001;14:514; copyright 2001, with permission from Elsevier.)

Table 30-6. Definition of Risk Groups According to Diagnostic Karyotype, Based on the Classification System Adopted by the UK Medical Research Council Group

Cytogenetic Risk Group	Karyotype	Comments
Favorable	t(8;21) t(15;17) inv(16)/t(16;16)	Whether alone or in conjunction with other abnormalities (including complex karyotype)
Intermediate	Normal Other abnormalities (not classified as favorable or adverse)	Cases lacking favorable or adverse cytogenetic features
Adverse	inv(3)(q21q26)/t(3;3)(q21;q26) abn(3q) –5/del(5q) –7 t(9;22) Complex (≥5 unrelated abnormalities)	Whether alone or in conjunction with intermediate-risk or other adverse-risk abnormalities

allogeneic bone marrow transplantation (BMT) in first CR has little to offer those patients with favorable cytogenetics, with any benefit in terms of reduced relapse risk being more than offset by treatment-related mortality and morbidity. Conversely, patients with adverse karyotypic features fare extremely poorly with conventional therapy and if no suitable allogeneic donor is available could be considered candidates for more experimental treatment approaches as part of first-line therapy.

Detection of Chromosomal Aberrations by Molecular Screening/FISH

Although karyotype assessment provides a valuable framework for design of risk-adapted treatment of AML, it has some limitations as a means of determining the most suitable treatment approach for individual patients. Indeed, information derived from karyotype assessment may be substantially embellished by the results of molecular analyses, which serve to further refine the diagnosis of AML, reliably identify subgroups of patients that require a specific treatment approach (e.g., molecularly targeted therapies), and define targets for subsequent MRD assessment.

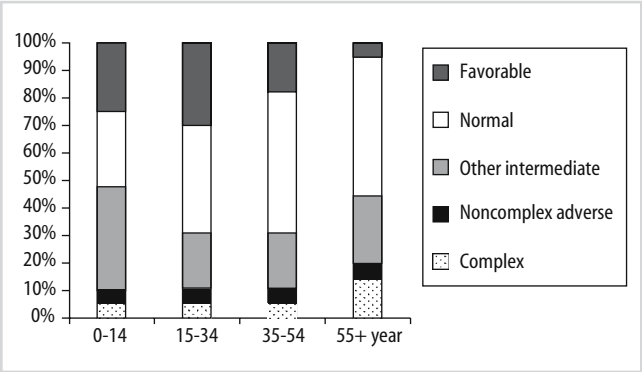


Figure 30-4. Variation in cytogenetic risk group according to age. Cytogenetic risk group (see Table 30-6) was determined in 2677 patients derived from Medical Research Council AML10 and AML11 trials. Balanced chromosomal rearrangements associated with favorable risk were relatively more frequent in children and younger adults, whereas adverse karyotypic features were more common in older patients. (Reprinted from Grimwade D, Walker H, Harrison G, et al. The predictive value of hierarchical cytogenetic classification in older adults with acute myeloid leukemia (AML): analysis of 1065 patients entered into the United Kingdom Medical Research Council AML11 trial. *Blood.* 2001;98:1312–1320. Copyright American Society of Hematology, used with permission.)

A critical failing of cytogenetic analysis is that it may yield misleading results due to sampling of residual normal marrow elements and is unsuccessful in approximately 10% of patients, with rates differing according to whether samples are analyzed in local or more remote laboratories, entailing substantial differences in transit time. In such cases, RT-PCR, FISH, or both readily detect the presence of fusion genes corresponding to the favorable cytogenetic risk group. For patients lacking molecular evidence of fusion genes associated with favorable risk, interphase FISH may prove helpful in identifying lesions that would assign cases to the adverse risk group. These include evidence of loss of chromosome 5 or 7 material, a complex pattern of losses and gains, or presence of the *BCR-ABL* fusion. In non-APL patients, screening for involvement of *MLL* in the pathogenesis of AML can effectively be performed by Southern blot analysis. Although this is more labor-intensive than RT-PCR or FISH, it provides additional information indicating the presence of *MLL* cleavage or duplication events, as well as occurrence of translocations. It is not clear that screening for *MLL* abnormalities currently has any significant bearing on management of AML, but it may be useful in identifying additional patients with *MLL*-associated fusion genes who could be monitored for MRD by PCR-based approaches.

For those cases of AML with evaluable metaphases at diagnosis, patients may not necessarily be assigned to the most appropriate risk group on the basis of karyotype assessment alone, due to the occurrence of cryptic rearrangements. Studies conducted by the European Working Party and UK Medical Research Council (MRC) have established that almost 10% of cases with morphologic APL and successful karyotype analysis actually lack the classical t(15;17) but nevertheless have an underlying *PML-RARA* fusion.[16,19] This is most commonly the result of insertion events, in which chromosomes 15 and 17 are typically of normal appearance. Such cases share the beneficial

response to targeted therapies, i.e., ATRA and arsenic trioxide. This is supported by outcome data from the UK MRC ATRA trial, which showed that patients in whom *PML-RARA* was identified solely by molecular means had a comparable outcome to those with t(15;17) documented by karyotype.[19] This would suggest that cases of AML with cryptic rearrangement of CBF genes (i.e., with *AML1-ETO* or *CBFB-MYH11*) also are likely to be biologically equivalent to and share the favorable prognosis of those with karyotype-documented t(8;21) or inv(16)/t(16;16).

This supports the adoption of more widespread molecular screening for CBF leukemias in AML, which is most efficiently undertaken by RT-PCR. Evidence to date suggests that at least 10% of cases of AML with CBF leukemia have a cryptic rearrangement and that such cases cannot be reliably identified on the basis of distinct morphological features (reviewed in Reference 17). Where evidence of CBF leukemia is revealed by RT-PCR screening in the absence of the typical cytogenetic lesion, the result should be independently confirmed using FISH to document fusion gene formation, for example, as a result of an insertion event (see Figure 30-5). FISH comfirmation serves to exclude the possibility that the result was spurious due to PCR contamination or was unlikely to be of clinical relevance, should the fusion gene have arisen in a minor clone.

The relative value of molecular screening is dependent on the age group of patients to be screened, with CBF leukemias representing a greater proportion of AML arising in children and younger adults than in the elderly, where they comprise less than 5% of cases. The relative benefit of molecular screening also will vary according to the success of cytogenetic assessment, being most rewarding where cytogenetics has a relatively high failure rate or when there are no facilities for central review of karyotypes. Molecular screening for CBF leukemias appears to be worthwhile because it enables the identification of additional patients who can be subject to MRD assessment using a fusion gene marker, as well as identifying individuals who can potentially be spared BMT in first CR. While cryptic *CBFB-MYH11* and *AML1-ETO* gene fusions have been observed in a wide range of FAB types, the *PML-RARA* fusion is very closely correlated with M3/M3v morphology. Therefore, routine molecular screening for the *PML-RARA* fusion among all newly diagnosed cases of AML is not considered worthwhile but should be restricted to infrequent cases with severe coagulopathy, an APL-like immunophenotype, or cases with cells suggestive of APL morphology even if occurring in a minor cell population.

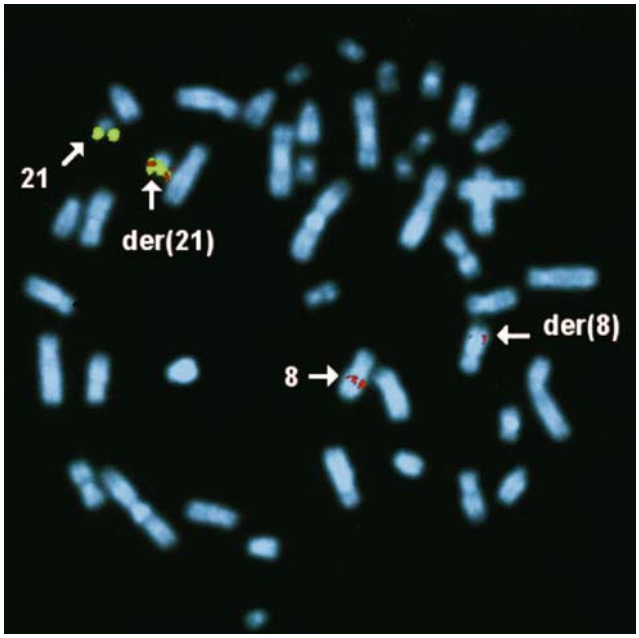


Figure 30-5. Characterization by FISH of AML with fusion gene formation due to cryptic rearrangements. Cryptic (8;21) rearrangement in an AML M2 case demonstrated by FISH with *AML1* (green) and *ETO* (red) probe set (Vysis, Downers Grove, IL). Fusion signals are seen on the derivative 21, and *ETO* signals are diminished on the derivative 8. (Figure prepared by Marina Lafage-Pochitaloff, Institut Paoli Calmettes, Marseille. Reprinted from Grimwade D. The clinical significance of cytogenetic abnormalities in acute myeloid leukaemia. *Baillieres Best Pract Res Clin Haematol*. 2001;14:510; copyright 2001, with permission from Elsevier.)

Detection of Mutations Implicated in Leukemogenesis and/or Predictive of Outcome

A further limitation of the use of karyotype to determine treatment approach in AML relates to the relatively large proportion of patients classified as intermediate risk. This group has been shown to be heterogeneous at the molecular level and includes cohorts of patients with markedly differing prognoses. A key challenge is to more precisely define the prognosis of rarer recurring cytogenetic abnormalities that are currently assigned to this group. There is interest in additional molecular assays that can distinguish patients with differing prognoses within cytogenetic risk groups, which could prove to be of value in improving risk stratification in AML. For patients with favorable risk cytogenetics, the most effective way of discriminating those at differing risk of relapse is through MRD assessment involving detection of leukemia-associated fusion genes (see Rationale for MRD Monitoring below).

Among patients with an intermediate risk karyotype, a number of molecular markers are helpful. Most notable is the prognostic impact of *FLT3* ITD (see Figure 30-6), the presence of which is associated with a substantial increase in relapse risk.[5] However, it is not clear that the poorer outcome in this group can be averted by treatment modification in the form of BMT in first CR.[20] Overexpression of *EVI1*, as determined by real-time quantitative RT-PCR (RQ-PCR), even in the absence of 3q abnormalities, has been found to predict a poor outcome.[21] FLT3 activation loop and *RAS* mutations appear to lack any independent prognostic significance,[5,22] whereas presence of *NPM1* or *CEBPA* mutation in the absence of *FLT3* ITD identifies further groups of patients with relatively favorable prognosis, who may not benefit from BMT in first CR.[8,13,14,23]

There also has been considerable interest in determining whether silencing of gene loci, for example, *P15*

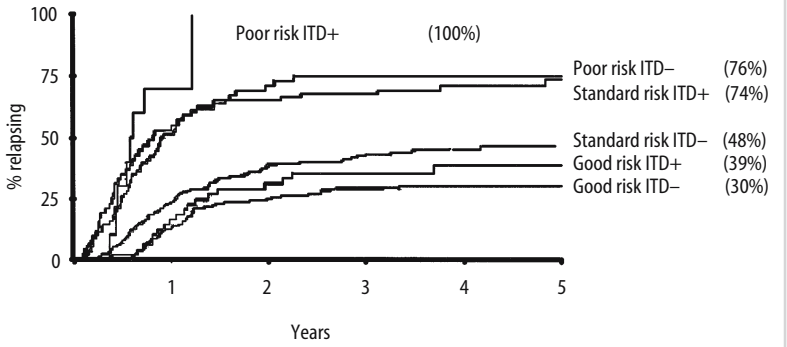


Figure 30-6. Prognostic impact of *FLT3* ITDs. Presence of a *FLT3* ITD is an independent predictor of increased relapse risk within UK MRC risk groups, as defined by diagnostic karyotype and bone marrow blast count following induction chemotherapy. Analysis was performed on 610 cases derived from MRC AML 10 and 12 trials for children and younger adults. (Reprinted from Kottaridis PD, Gale RE, Frew ME, et al. The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. *Blood.* 2001;98:1752–1759. Copyright American Society of Hematology, used with permission.)

(*INK4B*), by methylation contributes to the pathogenesis of AML and may be of prognostic value. These analyses have typically been based on methylation-specific PCR. However, such assays can be difficult to interpret, since results may be confounded by the presence of residual normal cell populations in which the gene in question also happens to be "switched off," a process that is associated with methylation of the regulatory elements.

The next few years are likely to see an expansion in the repertoire of molecular tests applied to diagnostic material from patients with AML as a means of determining suitability for novel targeted therapeutic approaches. One of the first examples was provided by molecular screening for activating mutations of the *FLT3* gene (see Other Mutations Implicated in the Pathogenesis of AML section above) to accompany the first wave of clinical trials evaluating targeted inhibitors of FLT3.

Rationale for Minimal Residual Disease Monitoring

Relapse remains a major cause of treatment failure for patients with AML. Since the group of patients that is destined to relapse cannot be reliably identified on the basis of pretreatment characteristics alone, there has been considerable interest in the development of assays to detect MRD as a means of more precisely tailoring therapy according to individual patient requirements (reviewed in Reference 24). MRD assessment might prove to be of value in a number of clinical scenarios. Reliable assessment of kinetics of response relatively early in the treatment course could help in identifying patients at differing risk of relapse. Patients at low risk could be spared excessive therapy with its inherent additional risks of treatment-related morbidity and mortality, including that relating to subsequent development of secondary MDS or AML. Conversely, the outcome of patients at high risk of relapse could potentially be improved with the administration of additional consolidation therapy, possibly in the form of BMT or more novel approaches. MRD assessment also may be of potential value in the management of patients undergoing transplant procedures, as a means of determining the most suitable type of transplant, as well as identifying the need for additional therapy in the posttransplant setting.

A number of methods have been evaluated for detection of MRD, including cytogenetics, FISH, immunophenotyping, and PCR-based approaches. Cytogenetics is too insensitive to provide a realistic option, although it may be of interest as a complement to morphological assessment to determine remission status following induction therapy.[25] Although FISH is applicable to at least 60% of AML cases, the number of nuclei or metaphases available for scoring limits its sensitivity, and interphase FISH is limited by the capacity to distinguish low levels of disease from background. Even with hypermetaphase FISH, which is extremely labor-intensive, sensitivities achieved are inferior to immunophenotype or PCR-based methods and hence attention has been focused on these approaches.

The use of flow cytometry to detect MRD depends on the identification of aberrant phenotypes among the leukemic blast cell population, and this approach is applicable to the majority of AML cases.[18,26] This technique is relatively sensitive (1 in 10^3 to 10^5), although concern has been raised regarding the frequent observation of "phenotypic switching" between diagnostic and relapse samples, which could lead to false-negative results. The occurrence of this phenomenon suggests that this approach to MRD assessment is likely to be most informative relatively early during the disease course, rather than during longitudinal testing after consolidation therapy has been completed. A number of studies have shown that kinetics of disease response as determined by flow cytometry has the potential to identify patients with differing risk of relapse,[18,26] although this approach has yet to be evaluated as the basis for risk-adapted therapy in large-scale AML clinical trials.

The majority of studies evaluating detection of MRD in AML have been based on RT-PCR (reviewed in Reference 24). This approach has been most extensively validated in the context of APL, providing an independent prognostic factor, which is now being used to direct treatment in individual patients. Previous studies have shown that patients in whom *PML-RARA* fusion transcripts are still detectable at the end of consolidation or reappear at a later stage are destined to relapse unless additional therapy is given.[24] Moreover, preliminary evidence from the Italian GIMEMA

group suggests that administration of preemptive therapy at the point of molecular relapse confers an overall survival advantage in comparison to that of patients treated in frank hematologic relapse.[27] MRD monitoring also has been shown to be of value in determining treatment strategy in patients with high-risk APL, as defined by persistent PCR positivity or those beyond first CR. Evidence to date suggests that patients with evidence of MRD, in whom harvested stem cells or bone marrow are also PCR positive, are unlikely to benefit from autologous transplant procedures, due to the high rate of relapse. Nevertheless, if a suitable donor is available, such patients can potentially be cured by allogeneic BMT. MRD assessment is also of value in the posttransplant setting to direct the need for additional therapy. The value of MRD assessment as determined by nested RT-PCR in other forms of AML is less clearly established; however, it seems likely that its role will be more precisely defined following more extensive evaluation of RQ-PCR approaches (see Real-time Quantitative RT-PCR section below).

Available Assays

RT-PCR

Leukemia-associated fusion transcripts are most conveniently detected by RT-PCR, involving initial conversion of RNA to complementary DNA (cDNA), followed by PCR. For molecular screening, a single round of PCR may suffice; however, to optimize specificity and sensitivity, two rounds of PCR using external and internal primer sets may be used (nested RT-PCR). Some laboratories avoid use of nested RT-PCR methods due to concern about increased risk of PCR contamination. Standardized methods for detection of AML-associated fusion genes by RT-PCR, most notably *PML-RARA*, *AML1-ETO*, *CBFB-MYH11,* and *BCR-ABL*, have been developed by the European BIOMED1 group.[28] The majority of studies using nested RT-PCR for MRD monitoring purposes have focused on patients with fusion genes associated with the favorable-risk cytogenetic abnormalities. Maximal sensitivities for nested RT-PCR lie between 1 in 10^4 and 1 in 10^6. Evaluation of primary diagnostic material by RQ-PCR has revealed that, as might be expected, the key determinant of the sensitivity of each respective assay is the level of the fusion gene expression in leukemic blasts in relation to that of endogenous control genes.[29] Indeed, RQ-PCR has revealed that *PML-RARA* is typically expressed at a lower level than *AML1-ETO*, thereby accounting for the difference in sensitivity observed with the respective qualitative RT-PCR assays (1 in 10^4 vs 1 in 10^5).

As discussed above, detection of MRD using conventional nested RT-PCR provides an independent prognostic factor among patients with a defined molecular lesion, in the form of *PML-RARA* in APL, who could not otherwise have been distinguished on the basis of pretreatment characteristics. Therefore, MRD monitoring is now firmly established as a means of shaping risk-adapted therapy in this subset of AML. A limitation to this approach relates to false-negative results, whereby relapses occur after negative RT-PCR tests. This could reflect relative insensitivity of the assay and insufficient frequency of MRD assessment with respect to the kinetics of relapse, but in some instances could be due to poor quality of RNA or relative inefficiency of the RT step.

For the CBF leukemias, the role of RT-PCR for molecular monitoring purposes has been more uncertain (reviewed in Reference 24). This reflects the detection of *AML1-ETO* and *CBFB-MYH11* transcripts in some patients in long-term remission using conventional nested RT-PCR, which is likely to be related to the greater sensitivity of these assays (as compared to those for detection of *PML-RARA*). Furthermore, patients with *CBFB-MYH11*-associated disease are often slow to clear disease-related transcripts, with PCR positivity detected as long as 24 months following achievement of CR, therefore making it difficult to reliably distinguish patients likely to be cured from those destined to relapse. These potential shortcomings of nested RT-PCR have led to considerable interest in quantitative RT-PCR approaches. Indeed, studies using competitor-based assays have shown that patients with t(8;21)-associated leukemia can be distinguished in terms of prognosis on the basis of transcript levels detected in remission, with patients testing negative or with low transcript levels maintaining CR, while those with high or rising transcript numbers progressing to relapse. These competitor-based assays are highly labor-intensive and therefore not readily applicable to evaluation of large numbers of patients in clinical trials. However, over the last few years a variety of semiautomated quantitative PCR approaches have been developed that provide rapid, sensitive, and reproducible methods, combined with a capacity for high throughput. These attributes afford considerable advantages over conventional endpoint nested RT-PCR assays for MRD detection in the context of large-scale clinical trials as well as in individual institutions.

Real-Time Quantitative RT-PCR

For the detection of leukemia-associated transcripts by RQ-PCR, RNA is initially converted to cDNA using RT protocols identical to those involved in conventional RT-PCR (reviewed in Reference 30). However, in contrast to the latter technique, in which two rounds of PCR amplification are routinely performed, RQ-PCR approaches involve only a single round of PCR. The basis of this technology is the measurement of the number of amplicons generated with each PCR cycle in the exponential phase of the reaction, detected as a proportionate rise in fluorescence intensity. PCR products may be detected through use of the fluorescent dye SYBR Green I, which binds to the minor groove of double-stranded DNA, or through specific "hybridization" or "hydrolysis" probes. SYBR Green I is not favored for

MRD detection, given its lack of sensitivity or specificity in comparison to the use of specific probes, which are significantly more expensive.

A key advantage of RQ-PCR approaches is the capacity to measure transcript levels of the fusion gene of interest, relative to the expression of endogenous control transcripts. This enables precise evaluation of kinetics of molecular response to treatment, documentation of rising leukemia-associated transcripts prior to frank relapse, and identification of poor-quality samples or problems with the RT step, which could give rise to false-negative results had conventional endpoint assays been used. Changes in leukemic target transcript level normalized to that of the endogenous control gene may be calculated on the basis of differences in Ct (cycle threshold) values (so-called ΔΔCT method), providing the efficiencies of the PCR reactions are comparable. Alternatively, normalized MRD data may be reported in terms of absolute copy numbers derived from respective plasmid standard curves run in parallel (reviewed in References 29 and 30).

For reliable performance of RQ-PCR, testing of diagnostic material is critical to define breakpoint location, particularly for rearrangements in which variable breakpoints occur at the genomic level (e.g., bcr 1, 2, and 3 for *PML-RARA* and types A, D, and E, which account for the majority of cases with the *CBFB-MYH11* fusion), since RQ-PCR assays are transcript specific. Furthermore, availability of diagnostic material is important so that normalized fusion gene expression levels can be related to the pretreatment level. Extensive collaborative studies have been undertaken by the Europe Against Cancer (EAC) Group to optimize and standardize the methods for detection of leukemia-associated fusion transcripts and to define suitable endogenous control transcripts.[29]

Mutation Screening

There are a variety of methods available for detection of mutations, for example, involving genes encoding tyrosine kinases (e.g., *KIT*) or transcription factors (e.g., *CEPBA*). For small genes, PCR and direct sequencing may provide the most suitable approach. However, for larger genes, or in situations where large numbers of samples are to be tested, initial screening may be conveniently undertaken by methods such as single-strand conformation polymorphism (SSCP) analysis, chemical cleavage of mismatch, or denaturing high-performance liquid chromatography (DHPLC), with sequencing restricted solely to cases with a suspected mutation. Length mutations such as *FLT3* ITDs are readily detected by PCR using genomic or complementary DNA as a template (see Figure 30-7). For muta-

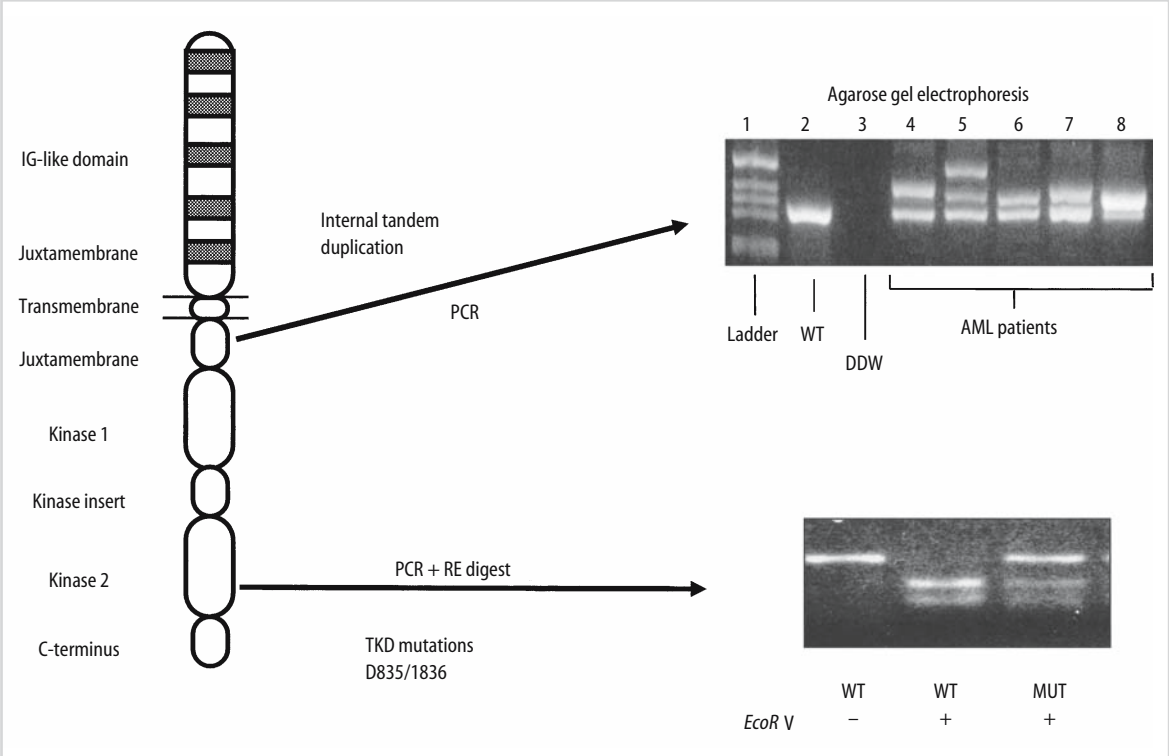


Figure 30-7. Detection of *FLT3* mutations in AML. ITDs affect the juxtamembrane region of the FLT3 receptor, while point mutations (e.g., D835/I836) involve the kinase loop. Both classes of mutation (MUT) lead to constitutive activation of the receptor, activating a number of potential downstream targets. Presence of *FLT3* ITD is revealed by the amplification of a larger-sized band than that associated with wild-type (WT) receptor following PCR of genomic or complementary DNA (see lanes 4–8, upper right panel). In some cases, more than one aberrant band may be observed (e.g., lane 5); this may be due to biallelic *FLT3* mutations in the leukemic clone, with the WT band being due to residual normal marrow elements. However, in some instances it may be due to leukemic subclones harboring different *FLT3* mutations. Tyrosine kinase domain (TKD) mutations may be detected by PCR and restriction enzyme RE digest, with presence of a mutation leading to loss of the *EcoR* V restriction site. DDW, water control. (Figure kindly prepared by Rosemary Gale, University College London.)

tions that target a particular region of the gene, a specific assay may be used involving PCR combined with a restriction enzyme digest, for example, to detect *FLT3* activation loop mutations (e.g., D835), or PCR using allele-specific primers, for example, for detection of *JAK2* mutation in myeloproliferative disorders (see chapter 35). Rapid screening for presence of an underlying *NPM1* mutation may be undertaken by immunohistochemistry using monoclonal antibodies against NPM that are in routine use for diagnosis of NPM-ALK-associated lymphoma, with presence of mutation being indicated by delocalization of NPM to the cytoplasm.[12]

Laboratory Issues

Transit of Samples to the Laboratory

A key issue in the provision of reliable laboratory results relates to sample collection procedures and transit time to the laboratory. Heparin can potentially interfere with PCR, and blood or bone marrow samples are more appropriately collected in EDTA. For assays that involve PCR amplification of genomic DNA, for example, detection of *FLT3* ITD/D835, relatively prolonged transit times (i.e., 3 to 7 days) are unlikely to be particularly problematic. However, for RNA-based assays, quantitative RT-PCR has revealed degradation equivalent to approximately 0.5 log per 24 hours. Appropriate quantitation of fusion transcripts is feasible if endogenous control transcripts with a comparable degradation rate are amplified in parallel. However, it is important to bear in mind that the control transcript will not correct for the loss in sensitivity resulting from prolonged transit time. Hence there is interest in the development of RNA-stabilizing agents that reduce RNA degradation during specimen transit; these have been most extensively evaluated in MRD monitoring in chronic myeloid leukemia.

RT-PCR Screening

An important problem with RT-PCR is the potential for false-positive results due to PCR contamination. To avoid this, stringent controls must be performed and ideally RNA extraction should be performed in a laboratory remote from that in which PCR is performed. Moreover, preparation of RT and PCR master mixes must be physically separated from areas in which PCR products are handled. Less well appreciated is the potential for molecular screening using RT-PCR to yield false-negative results in the presence of rare or atypical breakpoints. For example, molecular screening strategies for *CBFB-MYH11* typically involve detection of type A, D, and E breakpoints, which together account for approximately 90% of cases; hence, it is possible that in rare instances, cases with alternative breakpoints that lack inv(16)/t(16;16) could be missed. Similarly, infrequent cases of *PML-RARA*-associated APL with atypical breakpoints are not detected by some standard RT-PCR primer sets; however, such cases are unlikely to be missed, given the typical morphological features, and will give a positive result (microparticulate pattern) in the PML immunofluorescence test. False-negative results are also a potential problem in the presence of poor-quality RNA or inefficiency of the RT step.

Targets for MRD Detection in Cases Lacking Fusion Genes

A number of potential targets have been proposed for PCR-based approaches to MRD detection for cases of AML lacking a suitable fusion gene marker. These include detection of mutant forms of *NPM1*, which occur in over half of normal karyotype AML. Preliminary evidence using RQ-PCR assays indicates, that *NPM1* mutation is likely to provide a sensitive, specific and stable target for MRD detection.[31,32] This contrasts with *FLT3* mutations which are not always stable over the time course of the disease, raising concern over the value of these as reliable targets for MRD detection.[33] Nevertheless, these findings provide further evidence for the hypothesis that *FLT3* mutations are secondary events in the pathogenesis of AML. There is considerable interest in whether RNA-based assays that involve detection of transcripts that are overexpressed in leukemia, for example, *WT1,* could be of value as MRD markers in patients lacking a fusion gene target. However, to establish the role of such markers it is critical to determine the level of expression at diagnosis relative to endogenous control transcripts, for example, *ABL,* on an individual patient basis. Other key factors that must be taken into account are the range of expression levels documented in normal peripheral blood and bone marrow, as well as levels of expression in the context of regeneration following myelosuppressive therapy. Such information is vital to establish the sensitivity of the assay in any given patient and the likely clinical utility of any novel marker for MRD assessment.

Quality Assurance and Reporting of RQ-PCR Data

An important issue with regard to the provision of reliable results suitable for making treatment decisions is the adherence to rigorous internal quality control (QC) and participation in external quality assurance (EQA) programs. Such programs have been long established for morphological assessment, immunophenotyping, and cytogenetic analysis. However, they remain in their infancy for molecular detection of fusion genes. Since test methods are based on RT-PCR, it is important that QC schemes evaluate all steps of the procedure, particularly as the RT step is a considerable source of variability. For acute leukemias, distribution of primary patient material is unrealistic, whereas distribution of cell lines for QC purposes suffers from the problem that the heterogeneity encountered in

breakpoint patterns for patients with *PML-RARA* and *CBFB-MYH11* gene fusions, for example, is not represented among the available cell lines. Nevertheless, a number of European initiatives including the BIOMED-1, Europe Against Cancer (EAC) and European LeukemiaNet programs have been launched over the last few years to improve standardization of methods for the detection of leukemia-associated fusion transcripts by conventional and real-time RT-PCR.[28,29] Such initiatives are important to ensure reliability of diagnostic molecular screening and are absolutely fundamental if MRD data are to be used to predict outcome and modify treatment approach.

The EAC QC exercises have highlighted variations in performance between laboratories that are ostensibly carrying out RQ-PCR assays according to a standardized protocol. This reinforces the importance of developing and adhering to standard operating procedures for performance of RQ-PCR tests and generating clear guidelines for reporting of RQ-PCR results. In particular, each assay should be performed in triplicate, so that the degree of reproducibility of individual results may be readily appreciated and hence "outlying results" disgarded. The EAC program revealed that in some instances, fusion gene assays yield an amplification signal in one of the triplicate wells of negative control samples. Since in many instances the participating laboratories had not previously amplified the transcript in question, it is likely that such results reflect nonspecific amplification or cDNA carryover between adjacent wells. This is an intrinsic shortcoming of platforms using the 96-well plate format and has led us to report MRD analyses in which amplification is restricted to one of three wells as PCR negative (but recommending earlier repeat testing, should the result reflect possible residual disease at the limits of assay detection). To report assays as positive, we require specific amplification in at least two of three wells, with Ct values ≤40 (i.e., equivalent to at least one plasmid copy).[29] Since PCR amplification of primary patient samples is less efficient than plasmid controls, reproducible amplification at Ct values between 40 and 45 can occasionally be observed in follow-up patient samples, which most likely indicates low-level MRD; such results are of uncertain clinical significance and again would be seen as an indication for earlier repeat MRD assessment.

Future Directions

Research developments in terms of our understanding of the biology of AML and the clinical evaluation of novel targeted therapeutic approaches are clearly going to have an impact on the range of assays that will need to be developed and offered by molecular pathology laboratories in the future. There is considerable interest in determining whether gene expression profiling using microarrays[34–36] will yield further insights into the etiology of AML and lead to further refinement of the classification of the disease, enabling development of more rational treatment approaches. Large-scale gene expression profiling efforts may yield information that reliably distinguishes subgroups of patients with distinct prognoses and potentially identifies those who are likely to be highly sensitive or resistant to particular therapeutic agents. If these expectations are realized it could have major implications for hematology laboratories, leading to investigation of expression profiles of a more limited number of genes using Gene Chips and/or RQ-PCR as a key component of the diagnostic workup. Further information regarding individual variation in disease response according to polymorphisms in genes encoding components of drug-metabolizing pathways could see more widespread introduction of single nucleotide polymorphism (SNP) assays. Such approaches, in conjunction with flow cytometry, could be used to define resistance phenotypes, leading to the use of resistance modulators in suitable patients. Development of appropriate laboratory assays also is essential to determine whether novel therapeutic approaches designed, for example, to modulate drug resistance pathways or mediate gene reactivation through targeting histone deacetylases (HDACs) or DNA methyltransferases are indeed having their desired effect on the relevant pathways in vivo.

In addition to the expansion of the repertoire of assays used to evaluate newly diagnosed patients with AML, the next few years are likely to see more widespread use of MRD assessment to determine treatment approaches for patients with the disease. For patients with fusion gene transcripts, RQ-PCR is clearly the best approach, and international efforts are under way to establish optimal time points for MRD sampling and thresholds of fusion gene expression that are predictive of long-term outcome. A key challenge is to determine the most appropriate way to monitor patients who lack a suitable fusion gene marker, whether by RQ-PCR detection of genes that are typically overexpressed in AML or by use of flow cytometry. It is clear that the pathology laboratory is set to play an ever-increasingly important role in optimizing and individualizing the management of patients with AML.

Acknowledgments

I am grateful to Georgina Harrison for analyses relating to patients entered into the Medical Research Council AML trials. I also would like to thank Dr Barbara Bain, Dr Lyndal Kearney, Dr Monica Pallis, and Dr David Swirsky for critical reading of this manuscript and the Leukaemia Research Fund of Great Britain and European LeukemiaNet for grant support.

References

1. Grimwade D. The clinical significance of cytogenetic abnormalities in acute myeloid leukaemia. *Best Pract Res Clin Haematol.* 2001;14:497–529.

2. Kelly LM, Gilliland DG. Genetics of myeloid leukemias. *Annu Rev Genomics Hum Genet.* 2002;3:179–198.
3. Tenen DG. Disruption of differentiation in human cancer: AML shows the way. *Nat Rev Cancer.* 2003;3:89–101.
4. Mueller BU, Pabst T, Osato M, et al. Heterozygous PU.1 mutations are associated with acute myeloid leukemia. *Blood.* 2002;100:998–1007.
5. Kottaridis PD, Gale RE, Linch DC. Flt3 mutations and leukaemia. *Br J Haematol.* 2003;122:523–538.
6. Spiekermann K, Bagrintseva K, Schoch C, Haferlach T, Hiddemann W, Schnittger S. A new and recurrent activating length mutation in exon 20 of the FLT3 gene in acute myeloid leukemia. *Blood.* 2002;100:3423–3425.
7. Bain BJ. *Leukaemia Diagnosis.* 3rd ed. Oxford: Blackwell Publishing; 2003.
8. Preudhomme C, Sagot C, Boissel N, et al. Favorable prognostic significance of *CEBPA* mutations in patients with de novo acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA). *Blood.* 2002;100:2717–2723.
9. Wechsler J, Greene M, McDevitt MA, et al. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. *Nat Genet.* 2002;32:148–152.
10. Tartaglia M, Niemeyer CM, Fragale A, et al. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nat Genet.* 2003;34:148–150.
11. Loh ML, Reynolds MG, Vattikuti S, et al. *PTPN11* mutations in pediatric patients with acute myeloid leukemia: results from the Children's Cancer Group. *Leukemia.* 2004;18:1831–1834.
12. Falini B, Mecucci C, Tiacci E, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N Engl J Med.* 2005;352:254–266.
13. Döhner K, Schlenk RF, Habdank M, et al. Mutant *nucleophosmin* (*NPM1*) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics—interaction with other gene mutations. *Blood.* 2005;106:3740–3746.
14. Schnittger S, Schoch C, Kern W, et al. Nucleophosmin gene mutations are predictors of favorable prognosis in acute myelogenous leukemia with a normal karyotype. *Blood.* 2005;106:3733–3739.
15. Jaffe ES, Harris NL, Stein H, Vardiman JW, eds. *World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon, France: IARC Press; 2001.
16. Grimwade D, Biondi A, Mozziconacci M-J, et al. Characterization of acute promyelocytic leukemia cases lacking the classical t(15;17): results of the European working party. *Blood.* 2000;96:1297–1308.
17. Grimwade D. Screening for core binding factor gene rearrangements in acute myeloid leukemia. *Leukemia.* 2002;16:964–969.
18. San Miguel JF, Vidriales MB, Orfão A. Immunological evaluation of minimal residual disease (MRD) in acute myeloid leukaemia (AML). *Best Pract Res Clin Haematol.* 2002;15:105–118.
19. Burnett AK, Grimwade D, Solomon E, Wheatley K, Goldstone AH. Presenting white blood cell count and kinetics of molecular remission predict prognosis in acute promyelocytic leukemia treated with all-*trans* retinoic acid: result of the randomized MRC trial. *Blood.* 1999;93:4131–4143.
20. Gale RE, Hills R, Kottaridis PD, et al. No evidence that FLT3 status should be considered as an indicator for transplantation in acute myeloid leukemia (AML): an analysis of 1135 patients excluding acute promyelocytic leukemia from the UK MRC AML10 and 12 trials. *Blood.* 2005:106:3658–3665.
21. Barjesteh van Waalwijk van Doorn-Khosravani S, Erpelinck C, van Putten WLJ, et al. High *EVI1* expression predicts poor survival in acute myeloid leukemia: a study of 319 de novo AML patients. *Blood.* 2003;101:837–845.
22. Bowen DT, Frew ME, Hills R, et al. RAS mutation in acute myeloid leukemia is associated with distinct cytogenetic subgroups but does not influence outcome in patients <60 years. *Blood.* 2005:106:2113–2119.
23. Fröhling S, Schlenk RF, Stolze I, et al. *CEBPA* mutations in younger adults with acute myeloid leukemia and normal cytogenetics: prognostic relevance and analysis of cooperating mutations. *J Clin Oncol.* 2004;22:624–633.
24. Liu Yin JA, Grimwade D. Minimal residual disease evaluation in acute myeloid leukaemia. *Lancet.* 2002;360:160–162.
25. Cheson BD, Bennett JM, Kopecky KJ, et al. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. *J Clin Oncol.* 2003;21:4642–4649.
26. Kern W, Voskova D, Schoch C, Hiddemann W, Schnittger S, Haferlach T. Determination of relapse risk based on assessment of minimal residual disease during complete remission by multiparameter flow cytometry in unselected patients with acute myeloid leukemia. *Blood.* 2004;104:3078–3085.
27. Grimwade D, Lo Coco F. Acute promyelocytic leukemia: a model for the role of molecular diagnosis and residual disease monitoring in directing treatment approach in acute myeloid leukemia. *Leukemia.* 2002;16:1959–1973.
28. van Dongen JJM, Macintyre EA, Gabert JA, et al. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. *Leukemia.* 1999;13:1901–1928.
29. Gabert JA, Beillard E, van der Velden VHJ, et al. Standardization and quality control studies of "real-time" quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia—a Europe Against Cancer program. *Leukemia.* 2003;17:2318–2357.
30. Flora R, Grimwade D. Real-time quantitative RT-PCR to detect fusion gene transcripts associated with AML. *Methods Mol Med.* 2004;91:151–173.
31. Gorello P, Cazzaniga G, Alberti F, et al. Quantitative assessment of minimal residual disease in acute myeloid leukemia carrying *nucleophosmin* (*NPMI*) gene mutation. *Leukemia.* 2006;20:1103–1108.
32. Chou WC, Tang TL, Lin LI, et al. *Nucleophosmin* mutation in *de novo* acute myeloid leukemia: the age-dependent incidences and the stability during disease evolution. *Cancer Res.* 2006;66:3310–3316.
33. Kottaridis PD, Gale RE, Langabeer SE, Frew ME, Bowen DT, Linch DC. Studies of FLT3 mutations in paired presentation and relapse samples from patients with acute myeloid leukemia: implications for the role of FLT3 mutations in leukemogenesis, minimal residual disease detection, and possible therapy with FLT3 inhibitors. *Blood.* 2002;100:2393–2398.
34. Bullinger L, Döhner K, Bair E, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med.* 2004;350:1605–1616.
35. Valk PJM, Verhaak RGW, Beijen MA, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med.* 2004;350:1617–1628.
36. Haferlach T, Kohlmann A, Schnittger S, et al. Global approach to the diagnosis of leukemia using gene expression profiling. *Blood.* 2005;106:1189–1198.

Chapter 31

Acute Lymphoblastic Leukemia

Carlo Alberto Scrideli, Giovanni Cazzaniga, and Andrea Biondi

Acute lymphoblastic leukemia (ALL) is a heterogeneous group of disorders that originates from B- and T-cell progenitors.[1,2] Different B- and T-cell ALL can be recognized according to immunologic and molecular criteria.[3–5] The identification of the molecular events underlying the process of leukemia transformation has provided not only important biological information,[5–7] but also clinically relevant genetic markers for the identification of prognostically relevant ALL subgroups and for the molecular monitoring of minimal residual disease (MRD). For ALL, immunoglobulin (*IG*) and T-cell receptor (*TCR*) gene rearrangement studies are used as markers of clonality and for MRD detection, and the identification of different genetic variations is used to define different ALL subgroups.

Molecular Basis of Disease

IG and *TCR* Gene Rearrangements in ALL

The immune system can specifically recognize and distinguish an enormous variety of specific antigens and antigenic epitopes (at least 10^{10}). Each lymphocyte has a unique membrane receptor, the immunoglobulin receptor (IG) on B lymphocytes, and the T-cell receptor (TCR) on T lymphocytes, that recognizes specific antigens.

Rearrangements of gene segments that encode the variable regions of the *IG* and *TCR* genes are responsible for the enormous diversity of antigen-specific lymphocyte receptors.[8–12] This great variability of *IG* and *TCR* is mediated by several mechanisms: (1) the number of variable or V, diversity or D, and joining or J segments (germline genome diversity); (2) the number of possible V(-D)-J combinations (combinatory diversity); (3) deletions of nucleotides at the ends of rearranging V, (D), and J gene segments, as well as random insertion of noncoded nucleotides (N-region nucleotides) between V-D and D-J segments (junctional diversity); and (4) somatic mutations in the V gene segments of the *IG* genes. This last process has not been observed for *TCR* genes.[10,12,13]

The V, (D), and J segments of B and T precursor cells are rearranged during their differentiation, and each lymphocyte, in this way, contains a specific combination of V-(D-) and J segments.[10,14] Additional variability can be generated in the junctions V-D and D-J through the loss of nucleotides, as well as insertion of new nucleotides. Insertion of nucleotides (insertion of N regions) in the junctional sites is mediated by the enzyme terminal deoxynucleotidyl transferase (TdT), which is present in the nucleus of immature B and T cells. N-region insertion occurs in rearranged immunoglobulin heavy chain (*IGH*) genes, but is limited or absent in the immunoglobulin kappa and lambda light chain genes (*IGLK* and *IGLL*). Junctional diversity also occurs in *TCR* genes, primarily the T-cell receptor delta and gamma genes (*TCRD* and *TCRG*). The potential diversity introduced by the addition of N regions is very high and can increase the total combinatorial diversity to more than 10^{10} for the IGH molecules, 10^{12} for *TCRAB,* and more than 10^{9} for *TCRDG.*[10,12]

During B-cell differentiation, the *IGH* genes rearrange before the *IGL* genes, and of the *IGL* genes, *IGLK* is the first to rearrange; if the *IGLK* rearrangement is not functional, the *IGLL* gene will rearrange. Another type of rearrangement of the *IGH* gene that can occur is the process of class switching, from IgM and IgD to IgG, IgA, or IgE.[10,12]

During the differentiation of T cells, the *TCRD* gene is the first to rearrange, followed immediately by rearrangement of the *TCRG* gene and in many cells also the rearrangement of the *TCRB* gene. The *TCRD* gene can be deleted when *TCRA* rearrangement occurs.[10,12,15] Since ALL is derived from a single transformed lymphoid precursor cell, all ALL cells of a patient in principle have the same *IG* and *TCR* gene rearrangements with identical junctional region sequences that can be regarded as leukemia-specific fingerprints. Such targets can be identified at initial diagnosis in more than 95% of patients with ALL by using various polymerase chain reaction (PCR) primer sets.[16–20]

Studies of rearrangements of the *IG* genes in B-lineage ALL by Southern blot and PCR have shown that, despite an immature phenotype, more than 95% of B-precursor ALL

contain *IGH* rearrangements, approximately 60% have *IGLK* rearrangements or deletions, and approximately 20% have *IGLL* rearrangements.[12,16,17,19] Cross-lineage rearrangements of the *TCR* gene also have been observed in approximately 90% of B-lineage ALL pediatric patients.[20] Rearrangements of *TCRA*, *TCRB*, *TCRD*, and *TCRG* are found in 46% to 61%, 29% to 35%, 54% to 80%, and 49% to 70% of B-lineage ALL, respectively.[21–26] In B-precursor ALL, approximately 80% of all *TCRD* rearrangements are due to incomplete gene rearrangements *Vδ2–Dδ3* and *Dδ2–Dδ3*.[12,15,22] These *Vδ2–Dδ3* rearrangements also are prone to continuing rearrangements, particularly with *Jα* gene segments. *Vδ2-Jα* rearrangements are detected in approximately 40% of childhood B-precursor ALL and are rare or absent in normal lymphoid cells.[27]

In T-lineage ALL, *TCRA*, *TCRB*, *TCRD*, and *TCRG* gene rearrangements have been described in 17% to 67%, 85% to 89%, 68% to 90%, and 90% to 100% of cases, respectively.[12,19,23,24] Cross-lineage *IG* gene rearrangements are not very common in T-lineage ALL, occurring in less than 20% of cases, and involving only *IGH* rearrangements.[12,19,23,24]

The explanation for these cross-lineage rearrangements of the *IG* and *TCR* genes in non-B and non-T cells, respectively, is probably due to the fact that the leukemic cells originate from a less-committed lymphoid precursor or even a more-primitive pluripotential cell (lymphoid/myeloid), where both *IG* and *TCR* genes are accessible to a common recombinase enzyme. Such rearrangements occur before the final determination of lineage; as soon as this determination for the B, T, or myeloid lineage occurs, the cell adopts the lineage-specific phenotype with expression or differentiation of the lineage-specific antigen, while the gene of cross lineage persists as an "artifact." Due to this, rearrangements in the *TCR* gene are relatively common in B-precursor cell ALL; however, they are very rare in B-ALL and mature B-cell neoplasia (B-cell chronic lymphocytic leukemia [B-CLL], multiple myeloma, and B-cell non-Hodgkin's lymphoma [B-NHL]).[19,20] Similarly, rearrangements of *IG* genes are rare in mature T-cell neoplasms (T-cell chronic lymphocytic Leukemia [T-CLL] and T-cell non-Hodgkin's lymphoma [T-NHL]).[19,20]

Chromosomal Translocations Occurring in ALL

In recent years the genes involved in the most frequent leukemia chromosomal translocations have been identified, providing important insights into disease pathogenesis and normal cellular physiology.[7,8] Molecular assays for leukemia cells have been developed[24,28,29] and now permit more accurate diagnosis of leukemia subtypes with frequently recurring translocations. Somatically acquired chromosomal translocations or inversions have been found in up to 65% of the acute leukemias. Chromosomal translocations have two main consequences: activation of a protooncogene brought under the regulation of the promoter and enhancer elements of another gene that normally is active in the leukemia progenitor cell, and creation of a fusion gene from two distinct genes with expression of a chimeric protein (Figure 31-1). The products of the latter aberrant genes are most often transcription factors or tyrosine kinases.[30] The most common translocations occurring in ALL are shown in Table 31-1.

Abnormalities of the q23 region of chromosome 11 are seen with relative frequency in childhood ALL (5%), infant acute leukemia (70%), and secondary leukemias in patients who received topoisomerase II inhibitors (85%). Leukemias with translocations involving 11q23 have an early pre-B phenotype and express myeloid antigens. In the majority of cases, the translocation partner is chromosome 4, and less frequently chromosome 1, 10, or 19. At the molecular level, the t(4;11) fuses *MLL* on chromosome 11 band q23 to *AF4* on chromosome 4 band q21, resulting in an *MLL/AF4* chimeric gene. This translocation is

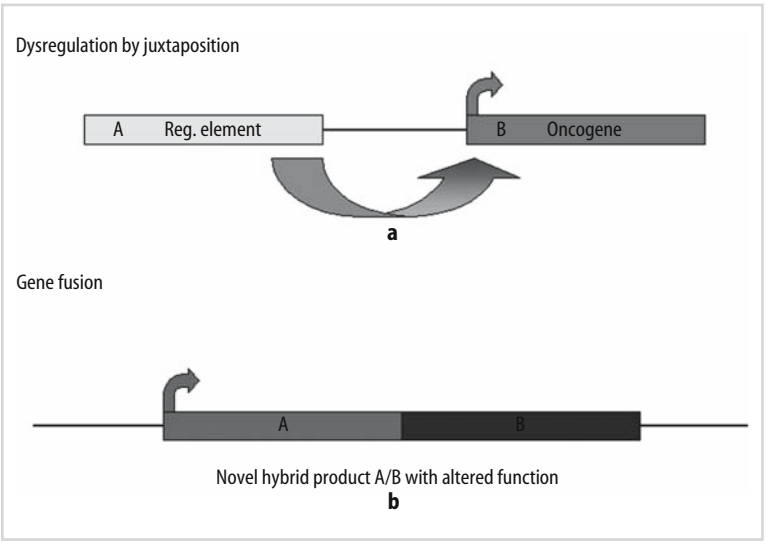


Figure 31-1. Schematic representation of the two types of gene alterations generated by chromosomal translocations. (a) A protooncogene is activated by the promoter and enhancer elements of a distinct translocation gene partner. (b) Discrete segments of two different genes are joined as a result of a translocation, creating a fusion gene encoding a chimeric protein.

Table 31-1. Common Genetic Lesions in ALL

Subgroup	Molecular Alteration	Leukemia Subtype
B-Cell Precursor ALL		
t (12;21)(p13;q22)	*ETV6/AML1*	pro-B
t(1;19)(q23;p13)	*E2A/PBX1*	pre-B
t(9;22)(q34;q11)	*BCR/ABL* (p190 and p210)	pro-B and pre-B
t(4;11)(q21;q23)	*MLL/AF4*	CD10 pro-B cell
t(8;14)(q24;q32)	*MYC* dysregulation	B-ALL
T-ALL		
1p32 deletion	*SIL/TAL1*	childhood T-ALL

related to aggressive clinical features such as hyperleukocytosis, organomegaly, frequent central nervous system involvement, and poor outcome.[31,32]

The t(1;19)(q23;p13) translocation is one of the most frequently observed translocations in childhood ALL, occurring in approximately 25% of the pre-B cytoplasmic (Cy) immunoglobulin (IG)-positive cases and in 1% of the early pre-B Cy IG-negative cases. This translocation most frequently creates a fusion of *E2A*, a helix-loop-helix (HLH) protein-coding gene on chromosome 19, with *PBX1*, a homeobox-containing gene from chromosome 1.[33] A few molecular variants of the t(1;19), leading to different species of *E2A/PBX1* mRNAs, have been described.[34] Although this translocation has been associated with a poor clinical outcome, the current chemotherapy protocols have nullified its adverse prognostic impact.

The t(9;22) translocation is identified in about 3% to 5% of children with ALL and about 25% of adults with ALL.[35] This translocation creates a novel chromosome (the Philadelphia chromosome, Ph^+) and creates a fusion gene between the protooncogene *ABL* on chromosome 9 and the "breakpoint cluster region" gene (*BCR*) on chromosome 22. While the *ABL* breakpoint on chromosome 9 occurs in the same region (between exons a1 and a2), the *BCR* breakpoints on chromosome 22 can occur in two different positions. The "minor" breakpoint cluster region (m-bcr), between exons e1 and e2, is present in about 90% of childhood Ph^+ ALL; by contrast, a "major" breakpoint cluster region (M-bcr), between exons b2 and b3 or b3 and b4, is a usual finding in chronic myelogenous leukemia (CML)[36] and occasionally in Ph^+ ALL. The m-bcr generates a fusion protein of 190 kDa (p190), whereas the M-bcr results in a fusion protein of 210 kDa (p210).[35–37]

The presence of t(9;22) is associated with a high risk of treatment failure in children and adults with ALL. However, recent retrospective studies showed that Ph^+ childhood ALL is a heterogeneous disease with regard to treatment response. High leukocyte count, old age, and poor response to the prephase treatment with prednisone and intrathecal methotrexate influence treatment outcome. Patients with a good prednisone response (PGR) have a significantly lower risk of treatment failure compared to those with a poor prednisone response (PPR), when treated with intensive Berlin-Frankfurt-Münster (BFM) chemotherapy, whether associated with bone marrow transplantation (BMT) or not. These PPR children have a clinical response and prognosis as poor as those of Ph^+ ALL adults.[35,38]

The chromosomal translocation t(12;21)(p13;q22) results in the *ETV6(TEL)/AML1* fusion gene. This is a cryptic translocation, which cannot be detected by cytogenetic analysis; thus, fluorescence in situ hybridization (FISH) or PCR must be used for detection. It is the most common chromosomal rearrangement in childhood ALL, occurring in about 25% of cases;[39] by contrast, it is rarely found in adult ALL (1–3%).[40] The t(12;21) translocation is almost exclusively associated with a young age (<10 years) and a precursor B-cell phenotype.[39]

Whether t(12;21) is associated with a good prognosis or with only a marginally better outcome compared to other cytogenetic ALL subgroups remains controversial.[39] *ETV6/AML1* rearrangement has been observed in about 25% of relapsing cases treated within European BFM-based protocols.[41] The duration of remission appears to be longer (>2 years) in patients carrying this translocation compared to other childhood ALL cytogenetic groups; however, early relapses also have been observed.[32]

The t(8;14)(q24;q32), involving *MYC* on 8q24 and *IGH* on 14q32, occurs in 85% to 90% of cases of surface(s) IG-positive B-cell ALL. Due to the great variability of the breakpoint regions, a standard PCR assay at the DNA level is not sufficient for the detection of this chromosomal translocation.[32] The availability of new and more efficient DNA polymerases that allow the amplification of genomic fragments many kilobase pairs long makes it now possible to identify the t(8;14) by long-distance polymerase chain reaction (LD-PCR).[42] However, although of potential interest for MRD monitoring, no data are available on its clinical application in childhood ALL.

A DNA deletion between *SIL* and *TAL1/SCL* at the chromosomal region 1p32 (*TAL* deletion) occurs in 20% to 25% of childhood T-ALL.[32] This rearrangement can be detected at the DNA level because the breakpoints are clustered within a restricted region, within a few hundred base pairs. Marked heterogeneity occurs at the junctional region, due to random deletion and insertion of nucleotides mediated by illegitimate V(D)J recombination.[43] The *TAL* deletion leads to expression of a *SIL-TAL* chimeric transcript, resulting from the fusion of the 5′ end of *SIL* to the 3′ end of *TAL1*.

Indications for Testing

Many of these genetic alterations have important prognostic implications that can guide the selection of therapy. Treatment of the acute leukemias has progressed from uniform strategies devised for large groups of patients to more-refined protocols tailored to the risk of relapse in discrete subgroups.[1,2] Although routinely recorded features, such as the blast cell immunophenotype and the

presenting white blood cell count, provide useful criteria for risk assessment, molecular genetic changes appear to offer the most sensitive markers of potential leukemia cell aggressiveness and hence are the best guides to treatment. *IG* and *TCR* gene rearrangement studies are useful both for the diagnosis of childhood ALL and for MRD assessment because these rearrangements occur in the vast majority of ALL patients.[13–17,44–46]

Molecular detection of chromosomal abnormalities in blast cells of leukemia patients, as described above, have important prognostic implications that can guide staging and selection of treatment for these patients. In addition, detection of specific translocations is used to assess MRD in leukemia patients.[47] The term "minimal residual disease" (MRD) has been used to define the lowest level of disease detected in patients in complete continuous remission by conventional methods of analyses.

At clinical presentation, the number of leukemic cells is approximately 10^{11} to 10^{12}; if the patient is not treated, the clone continues to expand, and death occurs with approximately 10^{13} total leukemic cells. With cytotoxic treatment, the number of neoplastic cells is diminished and when less than 5% of leukemia blasts are identified in the bone marrow by conventional cytology, the disease is considered in complete remission. Thus, patients considered in complete remission have from zero to 10^{10} leukemic cells. The detection of MRD when clinical remission is achieved or during the treatment can help direct adjustments to therapeutic strategies and identify patients with a higher risk of relapse.[44–46,48,49] In patients undergoing allogeneic hematopoietic stem cell transplantation (HSCT), detection of MRD may lead to tapering of immunosuppressive treatment[50,51] to prevent clinical relapse. MRD detection is also used to assess the effectiveness of purging procedures in the context of autologous HSC rescue.[52]

Available Assays

PCR Detection of Clonality

Monoclonality in ALL is detected by PCR amplification of the *IG* and/or *TCR* VDJ region using consensus primers most often to conserved sequences within the framework region of the V region and the D and J regions (Figure 31-2). The VDJ region varies in size and sequence across a population of B or T cells and allows for detection of a monoclonal leukemic cell population by PCR amplification of a unique PCR product specific to the leukemic clone. For a population of nonlymphoid cells, where the *IG* and *TCR* genes are in the germline configuration, amplification will not occur, due to the great distance between the closest V and J regions.[10,12,17,20]

RT-PCR for Leukemia-Specific Translocations

PCR analysis of fusion genes is based on the design of oligonucleotide primers within exons at the opposite sides of the breakpoint fusion regions, so that the PCR product contains the tumor-specific fusion sequence. The precise breakpoint recombination site at the DNA level is specific for each patient and difficult to determine. In the majority of the ALL translocations, the breakpoint regions span greater than 2 kilobases (kb; limit of amplification by conventional PCR), making DNA PCR difficult. However, since the breakpoints are mostly intronic, and spliced out at the mRNA level, the preferred target for detection of translocations is the chimeric messenger RNA (mRNA).[43]

Reverse transcription-PCR (RT-PCR) requires extraction of total RNA or mRNA from mononuclear cells, reverse transcription of RNA into cDNA, and then PCR, followed by electrophoresis. The sensitivity of the method is specific for each target and can be assessed by amplification of serial dilutions of RNA from diagnostic specimens or cell lines containing the specific translocation into RNA from healthy individuals that do not have the translocation. The presence of a very small number of abnormal cells, in the range of 1 in 10^5 to 1 in 10^6, has been consistently detected using appropriate conditions.[53] This sensitivity is approximately 1000 to 10,000 times more than Southern blot analysis. A PCR test that is sufficiently sensitive to detect one leukemia cell in 10^2 to 10^3 normal cells is acceptable for diagnostic testing. A higher sensitivity is required for MRD assessment during follow-up analyses and should achieve a sensitivity of one leukemia cell in 10^4 to 10^5 normal cells.[18] More than one primer set may need to be used to detect all fusion transcripts when the translocation breakpoints can be in different introns of one or both of the fusion genes, as shown in Figure 31-3.

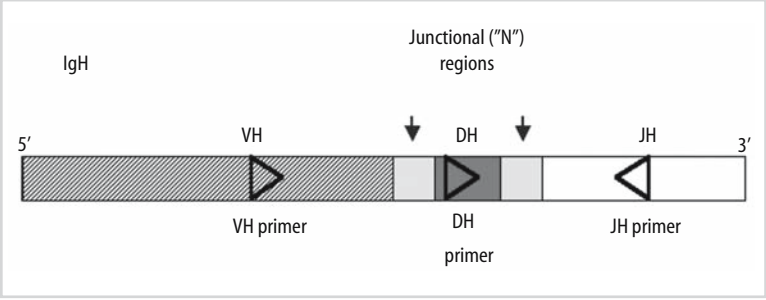


Figure 31-2. Schematic representation of an *IGH* VDJ recombination region and the strategy used to amplify the N regions. Probes for the N regions can be used as patient-specific probes (shaded boxes with arrows above = N regions).

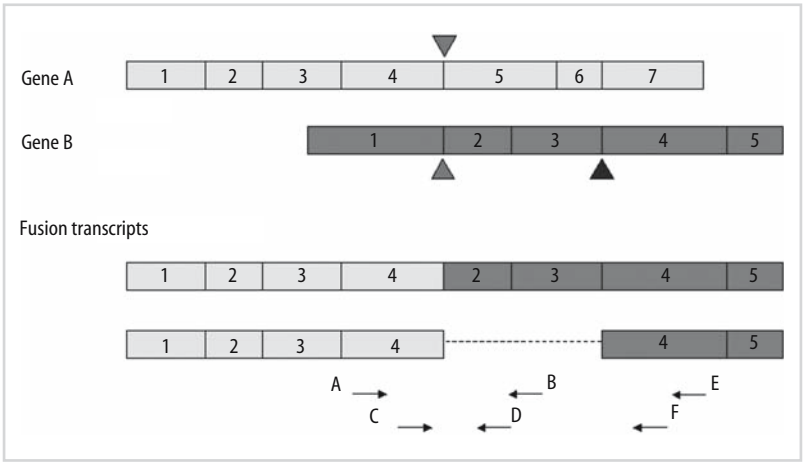


Figure 31-3. RT-PCR detection of different fusion transcripts generated by the same translocation. The presence of different fusion transcripts requires the use of multiple primer pairs to detect fusion transcripts formed by breakpoints in different introns of the genes, depending on the size of the transcript.

Quantitative MRD Analysis

In the past, most PCR-based MRD tests used semiquantitative methods for the detection of clone-specific translocations or *IG* or *TCR* gene rearrangements.[44–46] Standard PCR techniques have the ability to amplify target DNA to a plateau of amplification, so that after 35 to 40 cycles it is not possible to precisely define the initial amount of target DNA. Semiquantitative methods, such as dot-blot hybridization using a patient-specific VDJ region probe, competitive PCR, and limiting dilution PCR, are similarly based on post-PCR endpoint analysis.[54] These techniques require serial dilutions and the analysis of multiple replicates, both of which introduce variability and cost, and are too difficult and too time-consuming to be performed routinely in the clinical laboratory. Real-time quantitative PCR technology (RQ-PCR) circumvents these problems and allows for quantitative assessment of residual disease. The amount of DNA can be normalized to quantitative amplification of a control gene.[54]

MRD Quantification of *IG* and *TCR* Gene Rearrangements

RQ-PCR of *IG* and *TCR* gene rearrangements can be used to quantify MRD levels by using allele-specific oligonucleotide (ASO) probes. Sensitivities of 1×10^{-3} to 1×10^{-5}, comparable to results with the dot-blot method (Figure 31-4),[55–61] are achievable with this strategy. Although initial assays used an ASO fluorescent probe to the junctional region, a more-useful approach is to use a fluorescent probe complementary to the germline *IGH* and *TCR* gene segments, in combination with an ASO primer complementary to the junctional region.[56,57] The ASO primer approach theoretically results in more-sensitive MRD detection compared with use of germline primers, because no competition can occur with the amplification of similar rearrangements in normal cells. Although specific amplification can be easily distinguished from incidental nonspecific amplification, conditions with higher stringency of amplification may need to be used to overcome nonspecific amplification while maintaining the efficiency of the method.[54]

MRD Quantification of Leukemic-Specific Translocations

Numerous publications have demonstrated the feasibility of the RQ-PCR approach to quantify chimeric transcripts resulting from chromosomal translocations occurring in ALL.[47,61–63] Although the principles of RQ-PCR are the same whether DNA or RNA is being analyzed, the RT step represents a major assay variable for accuracy of quantification and sensitivity when RNA is used. In fact, it is necessary to correct variations linked to differences in the RNA amount taken for the reaction or, more importantly, in efficiency (or inhibition) during reverse transcription. For this reason, the number of target gene copies has to be normalized using a ubiquitously and constantly expressed housekeeping gene as a reference (e.g., *ABL*, *B2M*, and *PBGD*).[62–65] Thus, the number of chimeric transcripts will be expressed according to the number of copies of the reference gene transcripts.

Interpretation of Test Results

PCR for *IG* and *TCR* Gene Rearrangements

Leukemia-specific *IG* and *TCR* gene rearrangements identified at diagnosis can be used for MRD assessment of bone marrow (BM) or peripheral blood (PB) samples during and following treatment. However, similar *IG* and *TCR* gene rearrangements from normal lymphocytes in these specimens also are amplified. To discriminate between the leukemia-derived PCR products and PCR products of normal cells with comparable rearrangements, the amplified bands are subjected either to fingerprint[66–68] or homo-heteroduplex analysis.[69] Fingerprint analysis

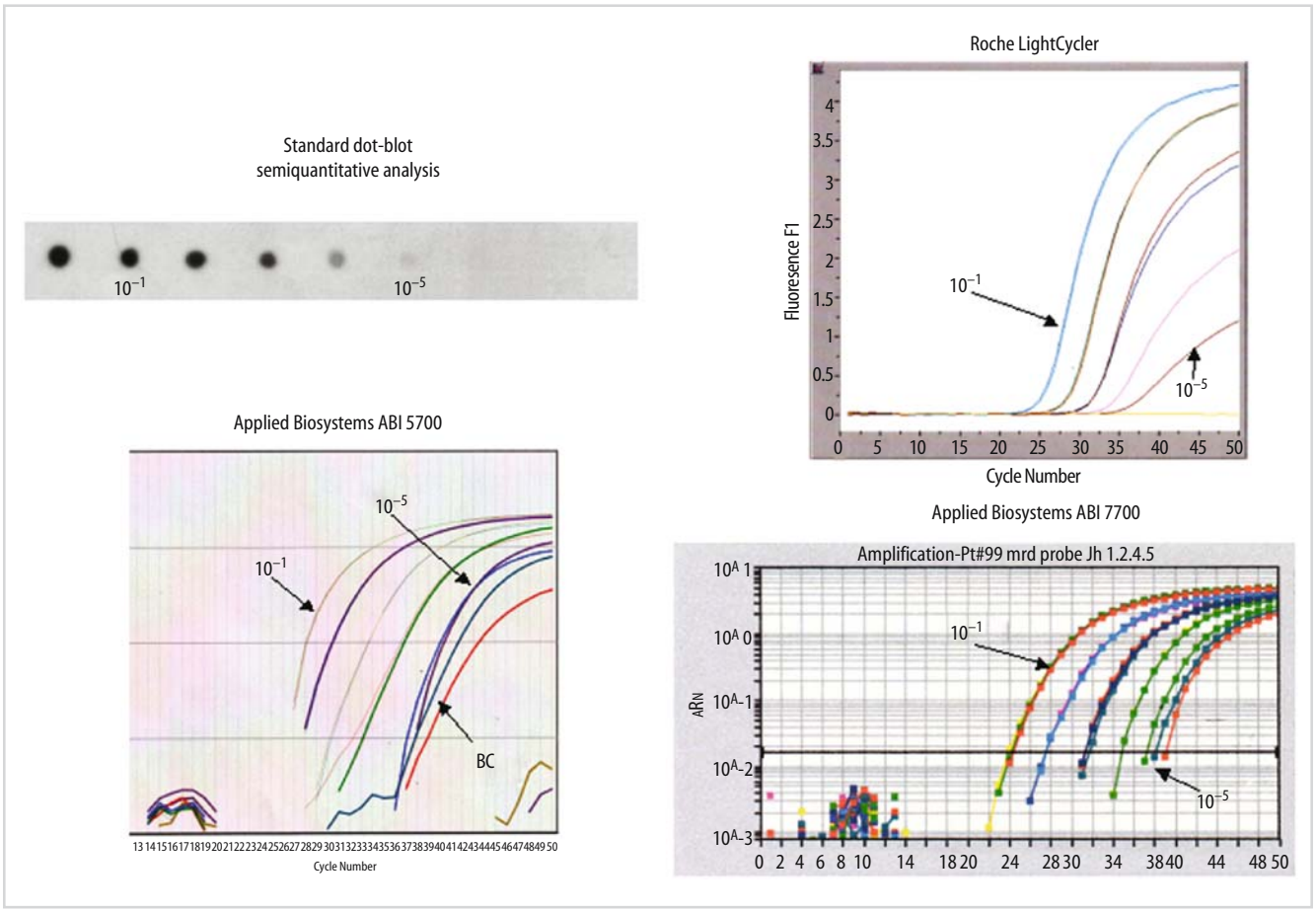


Figure 31-4. Assessment of the sensitivity of different methods for patient-specific *IGH* gene rearrangement assays.

consists of PCR amplification with a fluorescent primer and analysis by capillary electrophoresis, where clonal amplification results in a single peak within a background of polyclonal, constitutional amplification products.[66–68] The homo-heteroduplex analysis takes advantage of the different migration properties in a polyacrylamide gel of V-J rearrangements containing a few mismatches (heteroduplex) compared with fully matched V-J junctions (homoduplex).[69] It has been recently reported that the use of a semiautomated electrophoretic technique (Phast-System) gives comparable results to traditional analysis for detection of clonality but is faster, simpler, and highly reproducible.[70]

Several studies have shown multiple *IG* and *TCR* gene rearrangements at diagnosis in childhood B-precursor ALL. Oligoclonal *IGH* rearrangements have been observed in 30% to 40% of cases and *IGLK* rearrangements in 5% to 10%, and are due to VH replacements, VH-DJH junction, or de novo *IGH* gene rearrangements.[71–74] The presence of oligoclonality is supposed to be a rare event in *TCRB* and *TCRG* gene rearrangements in B-lineage ALL.[75–78] One exception was reported by Szczenpanski et al.,[20] who observed *TCRG* oligoclonality in 38% of patients with B-lineage ALL. Oligoclonality of *TCRD* frequently are observed in incomplete rearrangements, *Vδ2Dδ3* and *Dδ2Dδ3*.[12,25,78] In T-lineage ALL, oligoclonality is rarely seen at diagnosis.[75,76,79]

Clonal evolution, that is, a difference from the original clone at relapse, has been detected in rearrangements of both the *IG* and *TCR* genes. Rearrangements of the *IG* genes are particularly sensitive to clonal evolution that occurs in more than 30% to 40% of cases. Clonal evolution of rearranged *TCR* genes is less frequent, being observed in 10% to 25% of cases. The risk of change in gene rearrangements seems to increase with time, but further change is not commonly observed in the second or third relapses.[20,80–82] This clonal evolution may be due to the emergence of a new independent clone, to secondary leukemia, or, more commonly, to the subclones formed by continuous and secondary rearrangements of these genes.[83–86]

RT-PCR for Leukemia-Specific Translocations

Several potential pitfalls must be avoided to prevent false-positive and false-negative results with translocation RT-PCR assays. First, RNA degradation must be prevented by careful specimen handling. Second, amplification of a control mRNA from the same cDNA preparation used for amplification of the fusion transcript is used to control for the integrity of the RNA, the quality of the RT step, and the presence of inhibitors.

A major concern has been raised about the clinical utility of RT-PCR MRD translocation testing in ALL

patients because fusion transcripts can be detected in the hematopoietic tissues of healthy individuals. Low levels of *BCR/ABL* transcripts have been detected in a sizable proportion of healthy individuals.[87] Using a two-round (nested) RT-PCR method, the *MLL/AF4* fusion transcript has been detected in the fetal liver and BM of normal infant BM samples and in 12% of pediatric ALL with no cytogenetic or genomic evidence of 11q23 alterations.[88] The significance of these findings is still elusive.

Molecular Detection of MRD and Clinical Studies

The study of MRD has drawn great interest in clinical oncology because of the potential for tailoring treatment and the possibility of gaining insight into the nature of a cure. Several parameters are critical for the interpretation of MRD studies, including the type of disease (ALL or AML), therapeutic context, timing of sampling, target gene, sensitivity of the assay, interlaboratory standardization (particularly relevant in multicenter studies), retrospective or prospective nature of the study, and number of tests conducted for each patient.

A prerequisite for applying MRD measurements in clinical studies is that the data should be available for all patients. In ALL, PCR-based MRD detection with *IG* and *TCR* gene rearrangements can be applied in more than 90% to 95% of childhood and adult ALL cases. Accordingly, most of the clinical studies of MRD in ALL have used one of the different PCR approaches for the detection of antigen receptor gene rearrangements.[18,55]

Once the clonal rearrangements have been recognized at diagnosis, several methods can be applied to specifically detect the leukemia-derived PCR products during the follow-up of patients who have undergone therapy. The major variable lies in the sensitivity of the test, which can significantly influence the interpretation of the assay's results. Low-sensitivity approaches, with a detection sensitivity of one leukemic cell among 10^2 to 10^3 normal cells,[49,64,89] can be considered when the aim of the clinical protocol is to identify patients with high residual tumor load, likely to be at very high risk of relapse. By contrast, more sensitive assays including RQ-PCR assays, with a detection sensitivity of one leukemic cell among 10^4 to 10^5 normal cells, can allow identification of the subgroup of patients with very good response to early therapy. These highly sensitive RQ-PCR tests also allow precise quantification of intermediate MRD levels.[55,61,90] Whether patients in these risk groups could profit from treatment reduction or intensification is still unknown.

More recently, several prospective studies in childhood ALL have shown that MRD analysis by molecular or highly sensitive immunologic methods can predict outcome on the basis of the reduction of the leukemic cell burden during the first months of therapy.[3,44–47,91–95] The value of MRD detection at the end of treatment is questionable, because a negative result does not preclude subsequent relapse, thus reducing the utility of MRD detection as a criterion for elective cessation of treatment.[44–47]

Now that the cure rates of childhood ALL are approaching 80%, the challenge is the effective incorporation of MRD information into new therapeutic studies. Highly sensitive PCR techniques (detection limit less than 1×10^{-4}) allow the identification of a significant proportion of ALL cases with excellent clinical outcomes in the presence of negative MRD findings at early time points in treatment. By contrast, patients with 10^{-2} or more leukemic cells during any phase of remission induction should be regarded as having a very high risk of relapse, thus becoming eligible for early HSCT or experimental treatment. The clinical response to the presence of intermediate-range positive MRD findings (between 1×10^{-4} and 1×10^{-2}) remains unclear.

Some reports have addressed the question of whether the assessment of MRD at an earlier time point than the end of induction treatment could be equally satisfactory or even better for identifying ALL patients with different treatment outcomes. Two different studies[96,97] suggest that there are ALL patients with very early (day 15 to day 19) and profound cytoreduction, who are therefore candidates for future studies designed to test less-intense and hence less-toxic regimens of chemotherapy; however, data from larger series of patients will be needed to confirm these promising preliminary observations. Such patients may benefit from further intensification of therapy, but this has not yet been substantiated by randomized clinical studies. Thus, the German-Austrian BFM and Italian AIEOP study groups have adopted an MRD-based risk-group classification for treatment stratification in their ongoing clinical studies. It is hoped that a more sensitive and specific evaluation of remission and early response to treatment will improve the cure rates for children with ALL.

RT-PCR of fusion transcripts generated by the t(9;22), t(4;11), t(1;19), and t(12;21) has been used to assess MRD.[43,47,63,65] The data reported so far have been controversial, leaving unanswered the question of whether MRD assessment based on tumor-specific translocations would help to identify patients with different outcomes within the same genetically defined subgroup.

Fusion transcripts almost invariably persist in chemotherapy-treated patients with ALL bearing the t(9;22), and detection of these transcripts correlates with the high relapse rate of this disease subtype in both adults and children.[35,38] In studies of a large series of children, Ph^+ ALL appeared to be quite heterogeneous with regard to treatment sensitivity.[38] Thus, MRD evaluation may help to identify patients persistently negative for the *BCR/ABL* fusion gene among those with good early responses to treatment, who can be cured with intensive chemotherapy alone,[35] and be used as a surrogate marker to monitor in vivo response to new drugs in childhood Ph^+ ALL. For example, a randomized phase II/III study (*EsPhALL*) has been recently opened to compare the safety and efficacy of imatinib with chemotherapy in pediatric patients with Ph^+ ALL.

Studies using primer sets for the t(4;11) translocation of the *MLL* and *AF4* oncogenes suggest that early conversion to or persistently negative RT-PCR results (particularly after 3 months) are associated with prolonged complete clinical remission (CCR).[89]

In ALL patients with the t(1;19) translocation, persistently negative RT-PCR tests are a good indicator of CCR, but positive tests are not necessarily an accurate predictor of relapse, thus reducing the utility of this test for clinical decision making.[34,81] Continuous monitoring with accurate quantification may represent the most reliable approach.

Residual disease has been investigated in a limited number of patients with ALL carrying the t(12;21). Relapse was observed in cases with persistently positive MRD detection at greater than 1×10^{-3}.[90] However, relapse has been reported to occur even in patients with previously negative tests. Larger prospective studies are needed to fully assess the prognostic value of RT-PCR testing for the t(12;21) translocation, as well as its value as a marker for monitoring MRD in childhood ALL.[32,98]

Laboratory Issues

Although over the last decade numerous methods to monitor MRD in acute leukemias have been developed and new technologies are being used, standardization and quality control still are needed to apply molecular diagnostic procedures in hematopathology. This is particularly true in efforts to assure reproducible results within multicenter international studies.

The molecular approach for detection of antigen receptor gene rearrangements requires standardization and quality control, which has been addressed within the same BIOMED-1 framework. One of the major problems resides in the fact that each patient's leukemia has a unique rearrangement, requiring the use of patient-specific assays and individual optimization. Use of patient-specific assays represents a major difficulty in transferring this method to clinical practice. In an effort to maximize detection of virtually all ALL patients and to prevent false-negative results, study PCR primers have been designed for multiple targets: *TCRD*, *TCRG*, and *IGLK* rearrangements, as well as the *SIL/TAL* rearrangement.[18] The method uses a total of 54 primers with ASO probes in single or nested PCR for target identification at diagnosis, sequence analysis of junctional regions, and MRD detection in follow-up samples. A total of 25 PCR reactions are performed at diagnosis to identify the PCR targets. This standardized approach allows rapid detection of clonal *IG* and *TCR* rearrangements in ALL with high sensitivity and high specificity, and enables discrimination between mono- and oligoclonal gene rearrangements. The combination of the four PCR target types allowed PCR monitoring in more than 90% of B-cell precursor ALL and 95% of T-ALL cases. In the vast majority of childhood and adult ALL cases, two or more PCR targets were available for MRD monitoring. The sensitivity of detecting PCR targets depended at least partially on the size of the junctional region, with a level of 1×10^{-4} reached in most cases. Recently, to increase the percentage of cases successfully stratified by MDR, newly identified molecular targets have been incorporated. The monoclonal *Vδ2-Jα* rearrangements in precursor B-ALL were used as patient-specific targets for MDR detection, because they show high sensitivity and good stability.[27] A sensitive *TCRB* RQ-PCR assay was developed[99] with utility for MDR studies in T-ALL, in which the repertoire of *IG*/*TCR* rearrangements is limited and less sensitive.

Careful standardization and quality control of MRD techniques were also the aims of the European BIOMED-1 Concerted Action ("Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease"), with participants from several laboratories in European countries.[43] In particular, the five most-frequent, well-defined ALL chromosomal aberrations with fusion gene transcripts were selected: t(1;19) with *E2A-PBX1*, t(4;11) with *MLL-AF4*, t(9;22) with *BCR-ABL* (p190 and p210), t(12;21) with *TEL-AML1*, and microdeletion 1p32 with *SIL-TAL1*. PCR primers, positioned to cover several different transcript versions, were designed according to predefined criteria for single and nested PCR (Figure 31-3).

More recently, a Europe Against Cancer program was established to achieve standardization and quality control on the new RQ-PCR technique for detection and quantification of fusion gene transcripts. Twenty-five European laboratories in ten countries have collaborated and established consensus standards for RQ-PCR (Taqman technology) for the main translocations seen in a spectrum of hematologic malignancies, including ALL. A set of 12 primers and 9 probes has been selected to cover the most frequent chimeric transcripts, with a threshold of detection of 100 molecules and/or a 1×10^{-4} dilution.[47] A representative experiment of RQ-PCR detection of the fusion transcript generated by the translocation t(9;22)is shown in Figure 31-5.

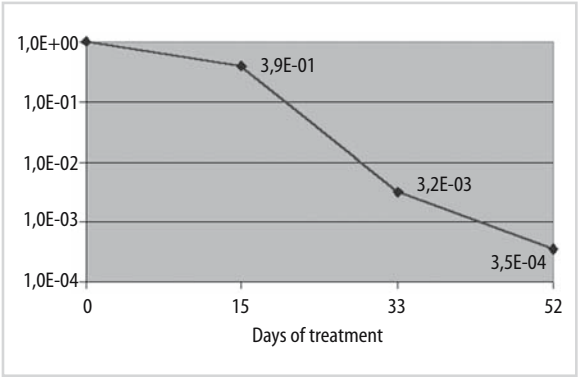


Figure 31-5. MRD monitoring by quantitative PCR for a patient with ALL carrying the t(9;22) translocation during the course of treatment.

This work is the first coordinated international effort on standardization and quality control methods for a molecular diagnosis procedure in hematopathology across therapeutic protocols. It should allow accurate quantitative measurement of fusion transcripts with an international consensus protocol for diagnosis and MRD assessment in follow-up samples from leukemia patients.

Conclusions

PCR amplification of fusion transcripts resulting from chromosomal translocations and of *IG* and *TCR* gene rearrangements has emerged as a sensitive and reproducible method to monitor MRD in ALL. The measure of the initial response to therapy in patients who have achieved complete remission by morphologic standards can dissect clinical heterogeneity within a genetically homogeneous childhood ALL subgroup. Moreover, MRD monitoring can be applied to the appropriate patient subgroups for early prediction of impending relapses.

However, before any application in clinical decision making, the MRD value must be studied in the context of prospective controlled trials. Moreover, high levels of confidence in interpreting MRD results are needed to rule out false-positive and false-negative tests, considering the problems and pitfalls of the current technology. In this context, international standardization is needed for the interpretation of MRD results produced by different studies.

MRD analysis should be incorporated into any future clinical studies investigating a therapeutic question. Moreover, only the combination of simplification and reliability of MRD methods will allow the potential benefits of MRD monitoring to be extended to all children with leukemia. In addition, new technologies such as microarrays can be of great value to identify new genetic lesions in patients carrying the same genetic subtypes of ALL but with different MRD clearance and to identify new specific markers for MRD monitoring.

References

1. Pui CH. Childhood leukaemias. *N Engl J Med.* 1995;332:1618–1630.
2. Pui CH, Evans WE. Acute lymphoblastic leukaemias. *N Engl J Med.* 1998;339:605–615.
3. Campana D, Coustan-Smith E. Detection of minimal residual disease in acute leukemia by flow cytometry [review]. *Cytometry.* 1999;38: 139–152.
4. Neale GA, Coustan-Smith E, Pan Q, et al. Tandem application of flow cytometry and polymerase chain reaction for comprehensive detection of minimal residual disease in childhood acute lymphoblastic leukemia. *Leukemia.* 1999;13:1221–1226.
5. Yeoh EJ, Roos ME, Shurtleff AS, et al. Classification, subtype discovery, and prediction outcome in pediatric acute lymphoblastic leukemia by gene expression profile. *Cancer Cell.* 2002;1:133–143.
6. Pui CH, Behm FG, Crist WM. Clinical and biological relevance of immunologic marker studies in childhood acute lymphoblastic leukemia. *Blood.* 1993;82:343–362.
7. Rabbitts TH. Chromosomal translocations in human cancer. *Nature.* 1994;372:143–149.
8. Cline MJ. The molecular basis of leukaemia. *N Engl J Med.* 1994;330:328–336.
9. Tonegawa S. Somatic generation of antibody diversity. *Nature.* 1983;302:575–581.
10. Van Dongen JJM, Wolvers-Tettero ILM. Analysis of immunoglobulin and T cell receptor genes. Part I: basic and technical aspects. *Clin Chim Acta.* 1991;198:1–92.
11. Schwartz RS. Jumping genes and the immunoglobulin V gene system. *N Engl J Med.* Cambridge, UK, 1995;333:42–44.
12. Van Dongen JJM, Langerak AW. Immunoglobulin and T-cell receptor rearrangements. In: Pui CH, ed. *Childhood Leukemias*. Cambridge University Press, Cambridge, UK, 1999:145–167.
13. Potter MN, Steward CG, Oakhill A. The significance of detection of minimal residual disease in childhood acute lymphoblastic leukemia. *Br J Haematol.* 1993;83:412–418.
14. Van Dongen JJM, Szczepanski T, De Bruijn MAC, et al. Detection of minimal residual disease in acute leukemia patients. *Cytokines Cell Mol Ther.* 1996;2:121–133.
15. Yano T, Pullman A, Andade R, et al. A common Vd2-Dd2-Dd3 T cell receptor gene rearrangement in precursor B acute lymphoblastic leukemia. *Br J Haematol.* 1991;79:44–49.
16. Brisco MJ, Tan LW, Orsborn AM, Morley AA. Development of a highly sensitive assay, based on the polymerase chain reaction, for rare B-lymphocyte clones in a polyclonal population. *Br J Haematol.* 1990;75:163–167.
17. Scrideli CA, Simoes AL, Defavery R, Bernardes JE, Duarte MH, Tone LG. Childhood B lineage acute lymphoblastic leukemia clonality study by the polymerase chain reaction. *J Pediatr Hematol Oncol.* 1997;19:516–522.
18. Pongers-Willemse MJ, Seriu T, Stolz F, et al. Primers and protocols for standardized MRD detection in ALL using immunoglobulin and T cell receptor gene rearrangements and TAL1 deletions as PCR targets. Report of the BIOMED-1 Concerted Action: Investigation of minimal residual disease in acute leukemia. *Leukemia.* 1999;13:110–118.
19. Van Dongen JJM, Wolvers-Tettero ILM. Analysis of immunoglobulin and T cell receptor genes. Part II: possibilities and limitations in the diagnosis and management of lymphoproliferative diseases and related disorders. *Clin Chim Acta.* 1991;198:93–174.
20. Szczepanski T, Beishuizen A, Pongers-Willemse MJ, et al. Cross-lineage T-cell receptor gene rearrangements occur in more than ninety percent of childhood precursor-B-acute lymphoblastic leukaemias: alternative PCR targets for detection of minimal residual disease. *Leukemia.* 1999;13:196–205.
21. Tkachuck DC, Griesser H, Takirara Y, et al. Rearrangement of T-cell delta locus in lymphoproliferative disorders. *Blood.* 1988;72:353–357.
22. Biondi A, Francia de Celli P, Rossi V, et al. High prevalence of T-cell receptor V delta 2-(D)-D delta 3 or D delta 1/2-D delta 3 rearrangements in B-precursor acute lymphoblastic leukemias. *Blood.* 1990;75:1834–1840.
23. Trainor KJ, Brisco MJ, Wan JH, et al. Gene rearrangement in B- and T-lymphoproliferative disease detected by the polymerase chain reaction. *Blood.* 1991;78:192–196.
24. Campana D, Pui CH. Detection of minimal residual disease in acute leukaemia: methodological and clinical significance. *Blood.* 1995;85: 1416–1434.
25. Ghali DW, Panzer S, Fischer S, et al. Heterogeneity of the T cell receptor delta gene indicating subclone formation in acute precursor B cell leukemias. *Blood.* 1995;85:2795–2801.
26. Szczepanki T, LangeraK AW, Wolvers-Tettero IL, et al. Immunoglobulin and T cell receptor gene rearrangement patterns in acute lymphoblastic leukemia are less mature in adults than in children: implications for selection of PCR targets for detection of minimal residual disease. *Leukemia.* 1998;12:1081–1088.
27. Szczepanski T, van der Velden VH, Hoogeveen PG, et al. Vdelta2-Jalpha rearrangements are frequent in precursor-B-acute lym-

phoblastic leukemia but rare in normal lymphoid cells. *Blood.* 2004;103:3798–3804.
28. Foroni L, Harrison CJ, Hoffbrand AV, Potter MN. Investigation of minimal residual disease in childhood and adult acute lymphoblastic leukaemia by molecular analysis. *Br J Haematol.* 1999;105:7–24.
29. Yin JA, Tobal K. Detection of minimal residual disease in acute myeloid leukaemia: methodologies, clinical and biological significance. *Br J Haematol.* 1999;106:578–590.
30. Look AT. Oncogenic transcription factors in the human acute leukaemias. *Science.* 1997;278:1059–1064.
31. Biondi A, Cimino G, Pieters R, Pui CH. Biological and therapeutic aspects of infant leukemia. *Blood.* 2000;96:24–33.
32. Cazzaniga G, Rossi V, Biondi A. Monitoring minimal residual disease using chromosomal translocations in childhood ALL. *Best Pract Res Clin Haematol.* 2002;15:21–35.
33. Hunger SP. Chromosomal translocations involving the E2A gene in acute lymphoblastic leukemia: clinical features and molecular pathogenesis. *Blood.* 1996;87:1211–1224.
34. Privitera E, Luciano A, Ronchetti D, et al. Molecular variants of the 1;19 chromosomal translocation in pediatric acute lymphoblastic leukemia (ALL). *Leukemia.* 1994;8:554–559.
35. Cazzaniga G, Lanciotti M, Rossi V, et al. Prospective molecular monitoring of BCR/ABL transcript in children with Ph+ acute lymphoblastic leukaemia unravels differences in treatment response. *Br J Haematol.* 2002;119:445–453.
36. Deininger MW, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. *Blood.* 2000;96:3343–3356.
37. Scrideli CA, Cazzaniga G, Fazio G, et al. Gene expression profile unravels significant differences between childhood and adult Ph+ acute lymphoblastic leukemia. *Leukemia.* 2003;17:2234–2237.
38. Schrappe M, Arico M, Harbott J, et al. Philadelphia chromosome-positive (Ph+) childhood acute lymphoblastic leukemia: good initial steroid response allows early prediction of a favorable treatment outcome. *Blood.* 1998;92:2730–2741.
39. Borkhardt A, Cazzaniga G, Viehmann S, et al. Incidence and clinical relevance of TEL/AML1 fusion genes in children with acute lymphoblastic leukemia enrolled in the German and Italian multicenter therapy trials. Associazione Italiana Ematologia Oncologia Pediatrica and the Berlin-Frankfurt-Munster Study Group. *Blood.* 1997;90:571–577.
40. Raynaud S, Mauvieux L, Cayuela JM, et al. TEL/AML1 fusion gene is a rare event in adult acute lymphoblastic leukemia. *Leukemia.* 1996;10:1529–1530.
41. Seeger K, Adams HP, Buchwald D, et al. TEL-AML1 fusion transcript in relapsed childhood acute lymphoblastic leukemia. The Berlin-Frankfurt-Munster Study Group. *Blood.* 1998;91:1716–1722.
42. Basso K, Frascella E, Zanesco L, Rosolen A. Improved long-distance polymerase chain reaction for the detection of t(8;14)(q24;q32) in Burkitt's lymphomas. *Am J Pathol.* 1999;155:1479–1485.
43. van Dongen JJ, Macintyre EA, Gabert JA, et al. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: investigation of minimal residual disease in acute leukemia. *Leukemia.* 1999;13:1901–1928.
44. Cave H, van der Werff Ten Bosch J, Suciu S, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia. *N Engl J Med.* 1998;339:591–598.
45. Biondi A, Valsecchi MG, Seriu T, et al. Molecular detection of minimal residual disease is a strong predictive factor of relapse in childhood B-lineage acute lymphoblastic leukemia with medium risk features. A case control study of the International BFM Study Group. *Leukemia.* 2000;14:1939–1943.
46. Van Dongen JJ, Seriu T, Panzer-Grumayer ER, et al. Prognostic value of minimal residual disease in acute lymphoblastic leukaemia in childhood. *Lancet.* 1998;352:1731–1738.
47. Gabert J, Beillard E, van der Velden VH, et al. Standardization and quality control studies of "real-time" quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia—a Europe Against Cancer program. *Leukemia.* 2003;17:2318–2357.
48. Scrideli CA, Kashima S, Cipolloti R, Defavery R, Bernardes JE, Tone LG. Minimal residual disease in Brazilian children with acute lymphoid leukemia: comparison of three detection methods by PCR. *Leuk Res.* 2002;26:431–438.
49. Scrideli CA, Queiroz RG, Bernardes JE, Valera ET, Tone LG. PCR detection of clonal IgH and TCR gene rearrangements at the end of induction as a non-remission criterion in children with ALL: comparison with standard morphologic analysis and risk group classification. *Med Pediatr Oncol.* 2003;41:10–16.
50. Knechtli CJ, Goulden NJ, Hancock JP, et al. Minimal residual disease status before allogeneic bone marrow transplantation is an important determinant of successful outcome for children and adolescents with acute lymphoblastic leukemia. *Blood.* 1998;92:4072–4079.
51. Knechtli CJ, Goulden NJ, Hancock JP, et al. Minimal residual disease status as a predictor of relapse after allogeneic bone marrow transplantation for children with acute lymphoblastic leukaemia. *Br J Haematol.* 1998;102:860–871.
52. Fardel S, Kurzrock R, Estrov Z. Minimal residual disease in hematologic disorders. *Arch Pathol Lab Med.* 1999;123:1030–1034.
53. Sklar J. Polymerase chain reaction: the molecular microscope of residual disease. *J Clin Oncol.* 1991;1:1521–1523.
54. Cross NC. Quantitative PCR techniques and applications. *Br J Haematol.* 1995;89:693–697.
55. Cazzaniga G, Biondi A. Molecular monitoring of childhood acute lymphoblastic leukemia using antigen receptor gene rearrangements and quantitative polymerase chain reaction technology. *Haematologica.* 2005;90:382–390.
56. Verhagen OJ, Willemse MJ, Breunis WB, et al. Application of germline IgH probes in real-time quantitative PCR for the detection of minimal residual disease in acute lymphoblastic leukemia. *Leukemia.* 2000;14:1426–1435.
57. Pongers-Willemse MJ, Verhagen OJ, Tibbe GJ, et al. Real-time quantitative PCR for the detection of minimal residual disease in acute lymphoblastic leukemia using junctional region specific TaqMan probes. *Leukemia.* 1998;12:2006–2014.
58. Germano G, Rossi V, Poli A, et al. Use of T-cell receptor g germline "TaqMan" probes for the analysis of minimal residual disease in childhood T-ALL by Real-Time Quantitative PCR. *Blood.* 1999;94: 204b.
59. Willems P, Verhagen O, Segeren C, et al. Consensus strategy to quantitate malignant cells in myeloma patients is validated in a multicenter study. Belgium-Dutch Hematology-Oncology Group. *Blood.* 2000;96:63–70.
60. Eckert C, Scrideli CA, Taube T, et al. Comparison between TaqMan and LightCycler technologies for quantification of minimal residual disease by using immunoglobulin and T-cell receptor genes consensus probes. *Leukemia.* 2003;17:2517–2524.
61. van der Velden VH, Hochhaus A, Cazzaniga G, Szczepanski T, Gabert J, van Dongen JJ. Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. *Leukemia.* 2003;17:1013–1034.
62. Hochhaus A, Weisser A, La Rosee P, et al. Detection and quantification of residual disease in chronic myelogenous leukaemia. *Leukemia.* 2000;14:998–1005.
63. Pallisgaard N, Clausen N, Schroder H, Hokland P. Rapid and sensitive minimal residual disease detection in acute leukaemia by quantitative real-time RT-PCR exemplified by t(12;21) TEL-AML1 fusion transcript. *Genes Chromosomes Cancer.* 1999;26:355–365.
64. Krauter J, Wattjes MP, Nagel S, et al. Real-time RT-PCR for the detection and quantification of AML1/MTG8 fusion transcripts in t(8;21)-positive AML patients. *Br J Haematol.* 1999;107:80–85.
65. Beillard E, Pallisgaard N, van der Velden VH, et al. Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using "real-time" quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR)—a Europe Against Cancer program. *Leukemia.* 2003;17:2474–2486.

66. Guidal C, Vilmer E, Grandchamp B, Cave H. A competitive PCR-based method using TCRD, TCRG and IGH rearrangements for rapid detection of patients with high levels of minimal residual disease in acute lymphoblastic leukemia. *Leukemia.* 2002;16:762–764.
67. Landman-Parker J, Aubin J, Delabesse E, et al. Simplified strategies for minimal residual disease detection in B cell precursor acute lymphoblastic leukaemia. *Br J Haematol.* 1996;95:281–290.
68. Evans PA, Short MA, Owen RG, et al. Residual disease detection using fluorescent polymerase chain reaction at 20 weeks of therapy predicts clinical outcome in childhood acute lymphoblastic leukemia. *J Clin Oncol.* 1998;16:3616–3627.
69. 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;83:3271–3278.
70. Germano G, Songia S, Biondi A, Basso G. Rapid detection of clonality in patients with acute lymphoblastic leukemia. *Haematologica.* 2001;86:382–385.
71. Kitchingman GR, Mirro J, Stass S, et al. Biologic and prognostic significance of the presence of more than two mu heavy-chain genes in childhood acute lymphoblastic leukemia of B precursor cell origin. *Blood.* 1986;67:698–703.
72. Kuang S, Gu L, Dong S, et al. Long-term follow-up of minimal residual disease in childhood acute lymphoblastic leukemia patients by polymerase chain reaction analysis of multiple clone-specific or malignancy-specific gene markers. *Cancer Genet Cytogenet.* 1996;88:110–117.
73. Green E, McConville CM, Powell JE, et al. Clonal diversity of Ig and T-cell-receptor gene rearrangements identifies a subset of childhood B-precursor acute lymphoblastic leukemia with increased risk of relapse. *Blood.* 1998;92:952–958.
74. Scrideli CA, Defavery R, Bernardes JE, Tone LG. Prognostic significance of bi/oligoclonality in childhood acute lymphoblastic leukemia as determined by polymerase chain reaction. *Sao Paulo Med J.* 2001;119:175–180.
75. Brumpt C, Delabesse E, Beldjord K, et al. The incidence of clonal T-cell receptor rearrangements in B-cell precursor acute lymphoblastic leukemia varies with age and genotype. *Blood.* 2000;96:2254–2261.
76. Hara J, Benedict SH, Yumura K, Ha-Kawa K, Gelfand EW. Rearrangement of variable region T cell receptor γ genes in acute lymphoblastic leukemia. Vγ gene usage differs in mature and immature T cells. *J Clin Invest.* 1999;83:1277–1283.
77. Beishuizen A, Verhoeven MA, Van Wering ER, et al. Analysis of Ig and T cell receptor genes in 40 childhood acute lymphoblastic leukemia at diagnosis and subsequent relapse: implications for the detection of minimal residual disease by polymerase chain reaction analysis. *Blood.* 1994;83:2238–2247.
78. Steward CG, Goulden NJ, Katz F, et al. A polymerase chain reaction study of the stability of Ig heavy-chain and T cell receptor delta gene rearrangements between presentation and relapse of childhood B-lineage acute lymphoblastic leukemia. *Blood.* 1994;83:1355–1362.
79. Scrideli CA, Queiroz RG, Kashima S, Sankarankutty BO, Tone LG. T cell receptor gamma (TCRG) gene rearrangements in Brazilian children with acute lymphoblastic leukemia: analysis and implications for the study of minimal residual disease. *Leuk Res.* 2004;28:267–273.
80. Campana D, Van Dongen JJM, Pui CH. Minimal residual disease. In: Pui CH, ed. *Childhood Leukemias*. Cambridge University Press, Cambridge, UK, 1999:413–442.
81. Marshall GM, Kwan E, Haber M, et al. Characterization of clonal immunoglobulin heavy chain and I cell receptor gamma gene rearrangements during progression of childhood acute lymphoblastic leukemia. *Leukemia.* 1995;9:1847–1850.
82. Scrideli CA, Kashima S, Cipolloti R, Defavery R, Tone LG. Clonal evolution as the limiting factor in the detection of minimal residual disease by polymerase chain reaction in children in Brazil with acute lymphoid leukemia. *J Pediatr Hematol Oncol.* 2002;24:364–367.
83. Steenbergen EJ, Verhagen OJ, Van Leeuwen EF, et al. Distinct ongoing Ig heavy chain rearrangement processes in childhood B-precursor acute lymphoblastic leukemia. *Blood.* 1993;82:581–589.
84. Lo Nigro L, Cazzaniga G, Di Cataldo A, et al. Clonal stability in children with acute lymphoblastic leukemia (ALL) who relapsed five or more years after diagnosis. *Leukemia.* 1999;13:190–195.
85. Choi Y, Greenberg SJ, Du TL, Ward PM, et al. Clonal evolution in B-lineage acute lymphoblastic leukemia by contemporaneous VH–VH gene replacements and VH–DJH gene rearrangements. *Blood.* 1996;87:2506–2512.
86. Height SE, Swanbury GJ, Matute E, et al. Analysis of clonal rearrangements of the Ig heavy chain locus in acute leukemia. *Blood.* 1996;87:5242–5250.
87. Bose S, Deininger M, Gora-Tybor J, et al. The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance and implications for the assessment of minimal residual disease. *Blood.* 1998;92:3362–3367.
88. Uckun FM, Herman-Hatten K, Crotty ML, et al. Clinical significance of MLL-AF4 fusion transcript expression in the absence of a cytogenetically detectable t(4;11)(q21;q23) chromosomal translocation. *Blood.* 1998;92:810–821.
89. Germano G, Songia S, Biondi A, Basso G. Rapid detection of clonality in patients with acute lymphoblastic leukemia. *Haematologica.* 2001;86:382–385.
90. Szczepanski T, Orfao A, van der Velden VH, San Miguel JF, van Dongen JJ. Minimal residual disease in leukaemia patients. *Lancet Oncol.* 2001;2:409–417.
91. Cimino G, Elia L, Rapanotti MC, et al. A prospective study of residual-disease monitoring of the ALL1/AF4 transcript in patients with t(4;11) acute lymphoblastic leukemia. *Blood.* 2000;95:96–101.
92. Zuna J, Hrusak O, Kalinova M, Muzikova K, Stary J, Trka J. Significantly lower relapse rate for TEL/AML1-positive ALL. *Leukemia.* 1999;13:1633.
93. Kerst G, Kreyenberg H, Roth C, et al. Concurrent detection of minimal residual disease (MRD) in childhood acute lymphoblastic leukaemia by flow cytometry and real-time PCR. *Br J Haematol.* 2005;128:774–782.
94. Laughton SJ, Ashton LJ, Kwan E, Norris MD, Haber M, Marshall GM. Early responses to chemotherapy of normal and malignant hematologic cells are prognostic in children with acute lymphoblastic leukemia. *J Clin Oncol.* 2005;23:2264–2271.
95. Neale GA, Coustan-Smith E, Stow P, et al. Comparative analysis of flow cytometry and polymerase chain reaction for the detection of minimal residual disease in childhood acute lymphoblastic leukemia. *Leukemia.* 2004;18:934–938.
96. Panzer-Grumayer ER, Schneider M, Panzer S, Fasching K, Gadner H. Rapid molecular response during early induction chemotherapy predicts a good outcome in childhood acute lymphoblastic leukemia. *Blood.* 2000;95:790–794.
97. Coustan-Smith E, Sancho J, Behm FG, et al. Prognostic importance of measuring early clearance of leukemic cells by flow cytometry in childhood acute lymphoblastic leukemia. *Blood.* 2002;100:52–58.
98. Taube T, Eckert C, Korner G, Henze G, Seeger K. Real-time quantification of TEL-AML1 fusion transcripts for MRD detection in relapsed childhood acute lymphoblastic leukaemia. Comparison with antigen receptor-based MRD quantification methods. *Leuk Res.* 2004;28:699–706.
99. Bruggemann M, van der Velden VH, Raff T, et al. Rearranged T-cell receptor beta genes represent powerful targets for quantification of minimal residual disease in childhood and adult T-cell acute lymphoblastic leukemia. *Leukemia.* 2004;18:709–719.

Chapter 32

B-Cell Lymphomas

Rita M. Braziel and Guang Fan

Molecular Basis of Disease

Introduction

The B-cell lymphomas (BCLs) represent 80% to 90% of non-Hodgkin lymphomas in the Western world and include multiple lymphoma subtypes with different biologies, natural histories, morphologic characteristics, immunophenotypes, genetic features, prognoses, and responses to therapy.[1] Numerous subtypes of B-cell malignancies are defined according to the World Health Organization (WHO) classification (Table 32-1). Accurate subclassification of these BCLs has always been a challenge for pathologists, resulting in early application of new techniques in genetic analysis to these tumors to improve diagnostic accuracy. Today, the genetic features of BCLs are used not only to aid in rendering an accurate primary diagnosis, but also to predict prognosis, to assess for minimal residual disease after therapy, and even to help determine optimal therapy.

Two types of molecular abnormalities in BCL have commonly been evaluated for clinical purposes,[2] those occurring in genes coding for antigen receptor (AgR) molecules and those occurring in oncogenes and tumor suppressor genes. However, it must be noted that the recent development of new research techniques such as microarray gene expression profiling, comparative genomic hybridization, and proteomics has resulted in such a rapid expansion of new knowledge about the biology of these different BCLs that the list of clinically important molecular markers is likely to change considerably over the next few years, presenting a daunting challenge to clinical molecular laboratories.

B-Cell Antigen Receptor Genes

Antigen receptors are the primary effector molecules of the adaptive immune system and consist of multisubunit glycoprotein molecules present on all B and T lymphocytes.[3] In B cells, the AgRs are membrane-bound immunoglobulins (Ig). B-cell AgR genes include the heavy-chain gene (*IGH*) at 14q32, the kappa light-chain gene (*IGLK*) at 2p11, and the lambda light-chain gene (*IGLL*) at 22q11. In a germline, or nonrearranged state, AgR genes are composed of separate coding segments distributed in a region of DNA estimated to be several hundred kilobases (kb) in length. During B-cell development, one *IGH* allele undergoes rearrangement, during which double-stranded breaks in DNA occur with subsequent deletion of large intervening DNA segments. This is followed by DNA repair at the site of recombination to bring the separated AgR gene segments together, to form a functional *IGH* gene. Each AgR gene in germline configuration consists of variable (V), joining (J), constant (C), and in the case of the *IGH* gene, diversity (D) regions. There are approximately 45 V, 23 D, and 6 J regions. AgR rearrangement occurs in a specific order: one *IGH* D segment is fused to one J segment first, and this DJ segment is then fused with a V segment to form a functional VDJ exon that encodes the variable antigen recognition site of the IGH protein. Following *IGH* rearrangement, the final product is a gene containing one V region (plus all upstream V regions not involved in the rearrangement), one J region (plus all the downstream J regions not involved in the rearrangement), one D region, and one C region, arranged in the following order on the chromosome: V-D-J-C (Figure 32-1).

The large variety of different recombinations that can occur within these genes creates the diversity and specificity of the receptors necessary for recognition of a vast number of antigens. However, recom-binational diversity alone does not account for the full range of diversity of the Ig repertoire. Nucleotide loss and random nucleotide addition between the D-J and V-D junctions by the enzyme terminal deoxynucleotidyl transferase (TdT) provides additional diversity, as does the variability of combinations among the V, D, and J segments of the *IGH*, *IGLK,* and *IGLL* genes in the formation of the antigen recognition site.

A hierarchy for Ig AgR gene rearrangements exists.[4] Ig gene rearrangement is an error-prone process, and many

Table 32-1. World Health Organization Classification of B-Cell Lymphomas

Lymphoma Subtype	Abbreviation Used in Text
Precursor B-cell lymphoblastic leukemia/lymphoma	Pre-B ALL
Mature B-Cell Lymphomas	
Chronic lymphocytic leukemia/small lymphocytic lymphoma	CLL/SLL
Lymphoplasmacytic lymphoma	LPL
Splenic marginal zone lymphoma	SMZL
Extranodal marginal zone B-cell lymphoma of MALT	MALToma
Nodal marginal zone B-cell lymphoma	Nodal MZL
Follicular lymphoma	FL
Mantle cell lymphoma	MCL
Diffuse large B-cell lymphoma	DLBCL
Mediastinal large B-cell lymphoma	MBCL
Intravascular large B-cell lymphoma	ILBCL
Primary effusion lymphoma	PEL
Burkitt lymphoma	BL
Plasmacytoma/plasma cell myeloma	PCT/MM

attempted rearrangements result in nonfunctional AgR genes, mostly due to a disruption of the open reading frame for translation of the encoded transcript by the N-region nucleotides. If the first attempt at *IGH* rearrangement fails, the cell can try to rearrange the second *IGH* allele to produce a functional gene. A single B cell may therefore have two *IGH* rearrangements, one nonfunctional and one functional. The *IGLK* alleles rearrange in a similar fashion, but only after the successful rearrangement of one of the *IGH* alleles. The *IGLL* genes rearrange in most cases only if the *IGLK* genes have been deleted on both alleles, following their unsuccessful rearrangement. Normal B cells that fail to produce a functional *IGH* or light chain rearrangement are usually eliminated through apoptosis.

Even after successful rearrangement, the Ig genes in mature B cells often undergo additional changes, including *IGH* isotype switching, V-segment substitution due to a second round of rearrangement in the germinal center, and somatic hypermutation. When somatic hypermutation of the *IGH* V region (IgV_H) occurs in germinal center–derived B cells in response to antigen exposure, affinity maturation of an already rearranged *IGH* gene occurs through point mutations, small insertions, or deletions, or some combination of these. The term "postgerminal center" is used to refer to B cells that have been exposed to antigen in the germinal center and undergone additional somatic mutations of IgV_H.

Structurally and genetically unique AgR rearrangements occur in every B cell and have been exploited for years as markers of cell lineage and clonality. B cells with these unique gene rearrangements may undergo limited clonal expansion as a part of a normal immune response, but uncontrolled clonal expansion occurs in BCLs (and other B-cell malignancies). In this setting, the unique AgR rearrangement can serve as a diagnostic marker of B-cell clonality and a marker for detection of minimal residual disease following therapy. Clonal rearrangements of *IGH* and *IGLK* can be detected in essentially all malignancies of mature B cells, but many precursor B-cell malignancies will have only *IGH* rearrangements, since malignant transformation occurs before rearrangement of the *IGLK* gene.

Oncogenes/Tumor Suppressor Genes in B-Cell Lymphomas

Most BCLs are associated with characteristic chromosomal abnormalities that may help to establish an accurate primary diagnosis, have specific prognostic significance, or serve as targets for minimal residual disease (MRD) detection. Table 32-2 lists the recurrent cytogenetic abnormalities and the molecular correlates for most subtypes of BCL. This list of genetic abnormalities illustrates the large number of potential targets for molecular testing, but also highlights the increasing challenge for clinical molecular laboratories to provide relevant genetic information to clinicians. Because most chromosomal abnormalities are associated with specific BCL subtypes, diagnostic and prognostic molecular testing is often necessarily different in different BCLs. The BCL subtype-specific molecular abnormalities with diagnostic/prognostic significance are therefore discussed under the individual BCL categories below, and more general considerations are discussed later in the sections on diagnostic and prognostic indications for molecular testing in BCL.

Chronic Lymphocytic Leukemia/ Small Lymphocytic Lymphoma

Although morphologically and immunophenotypically similar, genetic testing has proven that several different

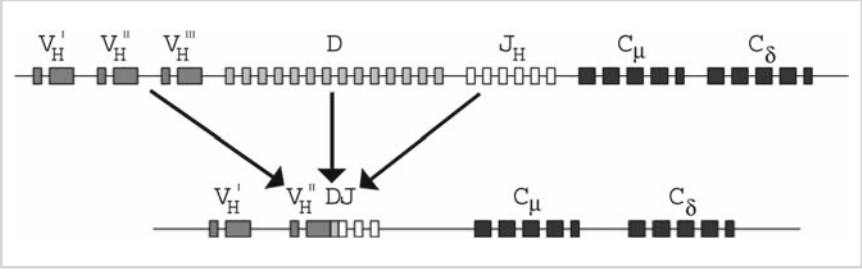


Figure 32-1. Configuration of the *IGH* gene on chromosome 14q32. DJ rearrangement occurs initially, followed by VD rearrangement. The rearrangements occur in a random fashion, resulting in a pattern of rearrangements that is unique to each normal B cell. The variable region of the rearranged *IGH* gene contains one V_H segment (plus all the upstream V_H segments), one D_H segment, one J_H (plus all the downstream J_H segments), and the C_H segments, in that order on the chromosome.

Table 32-2. Nonrandom Chromosomal Abnormalities in B-Cell Lymphomas

Lymphoma Subtype	Nonrandom Chromosomal Alterations	Genes Involved	Assay Used for Diagnosis/Prognosis*
CLL/SLL	Del 13q14	Unknown	FISH
	Trisomy 12	Unknown	FISH
	Del 11q22–23	*ATM*	FISH,
	Del 17p13	*TP53*	FISH, karyotyping
LPL†	t(9;14)(p13;q32)	*PAX5*/*IGH*	FISH, karyotyping
MZL	t(11;18)(q21;q21)	*API2*/*MALT1*	FISH, PCR
	t(1;14)(p22;q32)	*BCL10*/*IGH*	FISH, PCR
	t(14;18)(q32;q21)	*IGH*/*MALT1*	FISH
	Trisomy 3	?*BCL6*	FISH
	Trisomy 18	Unknown	FISH
FL	t(14;18)(q32;q21)	*IGH*/*BCL2*	FISH, PCR
MCL	t(11;14)(q13;q32)	Cyclin D1/*IGH*	FISH for t(11;14)
DLBCL	3q27	BCL6	FISH, SBA
	t(14;18)(q32;q21)	*IGH*/*BCL2*	FISH, PCR
BL/BLL	t(8;14)(q24;q32)	*MYC*/*IGH*	FISH, SBA
	t(2;8)(p11;q24)	*IGLK*/*MYC*	
	t(8;22)(q24;q11)	*MYC*/*IGLL*	

*The assays listed are those most commonly used today for clinical testing, but many of these abnormalities also can be detected by other techniques.
†The involvement of the *PAX5*/*IGH* translocation in LPL has been reported by only one laboratory and has not been confirmed subsequently, so it may not be a relevant translocation in this disease. The other molecular abnormalities listed have been confirmed by multiple sources.

diseases have been included in chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL), with different chromosomal abnormalities, natural histories, and responses to therapy. Established adverse prognostic biologic factors in CLL/SLL include CD38 antigen expression (detected by flow cytometry), lack of significant IgV_H mutation, and the presence of *TP53* or *ATM* gene defects. Many of the relevant chromosomal abnormalities in CLL/SLL are not detected by karyotyping; however, these abnormalities usually are detected by fluorescence in situ hybridization (FISH). Evaluation at diagnosis for these abnormalities is clinically indicated in most cases, as therapeutic decisions are affected by the genetic profile identified. Important specific genetic alterations in CLL/SLL are discussed below.

Somatic Hypermutation of IGH Variable Gene Region

Seen in ~50% of CLL/SLL patients, unmutated IgV_H status has been reported to predict poor clinical outcome. However, recent studies have suggested that the poor outcome associated with IgV_H mutation levels of <5% may, in fact, be due to overrepresentation of patients with TP53 dysfunction or CD38 expression, or both, in the unmutated IgV_H group.[5–7] Patients with intact TP53 function and absence of CD38 expression appear to have prolonged survival regardless of the extent of IgV_H mutation. As the measurement of IgV_H mutation status requires multiple polymerase chain reactions (PCRs) and sequencing of PCR products, this test is currently performed for research purposes only.

Deletions of 13q

As a single abnormality in CLL/SLL, 13q14 deletions convey the most favorable prognosis. CLL/SLL patients with 13q14 abnormalities most often also have IgV_H gene hypermutation. Detected by FISH, 13q14 deletions are seen in 40% to 55% of CLL/SLL cases and are thought to affect one or more tumor suppressor genes. Frequent deletions and downregulation of microRNA genes miR15 and miR16 at 13q14 have been described recently in the majority of CLL/SLL cases.[8] 13q14 deletions are sometimes seen in other BCLs, but the specific gene(s) involved may not be the same.

Trisomy 12

Detected by FISH or routine karyotyping in 16% to 20% of CLL/SLL cases (FISH illustration in Figure 32-2b), trisomy 12 usually is associated with unmutated IgV_H genes and also has been associated with atypical CLL. Patients with trisomy 12 have an intermediate prognosis, with a median survival exceeded only by patients with single 13q14 abnormalities.

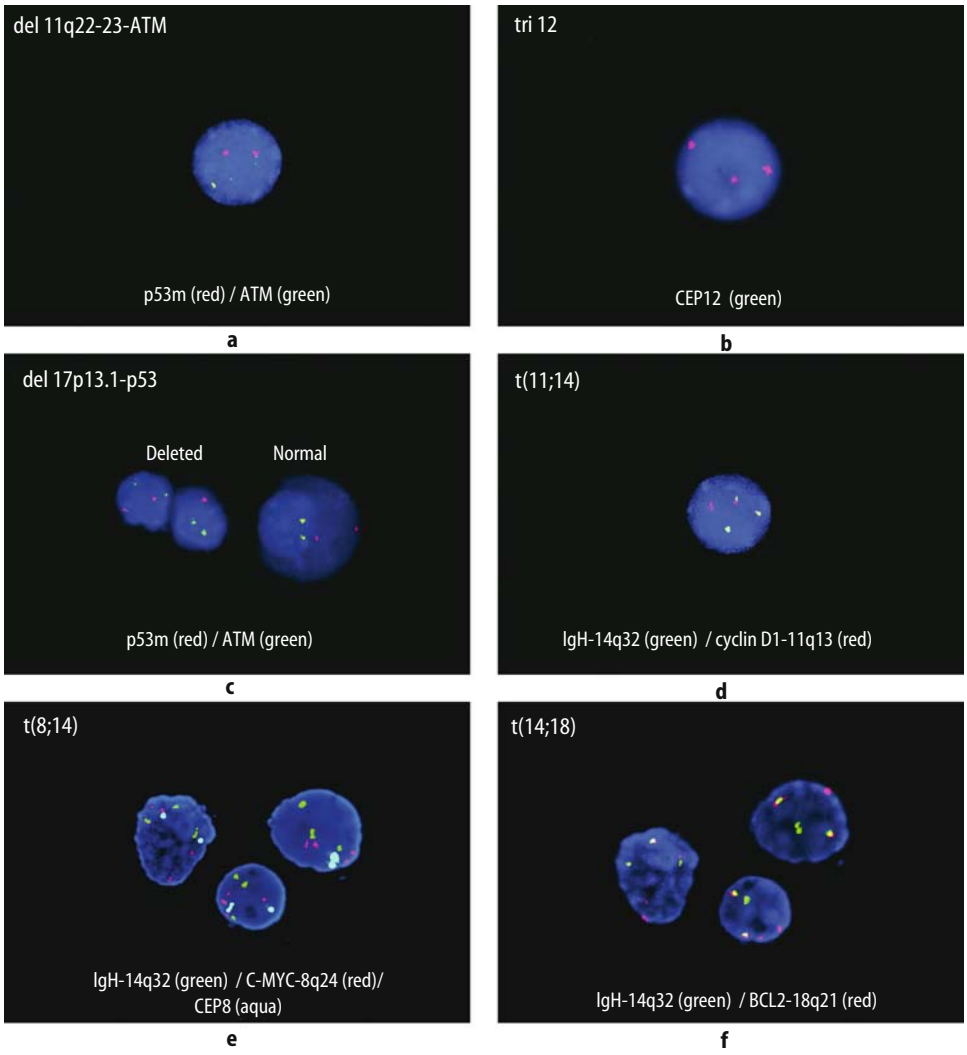


Figure 32-2. Illustration of some typical genetic abnormalities in BCL, detected by interphase FISH. (a) CLL/SLL lymphocyte with an ataxia telangiectasia (*ATM*) (11q22.3) deletion. The cell shows two red *TP53* (17p13.1) signals but only one green *ATM* signal (RRG). A normal cell would have two red and two green signals (RRGG). (b) A centromeric probe for chromosome 12 demonstrates trisomy 12 (RRR) in this cell from the blood of a CLL/SLL patient. A normal cell would have two red signals (RR). (c) In contrast to the first image, the abnormal cell in this CLL/SLL blood smear shows a deletion of the *TP53* gene, but normal *ATM* signals (RGG). The normal cells show an RRGG signal pattern. (d) A translocation pattern (RGYY) is seen in this cell, using probes that detect both derivative chromosomes resulting from the t(11;14)(q13;q32) *BCL1*/*IGH* translocation in a patient with mantle cell lymphoma (MCL). Both derivative chromosomes produce a yellow fusion signal, while the normal alleles produce separate red and green signals. (e) A complex interphase FISH pattern with the *IGH*, *MYC*, and centromeric chromosome 8 probes in cells from a lymph node biopsy in a patient with a Burkitt-like lymphoma (BLL). FISH testing with a probe set for the t(8;14) *MYC*/*IGH* translocation and a centromeric probe for chromosome 8 showed an abnormal signal pattern for both *IGH* and *MYC*, but no translocation fusion signal. Instead, the cells have four red *MYC* signals, three green *IGH* signals, and two aqua centromeric 8 signals (RRRRGGGAA); there are no yellow fusion signals. This patient was subsequently found to have an antecedent follicular lymphoma (FL), with transformation to a BLL. (f) Subsequent FISH analysis of cells from the same lymph node in (e) showed a typical translocation pattern (RGYY) with the probe set for the *IGH*/*BCL2* translocation characteristic of FL. The dual fusion probe set detected both derivative chromosomes from the t(14;18)(q21;q32) *IGH*/*BCL2* translocation. No fusion partner was identified for the *MYC* rearrangement.

Deletions in 11q22–23

Seen in 14% to 20% of CLL/SLL cases, deletions in 11q22–23 are associated with mutation and inactivation of the ataxia telangiectasia (*ATM*) gene. The ATM protein kinase is involved in TP53 regulation, and *ATM* deletions produce TP53 dysfunction in CLL/SLL cells. *ATM* deletions most often are detected by FISH (illustration in Figure 32-2a) and are associated with a poor prognosis.

Deletions of 17p13.3

The *TP53* gene at 17p13.3 is deleted in 7% to 11% of CLL/SLL cases; it is best detected by FISH (illustration in Figure 32-2c) but is sometimes also detected by karyotyping. TP53 maintains genome integrity by orchestrating the repair or elimination of cells with damaged DNA and contributes to the cytotoxicity of many anticancer agents. It is therefore not surprising that TP53 dysfunction is associ-

ated with an adverse clinical outcome in CLL/SLL patients, as well as other BCLs.

Lymphoplasmacytic Lymphoma

By one report, approximately 50% of lymphoplasmacytic lymphomas (LPLs) carry the t(9;14)(p13;q32) translocation, in which *PAX5* on 9p13 is juxtaposed with the constant switch (μ) region of the *IGH* locus on 14q32.[9] However, other investigators have not confirmed this observation. The translocation purportedly results in dysregulation and overexpression of a normal PAX5 protein, a crucial regulatory protein in normal B-cell differentiation. *PAX5* encodes a B-cell-specific transcription factor known as B-cell-specific activator protein (BSAP), thought to regulate B-cell genes such as *CD19* and *IGH*. PAX5 also is thought to stimulate B-cell proliferation and transcriptionally repress TP53 in vitro. Breakpoints involved in the *PAX5*/*IGH* translocation are widely dispersed, making it difficult to design PCR assays that will detect many translocations, so detection is by FISH, reverse transcription–polymerase chain reaction (RT-PCR), or karyotyping.

Marginal Zone B-Cell Lymphomas

Multiple chromosomal abnormalities have been found in extranodal marginal zone B-cell lymphomas (MZLs) (low-grade BCLs of mucosa-associated lymphoid tissue [MALTomas]), and recent reports suggest that low-grade MALTomas of different sites have different chromosomal abnormalities. Testing for these abnormalities is helpful in confirmation of primary diagnosis and prediction of prognosis. Diagnostic and prognostic testing for these chromosomal abnormalities usually is performed by multicolor FISH assays.

Trisomy 3

Trisomy 3 occurs in more than 50% of low-grade MALTomas. Comparative genomic hybridization studies have shown gains at 3q21–23 and 3q25–29, suggesting that *BCL6* is involved in some cases.

Trisomy 18

Trisomy 18 occurs in approximatey 30% of low-grade MALTomas. The genes involved are not well understood and the clinical relevance is unknown.

t(11;18)(q21;q21)

The most common translocation in gastric, intestinal, and lung MALTomas is t(11;18)(q21;q21) and occurs in 30% to 50% of cases.[10–14] It is associated with a worse prognosis, and is seen most frequently in *H. pylori*–positive, antibiotic-resistant, or advanced gastric MZL. It is not seen in nodal or splenic MZL, or in extranodal MZL with increased large cells or large-cell transformation. The genes involved are the apoptosis inhibitor gene (*API2*) on 11q21, a member of the IAP family with caspase-inhibitory functions, and the *MALT1* gene on 18q21, encoding a human paracaspase protein. The *API2*/*MALT1* translocation results in a chimeric transcript and a fusion protein, which leads to inhibition of apoptosis and confers a survival advantage on the MZL cells. FISH is the most sensitive and specific method for detection of this translocation; however, RT-PCR also is used for detection.

t(1;14)(p22;q32)

An apoptosis regulatory molecule, BCL10, is overexpressed as a result of the t(1;14)(p22;q32) translocation, which involves *BCL10* on 1p22 and the *IGH* gene on 14q32.[15] Advanced MALTomas sometimes have both *API2*/*MALT1* and *BCL10*/*IGH* translocations, and BCL10 appears to interact with API2/MALT1 fusion products to synergize activation of NF-κB, suggesting that they are part of a common pathway. Not many molecular laboratories offer testing for this translocation, because abnormal (nuclear) BCL10 protein expression can be used as a surrogate assay for this translocation. Nuclear BCL10 protein expression also may be seen in MZL with the t(11;18) translocation. The *BCL10*/*IGH* translocation is associated with a worse prognosis in low-grade MALTomas.

t(14;18)(q32;q21)

Recently described, the t(14;18)(q32;q21) translocation, detected by FISH, involves *IGH* on 14q32 and *MALT1* on 18q21.[16] It has been found in most liver MALTomas, as well as some cutaneous, ocular adnexal, and salivary gland lesions, but is rare in MALTomas of the stomach, intestine, lung, thyroid, and breast. It does not occur in nodal or splenic MZL. In a series of extranodal MALTomas, the *IGH*/*MALT1* translocation was not found to coexist with the *API2*/*MALT1* translocation but was sometimes found with trisomy 3 or trisomy 18 or both. The clinical significance of this translocation is not yet established. This translocation should not be confused with the cytogenetically identical, but molecularly distinct, t(14;18)(q32;q21)-*IGH*/*BCL2* translocation associated with follicular lymphoma.

Follicular Lymphoma

The t(14;18)(q32;q21) translocation, detected in about 80% of follicular lymphomas (FLs) at the time of initial diagnosis, is the cytogenetic hallmark of FL and juxtaposes the *BCL2* oncogene on 18q21with a J_H segment of *IGH* on

14q32. Overexpression of a normal *BCL2* protein results from this translocation and protects the cells from apoptosis. Molecular analysis for the *IGH*/*BCL2* translocation using FISH[17] (illustrated in Figure 32-2f) or PCR[18] is occasionally needed for confirmation of the initial diagnosis of FL in cases with equivocal histologic and immunologic findings. However, the vast majority of molecular testing in FL patients today is performed to evaluate the course of disease and the impact of therapy[19,20] (see section on *IGH*/*BCL2* MRD detection, below). Most *BCL2* breakpoints occur in the 3′ untranslated region, 60% in a 150 base pair (bp) span termed the major breakpoint cluster region (M-bcr) and 20% in the minor breakpoint cluster region (m-bcr). In the remaining cases, the breakpoint occurs outside of these areas, either in the variable cluster region (vcr) or the intermediate cluster region (icr). The insertion of variable numbers of random extra nucleotides at the breakpoint junction during crossover and the variability of the breakpoints in the *BCL2* fusion gene and the six J_H fragments result in considerable variation in the length of PCR products when testing is performed for this translocation. PCR fragment length usually ranges from 120 to 270 bp for the M-bcr, but occasional breakpoints as much as 800 bp downstream of the M-bcr region have been detected with standard M-bcr primers, resulting in PCR products of more than 1000 bp. The clustering of breakpoints on chromosome 18 and the high degree of sequence homology among the 3′ portions of the J_H segments make the *IGH*/*BCL2* translocation very amenable to PCR detection and a good PCR target for MRD detection.

Mantle Cell Lymphoma

Molecular testing may be performed initially to confirm the diagnosis of mantle cell lymphoma (MCL), as the diagnosis confers a very poor prognosis and mandates aggressive therapy.[21] Diagnostic testing generally targets the t(11;14)(q13q32) translocation, which juxtaposes the *BCL1*/cyclin D1/*PRAD1* gene on chromosome 11q13 with an *IGH* enhancer, resulting in overexpression of normal cyclin D1 protein and increased cell cycling. The translocation is not specific for MCL, as it also occurs in many multiple myelomas, but no diagnostic problems are caused by this overlap. Multiple methods have been used to detect the t(11;14), including karyotyping, Southern blot, PCR for t(11;14), RT-PCR for cyclin D1 messenger RNA (mRNA), and FISH.[22] The optimal diagnostic method has proven to be FISH (illustration in Figure 32-2d), which detects ≥90% of cases, even in paraffin-embedded biopsy tissue. PCR detection is less successful because the 11q13 breakpoints are widely distributed; approximately 30% to 50% are localized to a 1 kb DNA segment called the major translocation cluster (MTC), but the remaining translocations involve many different sites not easily detectable by PCR analysis. Quantitative RT-PCR assays on fresh or frozen tissue detect increased cyclin D1 mRNA in a high percentage of MCL,[23] but there is little application for this cumbersome RT-PCR assay when FISH is faster, is more specific, and can be performed on paraffin-embedded tissue. *ATM* deletions like those in CLL/SLL also are seen in many MCL cases, and progression or aggressive clinical behavior in MCL is reported to be associated with *P16* and *TP53* abnormalities. MCLs rarely have somatic mutation of IgV_H.

Diffuse Large B-Cell Lymphoma

Diffuse large B-cell lymphoma (DLBCL) is the most common type of lymphoma, and this category of the WHO classification represents another heterogeneous group of BCLs. The most common chromosomal abnormalities identified thus far in DLBCL are *BCL6* (3q27) abnormalities,[24,25] t(14;18)(q32;q21) *IGH*/*BCL2* translocations,[26] and *REL* (2p12–16) amplifications, but routine diagnostic testing of DLBCL usually is not performed for these chromosomal abnormalities. *IGH*/*BCL2* translocations occur in approximately 20% to 30% of DLBCL and are identical to those occurring in FL. The clinical significance of the *IGH*/*BCL2* translocations in DLBCL is not clear. *BCL6* translocations occur in about 40% of DLBCL, with numerous translocation partners. It is likely that the particular translocation partner dictates the prognostic significance of the *BCL6* translocation; those involving the *IGH* locus appear to portend a more favorable prognosis than other partner genes. Pathologic mutations of *BCL6* also occur in about 10% of DLBCL and may contribute to lymphomagenesis. Clinical testing usually is not done for *BCL6* at diagnosis because of the lack of confirmed prognostic significance. If done, Southern blotting is usually the method of choice.

A promising new method of identification of prognostic subgroups of DLBCL is microarray gene expression profiling (GEP).[27–29] By this method, a germinal center B-cell-like DLBCL subgroup (GCB) with a favorable prognosis and activated B-cell-like and unclassified DLBCL subgroups with a poorer prognosis have been identified. The GCB subtype of DLBCL expresses genes typical of normal germinal center B cells; *REL* amplification and t(14;18) translocations are specific to this subgroup. DLBCL is the most likely category of BCL in which miniarrays first will be used for detection of patterns of GEP at the time of initial diagnosis to stratify patients for different treatment protocols.

Burkitt Lymphoma

Burkitt lymphoma (BL) is a high grade BCL that is characteristically associated with the t(8;14)(q24;q32) translocation, involving the *MYC* gene on 8q24 and the *IGH* gene on 14q32 (illustrated in Figure 32-2e).[30] Occasional BLs demonstrate variant t(2;8)(p11;q24) or t(8;22)(q24;q11) translocations involving *MYC* and the kappa and lambda

light chain loci, respectively.[31] These translocations result in increased expression of MYC activity and markedly increased cell proliferation. The site of translocation of *MYC* is reported to vary between sporadic and endemic BL; in endemic disease, the 8q24 breakpoint occurs up to 300 kb 5′ from the coding region of the *MYC* gene, while sporadic BL characteristically involves a breakpoint in *MYC* exon 1.[32] There also is variable involvement of the *IGH* gene, with endemic cases involving the J region, while sporadic cases involve the C regions. The degree of molecular variability in these translocations is the reason they are not amenable to detection by standard PCR assays. Routine karyotyping, Southern blot, and FISH analysis using *MYC* probes are successful in detecting almost all *MYC* translocations in BL.

Multiple Myeloma

Recent studies have shown that molecular events involving the *IGH* gene at 14q32 are important and initiating events in multiple myeloma (MM).[33] Most of the translocations seen in MM are karyotypically silent but readily are identified by FISH studies utilizing molecular probes for 14q32 and its translocation partners. Translocations involving the *IGH* locus are believed to occur at the time of isotype class switching, in contrast to other B-cell malignancies. Multiple 14q32 translocation partners have been identified in MM, with different prognostic significance. The t(11;14)(q13;q32) translocation involves the cyclin D1 gene in ~25% of MMs and is associated with long survival. The t(4;14)(p16.3;q32) translocation portends the worst prognosis in MM and is seen in ~25% of cases; it involves both the fibroblast growth factor receptor-3 (*FGFR3*) and *MMSET* genes at 4p16.3. The t(14;16)(q32;q23) translocation involving the *MAF* gene conveys an intermediate prognosis, as does the t(6;14)(p25;q32) translocation involving the *MUM1/IRF4* gene. The t(14;16) occurs in ~20% of MM. Monosomy 13 is seen in ~50% of MM patients and is associated with an adverse prognosis.

Indications for Diagnostic and Prognostic Testing

Testing strategies for the molecular abnormalities of BCL is evolving as the number of genes important for diagnosis and prognosis of BCL increases almost daily. Thus, the molecular testing performed to confirm a primary diagnosis, predict BCL prognosis, or identify appropriate targeted therapies for subtypes of BCL often will not be the same as testing performed for detection of MRD after therapy. At initial diagnosis or relapse of BCL, the laboratory often needs to rapidly evaluate multiple genetic targets, but with only a low level of assay sensitivity. This is in distinct contrast to testing for residual disease after therapy, for which the *IGH* gene is used as the sole molecular target in most BCL subtypes and for which a very sensitive assay is needed. The specific chromosomal abnormalities and optimal molecular approaches for identification of clinically important diagnostic or prognostic factors for each BCL were presented in the previous section under the different BCL headings. The reader is referred to these sections and to the summary in Table 32-2, as these abnormalities only are discussed here in a general sense.

There are several major clinical indications for molecular testing in BCL: (1) to distinguish between a reactive and neoplastic proliferation of B cells, (2) to identify specific chromosomal abnormalities to aid in accurate subclassification of a BCL, (3) to predict the prognosis within a BCL subtype when alternative therapies related to specific molecular abnormalities are defined, and (4) to evaluate response to therapy by molecular analysis for detection of MRD. When there is a need to distinguish between reactive and neoplastic B-cell proliferations, assays for detection of B-cell clonality are most frequently used because they are applicable to virtually all BCLs. The *IGH* gene is the most common target evaluated for confirmation of B-cell clonality at diagnosis; the *IGH* gene is also the most common target for detection of persistent BCLs following therapy. The specific features of different molecular methods used for detection of B-cell clonality are discussed later (see "Assay Methods for Molecular Analysis of BCL"), as is the interpretation of the testing (see "Interpretation of Test Results").

Because substantial morphologic and immunologic overlap exists between different BCL subtypes, subclassification is sometimes difficult with routine histologic and phenotypic evaluation alone, and accurate subclassification may require evaluation for characteristic chromosomal abnormalities. Assays for specific BCL-associated chromosomal alterations have limited clinical utility in that they are applicable to only a subset of BCL patients, but they may provide essential information for accurate subclassification and prediction of prognosis. Assessment of BCL for molecular features related to adverse prognosis is being performed on a more frequent basis than in the past, due to not only the identification of increased numbers of significant molecular factors but also to the development of many new alternative therapies. Some prognostic factors for poor outcome appear to be relevant only to a particular subtype of BCL, but other factors, such as TP53 dysfunction, seem to be associated with poor prognosis or progressive disease in many BCL subtypes. Patients with BCL manifesting molecular markers of poor prognosis often are treated more aggressively at the time of initial diagnosis.

Maintaining multiple PCR-based molecular assays for detection of the many BCL-subtype-specific chromosomal abnormalities is impractical for clinical molecular laboratories, since assays other than for *IGH* gene rearrangements are applicable to only a small subset of BCL patients. The development of molecular probes against many of the

relevant genes involved in BCL for use in FISH assays has obviated this problem and provided a valuable alternative method for the rapid evaluation of multiple genes. The use of FISH assays for molecular evaluation of BCL at diagnosis and relapse has expanded remarkably over the past several years and will likely continue to expand as new molecular probes are identified. However, it must be stressed that the sensitivity of even the best FISH assay is not sufficient for detection of MRD in treated BCL patients. There is an increasing need for the development and implementation of standardized, sensitive, reproducible, and quantitative PCR testing for detection of MRD in BCL.

Minimal Residual Disease Detection

In the past, MRD testing was infrequently performed for most BCLs because the diseases were thought to be incurable in all but the very early stages. However, recent use of new therapeutic modalities such as monoclonal antibodies and vaccine therapies for treatment of BCL has resulted in improved clinical outcomes. Multiple large prospective studies have clearly demonstrated the high prognostic value of MRD monitoring in children with acute lymphoblastic leukemia,[34,35] and molecular remissions associated with prolongation of progression-free survival and possibly overall survival have been seen in some clinical trials of new therapies for BCL patients. With improvement of therapies, new methods for MRD detection are desperately needed as BCL patients achieve complete clinical, morphologic, and immunophenotypic responses. Unfortunately, the optimal methodology and timing for detection of MRD has yet to be determined for any subtype of BCL, and inexpensive and easily standardized methods for MRD detection of BCL have not been identified.

MRD detection in BCL is at once simpler than initial diagnostic or prognostic testing and more complex. The number of targets for which testing must be performed is limited in comparison to initial diagnostic testing, but the need for markedly increased sensitivity makes the testing more demanding for the laboratory and significantly increases the possibility of false-positive and false-negative results. Ideally, techniques used for MRD detection should have a sensitivity level to detect one BCL cell in 10^5 to 10^6 normal cells, be applicable to almost all patients with the disease, provide some quantification of the target, and be rapid, inexpensive, readily standardized, and disease specific. MRD assays also should have good intralaboratory and interlaboratory reproducibility. In reality, most commonly used molecular analyses for MRD detection in BCL, which target *IGH* rearrangements, do not meet many of these criteria. Assays closest to the ideal are quantitative or nested PCR detection of patient-specific, junctional region *IGH* gene rearrangements or *IGH*/*BCL2* chromosomal translocations. For most subtypes of BCL, either no translocation-associated molecular event is available for MRD testing or the recurrent translocations occur in too low a proportion of the BCL subtypes to be clinically useful.

IGH gene rearrangements are typically targeted for MRD detection in these diseases. The only sufficiently sensitive and specific method of testing for MRD detection using *IGH* gene rearrangements is patient or clone specific. This method takes advantage of the fingerprintlike sequences of the junctional regions of the rearranged *IGH* gene, which differ in length and composition for each B-cell clone. Multiple *IGH* PCR analyses of BCL biopsies at diagnosis or relapse are performed, the PCR products are cloned, and the junctional regions of the clonal *IGH* rearrangement are sequenced. The clone-specific *IGH* rearrangement sequence is then used for design of patient-specific PCR primers, which are subsequently used in standard or real-time PCR assays to assess MRD for that specific patient. The usual target for MRD detection is the VH-JH rearrangement, but the use of two MRD targets also has been recommended for reliable and sensitive MRD detection, if possible. Patient-specific *IGH* primers increase *IGH* PCR sensitivity up to 1000-fold (1 positive cell in 10^5 normal cells) compared to consensus primers for *IGH* gene rearrangements (1 positive cell in 10^2 to 10^3 normal cells). In patient-specific *IGH* PCR, the background signal from polyclonal B cells does not obscure the clonal PCR products.

If clonal junctional region sequences are not identified for generation of patient-specific primers, differences in the length of rearranged *IGH* junctional regions can be evaluated only by PCR product length assessment. The sensitivity of this technique, even in fresh material, is not sufficient for MRD detection. The consensus *IGH* primers used in standard *IGH* PCR assays result in less-efficient amplification than occurs with better-matched primers. Thus, standard *IGH* PCR tests can detect only one malignant cell in 10^2 to 10^3 normal cells. Real-time quantitative PCR (RQ-PCR) techniques using standard *IGH* primers do not increase the sensitivity of detection of *IGH* rearrangements to the optimal level for MRD detection but may provide some information about disease course over time. The combination of patient-specific *IGH* primers and RQ-PCR could make a major contribution to the achievement of standardized MRD detection in BCL.

For RQ-PCR MRD detection, a standard curve can be made from a dilution series of either a diagnostic lymphoma sample for patient-specific PCR or a cell-line dilution if patient-specific sequences are not known, and the amount of residual lymphoma cells relative to normal cells in a sample can be calculated by using this standard curve.[36] This technology may be used with fusion gene translocation targets as well as the *IGH* gene and is amenable to interlaboratory standardization. The determination of the trend in the quantitative numbers of residual BCL cells over time may provide important therapeutic information in the follow-up of BCL patients. Optimal primer sequences and methods have been reported by the European BIOMED-2 Concerted Action for RQ-PCR

detection of MRD in patients with precursor B leukemia/lymphoma and would likely be useful in BCL patients also. This type of MRD testing is currently performed mainly in specific research laboratories for patients in lymphoma clinical treatment trials. As therapeutic options increase for BCL patients, quantitative and sensitive MRD testing will become more routine.

Other than *IGH* PCR, the only other molecular target in common usage for MRD detection today is the t(14;18)(q32;q21) translocation in FL. Two types of PCR assays are commonly used for MRD detection in FL. Nested PCR assays have been used historically and remain the most sensitive method for MRD detection; nested PCR can detect one translocation-carrying cell in 10^5 to 10^6 normal cells (illustrated in Figure 32-3). However, RQ-PCR is less labor-intensive and lacks the risk of contamination of standard nested PCR, so many laboratories have already or are now trying to adapt this technology. Unfortunately, the analytical sensitivities of many reported RQ-PCR assays for *IGH*/*BCL2* translocations are less than that of nested PCR, at only one positive cell in 10^4 normal cells. Rare laboratories report an RQ-PCR analytical sensitivity of one positive cell in 10^5 normal cells. The clinical relevance of the difference in sensitivity between nested PCR and RQ-PCR is not clear, and it is possible that the measurement of serial quantitative levels of cells carrying the *IGH*/*BCL2* translocation, that is, the trend over time, may be more clinically important than having a higher level of sensitivity.

An underrecognized problem in interpretation of RQ-PCR analyses for MRD in FL is the inability of these assays to compare the PCR product size to that of the previous diagnostic or MRD FL sample. With a highly sensitive test capable of detecting MRD in FL, a rare benign *IGH*/*BCL2* translocation-carrying cell could produce a false-positive result if a comparison to the original clone is not made (Figure 32-3). This comparison is readily performed with nested PCR but requires substantial additional molecular analysis with RQ-PCR. However, using an RQ-PCR method for MRD detection and providing information about the trend over time for a given FL patient may obviate the necessity for this additional

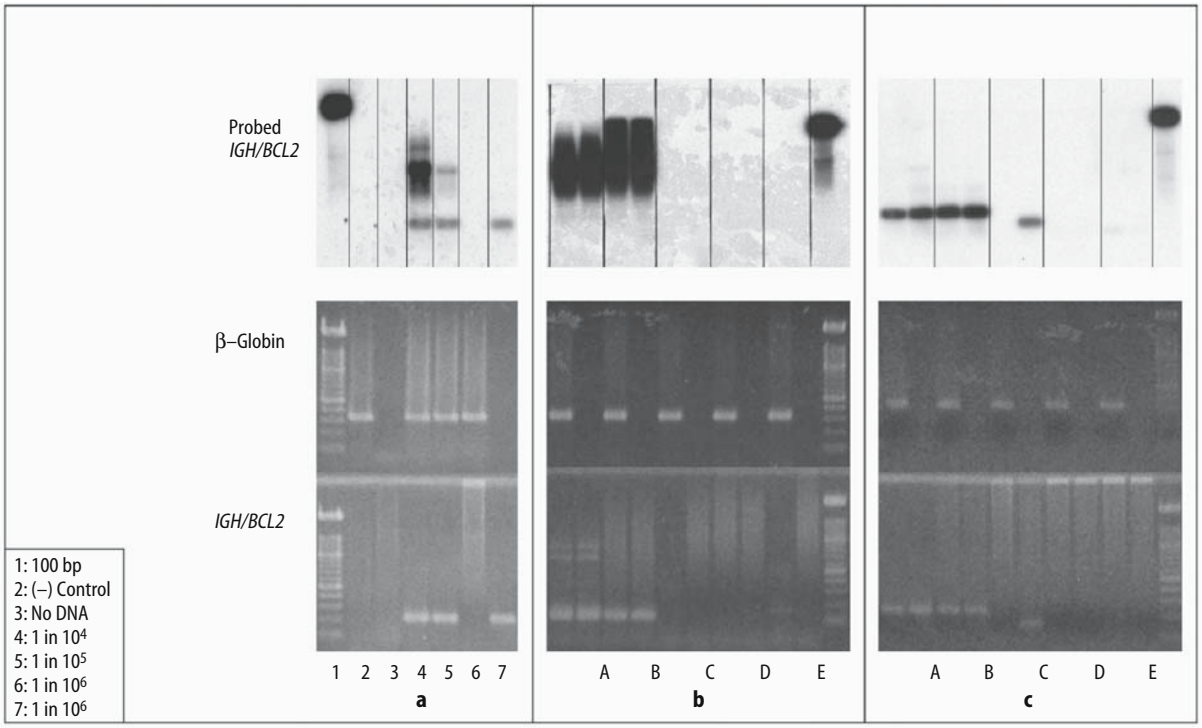


Figure 32-3. Illustration of PCR analysis for detection of MRD in sequential marrows from three patients with FL. Each marrow is subjected to PCR for β-globin housekeeping gene and *IGH*/*BCL2* M-bcr analysis; PCR products undergo gel electrophoresis and positive bands are detected by ethidium bromide staining (β-globin PCR results shown in top position, *IGH*/*BCL2* results in center position in each panel). A positive band is seen only if an appropriate product is detected; specificity of bands is confirmed by the use of multiple controls and also by probing of the Southern blotted *IGH*/*BCL2* PCR products with a *BCL2* probe (shown in bottom position in each panel). (a) Controls run in parallel with patient samples are critical for interpretation of this assay (see box insert). Positive controls are serial dilutions of an *IGH*/*BCL2*-positive cell line (lanes 4–7). The 10^6 dilution is run twice; as shown, a positive result is often detected in only one of these samples at the lower limit of analytical test sensitivity. (b, c) Parallel PCR analyses of five sequential marrows from two different FL patients. For each patient, sample A is pretreatment, B is 1–2 months after CHOP chemotherapy but before treatment with anti-CD20 monoclonal antibody (Ab), and marrows C, D, and E were obtained 2 months, 6 months, and 12 months after anti-CD20 monoclonal Ab therapy. All clinical samples from each patient were run in duplicate. The patient in (b) is positive for *IGH*/*BCL2* at diagnosis and after CHOP, but the marrow becomes PCR negative by 2 months after Ab therapy and remains negative at 6 and 12 months. The patient in (c) is also PCR positive at diagnosis and after CHOP, but has a positive band in one lane at 2 months after CD20 antibody therapy. This band is confirmed to represent a benign *IGH*/*BCL2* positive cell, not MRD, as it is not present in the duplicate sample and is clearly a different size from the patient's previously identified FL clone. This patient has achieved a molecular remission, but would be called positive by assays that do not evaluate the size of the product in comparison to the original FL. In contrast, the patient in (d) also has a positive result in only one duplicate at both 2 months and 6 months following Ab therapy, but these results clearly represent MRD; the abnormal bands are the same size as the original FL clone.

testing. Additional studies are needed to evaluate the clinical efficacy of MRD detection in FL patients in general, and to compare the relative clinical value of the different nested and RQ-PCR MRD detection methods.

PCR analysis using patient-specific primers is seldom performed in the United States but has been extensively used in Europe for molecular monitoring of MRD in patients with acute leukemias treated on different clinical trials. In addition, the Europe Against Cancer Program has reported a standardized approach for the most common fusion gene transcripts in ALL and BCL, developed with substantial government clinical trial support.

Assay Methods for Molecular Analysis of BCL

A variety of methods are employed today by clinical laboratories to detect the known molecular abnormalities in BCL. These include the "gold standard" Southern blot analysis (SBA), PCR analysis, RT-PCR analysis, and FISH. Each method has advantages and disadvantages for detection of molecular abnormalities in BCL.[37]

Southern Blot Analysis

The first widely used molecular method for detection of B-cell clonality was SBA, and it remains the molecular gold standard for detection of B-cell clonality. This is because the results of interlaboratory surveys testing for *IGH* clonality by SBA have consistently shown close to 100% agreement among participating laboratories, while marked interlaboratory variability has been seen when the same samples were analyzed by *IGH* PCR.[38] SBA also is useful in testing for genes with extensive breakpoint heterogeneity (*MYC*) and for genes with multiple translocation partners (*IGH*, *BCL6*). SBA may be used to identify comigrating translocation partners, such as the *IGH* and *BCL2* genes involved in t(14;18)(q32;q21) translocations (illustrated in Figure 32-4).

SBA for BCL detects changes that occur in a gene's germline structure during Ig gene rearrangement, which delete or create restriction endonuclease sites and lead to changes in the bands on the Southern blot. SBA of cells with nonrearranged Ig genes produces bands of characteristic size for a given restriction enzyme (germline bands). Cells with rearrangements or mutations in restriction sites will have bands of different sizes from the germline bands due to alterations of the relative positions of restriction sites (rearranged bands)(Figure 32-4).

The criteria existing for interpretation of antigen receptor SBA require that rearrangements are seen with two of three restriction enzymes for a positive monoclonal interpretation. Up to two rearrangement bands can be seen in any one digest lane, which can represent two rearranged *IGH* alleles (if the first rearrangement was unsuccessful and the second allele was also rearranged), or can be due to the creation of a restriction site within the region of the probe so different ends of the probe hybridize to different fragments of genomic DNA. Complete digestion of the genomic DNA is essential to proper interpretation, since partial digestion can create bands at nongermline positions.

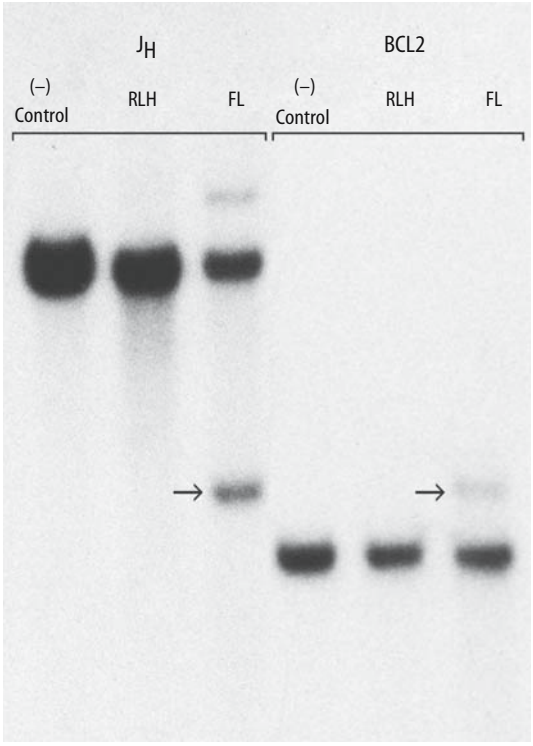


Figure 32-4. Parallel autoradiographs of SBA of *IGH* (J_H) and *BCL2* gene rearrangements in a patient with follicular lymphoma (FL). The negative control is a T-lymphoblastic cell line and shows a germline configuration of the *IGH* gene. A reactive lymphoid hyperplasia (RLH) shows a polyclonal smear of variably sized *IGH* rearrangements in normal B cells, while the biopsy from the patient with FL shows a decrease in the polyclonal background and the presence of two clonal *IGH* rearrangement bands, one above and one below the germline bands seen in all three specimens. The same membrane was stripped and reprobed with a *BCL2* probe, and the arrows show the comigrating *IGH* and *BCL2* rearrangement bands in the FL, indicating a translocation between these two genes.

The probes used for *IGH* SBA are usually against the *IGH* joining region (J_H), rather than the constant region (Cμ), because heavy-chain switching deletes Cμ but not J_H. For light-chain genes, Jκ is also less frequently deleted than Cκ in λ-expressing BCL. There is seldom a clinical use for Cλ probes. Rare BCL will lack a detectable J_H rearrangement on *IGH* SBA, but these will typically show Cκ rearrangement. Polyclonal B cells, each with a unique *IGH* rearrangement, produce a weak background smear of different-sized rearrangement bands on the SBA (Figure 32-4). The sensitivity of SBA for a monoclonal population is approximately 3% to 5% of the total cell population.

Successful SBA also requires high-molecular-weight DNA and thus can be reliably performed only on fresh or frozen tissue or cells; DNA from paraffin-embedded tissue is usually too degraded for successful SBA. SBA also requires large amounts of DNA, as 5 to 10 μg of DNA is used

per restriction endonuclease. Typical *IGH* SBA using three restriction endonucleases therefore requires 15 to 30 µg of DNA, which can be obtained from 10^6 cells or approximately 0.5 cm^3 of fresh or frozen tissue. Because tissue requirements are more strenuous and SBA is more labor-intensive, usually taking 2 to 3 weeks to obtain results, many clinical laboratories no longer perform SBA or use SBA infrequently, only as a backup for equivocal *IGH* PCR assays.

Polymerase Chain Reaction Analysis

Unlike SBA, which compares germline and rearranged alleles of the genes being analyzed, PCR amplifies rearranged alleles only because germline alleles have too great a distance between the PCR primer binding sites to allow for amplification. PCR can be performed on fresh, frozen, or paraffin-embedded tissue, as well as microdissected and cytology specimens, because it is not as affected by DNA fragmentation as SBA. PCR also requires much less DNA or RNA and is more rapid than SBA. Thus, PCR has versatility not achieved by SBA.

PCR amplification of genomic DNA is used for detection of AgR rearrangements as well as many BCL-associated translocations. In genes with numerous translocation partners, however, FISH is more useful than PCR. Translocations involving variable sites over an area of a chromosome too large for PCR may be amenable to RT-PCR. RT-PCR takes advantage of the splicing out of introns that occurs in the generation of mRNA. RT-PCR cannot reliably be performed on paraffin-embedded tissue because intact RNA is difficult to extract from paraffin-embedded tissues.

PCR detection of IgV_H rearrangements uses V- and J-region primers and relies on the V, D, and J segments being brought into close proximity during rearrangement so that the PCR reaction can amplify across these segments. The closest V and J segments are too far apart in the germline configuration for PCR amplification to occur. *IGH* clonality analysis by PCR uses consensus primers designed to anneal to conserved *IGH* V- and J-region sequences. For *IGH* PCR, one J-region primer will recognize all six J segments because there is a single well-conserved region among the six J regions, but there is no single V-region primer that will recognize all V segments. V regions have three more highly conserved framework regions (FR I, II, and III) and intervening highly variable sequences called complementarity-determining regions (CDR I and II). Because the FR sequences are more conserved across the different V regions, the V-region primers are designed to bind to FR sequences. Most commonly, FR III–region primers are used because the FR III region is closest to the J region in the rearranged state, thus resulting in a smaller PCR product that results in more efficient amplification. FR III– and J-region primers amplify the highly variable V-D junction (CDR III) and detect 60% to 70% of rearrangements. The sequence of the FR III V-region primer also affects the detection rate. One interlaboratory comparison showed a difference of from 55% to 70% based on the specific sequence used for the V-region primer.[39] The addition of a second PCR amplification using a FR II V-region primer with a J-region primer also increases the detection rate of the test significantly (80% to 90%).

IGH PCR amplifies any rearranged *IGH* allele, such that there is a background signal from the polyclonal B cells present in a specimen that may obscure the signal from a monoclonal B-cell population. Numerous strategies are used for PCR product detection, most commonly gel electrophoresis with colorimetric, fluorescent, or chemiluminescent labeling. Capillary electrophoresis (CE) with fluorescently labeled primers provides slightly enhanced sensitivity and higher throughput, with improved resolution, and is becoming the system of choice for many laboratories. Interpretation of CE results may be more objective than the standard PCR gel format for demonstration of B-cell clonality. A polyclonal B-cell population produces a smear or ladder on a gel or multiple small peaks on a CE instrument printout, while monoclonal B-cell populations produce one or two distinct bands on a gel, or one or two sharp peaks on CE (Figure 32-5). One criterion developed for determining whether a peak is monoclonal compared to the background polyclonal peaks is that the distinct peak should be more than two to three times the height of the adjacent polyclonal peaks. Specific criteria for interpretation of *IGH* PCR results have not been developed, especially with regard to the number of bands allowed per reaction. Sequence analysis of *IGH* PCR products can be done for identification of patient-specific sequences, which may then be used to design patient-specific primers for subsequent patient-specific PCR analyses for detection of MRD following therapy.

Fluorescence In Situ Hybridization

FISH is a very useful technique for detection of targeted BCL-associated chromosomal abnormalities and can detect both structural and numerical chromosomal abnormalities. FISH overcomes one of the problems with routine cytogenetics on BCL samples, that is, the need for metaphase cells, as it can be done with either metaphase or interphase preparations. Genomic probes for the breakpoints of many different BCL translocations and for gene deletions are now readily available. FISH assays are particularly useful in detection of chromosomal translocations in which the breakpoints are widely dispersed, because FISH probes are much larger than probes and primers used in SBA and PCR. For example, FISH probes can detect almost all of the *MYC* 8q24 breakpoints in the t(8;14)(q24;q32) translocations associated with BL. FISH detects some genetic abnormalities that are karyotypically silent.

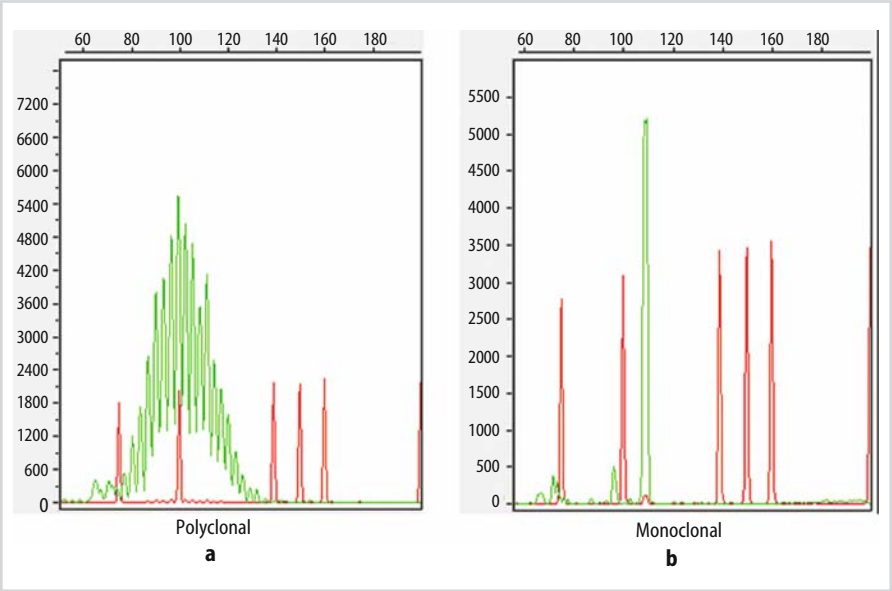


Figure 32-5. Capillary electrophoresis (CE) printouts of two *IGH* PCR analyses for B-cell clonality. (a) A polyclonal B-cell population produces multiple small peaks on the CE printout; to be considered monoclonal, a peak should be at least double the height of the peaks from the background polyclonal B cells. (b) A monoclonal B-cell population produces one sharp peak on CE (or two in some cases). The interpretation of the CE printouts is probably more objective than interpretation of routine agarose or polyacrylamide gels. (Black peaks are size marker standards.)

Interpretation of Test Results

Pitfalls in Interpretation of BCL Testing

It is important to remember that molecular genetic studies of BCL should always be interpreted in conjunction with routine histologic and immunophenotypic information and always with the knowledge of the range of genotypic features found in normal and hyperplastic lymphoid populations. Interpretation in the larger clinical and histopathologic context will avoid erroneous interpretations with the potential for adverse clinical consequences. Our knowledge of the genetic events occurring in benign and malignant lymphoid populations is rapidly expanding in this era of genomics and proteomics, such that the interpretation of a particular genetic finding may well change as the field evolves. There are a number of potential pitfalls in the interpretation of molecular analyses for BCL that may result in false-positive or false-negative results.

Molecular Abnormalities in Benign Lymphoid Proliferations

The presence of a clonal B-cell population, by itself, does not establish a diagnosis of B-cell malignancy. Small B-cell clones may be detected by *IGH* PCR and SBA in benign lymphoid hyperplasias in the absence of other criteria for malignancy.[40–42] This occurs in the setting of immune deficiencies, autoimmune diseases, and immunosuppression, and reinforces the critical necessity for interpretation of these tests for B-cell clonality in conjunction with clinical, morphologic, and immunophenotypic information. Patients with immune dysfunction have an increased risk of non-Hodgkin lymphoma (NHL), in particular BCL, but many of these patients will never develop lymphomas, even without correction of the abnormal immune status.

Other types of BCL-associated genetic alterations have been described in benign settings and could result in a false-positive result. Rare B cells carry a t(14;18)(q32;q21) *IGH*/*BCL2* translocation in normal individuals.[43–46] PCR assays capable of detecting one *IGH*/*BCL2* translocation-carrying cell in 10^5 to 10^6 normal cells will be positive in up to half of the tissue biopsies, bone marrow aspirates, and peripheral blood specimens from normal individuals. To date, there is no evidence that these individuals are at higher risk for development of FL. Diagnostic tests for FL must be designed with a sensitivity that avoids detection of the *IGH*/*BCL2* translocation in normal individuals.

Lineage Infidelity or Promiscuity

A potential pitfall in interpretation of molecular tests for AgR rearrangement in BCL is the occurrence of so-called lineage infidelity or promiscuity. This is particularly common in precursor B-cell malignancies but may also be seen in malignancies of mature B cells. The majority of precursor B-cell malignancies also show rearrangement of the *TCRG* gene. Only 5% to 10% of mature B-cell malignancies show *TCR* gene rearrangements, a finding that must be remembered to avoid erroneous conclusions about cell lineage.

Oligoclonality and Clonal Evolution

An oligoclonal pattern is defined as the presence of more bands than would be seen from a single monoclonal cell population, specifically, more than two bands per PCR. An oligoclonal pattern can be seen when there is expansion of several B-cell clones in reactive processes or in immunocompromised individuals with a reduced B-cell repertoire, as well as in specimens with very few B cells. Oligoclonal patterns must be interpreted with care. In addition, rare

patients will have two separate monoclonal cell populations; however, specific documentation of two separate B-cell populations by flow cytometry or immunophenotyping must be demonstrated to support the interpretation of two monoclonal populations.

During the course of the disease, the monoclonal population may develop additional *IGH* rearrangements that may alter the size of bands seen by *IGH* PCR or Southern blot (clonal evolution). If *IGH* PCR is used for MRD testing, clonal evolution may change or eliminate the diagnostic PCR band(s) and lead to false-negative results. *IGLK* gene rearrangements involving the kappa-deleting element (κde) appear to be more stable than *IGH* gene rearrangements, possibly because they usually delete the two enhancers for the *IGLK* gene.[45] However, these *IGLK*–κde rearrangements are present in only a minority of BCL, limiting their clinical utility.

Inadequate Test Sensitivity

Molecular assays generally do not achieve 100% analytical sensitivity, so a false-negative result must always be considered. SBA will not detect clonal B-cell populations representing less than 3% to 5% of total nucleated cells in a fresh or frozen sample. Standard *IGH* PCR of fresh or frozen tissue is slightly more sensitive than SBA, detecting one clonal cell in 10^2 to 10^3 normal cells. *IGH* PCR sensitivity in paraffin-embedded tissue is much lower, sometimes as low as 40% to 60%, and is highly subject to the conditions used for tissue fixation and processing. Surprisingly, many laboratories performing *IGH* PCR for clinical purposes today do not know the sensitivity level of their tests. It is imperative that every assay for *IGH* clonality include appropriate sensitivity controls, and it is essential to know the necessary level of sensitivity for the target of interest in each case. A test sensitivity of one positive cell in 20 normal cells (5%) is adequate for diagnostic or prognostic testing of almost all BCL, but there is no rationale for performing SBA or standard *IGH* PCR for MRD detection when the assay sensitivity is 5%. The minimal sensitivity for a test offered for MRD detection should be one positive cell in 10^4 normal cells, and a sensitivity of one positive cell in 10^5 to 10^6 normal cells is achievable and desirable for some assays. It is important to remember the difference in sensitivity of the *IGH* AgR and *IGH*/*BCL2* PCR assays; in most laboratories, the *IGH* AgR assay has a much lower sensitivity than the *IGH*/*BCL2* assay, which may lead to discrepant results when both tests are performed for the same case.

Poor Specimen Quality

The quality of extracted DNA or RNA is not usually a problem with fresh or frozen tissue or cells, if viable tissue has been obtained. However, poor DNA quality is a consistent problem with PCR assays of paraffin-embedded tissue, and this tissue is not suitable for RNA extraction and testing in most cases. Different fixatives will affect the success of DNA extraction; formalin or zinc-formalin fixation allows adequate DNA preservation for PCR in many cases, but it is most often not possible to obtain adequate DNA for PCR from most paraffin-embedded tissue fixed with mercury-based (B5), Zencker, or Bouin fixatives.

Primer Failure

Clearly, the most common cause of false-negative *IGH* PCR results for BCL is the failure of primer binding due to alteration of the sequence of the region amplified by the primer following somatic hypermutation of IgV_H regions in postgerminal center BCL. The false-negative rate with *IGH* PCR ranges from <5% (MCL) to >50% (MZL, FL, plasma cell dyscrasias). These BCLs that give false-negative results by PCR almost always will have clonal *IGH* rearrangements by SBA, which is used as a follow-up test for BCL with unexpectedly negative *IGH* PCR results in some clinical molecular laboratories. The possibility of a false-negative *IGH* PCR result in these postgerminal center BCLs must always be considered in the interpretation of *IGH* PCR results.

Unusual or Complex *IGH* Rearrangements

Occasional BCLs have very complex rearrangements of *IGH* and other associated genes that may confound standard *IGH* PCR testing. This has been seen most frequently in FL, but similar genetic events undoubtedly occur with other *IGH* translocation partners. For example, most FLs have a single *IGH*/*BCL2* rearrangement involving one chromosome 14 *IGH* allele, which is then unable to participate in a functional *IGH* rearrangement. An additional clonal rearrangement of the untranslocated *IGH* allele on the other chromosome 14 usually occurs and can be detected by *IGH* PCR. However, occasional FLs will have translocations of both *IGH* alleles to chromosome 18, resulting in the absence of a detectable clonal *IGH* rearrangement by PCR and often producing rearrangements of both the *BCL2* M-bcr and m-bcr regions.

Laboratory Issues

The variability of results of PCR testing for *IGH* rearrangements and *IGH*/*BCL2* translocations among clinical molecular laboratories has been documented multiple times over the past several years by the proficiency testing surveys of the College of American Pathologists,[47] as well as by other independent interlaboratory surveys.[48] It is of particular concern to clinicians, and to those setting up and running clinical treatment trials, that multiple interlaboratory surveys have documented a high level of false-negative results for these PCR assays, as well as a lower but still significant number of false-positive results. The

reasons for the lack of reproducibility among laboratories are multiple and not easily overcome. Some strategies for PCR testing and interpretation that have been shown to improve testing performance are discussed below.

Use of Appropriate Controls

Appropriate controls include positive, negative, sensitivity, and no DNA reactions. Many laboratories do not know or overestimate the diagnostic sensitivity of their *IGH* and *IGH*/*BCL2* PCR assays, possibly because the sensitivity of the assay is established at the time of test validation, but sensitivity controls are not included in every run. Interlaboratory surveys also indicate that some laboratories do not use adequate controls to allow accurate interpretation of results. An additional important PCR control, particularly with paraffin-embedded tissue, is the amplification of a nonrearranging gene to document the presence of adequate DNA in the tested sample and the absence of inhibitors of the Taq enzyme, to prevent reporting of a false-negative result.

Amplification of Duplicate Aliquots of DNA

In small samples or samples with few B cells, there are not always sufficient B cells to produce a polyclonal background. In this setting, PCR can amplify rare B cells, producing a misleading monoclonal or oligoclonal pattern of amplification for *IGH* PCR assays. Likewise, a sensitive *IGH*/*BCL2* assay capable of detecting MRD will also amplify a rare benign translocation-carrying cell, as previously discussed. Overinterpretation of a single band on *IGH* and *IGH*/*BCL2* PCR assays by laboratories not performing the assays in duplicate likely represents the most common cause of false-positive results for these assays. With amplification of duplicate aliquots side by side, only clearly visible bands of identical size in each duplicate should be called positive, not nonreproducible or weak bands. Bands seen in analyses of specimens containing rare B cells, or rare normal translocation-carrying cells, will not be reproducible.

Adherence to Strict Criteria for Interpretation and Reporting

The interpretation of assays for *IGH* rearrangements, the *IGH*/*BCL2* translocation, and other BCL-associated translocations must be performed with caution. It is apparent that even the most commonly performed PCR assays for *IGH* and *IGH*/*BCL2* have an unacceptably high level of erroneous results. Since no universal standards for interpretation of these assays have been delineated thus far, PCR results should be very carefully interpreted and reported, preferentially in conjunction with clinical history, morphology, and immunophenotyping data. Reports should clearly indicate that a false-negative result is a possibility. Use of SBA as a backup for negative or equivocal *IGH* PCR results is helpful in eliminating false-negative and false-positive *IGH* PCR results, but it may also substantially delay obtaining a final result. Clinical molecular laboratories must develop mechanisms to avoid reporting irrelevant *IGH*/*BCL2* translocations as MRD. Labor-intensive methods such as sequencing of PCR products and fluorescence-based PCR–single-strand conformation polymorphism (SSCP) analysis are not practical for routine clinical practice, but most false-positive *IGH*/*BCL2* PCR results can be avoided by performing the assay in duplicate. Parallel analysis of the original BCL with a posttreatment sample being evaluated for MRD will confirm that the rearrangement is detectable by the specific PCR assay and will allow comparison of PCR amplification product sizes, which is very helpful in ruling out false-negative and false-positive results (Figure 32-3). False-negative results for MRD detection occur most frequently with inadequate test sensitivity. Laboratories performing testing for MRD must offer a sensitivity level of 1 in 10^4 to 10^6.

Conclusion

Clinical demand for genetic characterization of BCL at the time of primary diagnosis has intensified as our understanding of NHL biology has increased and therapies targeting specific BCL have been developed. Today, genetic features of NHL are used to aid in rendering an accurate primary diagnosis, to predict prognosis, to evaluate for MRD after therapy, and even to help determine optimal therapy. The rapid acquisition of knowledge about BCL biology and translation of this knowledge into targeted therapeutic strategies is occurring today at an unprecedented rate. New testing strategies and technologies are essential for clinical laboratories to keep up with the growing information on BCL obtained from research methods such as gene expression profiling, comparative genomic hybridization, and proteomics.

Clinical laboratories doing molecular testing on BCL today are faced with two daunting tasks. First and foremost is the necessity of expanding test menus to meet the increasing clinical demand for testing for new genetic markers in NHL. Expanding test menus to meet clinical needs will require technical advances, as the current technology in clinical molecular laboratories does not allow rapid screening of BCL for a broad panel of relevant genes at a reasonable cost. Permutations of research methods are under development and show promise in alleviating the technological bottleneck preventing translation of this new genetic knowledge to the clinical laboratory. However, the lack of a reasonable level of reimbursement for molecular testing in general is a major roadblock to successful implementation of new techniques for molecular testing for important new genetic markers in BCL.

The second task facing clinical molecular laboratories, which is equally important to the credibility of the field, is the urgent need for standardization of molecular testing methods to eliminate the wide variability of results among molecular laboratories. Laboratories in Europe are ahead of those in the United States in standardizing molecular testing for lymphoma and leukemia clinical treatment trials, partly because of the substantial government support they have had to accomplish this task and partly because of their geography. It is likely that similar governmental support, probably linked to mandated proficiency testing, will be needed to achieve a comparable level of laboratory standardization of molecular testing in the United States. If molecular laboratories in the United States do not voluntarily develop a format for standardization of clinical molecular testing for leukemias and lymphomas, it is likely that hematopathology molecular testing in the future will be confined to large central reference laboratories.

The role of the molecular hematopathologist is clearly changing and expanding. As genomic profiling, comparative genomic hybridization, and other genomewide screening research identifies more new genes and biologic mechanisms that play important roles in the therapeutic responsiveness and overall survival of lymphoma patients, clinicians are beginning to incorporate this new knowledge into the selection of more targeted therapies. Clinical molecular laboratories performing hematopathology molecular diagnostic testing today must provide input at the time of initial diagnosis and during the course of treatment to ensure the performance of appropriate molecular testing and the accurate interpretation of results. The testing performed to confirm initial diagnosis and evaluate for prognostic genetic abnormalities in BCL is not the same testing that will be performed for detection of MRD after therapy. Furthermore, the lines between classical cytogenetics laboratories and molecular laboratories are blurring, due to the marked increase in the use of molecular probes for FISH analysis. Molecular pathologists increasingly will need to be very familiar with both lymphoma biology and different therapies for lymphoma, and will have to interact more closely with clinicians than ever before to accurately interpret the results of molecular tests.

References

1. Jaffe ES, Harris NL, Stein H, Vardiman JW, eds. *Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues (World Health Organization Classification of Tumors).* Lyon, France: IARC Press; 2001.
2. Arber DA. Molecular diagnostic approach to non-Hodgkin's lymphoma. *J Mol Diagn.* 2000;2:178–190.
3. Pan L, Cesarman E, Knowles DM. Antigen receptor genes: structure, function, and genetic analysis of their rearrangements. In: Knowles D, ed. *Neoplastic Hematopathology*, 2nd ed. Philadelphia: Lippincott, Williams and Wilkins; 2001:307–328.
4. Korsmeyer SJ, Hieter PA, Ravetch JV, et al. Developmental hierarchy of immunoglobulin gene rearrangement in human leukemic pre-B-cells. *Proc Natl Acad Sci U S A.* 1981;78:7096–7100.
5. Lin D, Sherrington PD, Dennis M, Matrai Z, Cawley JC, Pettitt AR. Relationship between p53 dysfunction, CD38 expression, and IgVH mutation in chronic lymphocytic leukemia. *Blood.* 2002;100:1404–1409.
6. Krober A, Seiler T, Benner A, et al. VH mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. *Blood.* 2002;100:1410–1416.
7. Rosenwald A, Alizadeh AA, Widhopt F, et al. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J Exp Med.* 2001;194:1639–1647.
8. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A.* 2002;99:15524–15529.
9. Idia S, Rao PH, Nallasivam P, et al. The t(9;14)(p13;q32) chromosomal translocation associated with lymphoplasmacytoid lymphoma involves the PAX-5 gene. *Blood.* 1996;88:4110–4117.
10. Baens M, Maes B, Steyls A, et al. The product of the t(11;18), an API2-MLT fusion, marks nearly half of gastric MALT type lymphomas without large cell proliferation. *Am J Pathol.* 2000;156:1433–1439.
11. Remstein ED, James CD, Kurtin. Incidence and subtype specificity of API2-MALT1 fusion translocation in extranodal, nodal, and splenic marginal zone lymphoma. *Am J Pathol.* 2000;156:1183–1188.
12. Ye H, Leu H, Raderer M, et al. High incidence of t(11;18)(q21;q21) in Helicobacter pylori-negative gastric MALT lymphoma. *Blood.* 2003;101:2547–2550.
13. Liu H, Ye H, Dogan A, et al. t(11;18)(q21;q21) is associated with advanced mucosa-associated lymphoid tissue lymphoma that expresses nuclear BCL10. *Blood.* 2001;98:1182–1187.
14. Okabe M, Inagaki H, Ohshima, et al. API2-MALT1 fusion defines a distinctive clinicopathologic subtype in pulmonary extranodal marginal zone B-cell lymphoma of mucosa-associated lymphoid tissue. *Am J Pathol.* 2003;162:1113–1122.
15. Zhang Q, Sieber R, Yan M, et al. Inactivation mutations and overexpression of BCL10, a caspase recruitment domain-containing gene in MALT lymphoma with t(1;14)(p22;q32). *Nat Genet.* 1999;22:63–68.
16. Streubel B, Lamprecht A, Dierlamm J, et al. t(14;18)(q32;q21) involving IGH and MALT1 is a frequent chromosomal aberration in MALT lymphoma. *Blood.* 2003;101:2335–2339.
17. Jiang F, Lin F, Price R, et al. Rapid detection of *IgH*/BCL2 rearrangement in follicular lymphoma by interphase fluorescence in situ hybridization with bacterial artificial chromosome probes. *J Mol Diagn.* 2002;4:144–149.
18. Albinger-Hagyi A, Hochreutener B, Abdou MT, et al. High frequency of t(14;18) translocation breakpoints outside of major breakpoint and minor cluster regions in follicular lymphomas: improved polymerase chain reaction protocols for their detection. *Am J Pathol.* 2002;160:823–832.
19. Gribben JG, Neuberg D, Freedman AS, et al. Detection by polymerase chain-reaction of residual cells with the *bcl*-2 translocation is associated with increased risk of relapse after autologous bone-marrow transplantation for B-cell lymphoma. *Blood.* 1993;81:3449–3457.
20. Lopez-Guillermo A, Cabanillas F, McLaughlin P, et al. Molecular response assessed by PCR is the most important factor predicting failure-free survival in indolent follicular lymphoma: update of the MDACC series. *Ann Oncol.* 2000;11(suppl 1):137–140.
21. Weisenburger DD, Vose JM, Greiner TC, et al. Mantle cell lymphoma: a clinicopathologic study of 68 cases from the Nebraska Lymphoma Study Group. *Am J Hematol.* 2000;64:190–196.
22. Belaud-Rotureau MA, Parrens M, Dubus P, et al. A comparative analysis of FISH, RT-PCR, PCR, and immunohistochemistry for the diagnosis of mantle cell lymphomas. *Mod Pathol.* 2002;15:517–525.
23. Hui P, Howe JG, Crouch J, et al. Real-time quantitative RT-PCR of cyclin D1 mRNA in mantle cell lymphoma: comparison with FISH and immunohistochemistry. *Leuk Lymphoma.* 2003;44:1385–1394.
24. Barrans SL, O'Connor SJ, Evans PA, et al. Rearrangement of the BCL6 locus at 3q27 is an independent poor prognostic factor in

nodal diffuse large B-cell lymphoma. *Br J Haematol.* 2002;117:322–332.
25. 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;98:945–951.
26. Barrans SL, Evans PA, O'Connor SJ, et al. The t(14;18) is associated with germinal center-derived diffuse large B-cell lymphoma and is a strong predictor of outcome. *Clin Cancer Res.* 2003;9:2133–2139.
27. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature.* 2000;403:503–511.
28. Shipp MA, Ross KN, Tamayo P, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med.* 2002;8:68–74.
29. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large B-cell lymphoma. *N Engl J Med.* 2002;346:1937–1947.
30. Neri A, Barriga F, Knowles DM, et al. Different regions of the immunoglobulin heavy-chain locus are involved in chromosomal translocations in distinct pathogenetic forms of Burkitt lymphoma. *Proc Natl Acad Sci U S A.* 1988;85:2748–2752.
31. Konigsberg R, Zojer N, Ackermann J, et al. Predictive role of interphase cytogenetics for survival of patients with multiple myeloma. *J Clin Oncol.* 2000;18:804–812.
32. Bellan C, Lazzi S, De Falco G, et al. Burkitt's lymphoma. New insights into molecular pathogenesis. *J Clin Pathol.* 2003;56:188–192.
33. Braziel RM, Arber DA, Slovak ML, et al. The Burkitt-like lymphomas: a Southwest Oncology Group study delineating phenotypic, genotypic, and clinical features. *Blood.* 2001;97:3713–3720.
34. Moreau P, Facon T, Leieu X, et al. Recurrent 14q32 translocations determine the prognosis of multiple myeloma, especially in patients receiving intensive chemotherapy. *Blood.* 2002;100:1579–1583.
35. Eckert C, Landt O, Seeger K, et al. Potential of LightCycler technology for quantification of minimal residual disease in childhood acute lymphoblastic leukemia. *Leukemia.* 2000;14:316–323.
36. Szczepanski T, Willemse MJ, Brinkhof B, et al. Comparative analysis of Ig and TCR gene rearrangements at diagnosis and at relapse of childhood precursor-B-ALL provides improved strategies for selection of stable PCR targets for monitoring of minimal residual disease. *Blood.* 2002;99:2315–2323.
37. Gabert J, Beillard E, van der Veldon VHJ, et al. Standardization and quality control studies of "real-time" quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia—A Europe Against Cancer Program. *Leukemia.* 2003:1–40.
38. O'Leary TJ, Brindza L, Kant JA, et al. *Immunoglobulin and T-Cell Receptor Gene Rearrangement Assays: Approved Guidelines.* National Committee for Clinical Laboratory Standards (NCCLS). Vol 15. 1995:1–30.
39. Bagg A, Braziel RM, Arber DA, et al. Immunoglobulin heavy chain gene analysis in lymphoma: a multi-center study demonstrating the heterogeneity of performance of polymerase chain reaction assays. *J Mol Diagn.* 2002;4:81–89.
40. Elenitoba-Johnson KSJ, Bohling SE, Mitchell RS, et al. PCR analysis of the immunoglobulin heavy chain gene in polyclonal processes can yield pseudoclonal bands as an artifact of low B cell number. *J Mol Diagn.* 2000;2:92–96.
41. Nihal M, Mikkola D, Wood GS. Detection of clonally restricted immunoglobulin heavy chain gene rearrangements in normal and lesional skin. *J Mol Diagn.* 2002;2:5–10.
42. Lee SC, Berg KD, Racke FK, et al. Pseudo-spikes are common in histologically benign lymphoid tissues. *J Mol Diagn.* 2000;2:145–152.
43. Summers KE, Goff LK, Wilson AG, et al. Frequency of the Bcl-2/*IgH* rearrangement in normal individuals: implications for the monitoring of disease in patients with follicular lymphoma. *J Clin Oncol.* 2001;19:420–424.
44. Limpens J, De Jong D, van Krieken JHJM, et al. Bcl-2/JH rearrangements in benign lymphoid tissue with follicular hyperplasia. *Oncogene.* 1991;6:2271–2276.
45. Limpens J, Stad R, Vos C, et al. Lymphoma-associated translocation t(14;18) in blood B cells in normal individuals. *Blood.* 1995;85:2528–2536.
46. Van der Velden VHJ, Willemse JM, van der Schoot CE, et al. Immunoglobulin kappa deleting element rearrangements in precursor B acute lymphoblastic leukemia are stable targets for detection of minimal residual disease by real-time quantitative PCR. *Leukemia.* 2002;16:928–936.
47. Hsi ED, Tubbs RR, Lovell, et al. Detection of bcl-2/JH translocation by polymerase chain reaction: a summary of the experience of the molecular oncology survey of the College of American Pathologists. *Arch Pathol Lab Med.* 2002;126:902–908.
48. Johnson PWM, Swinbank K, MacLennan S, et al. Variability of polymerase chain reaction detection of the bcl-2-IGH translocation in an international multicentre study. *Ann Oncol.* 1999;10:1349–1354.

Chapter 33

T-Cell Lymphomas

Iris Schrijver and Daniel A. Arber

Molecular Basis of Disease

There are two broad genetic changes that can be studied in T-cell lymphomas, namely, T-cell receptor (*TCR*) gene rearrangements and chromosomal translocations, deletions, or additions.

TCR Gene Rearrangements

Pluripotent bone marrow stem cells give rise to progenitor T cells, which migrate to the thymus for primary ontogeny. There, *TCR* genes undergo somatic rearrangement of germline gene sequences, similar to the process that occurs in immunoglobulin heavy (*IGH*) and light-chain (*IGLK* and *IGLL*) genes. The four *TCR* genes are rearranged in the following order: *TCRD* (on chromosome 14q11), *TCRG* (7p15), *TCRB* (7q34), and *TCRA* (14q11).[1] These rearrangements occur early in T-cell development and are unique to each cell, ensuring great receptor diversity (Figure 33-1). The genetic recombination occurs when the variable (V) segments are joined with a diversity (D) region (present only in the *TCRB* and *TCRD* genes) or with a joining (J) region by deletion of the intervening coding and noncoding DNA sequences. Thus, the *TCRB* and *TCRD* gene rearrangements result in V-D-J juxtaposition similar to the *IGH* gene, whereas the *TCRA* and *TCRG* genes contain only V-J rearrangements. In all cases, the V-(D)-J segment is apposed to the downstream constant (C) region by mRNA splicing (Figure 33-2). The *TCR* genes are translated into two types of receptors, which exist as heterodimers (αβ or γδ).[2] Approximately 95% of mature, circulating T cells express the αβ receptor because the initial δ or γ rearrangements failed to produce a functional receptor. In the skin, spleen, gastrointestinal tract, and other extranodal sites, γδ T cells are more commonly identified. T-cell neoplasms ensue after maturation arrest at one of the stages of T-cell development (Figure 33-1). They can originate from immature T cells as in lymphoblastic T-cell lymphoma, or from more mature T cells, as seen in peripheral T-cell lymphoma and mycosis fungoides. Related neoplasms of natural killer (NK) cells also occur, but because these are not true T-cell-derived tumors, they do not usually demonstrate T-cell receptor gene rearrangements. The T- and NK-cell neoplasms of the World Health Organization (WHO) classification are listed in Table 33-1.[3]

Translocations and Deletions

Clonal T-cell abnormalities can originate from a variety of molecular genetic events. Translocations can result in a fusion of (parts of) different genes if the breakpoint is within the affected genes. Upregulation of gene expression also can result, if the breakpoint is located outside the coding region. This can expose a gene to the enhancing effects of other genes and regulatory sequences that are in close proximity due to the translocation. Anaplastic large cell lymphoma, which is a T-cell or null-cell lymphoma, is associated with translocation t(2;5)(p23;q35). This translocation is characterized by a fusion between the nucleophosmin (*NPM*) gene on chromosome 5 and the anaplastic lymphoma kinase (*ALK*) gene on chromosome 2.[3,4] It is the most common cytogenetic abnormality in noncutaneous forms of anaplastic large cell lymphoma and the only recurring translocation that is routinely tested for in mature T-cell lymphomas. *ALK* can fuse with other genes as well, including *TPM3* at 1p21, *TFG* at 3q21, *ATIC* at 2q35, *CTLC* at 17q23, and *MSN* at Xq11-12.[4,5]

In addition to gene translocations, allelic loss of a tumor suppressor gene can contribute to the development of lymphoma. An example is loss of function of the *P15* and *P16* genes on the short arm of chromosome 9.[6] Allelic loss and gene silencing due to promoter methylation were identified in both early and advanced stages of Sézary syndrome and mycosis fungoides. Cytogenetic anomalies, including deletions and inversions on the long arms of chromosomes 8, 11, and 14, are associated with T-cell prolymphocytic leukemia. However, due to the variation in gene fusion

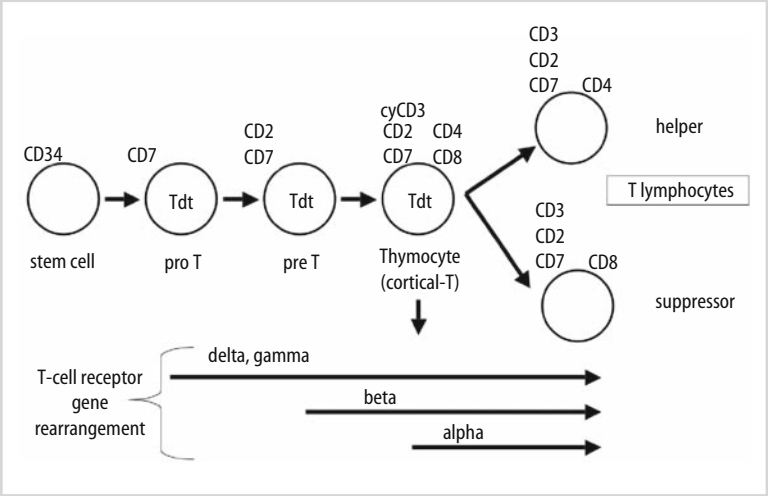


Figure 33-1. Various stages of T-cell development and an overview of the order of *TCR* gene rearrangements. (TdT, terminal deoxynucleotidyl transferase.)

sites, none of these translocations is currently analyzed routinely at the molecular level.

Indications for Testing

The diagnosis of a T-cell malignancy is best made in the context of clinical information, tissue morphology, immunohistochemical stains, and immunophenotypic analysis. In contrast to the restricted IGL protein expression in mature B-cell lymphomas, however, T cells do not have a definitive immunophenotypic marker of clonality. Therefore, *TCR* gene rearrangement studies can be essential to complement a diagnostic evaluation and to distinguish polyclonal from monoclonal lymphoproliferations. When clonality is assessed by *TCR* polymerase chain reaction (PCR), a visible band with products of exactly the same size is indicative of the presence of a clone, whereas polyclonal amplification generates a smear of nondistinct products of various sizes. A result is considered oligoclonal if more than two distinct bands are visible. The number of bands that would lead to a determination of oligoclonality varies by method, but the band size often will vary between duplicate PCR reactions in oligoclonal proliferations. In addition to the differentiation of reactive and neoplastic proliferations, molecular methods can be applied to identify disease-related findings, such as an associated virus or a specific gene fusion product, that enable subclassification of the malignancy.

Anaplastic large cell lymphoma, characterized by translocation t(2;5)(p23;q35), is the only translocation for which testing is widely available. Diagnostic testing is complicated by the variety of *ALK* fusion partners and inconsistency of genetic anomalies seen during tumor progression.

Adult T-cell leukemia/lymphoma (ATLL) is strongly associated with HTLV-1 infections.[7] The provirus is clonally integrated in the host DNA in virtually all ATLL patients, but in situ hybridization studies for this virus are difficult to perform and are not routinely offered. Serologic studies or PCR analysis are better suited to detect the virus.

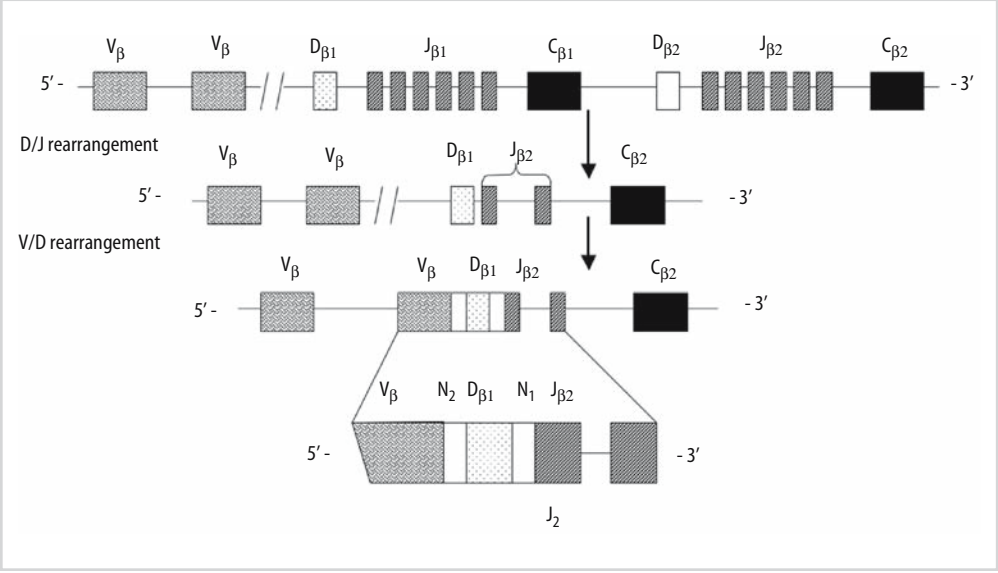


Figure 33-2. The *TCRB* locus on chromosome region 7q34 is used as an example to demonstrate the rearrangement of the variable (V), diversity (D), joining (J), and constant (C) regions. N, nucleotides added by the enzyme TdT.

Table 33-1. T-Cell and NK-Cell Neoplasms of the WHO Classification

Precursor T-lymphoblastic leukemia/lymphoma
T-cell prolymphocytic leukemia
T-cell large granular lymphocytic leukemia
Aggressive NK-cell leukemia
Adult T-cell leukemia/lymphoma
Extranodal NK/T-cell lymphoma, nasal type
Enteropathy-type T-cell lymphoma
Subcutaneous panniculitis-like T-cell lymphoma
Blastic NK-cell lymphoma
Mycosis fungoides/Sézary syndrome
Primary cutaneous CD30 positive T-cell lymphoproliferative disorders
Primary cutaneous anaplastic large cell lymphoma
Lymphomatoid papulosis
Borderline lesions
Angioimmunoblastic T-cell lymphoma
Peripheral T-cell lymphoma, unspecified
Anaplastic large cell lymphoma

Many NK-cell lymphomas are extranodal proliferations frequently associated with Epstein-Barr virus (EBV) infection within tumor cells.[8] In situ hybridization testing for EBER-1 RNA of EBV is offered in many laboratories and often is useful in the diagnosis of this tumor type. Because of the high frequency of latent infection by EBV in normal adults, serologic and PCR tests for EBV are of little value in determining a disease association, and the in situ hybridization procedure is the preferred method of testing.

Available Assays

Although no single methodology detects malignancy in all cases, several molecular methods can be applied with high sensitivity and specificity.

Southern blot analysis (SBA) requires fairly large amounts of fresh tissue and is both labor-intensive and time-consuming. Radioactive probes often are used, but fluorescent probes for SBA also are widely available. Furthermore, detection of a clonal rearrangement by SBA requires that at least 1% to 5% of the cells in the sample have the suspected abnormality. SBA remains a useful method for some testing, however, especially as an independent diagnostic method when clinical suspicion is high, and during optimization of *TCR* PCR tests.[9,10] SBA of the *TCRB* gene is most commonly applied, and although it will detect more than 90% of T-cell malignancies, it is generally unable to identify gene rearrangements in clonal γδ T cells or NK cells. SBA requires DNA digestion by restriction enzymes, followed by electrophoresis and transfer of the DNA to a membrane. The DNA on the blot is then hybridized with probes directed against the *TCRB* constant region (Cβ) or probes directed against one or both of the *TCRB* joining regions (Jβ1 and 2).[11] An example of a *TCRB* Southern blot is shown in Figure 33-3. Even though a large number of rearrangements are present in a lane, only the clonal rearrangement reaches the detection level.

As recommended in the recent Clinical and Laboratory Standards Institute (CLSI) guidelines,[12] SBA is performed using three restriction enzyme digests of genomic DNA. The presence of up to two nongermline rearrangement bands in at least two of the three digests is interpreted as a positive result. A nongermline band or bands in only one of the three digests may represent an individual polymorphism or a true rearrangement. Only two of the three digests are required to show rearrangement bands, since one of the digests may have a rearrangement that is the same size as the germline band. The allowance for up to two rearrangement bands per digest accounts for rearrangement of both TCR alleles, or the creation of a new restriction site within the probe region, by the rearrangement. The presence of more than two nongermline rearrangement bands is interpreted as an oligoclonal pattern.

PCR tests are now much more frequently utilized than SBA. Advantages of PCR include the short turnaround-time, minimal quantity requirements for DNA or RNA, and very high sensitivity (in the absence of a polyclonal T-cell background). Direct PCR amplification of genomic DNA can be used for assessment of T-cell clonality in fluids and tissue, including paraffin-embedded tissue. Usually the *TCRG* gene is amplified, which contains 11 variable regions, composed of only four V families, and five J-region exons and, therefore, is much less complex than the *TCRB* locus.[13–15] Because the *TCRG* locus is rearranged prior to the *TCRB* locus, clonal rearrangements can be detected even if the ultimate receptor pair expressed on a cell is Tαβ. The first-generation *TCRG* PCR tests employed consensus primers for at least a portion of the joining (J1/2) region and for the first family of the variable regions (Vγ1–8). However, false-negative reactions can be minimized with an expanded panel of primers directed against all the

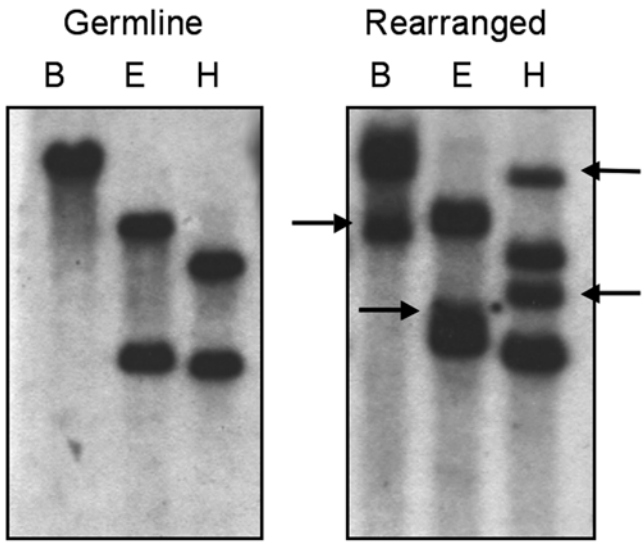


Figure 33-3. Example of a *TCRB* Southern blot. Three restriction enzymes (B, *Bam*H I; E, *Eco*R II; H, *Hind* III) were used on genomic DNA from germline (left) and T-cell lymphoma (right) samples. The lymphoma sample demonstrates one to two rearrangements (arrows) with all three restriction enzymes.

TCRG variable and joining regions. The primers for regions Vγ1–8, Vγ9, Vγ10, and Vγ11, coupled with a group of J-region primers, are most widely used.[16] After PCR amplification, the products can be visualized by methods including heteroduplex analysis, denaturing gradient gel electrophoresis, or capillary electrophoresis.[17,18] Figure 33-4 demonstrates examples of *TCRG* PCR test results with clonality assessment by heteroduplex analysis and by capillary electrophoresis.

Reverse transcription–polymerase chain reaction (RT-PCR) is used for the identification of fusion transcripts such as *NPM/ALK*, associated with t(2;5) in anaplastic large cell lymphoma (Figure 33-5). Because the fusion product is small, testing of RNA from paraffin-embedded tissue sections may be possible. This translocation results in expression of the ALK protein, which is not normally expressed in lymphocytes. ALK expression is a favorable prognostic marker and is associated with improved survival compared to ALK-negative anaplastic large cell lymphoma. This association is independent of the translocation partner. Because some *NPM/ALK* fusions as well as many *ALK* translocations to different fusion partners are missed by specific methods such as RT-PCR analysis, and because of the important prognostic significance of ALK expression, ALK immunohistochemistry is the most widely used test for this disease. RT-PCR, however, may have greater utility in minimal residual disease (MRD) monitoring, in select cases.[19]

Fluorescence in situ hybridization (FISH) on fresh and even paraffin-embedded tissues also are used to detect the t(2;5). This method offers an improved rate of detection in paraffin-embedded tissues when compared to RT-PCR for *NPM/ALK*, but t(2;5) FISH is not offered as a diagnostic tests in most laboratories.

Real-time quantitative PCR (RQ-PCR) is used to monitor patients for MRD after therapy, or for disease recurrence. The *NPM/ALK* product can be monitored with extremely high sensitivity (as low as 0.001% tumor burden), but RQ-PCR assays for this fusion transcript are generally not offered diagnostically. Quantitative *TCRG* detection of MRD has been studied in childhood acute lymphoblastic leukemia (T-cell ALL),[20] which may be used clinically upon future optimization.

In situ hybridization (ISH), other than FISH analysis, is commonly used for the detection and localization of EBV in tumor cells, particularly nasal-type NK/T-cell lymphomas. ISH can be performed on formalin-fixed, paraffin-embedded tissue sections and identifies evidence of EBV RNA within the nuclei of virtually all tumor cells. EBER-1 RNAs are most commonly targeted due to the very high copy number in EBV-infected tumor cells.

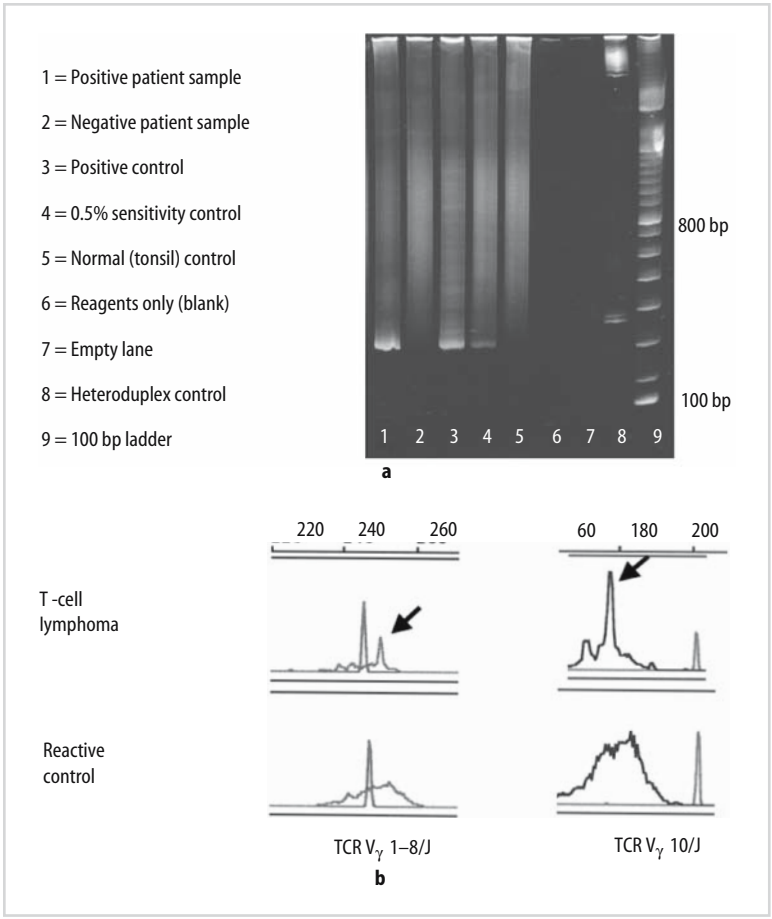


Figure 33-4. (a) Example of PCR-based analysis of the TCRγ locus with consensus primers for the joining (J1/2) region and for first family of the variable regions (V1), and visualization by heteroduplex analysis on a mutation detection enhancement (MDE) gel. After PCR amplification, the products are denatured and slowly allowed to reanneal (heteroduplex annealing), followed by electrophoresis. A clonal rearrangement will form a homoduplex band, which migrates faster than the heterogeneous polyclonal rearrangements. (b) Capillary electrophoresis of a clonal T-cell neoplasm and a reactive control. The arrows indicate clonal rearrangements detected with primers directed against Vγ1-8/J and Vγ10/J. Molecular weight standard peaks are present in the center of the left panels, and to the right of the right panels.

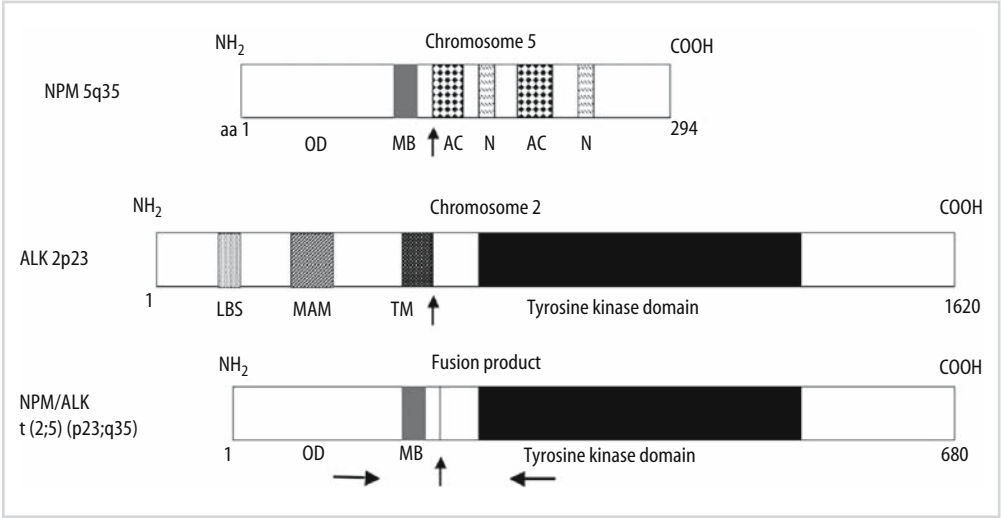


Figure 33-5. The *NPM/ALK* fusion of the t(2;5) is the most common *ALK* translocation in anaplastic large cell lymphoma. The *NPM* gene has an oligodimerization (OD) domain, a metal binding (MB) domain, 2 acidic amino acid clusters (AC) and 2 nuclear localization signals (N). The *ALK* gene has a putative ligand-binding site for pleiotrophin (LBS), a Meprin/A5/protein tyrosine phosphatase Mu (MAM) domain, and a transmembrane (TM) domain in addition to the tyrosine kinase domain. Horizontal arrows indicate the approximate location of primers used for the RT-PCR amplification of this fusion.

Interpretation of *TCR* PCR Test Results

Issues of *TCR* PCR test result interpretation are related to the method used to evaluate the PCR product. In gel-based assays, a positive result is based on finding a discrete band of the appropriate size in contrast to a smear or ladder of bands found with polyclonal T cells. With capillary electrophoresis instruments, each laboratory must establish and validate its own criteria for a positive result. The most commonly used criteria consider a peak that is two (or three) times larger than the third-largest peak to be evidence of a monoclonal population. This approach allows for biallelic rearrangements that may show two clonal peaks. Less-stringent criteria can be used with capillary electrophoresis in follow-up samples when the size of the expected clonal PCR product is known from prior testing.

Oligoclonal proliferations offer additional interpretive challenges. The presence of more than two bands on a gel (or more than two prominent peaks by capillary electrophoresis) should be considered evidence of an oligoclonal proliferation rather than monoclonality. Variation in the size of the bands or peaks on duplicate runs of the same sample should be interpreted as evidence of oligoclonality (or selective amplification of small numbers of polyclonal T cells in a paucicellular sample). For this reason, many laboratories test samples in duplicate to confirm the reproducibility of a monoclonal population before reporting a positive monoclonal result.

Sensitivity

For most T-cell lymphomas and acute lymphoblastic leukemias, *TCR* gene rearrangement studies are the only molecular assays available. The sensitivity of these assays varies. SBA remains the gold standard for *TCR* rearrangement testing because, in principle, all rearrangements can be detected. SBA, however, has a sensitivity of approximately 1% to 5%. Most PCR-based assays detect only ~80% of *TCR* rearrangements compared to SBA because PCR is primer dependent. However, application of multiple *TCRG* primer sets allows the detection of the vast majority of rearrangements.[21] Patient-specific primers have been developed on a research basis and allow a lower level of detection (around 0.001%), but this approach is labor-intensive and time-consuming, thus prohibitive for wide clinical application. Due to the great diversity in possible rearrangements, the vast majority of PCR tests use consensus primers, reaching a sensitivity of around 0.1% to 1% (in the absence of a prominent polyclonal background and depending on the method used for analysis of the PCR products). *TCR* gene rearrangements may remain obscure if the locus under investigation has been deleted in the rearrangement, or has remained in the germline configuration. It is also possible that the primer binding site itself was removed or changed during the rearrangement. Moreover, false-negative results may be caused by an insufficient number of clonal cells in the sample. For MRD testing using PCR, the level of detection varies markedly based on the test method and the exact molecular basis of the lymphoid proliferation.

Specificity

TCR gene rearrangement tests are helpful for lineage determination but are not always definitive due to "lineage infidelity," rearrangement of the *TCR* gene in non-T cells. In difficult cases of ambiguous lineage, both B- and T-cell clonality studies should be performed to ensure comprehensive assessment. *TCRG* rearrangements may occur in up to 55% of precursor B-ALL, although they are absent in most mature B-cell malignancies. Approximately 10% of patients with acute myelogenous leukemia (AML) demonstrate positivity in *TCRG* rearrangement studies. In these settings, the detection of a gene rearrangement should be viewed as evidence of clonality rather than of lineage, and correlation with immunophenotyping data is

used to assign lineage. Overall, more than 90% of T lymphoblastic lymphoma/leukemia cases have *TCR* gene rearrangements, with nearly 20% also having *IGH* gene rearrangements.

TCR gene rearrangement studies are typically performed to distinguish polyclonal from monoclonal lymphoproliferations, but the detection of a monoclonal *TCR* gene rearrangement in isolation is not always indicative of malignancy. There are several scenarios in which *TCR* gene rearrangements may be detected in nonneoplastic conditions. The most common examples of this phenomenon are in autoimmune disorders, infectious diseases, and cutaneous lesions.[22] Lymphomatoid papulosis is a CD30-positive cutaneous lymphoproliferative disorder that may progress to cutaneous anaplastic large cell lymphoma. A subset of cases of lymphomatoid papulosis has *TCR* gene rearrangements, but this finding is not predictive of progression to lymphoma. Therefore, in this setting, *TCR* gene rearrangement studies are of little value. Autoimmune disorders and some infectious diseases, including EBV infection, are associated with oligoclonal B- and T-cell proliferations. Such oligoclonal processes may result in selective amplification of one of the clones, mimicking a clonal gene rearrangement. In a similar fashion, small numbers of T cells, particularly in small skin biopsy specimens, may result in monoclonal or oligoclonal amplification patterns by PCR. Duplicate testing of these sample types will frequently reveal different-sized gene rearrangements, indicating "pseudoclonality."

Laboratory Issues

Controls

As for any clinical molecular test, the choice of controls is critical. Previously positive patient specimens and T-lymphocyte-derived cell lines can serve as positive controls. Sensitivity controls can be made by dilution of a positive control cell line in a polyclonal T-cell population, such as tonsil, to determine the level of detection in a polyclonal background. The sensitivity in a nonlymphoid background will be higher and may lead to inaccurate assessment of sensitivity. In addition to a positive control, and one or more sensitivity controls, PCR tests should always include a normal (negative) sample and a sample containing reagents but no DNA as controls for possible contamination with extraneous DNA. SBA controls include size markers, germline DNA (e.g. from placenta), and a sensitivity control.

Failure of PCR Tests

In contrast to general clinical molecular PCR tests, apparent failure of the reaction in T-cell clonality assays can be due to a paucity of T cells rather than poor-quality DNA or the presence of PCR inhibitors such as heparin. To rule out an intrinsic problem with the amplification itself, the quality of the DNA is assessed by amplification of an unrelated control gene. False-negative results can largely be avoided by optimal selection of primers, as well as amplification of a control sequence for each sample, in addition to clonality testing. The detection rate for *TCRG* PCR testing also depends on the specimen type and quality used for DNA extraction. A significantly lower number of true positives may be achieved using paraffin-embedded tissue, and laboratories often validate the sensitivity of the assay specifically for paraffin-embedded tissue specimens.

Selection of Primers and Probes for Clonality Tests

Each laboratory should individually evaluate which assays are most suitable to answer the diagnostic questions that arise in its patient population, and which set of probes and primers are most appropriate. A recent multicenter study[16] reports that laboratories routinely use two to five different restriction enzymes for SBA analysis of the *TCRB* chain gene. The majority of laboratories that participated in this study (9 of 13) used three enzymes, as recommended by CLSI.[12] Ten of 13 participating laboratories also used only a single hybridization probe, whereas the remaining three laboratories in the study used two probes. The use of the additional second probe may be a valid alternative to the use of a larger number of enzymes when the available tissue is minimal. Whereas each additional enzyme digest requires additional tissue, the same Southern blot can be stripped and evaluated with additional probes without the need for more sample. In such cases, however, hybridized material must be removed efficiently without significant loss of genomic DNA from the membrane.

As mentioned previously, the false-negative rate of *TCRG* PCR can be significantly reduced if primers directed against all of the V and J regions are used. Such an approach at a single institution was recently reported to detect 95% of clonal T-cell proliferations by PCR analysis.

Result Interpretation

In T-cell clonality assessment by PCR, bands may be difficult to interpret when they are weak or outside the expected size range. Weak results should be repeated to confirm reproducibility. Capillary electrophoresis instruments are useful for differentiating oligoclonal proliferations from clonal peaks due to their ability to separate PCR products at the single base pair level. Peaks or bands outside the expected diagnostic range could be definitively characterized by direct DNA sequencing. However, specific criteria for interpretation of PCR methods for clonality testing have not been established at this time.

References

1. Rezuke WN, Abernathy EC, Tsongalis GJ. Molecular diagnosis of B- and T-cell lymphomas: fundamental principles and clinical applications. *Clin Chem.* 1997;43:1814–1823.
2. Delfau M-H, Hance AJ, Lecossier D, Vilmer E, Grandchanp B. Restricted diversity of Vγ9-JP rearrangements in unstimulated human γ/δ T lymphocytes. *Eur J Immunol.* 1992;22:2437–2443.
3. Jaffe ES, Harris NL, Stein H, Vardiman JW, eds. *World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues.* Lyon, France: IARC Press; 2001.
4. Drexler HG, Gignac SM, von Wasielewski R, Werner M, Dirks WG. Pathobiology of *NPM-ALK* and variant fusion genes in anaplastic large cell lymphoma and other lymphomas. *Leukemia.* 2000;14: 1533–1559.
5. Duyster J, Bai R-Y, Morris SW. Translocations involving anaplastic lymphoma kinase (*ALK*). *Oncogene.* 2001;20:5623–5637.
6. Scarisbrick JJ, Woolford AJ, Calonje E, et al. Frequent abnormalities of the p15 and p16 genes in mycosis fungoides and Sezary syndrome. *J Invest Dermatol.* 2002;118:493–499.
7. Mortreux F, Gabet AS, Wattel E. Molecular and cellular aspects of HTLV-1 associated leukemogenesis in vivo. *Leukemia.* 2003;17:26–38.
8. Ohsawa M, Nakatsuka S-I, Kanno H, et al. Immunophenotypic and genotypic characterization of nasal lymphoma with polymorphic reticulosis morphology. *Int J Cancer.* 1999;81:865–870.
9. Moreau EJ, Langerak AW, van Gastel-Mol EJ, et al. Easy detection of all T cell receptor gamma (TCRG) gene rearrangements by Southern blot analysis: recommendations for optimal results. *Leukemia.* 1999;13:1620–1626.
10. Langerak AW, Wolvers-Tettero ILM, van Dongen JJM. Detection of T cell receptor beta (TCRB) gene rearrangement patterns in T cell malignancies by Southern blot analysis. *Leukemia.* 1999;13:965–974.
11. van Dongen JJM, Wolvers-Tettero ILM. Analysis of immunoglobulin and T cell receptor genes. Part I: basic and technical aspects. *Clin Chim Acta.* 1991;198:1–91.
12. Enns RK, Altmiller DH, Hinkel CS, Madej RM, O'Leary TJ. *Immunoglobulin and T-Cell Receptor Gene Rearrangement Assays: Approved Guideline.* 2nd ed. National Committee for Clinical Laboratory Standards (NCCLS) document MM2-A2; 2002.
13. Benhattar J, Delacretaz F, Martin P, Chaubert P, Costa J. Improved polymerase chain reaction detection of clonal T-cell lymphoid neoplasms. *Diagn Mol Pathol.* 1995;4:108–112.
14. Theodorou I, Bigorgne C, Delfau M-H, et al. VJ rearrangements of the TCRγ locus in peripheral T-cell lymphomas: analysis by polymerase chain reaction and denaturing gradient gel electrophoresis. *J Pathol.* 1996;178:303–310.
15. Theodorou I, Raphaël M, Biborgne C, et al. Recombination pattern of the TCRγ locus in human peripheral T-cell lymphomas. *J Pathol.* 1994;174:233–242.
16. Arber DA, Braziel RM, Bagg A, Bijwaard KE. Evaluation of T cell receptor testing in lymphoid neoplasms: results of a multicenter study of 29 extracted DNA and paraffin-embedded samples. *J Mol Diagn.* 2001;3:133–140.
17. Greiner TC, Raffeld M, Lutz C, Dick F, Jaffe ES. Analysis of T cell receptor-γ gene rearrangements by denaturing gradient gel electrophoresis of GC-clamped polymerase chain reaction products. Correlation with tumor-specific sequences. *Am J Pathol.* 1995;146: 46–55.
18. Greiner TC, Rubocki RJ. Effectiveness of capillary electrophoresis using fluorescent-labeled primers in detecting T-cell receptor γ gene rearrangements. *J Mol Diagn.* 2002;4:137–143.
19. Cataldo KA, Jalal SM, Law ME, et al. Detection of t(2;5) in anaplastic large cell lymphoma: comparison of immunohistochemical studies, FISH, and RT-PCR in paraffin-embedded tissue. *Am J Surg Pathol.* 1999;23:1386–1392.
20. van der Velden VH, Wijkhuijs JM, Jacobs DC, van Wering ER, van Dongen JJ. T cell receptor gamma gene rearrangements as targets for detection of minimal residual disease in acute lymphoblastic leukemia by real-time quantitative PCR analysis. *Leukemia.* 2002;16:1372–1380.
21. van Dongen JJ, Langerak AW, Bruggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98–3936. *Leukemia.* 2003;17:2257–2317.
22. Wood GS. Analysis of clonality in cutaneous T cell lymphoma and associated diseases. *Ann N Y Acad Sci.* 2001;941:26–30.

Chapter 34

Lymphoproliferations of Immunodeficiency

Ethel Cesarman and Amy Chadburn

Molecular Basis of Disease

Immunodeficient patients are at an increased risk for developing lymphoproliferative disorders (LPD), including lymphomas. The World Health Organization (WHO) classification recognizes four clinical settings associated with the development of immunodeficiency-related LPDs: (1) primary immune disorders, (2) HIV infection, (3) iatrogenic immunosuppression following solid organ or allogeneic bone marrow transplantation (posttransplant lymphoproliferative disorder [PTLD]), and (4) methotrexate therapy, usually for an autoimmune disorder.[1] These lesions are highly heterogeneous, largely due to the various underlying causes of the different immunodeficiencies; however, they share several features (see Table 34-1). In most instances, the LPD are related to Epstein-Barr virus (EBV or HHV-4) infection, and thus, in situations where immunocompetence can be reestablished, these EBV-driven proliferations may regress. However, the development of secondary genetic structural alterations in oncogenes and tumor suppressor genes, not all of which have been defined, results in transformation to a neoplastic process that is no longer responsive to immune modulation.

Immunodeficiency-related LPD are all of either germinal center or postgerminal center cell origin, so they have undergone somatic hypermutation (SHM) of the immunoglobulin variable to joining regions. Recent studies indicate that the process of SHM often misfires, so a relatively large proportion of these lymphomas (as well as those arising outside immunodeficiency) have mutations in the regulatory region of a number of protooncogenes, including *MYC*, *BCL6*, *PAX5*, *PIM1*, and *RHOH/TTF*.[2] Thus, in spite of aggressive therapeutic intervention and either due to the inability to reestablish normal immune function or to neoplastic transformation, these lesions may progress, leading to the patient's demise.

The morphologic diagnosis of LPD often is difficult. In some instances the lesions are clearly neoplastic, such as those arising in the setting of HIV infection; however, other lesions, such as many of the PTLD, are difficult to classify due to their polymorphic appearance. Thus, the accurate diagnosis and treatment of these immunodeficiency-related LPDs often requires not only morphologic examination but also the use of immunophenotypic and genotypic techniques.

The most frequently encountered immunodeficiency-related LPDs are those occurring in the setting of HIV infection and following organ transplantation. These lesions, because of their frequency, have been better characterized and thus are the focus of this chapter. The primary immune disorders and methotrexate-related LPDs are relatively rare and have not been as extensively studied. Other than a B-cell lineage and the documentation of a role for EBV in their development, there is not much information regarding specific molecular alterations (other than those causing the primary immunodeficiency).

HIV-Related Lymphomas

The incidence of non-Hodgkin lymphomas (NHL) in HIV-positive individuals is estimated to be between 4% and 10%, but the incidence of at least some subsets appears to be decreasing with combination antiretroviral therapy. The incidence of Hodgkin lymphoma is somewhat increased in HIV-infected individuals but is not considered to be an AIDS-defining condition. The pathogenesis of NHL in the context of AIDS is complex and thought to be related to disrupted immune surveillance, chronic antigenic stimulation, genetic alterations, cytokine dysregulation, and herpes virus infection.[3–5]

Although HIV-related lymphomas are almost always of B-cell origin, they are morphologically diverse. Several subtypes are similar to lymphomas occurring in immunocompetent patients, while others preferentially develop in the context of AIDS. HIV-related lymphomas can be classified by morphology (as in the WHO classification), primary site of presentation (i.e., systemic, primary central

Table 34-1. Common Features of Lymphoproliferations of Immunodeficiency

Typically extranodal
Diffuse aggressive histology
B cell lineage
Rapid clinical progression
Often related to Epstein-Barr virus
Hypermutation of immunoglobulin genes and regulatory region of protooncogenes

nervous system [CNS], body cavity), or both.[5,6] The major categories, with their viral associations and molecular features, are detailed in Tables 34-2 to 34-4.

Lymphomas Also Occurring in Immunocompetent Patients

HIV-related Burkitt lymphomas (BL) include cases exhibiting the features of classical BL, those showing plasmacytoid differentiation, and those exhibiting features of atypical Burkitt or Burkitt-like lymphoma. Translocation of *MYC* into one of the immunoglobulin loci is considered by some to be a prerequisite for classification of a lymphoma as BL or atypical BL. The most common translocation is a t(8;14), involving the *MYC* and immunoglobulin heavy-chain (*IGH*) genes, but in 10% the translocation involves *MYC* and one of the light-chain (*IGL*) genes. It is thought that this translocation leads to deregulated expression of the *MYC* gene. Mutation of the *MYC* locus also occurs in Burkitt lymphoma and also may lead to abnormal MYC expression.

Diffuse large B-cell lymphoma (DLBCL) can be divided into centroblastic (CB) and immunoblastic (IB) categories. The IB type is more frequently associated with EBV infection, and patients with these lymphomas usually have significant immunosuppression, with low CD4 counts (median $<100 \times 10^6$/L), and approximately one third have been previously diagnosed with an AIDS-defining illness. This degree of immune dysfunction allows EBV to be the driving proliferative force, with expression of the oncogenic but also immunogenic LMP1 and EBNA2 proteins.

Table 34-2. Major Categories of HIV-Associated Lymphomas and Approximate Frequencies

Lymphomas Also Occurring in Immunocompetent Patients	
Burkitt and Burkitt-like lymphoma	30%
Diffuse large B-cell lymphoma	
Centroblastic	25%
Immunoblastic	22%
Lymphomas Occurring Primarily in HIV-Positive Patients	
Primary effusion lymphoma	3%
KSHV-positive extracavity lymphoma	Unknown
Polymorphic B-cell lymphoma (PTLD-like)	Rare
Plasmablastic lymphoma of oral cavity	Rare
Primary CNS lymphomas	10–15%

Table 34-3. Correlation of Morphology and Viral Infection in HIV-Related Lymphomas

Virus	BL	CB	IB	PEL	Poly	PCNSL
EBV	25–75%	20–40%	80%	87%	40%	100%
LMP1+	0%	Rare	65%	Rare	ND	90%
KSHV	0%	0%	15%	100%	ND	0%
HIV-1	0%	0%	0%	0%	0%	0%

BL, Burkitt and Burkitt-like lymphoma; CB, centroblastic diffuse large B-cell lymphoma; IB, immunoblastic diffuse large B-cell lymphoma; PEL, primary effusion lymphoma; Poly, polymorphic B-cell lymphoma (PTLD-like); PCNSL, primary CNS lymphoma; ND, not determined.

Lymphomas Occurring Primarily in HIV-Positive Patients

Primary effusion lymphomas (PEL) are characterized by the presence of Kaposi's sarcoma–associated herpesvirus (KSHV or HHV-8) within the tumor cells. PEL are rare tumors, accounting for only approximately 3% of all HIV-related NHLs. These neoplasms usually present as an effusion involving one or more of the pleural, peritoneal, and pericardial spaces. While the majority lack an associated tissue mass at presentation, in about one third of the cases dissemination to extracavitary sites occurs. In addition, KSHV-associated lymphomas can occur as primary solid tumor masses without an effusion. Some lymphomas first diagnosed as IB lymphoma have been found to contain KSHV, with immunophenotypic and molecular features similar to those of PEL.

Plasmablastic lymphomas, associated with multicentric Castleman's disease, occur in HIV-positive patients. While these plasmablastic lymphomas are KSHV positive, they differ from PEL in a number of ways. Plasmablastic lymphomas are EBV negative, do not contain mutations in the immunoglobulin genes, and are thought to arise from naive IgM lambda-expressing B cells rather than terminally differentiated B cells. Polymorphic B-cell lymphomas (PTLD-like) are extremely rare lesions but morphologically resemble polymorphic PTLD (see PTLD section below).

Table 34-4. Correlation of Morphology and Genotype in HIV-Related Lymphomas

Genetic						
Lesion	BL	CB	IB	PEL	Poly	PCNSL
IGH/IGL R	>95%	>95%	>95%	>95%	80%	>95%
BCL6 R	<5%	20%	15%	0%	0%	0%
BCL6 M	50–60%	70%	50%	80%	10%	50%
MYC R	>95%	30%	25%	Rare	0%	0%
RAS M	23%	Rare	Rare	0%	0%	ND
TP53 M	40–60%	10%	10%	Rare	20%	ND
BCL1, *BCL2* R	0%	0%	0%	0%	ND	ND

BL, Burkitt and Burkitt-like lymphoma; CB, centroblastic diffuse large B-cell lymphoma; IB, immunoblastic diffuse large B-cell lymphoma; PEL, primary effusion lymphoma; Poly, polymorphic B-cell lymphoma (PTLD-like); PCNSL, primary CNS lymphoma; M, mutations; R, rearrangements; ND, not determined.

Table 34-5. Major Differences Between PTLDs in Solid Organ Transplantation and Allogeneic Bone Marrow Transplantation

	Solid Organ Transplantation	Allogeneic Bone Marrow Transplantation
Usual origin of PTLD	Recipient	Donor
Course	Variable	Usually aggressive
Incidence	More common (0.3–12.5%)	Less common (~1%)
Prognostic association of molecular abnormalities, including *BCL6* mutations	Yes	No

Primary CNS lymphomas differ from systemic DLBCL, as noted in Tables 34-3 and 34-4, with the majority of cases exhibiting IB morphology and EBV-positivity. Plasmablastic lymphoma of the oral cavity, a recently described distinct entity, has features similar to IB lymphomas, but is less heterogeneous and polymorphic. Approximately 50% of cases are associated with EBV infection.

Posttransplantation Lymphoproliferative Disorders

Posttransplantation lymphoproliferative disorders (PTLD) develop in the setting of iatrogenic immunosuppression following solid organ transplantation (SOT) or allogeneic bone marrow transplantation (BMT), with some differences evident (Table 34-5). The incidence of these lesions varies based on the type of organ transplanted as well as the type and amount of immunosuppression employed. In these immunosuppressed patients, morphology does not accurately predict clinical course, making molecular analysis particularly useful. In some SOT recipients, a lesion may regress completely following a reduction in immunosuppression, while morphologically similar lesion(s) in other patients may progress despite aggressive clinical intervention, resulting in the patient's demise. Furthermore, institution of the correct therapy in patients with PTLD is crucial since a reduction in immunosuppression can potentially result in organ loss, while chemotherapy can lead to life-threatening infection in an already immunosuppressed individual. Although in some studies specific molecular events correlate well with biologic aggressiveness of the lesions, in other studies these findings cannot be confirmed. It may be that differences in the type of organ transplanted, the immunosuppressive regimens, as well as host and donor factors influence the pathogenesis and biologic behavior of PTLD. In contrast to SOT PTLD, patients who develop PTLD following allogeneic BMT usually have an aggressive, frequently fatal, clinical course. As with other immunodeficiency-related LPDs, the development of PTLD in both SOT and BMT recipients is highly associated with EBV infection. Furthermore, the relative incidence of these lesions is higher in patients who are EBV negative at the time of transplantation.

PTLD Arising in Solid Organ Transplant Recipients

The overall incidence of PTLD after SOT is ~1.5% but varies from 0.3% to 12.5% depending on the organ transplanted and the immunosuppressive regimen, with a higher incidence in children than adults. Extensive studies of PTLD occurring in SOT recipients (primarily heart, kidney, and lung recipients) have shown that most of these lesions can be separated into three categories based on morphologic and molecular genetic criteria. These categories correlate with the biologic behavior of the lesions. Because of the special clinical setting, that is, iatrogenic immunosuppression that can be modulated, a unique classification scheme was originally developed by Nalesnik et al.,[7] to describe the clinical course and outcome of SOT PTLD patients, and subsequently modified by Knowles, et al.[8] (Table 34-6).

This classification divides cases into three major categories (Table 34-6). Plasmacytic hyperplasia (PH) lesions show retention of the overall architecture of the tissue. Genotypically, these lesions are usually polyclonal based on *IGH* rearrangement studies, or have small monoclonal/oligoclonal populations. The majority of PH cases are EBV positive and do not contain structural alterations of known oncogenes or tumor suppressor genes. Furthermore, examination of EBV antigen expression shows that PH lesions express LMP1 and in some instances EBNA2, indicating

Table 34-6. Solid Organ Posttransplantation Lymphoproliferative Disorders

Category	Histopathology	Ig Genes	EBV	Onco/TS Genes	Clinical Outcome
Plasmacytic hyperplasia	Retention of architecture; plasma cells, lymphocytes	Polyclonal	Poly-, oligo-, monoclonal Latency II, III	None	Nonaggressive
Polymorphic PTLD	Disruption of architecture; hetergeneous cell population; +/– necrosis, atypia	Monoclonal	Monoclonal Latency II, III	None *BCL6*	Nonaggressive Aggressive
Non-Hodgkin lymphoma/ multiple myeloma	Malignant lymphocytes or plasma cells	Monoclonal	Monoclonal Latency I	Present	Aggressive

TS, tumor suppressor; Ig, immunoglobulin.

that they exhibit either the latency types II or III patterns of EBV gene expression.

The second category is the polymorphic PTLD lesions that histologically show destruction of the underlying architecture and are composed of a heterogeneous (polymorphic) cell population. At the genetic level these lesions are monoclonal based both on immunoglobulin studies and the presence of clonal EBV. Polymorphic PTLD lacks structural alterations in oncogenes and tumor suppressor genes except for the presence of *BCL6* gene mutations, which are present in approximately half of the cases studied.[9] Lesions containing the wild-type configuration of the *BCL6* gene regress following a reduction in immunosuppression, while those lesions containing *BCL6* gene mutations exhibit more aggressive behavior, requiring clinical intervention (Table 34-6). Polymorphic PTLD, like PH, exhibit the latency types II or III patterns of EBV gene expression. Type II and III latencies include expression of the transforming but immunogenic LMP and EBNA viral proteins, suggesting that reconstitution of an immune response in the host is more likely to result in elimination of these infected cells.

The third category is the monomorphic PTLD (malignant lymphoma/multiple myeloma) lesions that are composed of cytologically malignant cells and should be classified according to the WHO classification. They are monoclonal based both on *IGH* rearrangement and EBV studies. Additionally, they contain structural alterations in oncogenes and tumor suppressor genes frequently involved in lymphomagenesis, such as *TP53* and *MYC*.[8] In contrast to the other SOT PTLD, these lesions often exhibit the latency type I pattern of EBV gene expression, suggesting that they are less dependent on EBV and that they may no longer be recognized by the immune system even after immune reconstitution.

PTLD in Allogeneic Bone Marrow Transplant Recipients

PTLD occurring in allogeneic BMT recipients are biologically distinct from those occurring in SOT recipients. BMT PTLD are of donor origin, often have an explosive clinical presentation, and are frequently fatal. Furthermore, BMT PTLD tend to occur in the first 6 months following transplantation, since the level of anti-EBV cytotoxic T-cell precursors returns to normal only approximately 6 months after transplantation. In comparison to SOT PTLD, the incidence of BMT PTLD is relatively low (cumulative incidence of 1% at 10 years). As with SOT PTLDs, the vast majority of BMT PTLD cases are associated with EBV infection.

Morphologically, BMT PTLD consists of lesions exhibiting the histologic features of PH and polymorphic PTLD, as seen in SOT recipients.[10] However, none of the PTLD studied from 27 allogeneic BMT recipients showed the morphologic features of monomorphic PTLD,[11] although other investigators have identified such lesions.[12] Although morphology correlates with clinical outcome, B-cell clonality status does not correlate with morphology or with clinical outcome, based on polymerase chain reaction (PCR) analysis of the early-onset PTLD.[11] Specifically, both patients with monoclonal/oligoclonal PTLD as well as those with polyclonal PTLD died of PTLD. In contrast to SOT PTLD, examination of the BMT PTLD for *BCL6* mutations did not identify any correlation of this genetic alteration with patient outcome.[13] All BMT PTLD in this study exhibited EBV latency patterns type II or III (expressed LMP1 with or without EBNA2), indicating that these lesions are EBV driven.[11] The clinical outcome of BMT PTLD patients is largely related to the number and competence of EBV-specific cytotoxic T cells rather than to intrinsic differences in BMT PTLD subtypes.

Indications for Testing

HIV-Related Lymphomas

While many HIV-related lymphomas are obvious high-grade lesions that rarely need molecular confirmation for diagnostic purposes, situations arise when molecular confirmation of the diagnosis is important. Benign lymphadenopathies associated with HIV infection can be dramatic, with large hyperplastic follicles composed almost entirely of mitotically active centroblasts with only rare centrocytes. Furthermore, these follicles may contain a small number of B-cell clones. When a needle biopsy is performed, often only fragments of germinal centers are obtained without architectural references for histologic evaluation. Without these architectural references, the collections of germinal center cells can be confused with centroblasts of high-grade lymphoma. In addition, immunophenotypic analysis may show that the cells are monotypic. In this setting, analysis of *IGH* gene rearrangements may be very helpful for diagnosis. This type of clonality analysis also can distinguish a recurrence from a second independent lymphoma in patients with lymphoma that responded to therapy.

Molecular testing contributes to the subclassification of AIDS-related lymphomas. The presence of *MYC* translocations is supportive of BL. Accurate diagnosis of BL is particularly important because the treatment of BL is different from that of DLBCL and usually requires CNS prophylaxis. Identification of KSHV is indicative of a primary effusion lymphoma, or a KSHV-associated (PEL-like) extracavitary lymphoma. In addition, the finding of certain molecular markers is indicative of a particular biologic process and thus is informative. For example, the presence of EBV, and in particular the EBV gene expression profile, is reflective of the immune competence of the host, with patients that are more immunodeficient tend to have lymphomas with a type III latency pattern. As more biological therapies are developed, molecular classification will become increasingly important.

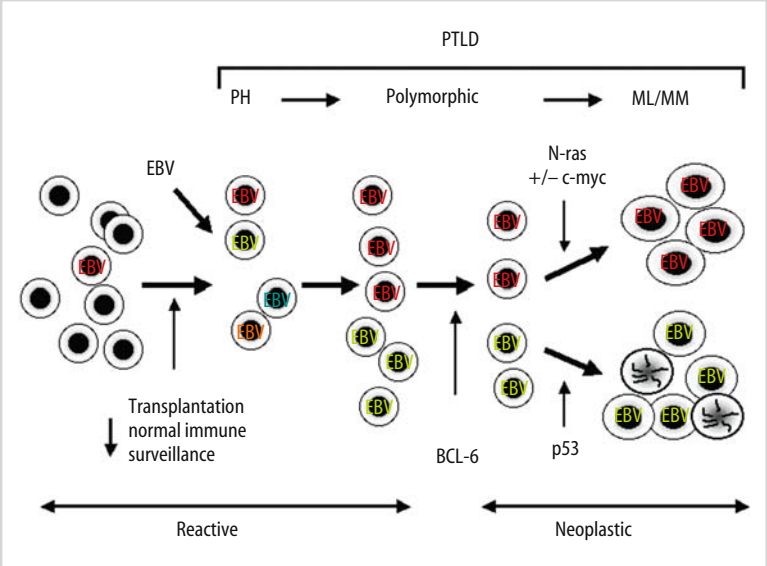


Figure 34-1. Model of tumor progression in EBV-positive solid organ posttransplantation lymphoproliferative disorders (PTLD). An initial step in the development of PTLD is the polyclonal expansion of EBV-infected B cells, as a result of either release from immune surveillance of latently infected cells or de novo infection. This expansion has a histologic appearance of a plasmacytic hyperplasia (PH). Subsequently, some clones with unknown selective advantages outgrow others, leading to oligoclonal and then monoclonal expansions, which are still responsive to immune responses if restored. Each individual clone is represented by a different color of EBV. These lesions have the histologic appearance of a polymorphic PTLD. However, if genetic alterations in oncogenes or tumor suppressor genes occur, a truly neoplastic process arises, which may no longer be responsive to immune surveillance mechanisms. Mutation in *BCL6* (BCL-6) alone does not appear to affect the morphologic features, but cases with structural alterations in *TP53* (p53), *NRAS* (N-ras), or *MYC* (c-myc), or some combination of these, have the histologic appearance of a monomorphic malignant lymphoma (ML) or, more rarely, a multiple myeloma/plasmacytoma (MM). (Reprinted from Cesarman E. Epstein-Barr virus (EBV) and lymphomagenesis. *Front Biosci.* 2002;7:e58–e65.)

Posttransplantation Lymphoproliferative Disorders

PTLD are frequently difficult to classify based on morphology alone, and frequently expert hematopathologists cannot agree on classification. Therefore, a strict molecular classification would be ideal. Correlation between molecular alterations and outcome appears to exist for lesions occurring in SOT recipients[14] but not in those occurring after BMT.[13] Assessment of the presence of EBV is an important first step for the diagnosis of PTLD, since this virus is present in approximately 95% of cases, and the pathogenesis and clinical outcome of EBV-negative posttransplant lymphomas appear to be different.[15,16] Study of clonality (by *IGH* rearrangement or EBV terminal repeat analysis) also is important for understanding the nature of the lesions and confirming a diagnosis of PTLD as opposed to reactive hyperplasia or rejection (if the infiltrate involves the allograft). Monoclonal rearrangements are common in both polymorphic and monomorphic PTLD, and their presence is sometimes useful in distinguishing these from PH, which usually shows a polyclonal pattern. Clonality analysis also is useful to determine whether multiple lesions are clonally related by having the same size rearrangement or EBV terminal repeats, which is not always the case in PTLD patients, who frequently have separate tumors that represent independent clonal expansions. The presence of structural alterations in oncogenes (*MYC*, *NRAS*) or tumor suppressor genes (*TP53*), or both, is usually indicative of a high-grade or monomorphic lesion. Analysis for mutations in the *BCL6* gene also predicts clinical outcome of PTLD in SOT patients but not in BMT patients. A model of tumor progression in SOT PTLD is shown in Figure 34-1.

Quantitative analysis of peripheral blood or serum EBV in transplant patients is helpful in identifying patients at risk for PTLD.[17–20] In addition, several recent studies suggest a correlation between tumor burden and viral loads, so quantitative EBV testing may be useful in monitoring response to therapy once PTLD develops.[21–23]

Available Assays

For a list of available assays, see Table 34-7.

Clonality

IGH Gene Rearrangement Analysis

IGH gene rearrangement analysis is used to confirm lineage and clonality. Clonality testing can be performed for *IGH*, *IGLK* or *IGLL* genes. The most commonly used method is PCR analysis for the *IGH* gene using primers for

Table 34-7. Tests for Lymphoproliferations of Immunodeficiency

Clonality
Immunoglobulin (*IGH*, *IGLK* or *IGLL*) gene rearrangement analysis
T-cell receptor (*TCR*) gene rearrangements
Epstein-Barr virus (EBV/HHV-4) long-terminal repeat analysis
Viral Associations
Epstein-Barr virus detection or quantitation
Kaposi sarcoma-associated herpesvirus detection
Oncogenes/Tumor Suppressor Genes (*MYC*, *RAS*, *TP53*, *BCL6*)
Fluorescence in situ hybridization (FISH)
Chromogenic in situ hybridization (CISH)
Southern blot analysis
Single-strand conformation polymorphism (SSCP)
Sequence analysis

framework 3 (FR3) of the variable region and for the joining region. Primers that recognize framework 1 or 2 of the variable region are useful in some instances to exclude false-negative results. PCR is followed by size fractionation by gel or capillary electrophoresis. Primers that recognize rearrangements for the *IGLK* or *IGLL* genes also can be used, although these have higher rates of false-negative results. The presence of *IGH* gene rearrangements also can be determined by Southern blot analysis (SBA), which is particularly useful for testing lymphoma types derived from germinal center or postgerminal center B cells, due to the high false-negative rate seen with PCR as a consequence of somatic hypermutation of the *IGH* VDJ junction regions. This high false-negative rate is particularly common in PTLD, PEL, and some HIV-related IB lymphomas. However, SBA is much more labor-intensive than PCR analysis, has a longer turnaround time, requires fresh or frozen tissue, and is no longer performed routinely in most clinical molecular laboratories. Alternative methods for immunoglobulin clonality assessment are in situ hybridization, immunohistochemistry, or flow cytometric analysis for monotypic light-chain expression.

TCR Gene Rearrangement Analysis

While the vast majority of immunodeficiency-associated lymphomas are of B-cell origin, rare cases with T- or null-cell phenotype occur, which warrant T-cell receptor (*TCR*) gene rearrangement analysis to define cell lineage. The most common assay is PCR for the *TCRG* gene rearrangements. See chapter 33 for more detail on these assays.

Clonality by Epstein-Barr Virus Analysis

Clonality of EBV-associated lymphomas can be determined using SBA for the terminal repeat region of the EBV genome. Like other herpesviruses, infectious EBV particles contain a linear double-stranded DNA viral genome that circularizes upon entry into the cell. The linear viral genome contains a variable number of tandem ~500 base pair repeats at each terminus, and upon circularization the number of tandem repeats is fixed. Circular episomes replicate with each cellular division, maintaining the same number of repeats in each infected cell. In a polyclonal population, where EBV independently infected many different cells, a range of repeat sizes will be present. In contrast, if all the cells in a population are derived from a single EBV-infected cell, only one size of terminal repeats will be present. A Southern blot assay that takes advantage of this biology was originally developed by Raab-Traub and Flynn.[24] A restriction enzyme that cuts outside the terminal repeat region is used to digest genomic DNA, followed by gel electrophoresis, transfer to a membrane, and hybridization with a probe to the terminal repeat region. A polyclonal population produces a ladder or smear, and a monoclonal population produces a single band. While SBA is not done routinely for diagnostic purposes, it can be extremely useful for determination of monoclonality in EBV-associated lymphomas, as it is more sensitive than SBA for immunoglobulin gene rearrangements because of the presence of multiple copies of the EBV genome in each infected cell.

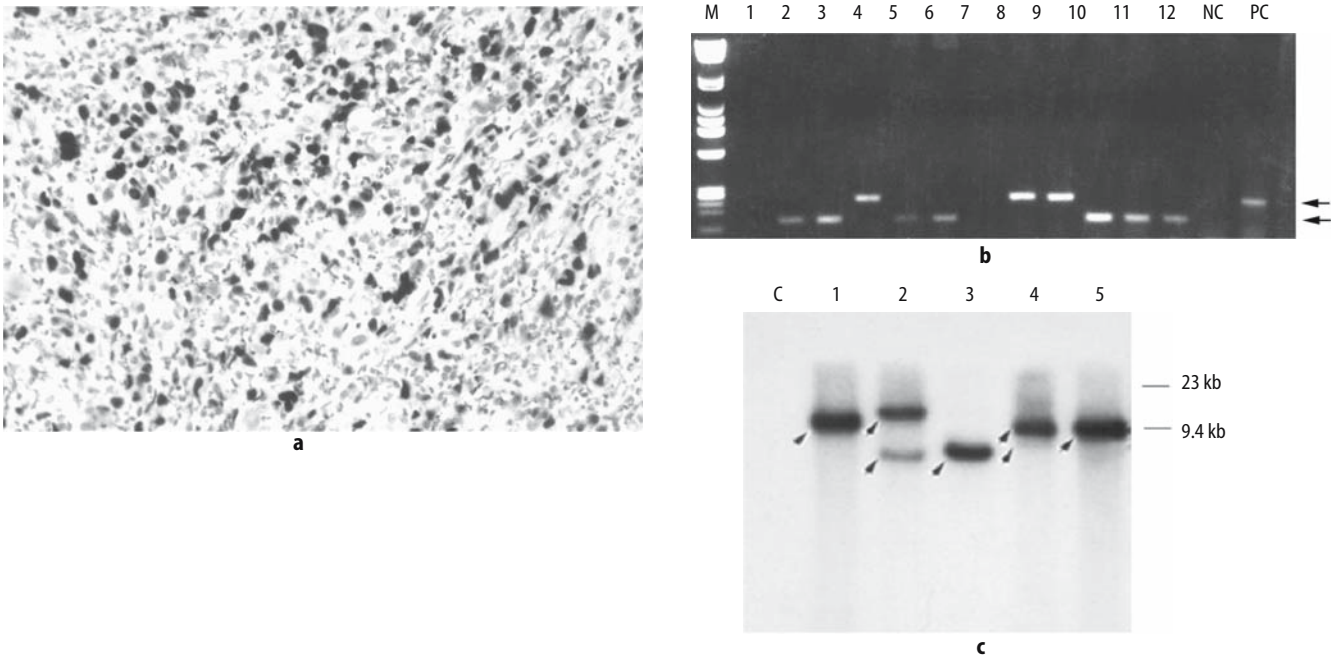
Viral Associations

Epstein-Barr Virus

The presence of EBV is most frequently determined by in situ hybridization (ISH) for EBERs which are abundantly expressed, small, untranslated RNAs.[20] Commercial kits are available for EBER ISH, which can be performed on routinely processed, formalin-fixed, paraffin-embedded tissue sections (Figure 34-2). An alternative or complementary approach is PCR analysis for the EBV genome. This can be qualitative or quantitative. Qualitative PCR can be performed for specific regions of the EBV genome that have strain-specific differences (type 1 or A vs type 2 or B), like EBNA2 and EBNA3C (Figure 34-2). Although Type 1 EBV is more transforming in vitro and is almost invariably associated with PTLD, AIDS lymphomas can be infected by either strain. Quantitative PCR is most frequently performed using Taqman chemistry or SYBR Green and melting-curve analysis using DNA from whole blood, serum, blood mononuclear cells, or cerebrospinal fluid to monitor EBV viral loads in patients with EBV-associated lymphomas.[19,25] Another useful method is RT-PCR for EBV gene expression characterization, which can distinguish among different types of latency.[26] This assay, however, is used more commonly for research than for clinical purposes. Alternatively, the different types of latency can be distinguished by immunohistochemistry using antibodies to several different EBV antigens, in particular to LMP1 and EBNA2.

Kaposi Sarcoma–Associated Herpesvirus

The most common method for detecting KSHV infection is PCR, although antibodies are now available for KSHV detection by immunohistochemistry. Although different PCR primers are used by different laboratories, the most commonly used primer set is called "KS330" and targets the ORF26 region of the viral genome.[27] This primer set was described with the original identification of KSHV in 1994 and used in most early studies. Another primer set with high sensitivity and specificity targets the K9 open reading frame of the KSHV genome.[28] While some studies suggest that quantitative KSHV assays are useful for monitoring Kaposi's sarcoma evolution, the utility for monitoring KSHV-associated lymphomas is still unclear.



a b c

Figure 34-2. Detection of Epstein-Barr virus. (a) In situ hybridization for EBER is positive in an AIDS-related immunoblastic lymphoma. Numerous large cells with black staining can be seen. (b) PCR analysis for the presence of EBV using primers for the EBNA3C region, which distinguishes type 1 and 2 (also called A and B) EBV in a panel of AIDS lymphomas. Type 1 isolates show a PCR band of 153 base pairs (bp; lower arrow), and type B isolates result in a band of 246 bp (upper arrow). Lanes 1 through 12 represent different AIDS-related lymphomas. Cases 1 and 7 were negative for EBV, and all other cases were positive. In AIDS lymphomas, both strains are common, while in PTLD, EBV type 1 is involved in the vast majority of cases. NC is a negative (uninfected) control, and PC is a positive control for Type 2 EBV. (c) EBV Southern blot showing genomic DNA digested with *Bam*H I, followed by electrophoresis, transfer and hybridization with a radiolabeled probe for the terminal repeat region of EBV. C represents the EBV negative cell line, HL-60. Lanes 1 through 5 are analysis of five cases of PTLD showing monoclonal rearrangements for EBV (arrowheads). Case two has two distinct bands, indicating two different EBV-positive clonal populations.

Oncogenes/Tumor Suppressor Genes

Testing for the presence of alterations in *MYC*, *TP53*, *NRAS*, and *BCL6* is relevant for immunodeficiency-associated lymphomas. Of these, probably the most important and feasible in a clinical molecular laboratory on a routine basis is the detection of *MYC* translocation using fluorescence (FISH) or chromogenic (CISH) in situ hybridization assays.

Fluorescence In Situ Hybridization

FISH is an excellent method for documenting specific chromosomal translocations. FISH can be performed using interphase nuclei or metaphase chromosomes and can be performed using formalin-fixed, paraffin-embedded tissue sections. As commercial probes are entering the market, FISH is being implemented by increasing numbers of clinical laboratories. In the setting of immunodeficiency lymphomas, FISH is the best assay to document *MYC* translocations when a Burkitt lymphoma is suspected or a large cell lymphoma has Burkitt-like features. FISH may be performed either with two different probes, one to *MYC* and the other to *IGH*, or with a single *MYC* "split apart" probe (Figure 34-3). The advantage of the second single *MYC* probe method is detection of variant translocations of *MYC* and the *IGLK* or *IGLL* genes.

Chromogenic In Situ Hybridization (CISH)

CISH is very similar to FISH, except the probes are labeled with two different chromogenic substrates, so their evaluation does not require a fluorescent microscope and the slides provide a permanent record of the assay. While somewhat less sensitive than FISH, commercial CISH probes to detect the *MYC-IGH* translocation are available, and are being used successfully in clinical laboratories. CISH may soon become standard of care for diagnosis of BL or atypical BL when histologic features are suggestive of this disease.

Southern Blot Analysis

SBA can be used to document translocations involving the *MYC* and *BCL6* genes, but as mentioned above it is labor-intensive and time-consuming, has long turnaround times, and requires fresh or frozen tissue. Thus, this technique is used primarily for research applications. A comparison of CISH and Southern blot for the *MYC-IGH* translocation is shown in Figure 34-3.

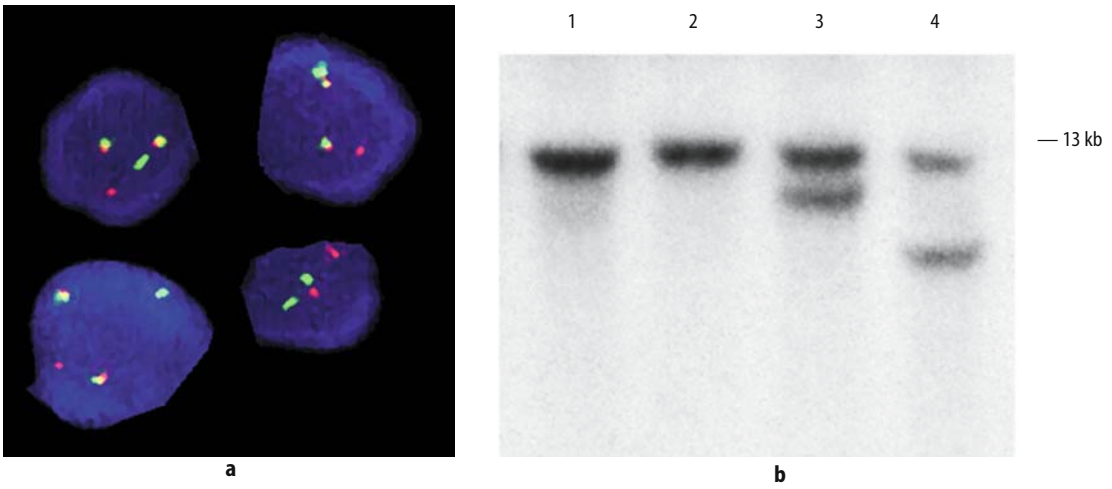


Figure 34-3. Detection of *MYC* translocations. (a) FISH analysis of Burkitt lymphoma cells demonstrating the *MYC-IGH* fusion signals observed in a case positive for t(8;14). *IGH* is labeled with Spectrum Green and *MYC* is labeled with Spectrum Orange. Two fusion signals show the reciprocal translocation. Single red and green signals correspond to the wild-type allele of *MYC* on chromosome 8 and *IGH* on chromosome 14. (Image provided by Dr. Susan Mathew, Weill Medical College of Cornell University.) (b) Southern blot analysis using a *MYC* probe in four cases of AIDS-related lymphoma. Cases 1 and 2 show a single band, corresponding to the germline configuration, and cases 3 and 4 have a *MYC* translocation, as indicated by a smaller second band. DNA was digested with *Hind* III, and after electrophoresis and transfer, was hybridized to a radiolabeled probe corresponding to the third exon of the *MYC* gene.

Single-Strand Conformation Polymorphism Analysis

Single-strand conformation polymorphism analysis (SSCP) is a useful assay to screen for point mutations of *RAS* and *TP53*, although it is not used routinely in clinical laboratories due to the complexity of the assay. The procedure should be carefully validated and the sensitivity of the assay determined to allow for adequate interpretation of results.

Sequence Analysis

DNA sequencing is the ultimate confirmation of point mutations and small deletions or insertions in relevant genes. Sequence analysis can be performed by direct sequencing of PCR products of the target region, or by cloning the PCR products into a plasmid vector followed by sequencing of multiple independent clones. Direct sequencing is the easiest, but it can be done only when the tissue is composed almost entirely of a pure tumor population, as the mutation may be masked by infiltrating reactive cells in addition to the normal allele of the gene of interest. Therefore, mutations of *TP53* are likely to be identifiable using direct sequencing on an AIDS-related BL, where the majority of the cells are neoplastic and one *TP53* allele is likely to be deleted. In contrast, *BCL6* mutations in PTLD are unlikely to be identifiable by sequencing. An alternative approach that can be used in a mixed cell population is PCR followed by cloning into plasmid vectors and sequencing of multiple clones. A mutant allele will consistently show up in a proportion of the clones, while the normal allele from tumor cells and normal infiltrating cells will represent the remainder of the sequenced clones. Of note, because DNA polymerases used in the PCR are prone to errors, the identical alteration has to be present in multiple clones to be scored as positive.

Interpretation of Results

Immunoglobulin Gene Rearrangement Analysis

IGH gene rearrangement is most frequently determined by PCR analysis, which has a high rate of both false-positive and false-negative results in the context of immunodeficiency-related lymphomas. In the setting of SOT, monoclonal populations are identified by PCR but not by SBA in the low-grade or early lesions (PH) that frequently regress following a reduction of immunosuppression. Similarly, in AIDS-related lymphadenopathies, monoclonal populations can be identified by PCR but usually not by SBA. These PCR-positive but SBA–negative results are probably due to strong antigenic stimuli from infectious agents such as EBV, giving rise to small mono-

clonal expansions that are identified by the more-sensitive PCR assay. False-negative results also are common, since many posttransplantation and AIDS-related lymphomas are of postgerminal center B-cell origin, where extensive somatic hypermutation of the *IGH* VDJ region has occurred. These mutations may be in the primer recognition site, precluding primer annealing and PCR amplification. SBA is sometimes helpful in eliminating false-negative and false-positive results but is impractical in a clinical molecular laboratory setting. SBA also can give false-negative results when the tumor cells comprise less than 1% to 5% of the population in the tissue being analyzed. It has been suggested that a significant proportion of AIDS lymphomas are polyclonal. By combining SBA and PCR, including analysis of *IGLK* and *IGLL* genes and EBV clonality, monoclonal populations can be documented in practically every case of histologically confirmed AIDS-related lymphoma. However, in the absence of such extensive analysis, the inability to demonstrate a monoclonal population should never be taken as an absolute criterion against the diagnosis of lymphoma. Similarly, the presence of a monoclonal population by PCR should not be considered diagnostic of lymphoma. Either a positive or a negative result should be correlated with clinical and pathologic information to establish the final diagnosis.

Viral Associations

PCR can be used for the identification of a virus in a specimen. However, a positive PCR can be seen from the presence of a given virus either in the tumor cells or in the infiltrating reactive cells in a latently infected individual. Thus, a positive PCR result does not necessarily indicate that the tumor cells contain a specific virus. This problem is more common for EBV than for KSHV, since latent EBV infection is ubiquitous, while latent KSHV infection is not.[29] Assays that are quantitative or semiquantitative can distinguish virus in tumor cells from infiltrating cells, where the viral copy number would be at least as high as the number of tumor cells. Infiltrating benign lymphocytes would correspond to a very small proportion of the total cell population unless active viral infection is present (such as infectious mononucleosis). Alternatively, for EBV, ISH for EBER can be used to confirm the presence of EBV in the tumor cells. For KSHV, the best approach for confirmation of a positive PCR result is immunohistochemistry with monoclonal antibodies for ORF73, which encodes the latency-associated nuclear antigen (LANA or LNA).

Oncogenes and Tumor Suppressor Genes

As each technique for analysis of oncogenes and tumor suppressor genes has its own advantages and disadvantages, the individual using this information for classification, diagnostic purposes, or both should be familiar with the method used, the laboratory performing the assays, and specific caveats for the given procedure. It is always recommended that the information be correlated with clinical and pathologic information for the proper interpretation of a particular result.

Laboratory Issues

The most important laboratory issue of concern, particularly with respect to PCR analysis, is the avoidance of contamination. When performing PCR for EBV or KSHV repeatedly, it is essential to use extreme care in separating the PCR products from the reagents and specimens. When using ultrasensitive techniques, such as nested PCR, the potential for contamination is intensified. Complete separation of pre- and post-PCR steps should be maintained. Cross contamination among patient samples and controls should be avoided. PCR primers for specific assays tend to differ among various laboratories, such that each laboratory validates and sets normal ranges for the assay.

An issue that is specific to AIDS-related lymphomas is the concern of working with HIV-infected tissues. While the lymphoma cells do not carry the HIV genome, the tissue may have infected T cells or macrophages. If DNA is being extracted from formalin-fixed, paraffin-embedded tissues, infectious HIV is not a concern, as any viral particles would have been inactivated. However, when extracting nucleic acids from fresh and frozen tissues, extreme care should be taken and the individual doing the extraction should be informed of the risks and be properly trained. Nucleic acid extraction from fresh or frozen tissue should be done at a BSL-2 biosafety level, which essentially requires the procedure to be done in a laminar flow hood and for the technologist to wear gloves and a laboratory coat. Since blood-borne pathogen exposure is a risk when working with any tissue, and HIV status is frequently unknown, the use of gloves and a lab coat are required for every extraction involving fresh or frozen tissues (universal precautions).

A specific difficulty for performing some of the oncogene and virus detection assays is obtaining good positive controls. For *MYC* translocations, any Burkitt lymphoma cell line can be used, and many are available from the American Type Culture Collection (ATCC). For *BCL6* mutations, controls are more difficult to obtain, as SSCP involves amplification of separate but overlapping regions, with different controls required for each region. Thus, controls will vary with the amplification fragment established by the individual laboratory. Several diffuse large B-cell lymphoma lines have *BCL6* mutations, or patient specimens with *BCL6* mutations can be used as controls. For virus detection, several EBV-positive lymphomas and lymphoblastoid cell lines, as well as KSHV-positive primary effusion lymphomas, are available from the ATCC and the AIDS and Cancer Specimen Bank (ACSB) repositories. PEL

cell lines are available through ATCC that contain both EBV and KSHV, which can be used as a positive control for multiple assays.

References

1. Jaffe ES, Harris NL, Stein H, Vardiman JW. *Pathology and Genetics of Tumours of Hematopoietic and Lymphoid Tissues.* Lyon, France: IARC Press; 2001.
2. Gaidano G, Pasqualucci L, Capello D, et al. Aberrant somatic hypermutation in multiple subtypes of AIDS-associated non-Hodgkin lymphoma. *Blood.* 2003;102:1833–1841.
3. Gaidano G, Carbone A, Dalla-Favera R. Genetic basis of acquired immunodeficiency syndrome-related lymphomagenesis. *J Natl Cancer Inst Monogr.* 1998;23:95–100.
4. Carbone A. Emerging pathways in the development of AIDS-related lymphomas. *Lancet Oncol.* 2003;4:22–29.
5. Knowles DM. *Neoplastic Hematopathology.* 2nd ed. Baltimore: Williams and Wilkins; 2001.
6. Raphael M, Borisch B, Jaffe ES. Lymphomas associated with infection by the human immune deficiency virus (HIV). In: Jaffe ES, Harris NL, Stein H, Vardiman JW, eds. *World Health Organization Classification of Tumors, Tumors of Haematopoietic and Lymphoid Tissues.* Lyon, France: IARC Press; 2001:260–263.
7. Nalesnik MA, Jaffe R, Starzl TE, et al. The pathology of posttransplant lymphoproliferative disorders occurring in the setting of cyclosporine A-prednisone immunosuppression. *Am J Pathol.* 1988; 133:173–192.
8. Knowles DM, Cesarman E, Chadburn A, et al. Correlative morphologic and molecular genetic analysis demonstrates three distinct categories of posttransplantation lymphoproliferative disorders. *Blood.* 1995;85:552–565.
9. Cesarman E, Chadburn A, Liu YF, Migliazza A, Dalla-Favera R, Knowles DM. BCL-6 gene mutations in posttransplantation lymphoproliferative disorders predict response to therapy and clinical outcome. *Blood.* 1998;92:2294–2302.
10. Shapiro RS, McClain K, Frizzera G, et al. Epstein-Barr virus associated B cell lymphoproliferative disorders following bone marrow transplantation. *Blood.* 1988;71:1234–1243.
11. Chadburn A, Hyjek E, Frizzera G, et al. Post-transplantation lymphoproliferative disorders (PT-LPDs) in bone marrow and solid organ transplant recipients differ. *Blood.* 2000;96(suppl 1):505a.
12. Abed N, Casper JT, Camitta BM, et al. Evaluation of histogenesis of B-lymphocytes in pediatric EBV-related post-transplant lymphoproliferative disorders. *Bone Marrow Transplant.* 2004;33:321–327.
13. Chadburn A, Hyjek E, Pan L, et al. Bcl-6 gene mutations in bone marrow transplantation (BMT). *Mod Pathol.* 2001;14:167A.
14. Chadburn A, Chen JM, Hsu DT, et al. The morphologic and molecular genetic categories of posttransplantation lymphoproliferative disorders are clinically relevant. *Cancer.* 1998;82:1978–1987.
15. Leblond V, Davi F, Charlotte F, et al. Posttransplant lymphoproliferative disorders not associated with Epstein-Barr virus: a distinct entity? *J Clin Oncol.* 1998;16:2052–2059.
16. Choquet S, Mamzer BM, Hermine O, et al. Identification of prognostic factors in post-transplant lymphoproliferative disorders. *Recent Results Cancer Res.* 2002;159:67–80.
17. Hoshino Y, Kimura H, Kuzushima K, et al. Early intervention in posttransplant lymphoproliferative disorders based on Epstein-Barr viral load. *Bone Marrow Transplant.* 2000;26:199–201.
18. Niesters HG, van Esser J, Fries E, Wolthers KC, Cornelissen J, Osterhaus AD. Development of a real-time quantitative assay for detection of Epstein-Barr virus. *J Clin Microbiol.* 2000;38:712–715.
19. Fan H, Gulley ML. Epstein-Barr viral load measurement as a marker of EBV-related disease. *Mol Diagn.* 2001;6:279–289.
20. Gulley ML. Molecular diagnosis of Epstein-Barr virus-related diseases. *J Mol Diagn.* 2001;3:1–10.
21. van Esser JW, Niesters HG, Thijsen SF, et al. Molecular quantification of viral load in plasma allows for fast and accurate prediction of response to therapy of Epstein-Barr virus-associated lymphoproliferative disease after allogeneic stem cell transplantation. *Br J Haematol.* 2001;113:814–821.
22. Clave E, Agbalika F, Bajzik V, et al. Epstein-Barr virus (EBV) reactivation in allogeneic stem-cell transplantation: relationship between viral load, EBV-specific T-cell reconstitution and rituximab therapy. *Transplantation.* 2004;77:76–84.
23. 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;40:351–358.
24. Raab-Traub N, Flynn K. The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation. *Cell.* 1986;47:883–889.
25. Leung E, Shenton BK, Jackson G, Gould FK, Yap C, Talbot D. Use of real-time PCR to measure Epstein-Barr virus genomes in whole blood. *J. Immunol. Methods.* 2002;270:259–267.
26. Brooks L, Yao QY, Rickinson AB, Young LS. Epstein-Barr virus latent gene transcription in nasopharyngeal carcinoma cells: coexpression of EBNA1, LMP1, and LMP2 transcripts. *J Virol.* 1992;66:2689–2697.
27. Chang Y, Cesarman E, Pessin MS, et al. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science.* 1994;266:1865–1869.
28. Pan L, Milligan L, Michaeli J, Cesarman E, Knowles DM. Polymerase chain reaction detection of Kaposi's sarcoma-associated herpesvirus-optimized protocols and their application to myeloma. *J Mol Diagn.* 2001;3:32–38.
29. Cesarman E. Epstein-Barr virus (EBV) and lymphomagenesis. *Front Biosci.* 2002;7:e58–e65.

Chapter 35

Myeloproliferative Disorders and Myelodysplastic Syndromes

Jo-Anne Vergilio and Adam Bagg

The myeloproliferative disorders (MPD) and myelodysplastic syndromes (MDS) encompass a pathologically and clinically heterogeneous group of hematologic entities that are united by their putative origin from pluripotent hematopoietic stem cells. The World Health Organization classifies these entities into three broad categories: (1) the chronic myeloproliferative diseases, (2) the myelodysplastic syndromes, and (3) the myelodysplastic/myeloproliferative diseases[1] (see Table 35-1). Though all are true hematopoietic stem cell disorders characteristically associated with bone marrow hyperplasia, they are divergent in that MPD typically are associated with *effective* hematopoiesis, while MDS are associated with *ineffective* hematopoiesis, reflected by high or low peripheral blood counts, respectively.

Of the 17 discrete diseases currently recognized, the molecular pathogenesis of chronic myelogenous leukemia (CML) is best understood and ultimately has resulted in the development and application of targeted molecular therapy. As a result, molecular techniques play an integral and well-established role in the diagnosis of CML, in the determination of therapeutic response and efficacy, and in the detection of minimal residual disease. Consequently, most of this chapter focuses on the role of molecular pathology in the laboratory assessment of CML. However, the molecular pathogenesis of some of the other MPDs has recently been elucidated, with the potential for the development of clinically useful molecular tests. By contrast, the molecular underpinnings of MDS have not yet been as well delineated; nonetheless, molecular pathology testing has emerging utility in the diagnosis and assessment of these disorders and also is discussed briefly.

CHRONIC MYELOGENOUS LEUKEMIA

Molecular Basis of Disease

CML comprises ~20% of all leukemias and is diagnosed at a median age of approximately 50 years. The disease originates from the transformation of a hematopoietic stem cell with resultant expanding myelopoiesis that characteristically evolves through three phases when untreated: (1) a chronic phase of 4 to 5 years' duration manifest by myeloid hyperplasia with circulating granulocytes that are present in all stages of maturation; (2) an accelerated phase of shorter duration during which myeloid elements begin to lose the ability to differentiate; and (3) inevitably, a blast phase of acute leukemia of myeloid (70%) or lymphoid (30%) phenotype.

The reciprocal t(9;22)(q34;q11) translocation is identified as the initial transforming event in the development of CML, although some data suggest that it may be preceded by clonal hematopoiesis. The translocation yields a shortened chromosome 22 called the Philadelphia chromosome (Ph)[2] (Figure 35-1). With the translocation, two distinct genes are fused: (1) *BCR* that encodes a cytoplasmic protein of uncertain function but with oligomerization, serine-threonine kinase, and GTPase-activating domains, and (2) *ABL1* that encodes a nonreceptor tyrosine kinase normally localized to the nucleus.[3] The resultant chimeric gene and fusion transcript yield a fusion protein with constitutively increased tyrosine kinase activity that is relocated from the nucleus to the cytoplasm and phosphorylates a variety of cellular substrates. The result is growth factor–independent proliferation, decreased apoptosis, and defective adhesion in the transformed cells.

The t(9;22) translocation can involve several different breakpoints in the *BCR* and *ABL1* genes, resulting in different chimeric fusion proteins that confer somewhat specific clinicopathologic features and highlight the fact that this translocation is not pathognomonic for CML[4] (Figure 35-2). These breakpoints are indistinguishable by traditional karyotyping and can be differentiated only using molecular techniques. For all fusions, most of *ABL1* is juxtaposed to variable 5′ portions of *BCR*. Whereas the breakpoint involving *ABL1* is relatively conserved, usually arising in the intron before exon 2 (a2), the breakpoints involving *BCR* are more variable.

BCR breakpoints arising in the major breakpoint cluster region (M-bcr) occur after either exon 13 (e13 or b2) or exon

Table 35-1. The World Health Organization Classification of MPD and MDS

Chronic Myeloproliferative Diseases
Chronic myelogenous leukemia
Chronic neutrophilic leukemia
Chronic eosinophilic leukemia/hypereosinophilic syndrome
Polycythemia vera
Chronic idiopathic myelofibrosis
Essential thrombocythemia
Chronic myeloproliferative disease, unclassifiable
Myelodysplastic Syndromes
Refractory anemia
Refractory anemia with ringed sideroblasts
Refractory cytopenia with multilineage dysplasia
Refractory anemia with excess blasts
Myelodysplastic syndrome, unclassifiable
Myelodysplastic syndrome associated with isolated del(5q) chromosome abnormality
Myelodysplastic/Myeloproliferative Diseases
Chronic myelomonocytic leukemia
Atypical chronic myelogenous leukemia
Juvenile myelomonocytic leukemia
Myelodysplastic/myeloproliferative diseases, unclassifiable

14 (e14 or b3). Fusion transcripts involving the M-bcr encode a p210 fusion protein and are present in the vast majority (~99%) of patients with CML as well as in ~40% of adult and ~10% of pediatric patients with t(9;22)-positive precursor B-cell acute lymphoblastic leukemia (ALL).[5]

In contrast, breakpoints in the minor breakpoint cluster region (m-bcr) are more 5′ in location, occurring after exon 1 (e1). Breakpoints involving the m-bcr, which ultimately result in a smaller (p190) fusion protein with enhanced transforming potential, are rarely present in CML. The m-bcr breakpoints and fusion transcripts are present in ~60% of adult and ~90% of pediatric patients with t(9;22)-positive precursor B-ALL.[5]

Overall, the presence of t(9;22) is quite common in adult patients with ALL (~25% ALL) but relatively less common in pediatric patients (~5% ALL). Infrequently, an extreme 3′ breakpoint occurs beyond exon 19 (e19) of *BCR* in the designated micro region (μ-bcr), resulting in a larger (p230) fusion protein characteristically associated with neutrophilic CML, in which mature granulocytes predominate, as well as in chronic neutrophilic leukemia (CNL), although recent data indicate a less well-substantiated association with the latter.[6,7] The Philadelphia chromosome also has been detected in rare cases of de novo acute myeloblastic leukemia and T-cell ALL.

The occurrence of additional specific cytogenetic and molecular genetic events subsequent to the initial t(9;22) translocation heralds disease progression prior to hematologic or clinical manifestations or both. The acquisition of such cytogenetic abnormalities as a second Philadelphia chromosome (Ph^+), isochromosome 17q, trisomy 8, trisomy 19, and others commonly indicates an impending blast crisis.[8,9] Other molecular abnormalities associated with disease progression include overexpression of *BCR-ABL1*, upregulation of *EVI1* gene expression, mutations in tumor suppressor genes such as *P16*, *TP53*, *CDKN2A*, and *RB1*, and aberrant DNA methylation of the translocated *ABL1* allele and of the calcitonin gene.[10] The presence of a derivative chromosome 9 deletion in addition to the t(9;22) translocation may serve as an independent prognostic factor predicting a rapid progression to blast crisis with a worse response to therapy and, hence, a shortened survival.[11]

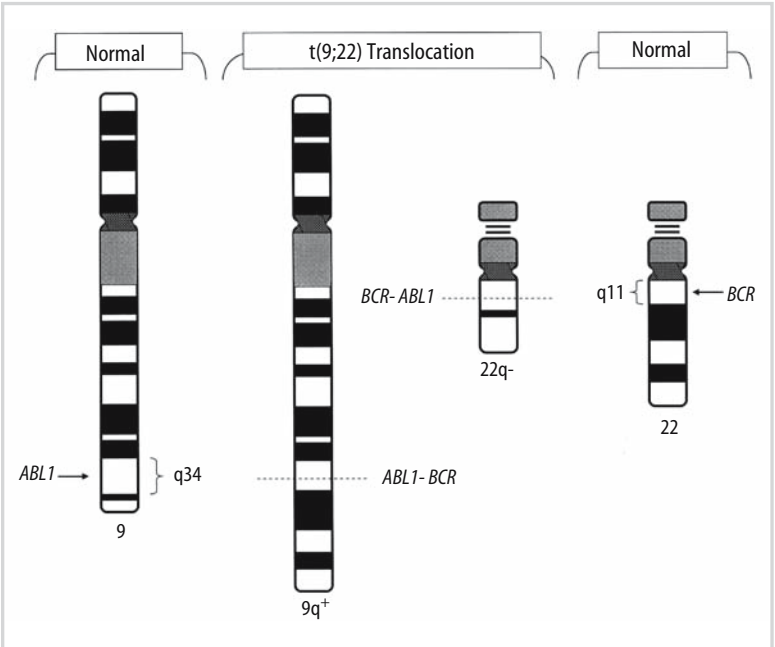


Figure 35-1. Schematic representation of the reciprocal t(9;22)(q34;q11) translocation with the disruption and fusion of *ABL1* from the long arm of chromosome 9 with *BCR* on the long arm of chromosome 22. The resultant 22q– is the Philadelphia chromosome and contains the *BCR-ABL1* fusion gene, while chromosome 9q+ contains the reciprocal *ABL1-BCR* fusion gene that apparently lacks biologic function, but which is expressed in approximately 60% of patients with CML.

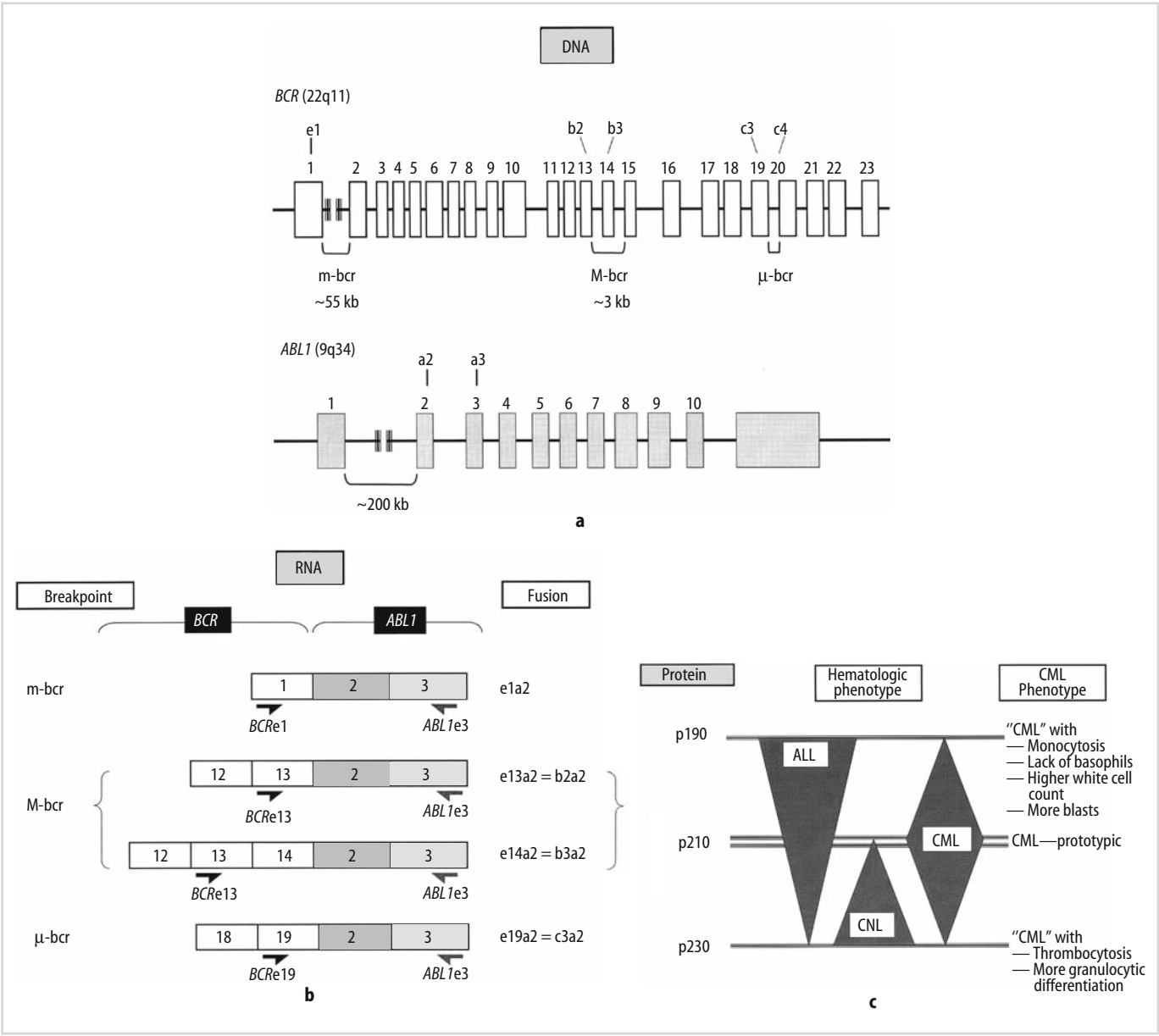


Figure 35-2. Schematic representation of (a) the *BCR* and *ABL1* genes demonstrating the primary breakpoint cluster regions (bcr), (b) the RNA/cDNA structure of the most common chimeric transcripts and the primers used for their RT-PCR amplification, and (c) the resultant oncoproteins with their specific clinicopathologic correlates. See text for details. The association of the p230 with CNL has been questioned, with this being more a feature of "neutrophilic CML." ALL, acute lymphoblastic leukemia; CML, chronic myelogenous leukemia; CNL, chronic neutrophilic leukemia. (Modified, with permission, from Vergilio JA and Bagg A. Chronic myeloid leukaemia—Molecular diagnosis and monitoring. In: Encyclopedia of Medical Genomics and Proteomics (EMGP). Jürgen Fuchs and Maurizio Podda, (eds.) Marcel Dekker 2005:252–258 and Bagg A. Chronic myeloid leukemia: A minimalistic veiw of post-therapeutic monitoring. *Journal of Molecular Diagnosis* 2002;4:1–10.

Indications for Testing

There are two broad scenarios in which molecular testing is indicated in CML, at diagnosis and for monitoring during and following therapy.

Initial Diagnosis

Cytogenetics serves to identify the presence of the t(9;22) translocation in only approximately 95% of cases of CML but can identify the presence of other chromosomal abnormalities. In one half of the patients with a normal karyotype, the *BCR-ABL1* fusion transcript is detectable at the molecular level. This discrepancy is due to a submicroscopic genetic fusion. Therefore, in those patients with a normal karyotype, who have the clinical and hematological profile of CML, molecular testing serves a primary role in CML diagnosis. In the remaining patients, who are negative for both the Philadelphia chromosome and the *BCR-ABL1* fusion transcript, alternative diagnoses such as atypical CML or chronic myelomonocytic leukemia should be considered (Figure 35-3).

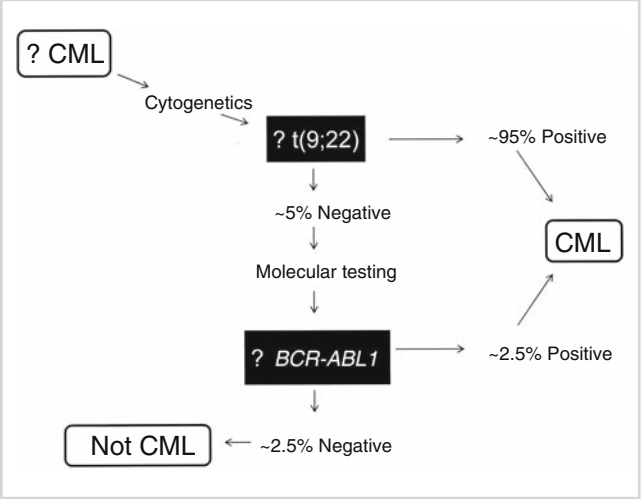


Figure 35-3. Broad diagnostic algorithm for the utilization of cytogenetic and molecular testing in the initial diagnosis of CML, highlighting the complementary roles of these technologies. Even if cytogenetic analysis is positive, molecular testing is required to determine the molecular target for therapy and monitoring.

Molecular testing serves to document the presence of *BCR-ABL1* fusion transcripts at diagnosis, both for the rational use of targeted therapy (imatinib mesylate and newer agents) and for subsequent molecular monitoring. Indeed, once the initial diagnosis of CML is established, periodic monitoring of both the therapeutic response and the level of residual disease becomes critical to evaluation and therapeutic decision-making.

Monitoring

Following therapy, disease response is assessed using three parameters: hematologic, cytogenetic, and molecular status. Hematologic remission is achieved when the blood counts and spleen size have normalized. Cytogenetic response is quantified and graded based on the percentage of residual Ph-positive cells, with a major complete response defined as 0% Ph-positive cells, a major-partial response as 1% to 34% Ph-positive cells, a minor response as 35% to 94% Ph-positive cells, and no response as ≥95% Ph-positive cells.[12] Traditionally, the cytogenetic response serves as the "gold standard" for assessment and is an important predictor of patient survival.

Once conventional hematologic remission and cytogenetic complete response have been achieved, monitoring relies on more-sensitive molecular techniques. When the *BCR-ABL1* fusion transcript is no longer detectable, molecular remission has been attained. However, a widely accepted definition of molecular remission does not yet exist, and its designation varies with test method and sensitivity. Nonetheless, a single negative or positive *qualitative* reverse transcription–polymerase chain reaction (RT-PCR) result is, in itself, of little clinical predictive value, since such assays may lack precision, depending on the sensitivity of the specific test. In contrast, *serial quantitative* tests that assess the kinetics of tumor clearance (response) or reappearance (relapse) have greater predictive value. Rising levels of *BCR-ABL1* messenger RNA (mRNA) have been shown to precede disease recurrence and may signal a need for therapeutic intervention, while low, diminishing, or stable fusion transcript levels identify patients in whom treatment has been effective.

A rational understanding of laboratory monitoring is predicated on an appreciation of therapeutic goals and options. Indeed, eradication of the neoplastic clone cannot always be achieved, in which case therapeutic intent focuses on tumor control with an attempt to forestall disease progression. Standard treatments available to patients in the chronic phase of CML include allogeneic stem cell transplantation (ASCT), imatinib mesylate, interferon-α, and hydroxyurea (Table 35-2), with imatinib mesylate having become the new gold standard for initial therapy.[13,20]

Molecular Monitoring in ASCT

Notwithstanding the recent excitement regarding imatinib mesylate, ASCT is the only proven curative therapy for CML, with cure rates of 70% to 80% in young (age <40 years), chronic-phase patients that have HLA-matched donors and undergo transplantation within 1 year of diag-

Table 35-2. Responses to Treatment for Early Chronic-Phase CML

	Therapeutic Response (% patients)					
		Cytogenetic		Molecular		
Standard Therapies	Hematologic (complete)	Major	Complete	Major	Complete	Cure
ASCT	~75	~75	~75	~75	~75	~75
Chemotherapy	~35	<1	<1	<1	<1	<1
Interferon-α	~80	~30	~15	<5–25	<5	<1
Imatinib mesylate	~95	~85	~75	~50	<5	<1

Data summarized and averaged from a variety of studies.[13,14–18,19]
With regard to molecular testing, only in ASCT is the endpoint complete molecular remission, while for interferon-α and imatinib mesylate, the goal is a response, with complete molecular remission rarely attained. However, reports of molecular responses vary widely with different definitions and test methods, with some studies[16] reporting a complete molecular response rate of up to 28%.
Major molecular response is currently best defined as a ≥3 log reduction in % *BCR-ABL/ABL*.

nosis.[13,20] Despite the curative success of ASCT, the associated morbidity and mortality are significant; furthermore, the majority (~65%) of young patients do not have a suitably matched donor, while older patients are often suboptimal candidates for transplantation.

After transplantation, molecular testing serves two functions, first to document remission and then to monitor for disease relapse. Molecular monitoring for relapse permits early disease detection (prior to hematologic or cytogenetic manifestations), when the tumor burden is low and presumably more amenable to treatment. At molecular relapse, therapeutic options include the withdrawal of immunosuppressive agents, the administration of donor lymphocyte infusions (DLI), or both, with an increased likelihood of response achieved when DLI are administered prior to overt hematologic relapse.[21]

Molecular Monitoring in Other Therapeutic Settings

For patients who are not suitable candidates for ASCT, alternative therapies have historically included hydroxyurea and interferon-α; however, these agents have largely been supplanted by imatinib mesylate (or Gleevec, STI571), a synthetic tyrosine kinase inhibitor.[22] Imatinib mesylate is one of the first examples of molecular-targeted drug therapy and has shown activity in all phases of CML, with substantial responses in newly diagnosed patients in chronic phase.[14–17] Thus far, imatinib has shown improved hematologic and cytogenetic responses, as well as a prolonged freedom from disease progression when compared to more traditional therapies. In an ongoing randomized Phase III trial (IRIS), the observed rate of complete cytogenetic remission in newly diagnosed CML patients was 76% with imatinib versus 15% with interferon-α plus Ara-C (median follow-up, 18 months). In addition to its apparent clinical superiority, imatinib also is tolerated well, a therapeutic aspect that is in marked contrast to interferon-α, with its numerous prohibitive toxicities. Unlike imatinib and interferon-α, hydroxyurea successfully lowers the peripheral leukocyte count but neither induces a cytogenetic response nor delays disease progression, thereby limiting its role in current therapy.

Although long-term outcome data are not yet available, imatinib has emerged as the primary therapy of choice for newly diagnosed CML, though it is probably still not curative since the *BCR-ABL1* fusion transcript typically persists at low levels. Persistence of molecular disease *without* relapse suggests attainment of a "dormant" state that may be induced and regulated by various mechanisms, including immune surveillance and impaired proliferation, that confer clonal quiescence. Recent reports suggest that high doses of imatinib (800 mg vs 400 mg) may indeed induce molecular remission in up to 28% of patients.[18]

Therapeutic resistance to imatinib has been documented in some patients, particularly in those with more advanced stages of disease. Mechanisms of resistance include (1) amplification of the *BCR-ABL1* chimeric gene at the genomic or transcript level, (2) point mutations in the kinase domain of *BCR-ABL1* that alter its responsiveness to imatinib, and (3) overexpression of multidrug resistance P-glycoprotein that transports imatinib from cells.[23] Monitoring of the fusion transcript levels in this scenario allows for the early recognition of therapeutic resistance with as little as a two-fold increase predictive of resistance,[24] which might prompt the implementation of alternative treatment strategies. Future molecular testing might be directed toward detection of resistance, for example, by screening for gene-specific point mutations.

Molecular techniques, therefore, encompass many aspects of laboratory testing that are critical to the overall evaluation of patients with CML and contribute to the diagnosis, assessment of therapeutic efficacy, and evaluation of minimal residual disease in these patients. The clinical context of the patient and the specific treatment modality determine the specific tests that are appropriate, and algorithms have been proposed to effectively and economically guide both laboratory diagnosis and monitoring[25,26] (Figures 35-4a and b). These algorithms appropriately incorporate various test methods including cytogenetic and fluorescence in situ hybridization (FISH) analysis in addition to molecular techniques. Though this chapter focuses on molecular methodologies, particularly qualitative and quantitative RT-PCR, it is necessary to recognize the distinct, yet complementary, role that each method plays.

Available Assays

Molecular testing for CML can be directed at the chimeric gene (e.g., genomic PCR, Southern blot), the fusion mRNA transcript (e.g., RT-PCR, northern blot), or the oncoprotein (e.g., Western blot). Given the widely dispersed intronic breakpoints in CML, genomic DNA is not amenable to routine PCR analysis; furthermore, Southern and northern analyses require relatively large amounts of high-molecular-weight DNA or RNA, typically entail the use of radioactivity, and are labor-intensive and have extended turnaround times. Protein-based analysis lacks sensitivity, a feature that ultimately limits its utility. RT-PCR is, therefore, the preferred testing method for evaluation and may be either qualitative or quantitative.

Qualitative RT-PCR

Qualitative RT-PCR has several attributes that make it well suited for initial diagnosis. Bone marrow aspirate and peripheral blood are both suitable and comparable specimens,[27] obviating the need for frequent invasive procedures, while sensitivity and rapid turnaround facilitate initiation of therapy.

Methodologically, qualitative RT-PCR can be performed with a simple, nested, or multiplex approach. However, nested methods are suboptimal due to a significant risk of PCR contamination and consequent false-positive results. With nonnested, simplex RT-PCR, one assay is performed using a single pair of primers. Primers for *BCR* exon 13 (b2) and *ABL1* exon 2 (a2) identify both the e13a2 (b2a2) and e14a2 (b3a2) fusion transcripts that differ in size by 75 base pairs (bp) (Figure 35-5a, lanes 1 and 6, respectively). Using PCR primers to these regions allows for the molecular detection of CML in nearly all cases (~99%); however, variant breakpoints occur. When variant breakpoints fall outside of the region recognized by the primers, a false-negative RT-PCR result occurs, whereas if the variant breakpoints fall within the primer region, an RT-PCR product of unexpected size results. Variant breakpoints are diverse and often complex and cannot be globally assessed with any one pair of PCR primers; however, they are quite rare and are thus not a major cause of false-negative results. The incorporation of an *ABL1* exon 3 (a3) primer in place of an exon 2 primer permits the routine detection of rare *ABL1* intron 3 breakpoints and has been recommended by some for routine evaluation.[28] However, there have been fewer than ten patients with CML reported with this variant breakpoint, which is similarly rare in ALL.

Multiple fusion transcripts may be detectable with a single *BCR* intron 14 breakpoint due to alternative splicing of the primary transcript. With this breakpoint, the resultant chimeric gene is always directly transcribed to produce the e14a2 (b3a2) transcript as described above; however, concomitant alternative splicing may yield an accompanying e13a2 (b2a2) (Figure 35-5a, lane 2) and even an e1a2 (Figure 35-5a and b, lane 6) transcript.[29,30] e1a2 alternative transcripts may also be seen with breakpoints following e13 (b2). Identification of the e1a2 transcript requires a separate *BCR* primer targeting *BCR* exon 1 (e1) (Figure 35-5b, lane 6). Although the e1a2 primer set usually is not essential for the diagnosis of CML, some studies suggest that the presence of this alternatively spliced transcript is associated with a poor prognosis and transformation to accelerated phase.[8]

In the diagnostic setting, qualitative RT-PCR serves several functions. The presence or absence of *BCR-ABL1* transcripts distinguishes CML from leukemoid reactions

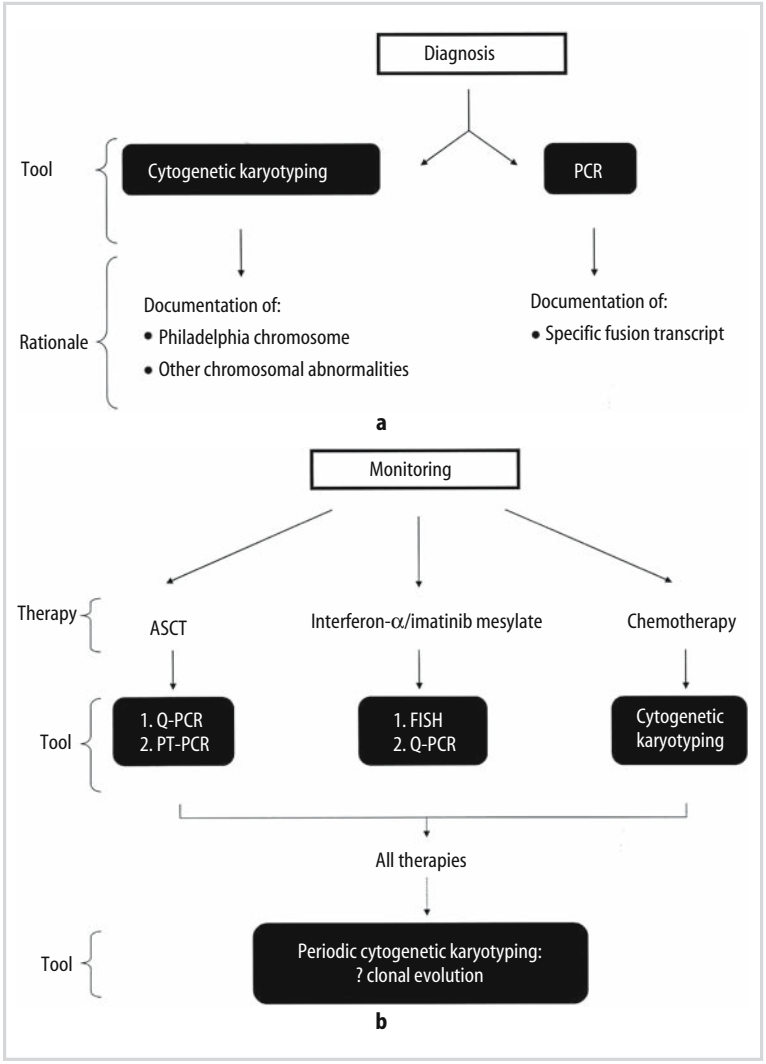


Figure 35-4. Specific laboratory approaches to the (a) diagnosis and (b) therapeutic monitoring of CML. See also Figure 35-7. ASCT, allogeneic stem cell transplantation; FISH, fluorescence in situ hybridization; RT-PCR, qualitative reverse transcription–polymerase chain reaction; Q-PCR, quantitative reverse transcription–polymerase chain reaction. Recent recommendations propose that FISH is no longer an appropriate method for monitoring patients in complete cytogenetic remission.[78]

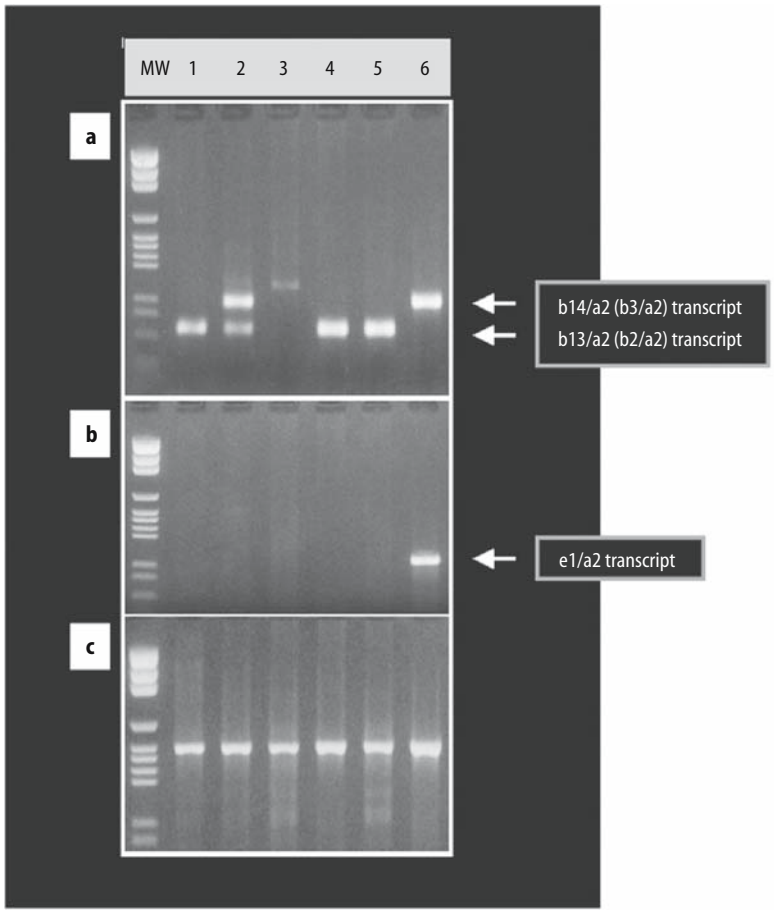


Figure 35-5. Conventional qualitative RT-PCR for detection of the most common *BCR-ABL1* fusion transcripts using gel electrophoresis. (a) *BCR* exon 13 (b2) and *ABL1* exon 2 (a2) primers. (b) *BCR* exon 1 (e1) and *ABL1* exon 2 (a2) primers. (c) β2-microglobulin primers. Lanes: (1) *BCR* intron 13 breakpoint, (2) *BCR* intron 14 breakpoint with alternative splicing to *BCR* exon 13, (3) possibly novel breakpoint or nonspecific band, (4) *BCR* intron 13 breakpoint, (5) *BCR* intron 13 breakpoint, (6) *BCR* intron 14 breakpoint with alternative splicing to *BCR* exon 1. See text for details. (Used with permission from Vergilio JA and Bagg A. Chronic myeloid leukaemia—Molecular diagnosis and monitoring. In: Encyclopedia of Medical Genomics and Proteomics (EMGP). Jürgen Fuchs and Maurizio Podda, (eds.) Marcel Dekker, 2005:252–258.)

or other myeloproliferative disorders. A positive qualitative result is critical for the diagnosis of CML, particularly in those patients lacking the Philadelphia chromosome on routine karyotypic analysis. The presence of *BCR-ABL1* transcripts validates the implementation of targeted therapy with imatinib mesylate, while identification of the type of fusion transcript that is specific to the neoplastic clone is important for subsequent monitoring. Although the majority of patients with CML possess the e13a2 (b2a2) or e14a2 (b3a2) transcript, identification of the e1a2 or e19a2 forms may explain unusual hematologic phenotypes (Figure 35-2c).

Thus, detection of the most common *BCR* breakpoints (e1, e13, and e14) and the common (a2) as well as variant (a3) breakpoints involving *ABL1* can be accomplished with two parallel conventional RT-PCR assays that use *BCR*-exon 1/*ABL1*-exon 3 and *BCR*-exon 13/*ABL1*-exon 3 primers. An independent RT-PCR assay for an unrelated mRNA (e.g., *B2M*, *GAPDH*, *BCR*, or *ABL1*) is also necessary as a quality control to assess the integrity of the RNA and the presence of RT-PCR inhibitors (Figure 35-5c). These three reactions may be multiplexed to allow for the simultaneous detection of the various transcripts, but multiplexing may reduce the sensitivity of the individual reactions. While reduced sensitivity is not relevant for diagnostic testing, it is very relevant for monitoring residual disease.

In addition to its role at diagnosis, qualitative RT-PCR had been the mainstay of minimal residual disease assessment. As noted, however, a single positive (or negative) qualitative result is of little clinical predictive value. Consequently, quantitative RT-PCR has emerged as a preferred modality for posttherapeutic monitoring and has further refined the prognostic significance of standard qualitative results.

Quantitative RT-PCR

Technologically, quantitative RT-PCR has evolved from the initial use of competitive and limiting dilution RT-PCR methods to real-time RT-PCR methods. Competitive RT-PCR is labor-intensive, technically demanding, time-consuming and lacks standardization. Therefore, real-time RT-PCR is the most common method currently used (see chapter 2).

Real-time RT-PCR affords sensitive, rapid, and reproducible quantification of the *BCR-ABL1* fusion transcripts. A closed-tube system is employed, which eliminates the need for post-PCR processing and minimizes the potential for contamination while simultaneously decreasing testing time.

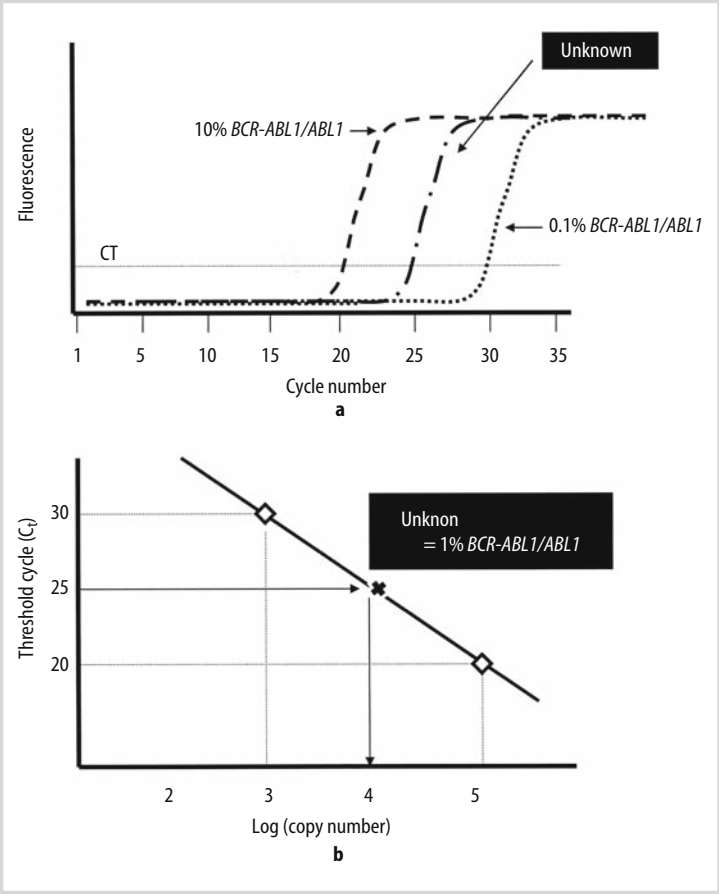


Figure 35-6. Schematic representation of real-time RT-PCR. (a) Amplification plot with known *BCR-ABL1* standards and an unknown patient sample. The increase in fluorescence intensity reflects the increase in PCR products generated during amplification. The initial template concentration dictates the cycle number at which fluorescence increases above the threshold level, C_T. (B) Standard curve from which the unknown *BCR-ABL1* copy number is determined.

Although the sensitivity of real-time RT-PCR methods is somewhat less than conventional or nested RT-PCR, the dynamic range of fluorescent detection is much broader, spanning 5 to 6 orders of magnitude, with a lower limit of detection of <0.01%.[26] Most important, real-time RT-PCR has the precision required in clinical diagnostic applications.

Various fluorescent detection systems have been used to quantify *BCR-ABL1* transcripts. The major chemistries employed are the Taqman[31] or FRET probe[32] methods. For both of these, RNA or complementary DNA (cDNA) standards of known concentration are used to generate a standard curve (log copy number versus threshold cycle [Ct]), from which the unknown sample quantity is determined (Figure 35-6) and then normalized against an internal reference (e.g., *ABL1* transcripts). The final result is reported as a percentage ratio (e.g., *BCR-ABL1*/*ABL1*), although alternative methods of reporting include copy number or micrograms of RNA, or both. Other methods include molecular beacons and melting-curve analysis (see chapter 2).

Bone Marrow Engraftment Analysis

Several other molecular tests have a role in the laboratory evaluation of CML patients. In the setting of ASCT, chimerism analysis by variable number tandem repeats (VNTR), short tandem repeats (STR), or single nucleotide polymorphisms (SNP) is used for assessment of bone marrow engraftment of donor cells and to detect the presence of residual recipient cells, so-called mixed chimerism (see chapter 46). After successful transplantation with full donor engraftment, the presence of mixed chimerism is associated with a higher risk of disease relapse.[33] Increasing levels of the *BCR-ABL1* transcripts may be associated with mixed chimerism such that quantitative real-time RT-PCR and chimerism analysis together may provide information to guide early therapeutic intervention.[34]

Interpretation of Test Results

The interpretation of test results is influenced by the clinical context in which the test is performed. In the diagnostic setting, a negative RT-PCR result renders a diagnosis of CML unlikely. However, it does not definitively exclude a diagnosis of CML since rare variant breakpoints can occur that may not be detected by conventional PCR primer sets. Similarly, a positive qualitative RT-PCR result does not

definitively invoke a diagnosis of CML. Using an extremely sensitive nested RT-PCR technique, *BCR-ABL1* fusion transcripts have been identified in the blood of as many as two thirds of healthy adults.[35,36] Whereas the vast majority of CML patients express either the e13a2 (b2a2) or the e14a2 (b3a2) fusion transcripts, healthy individuals express primarily the e1a2 type. However, the extreme level of sensitivity achieved (~10^{-8}) with this nested technique is neither required nor recommended for routine clinical testing, and thus misdiagnosis should not be an issue.

For therapeutic monitoring, a single positive qualitative RT-PCR result is not predictive of relapse in any individual patient regardless of therapeutic context; nonetheless, therapeutic context ultimately dictates the clinical ramifications of any test result. For instance, in the setting of bone marrow transplantation, the interval after transplant, the type of transplant (i.e., unrelated vs related matched), and the presence or absence of stem cell product manipulation (e.g., T-cell depletion) may impact the prognostic relevance of qualitative test results. Most CML patients have a positive qualitative RT-PCR result in the first 6 months following ASCT that is not associated with a higher risk of relapse. Six to 12 months after ASCT, however, positive qualitative results are highly associated with, and considered to be an independent predictor of, subsequent relapse.[37] In this setting, positive RT-PCR detection of *BCR-ABL1* fusion transcripts precedes cytogenetic and hematologic relapse by several months. Furthermore, detection of fusion transcripts by RT-PCR 6 to 12 months after transplant with an allogeneic related matched donor or with T-cell-depleted stem cell products is associated with a higher rate of relapse than in patients receiving unrelated donor or nondepleted stem cell products.[38]

Despite attainment of complete cytogenetic remission, patients treated with interferon-α can continue to be qualitatively RT-PCR positive for several years. Furthermore, in the small number of patients treated with interferon-α who do become RT-PCR negative, FISH analysis can remain positive for the *BCR-ABL1* gene fusion, supporting the notion that interferon-α may decrease chimeric gene transcription without successful eradication of the neoplastic clone.[39] Consequently, qualitative RT-PCR has traditionally been of little clinical utility in the posttherapeutic monitoring of patients treated with interferon-α, while FISH analysis has provided more meaningful clinical data. A long-term outcome study (with an approximately 10-year follow-up), however, has documented that repeated negative qualitative RT-PCR tests, or even the attainment of at least one documented molecular remission after interferon-α therapy, is associated with improved long-term event-free survival as well as the prolonged durability of a major cytogenetic response.[19] Although this finding offers validity to qualitative RT-PCR monitoring during or after interferon-α therapy, it is relevant to few CML patients.

In contrast to qualitative RT-PCR analysis, quantitative RT-PCR shows seemingly broad relevance to posttherapeutic monitoring whether after transplantation, or after therapy with interferon-α or imatinib mesylate. Low or falling transcript levels correlate with successful treatment or continued remission, and elevated or rising transcript levels predict relapse.[14,16,17] The establishment of specific criteria, critical thresholds to define molecular relapse, or both has, however, been difficult in light of the ongoing evolution and continued sophistication of quantitative RT-PCR technology. The European Investigators on CML (EICML) issued recommendations in 1994 that defined a quantitative RT-PCR relapse as a 10-fold or greater increase in the level of *BCR-ABL1* transcripts, as determined by a minimum of three consecutive quantitative analyses.[40] Although these guidelines offer a framework for monitoring, they were developed over a decade ago and prior to the advent of real-time RT-PCR technology.

Since then, much data have been generated though still compromised by the absence of large prospective clinical trials and the lack of technical standardization; nevertheless, pertinent information can be ascertained. With respect to monitoring after transplantation, several quantitative indicators emerge. Thus, a modified definition of molecular relapse has been proposed as rising or persistently high levels of *BCR-ABL1*, delineated as a *BCR-ABL1/ABL1* ratio of >0.02% (~10^{-4}), in two sequential specimens procured more than 4 months after ASCT. Although imatinib has shown encouraging results in the treatment of CML, the long-term correlates of clinical outcome remain to be defined. Nonetheless, recent studies suggest that quantitative RT-PCR values correlate with established cytogenetic response criteria and are associated with a differential short-term outcome.[41] In the IRIS study, a "major molecular response (MMR)" is defined as a ≥3 log reduction in *BCR-ABL1/BCR* levels compared to median pretreatment levels. Here, MMR was achieved in 39% of newly diagnosed CML patients after 12 months of imatinib as compared to only 2% of patients on interferon-α plus Ara-C.[17] Patients with this degree of molecular response to imatinib had a negligible risk of disease progression in the short term. In addition, the patterns of residual disease levels after attaining MMR (such as continuing decline in levels, plateau, or increasing levels) can predict cytogenetic relapse.[42]

One additional advantage of quantitative BCR/ABL1 monitoring with imatinib therapy is that it allows for the early recognition of therapeutic resistance. In this scenario, patients with a suboptimal therapeutic response may undergo screening for causative point mutations (e.g., using allele-specific oligonucleotide PCR, denaturing HPLC, or heteroduplex formation), the identification of which could prompt implementation of alternative treatment strategies. Preliminary work with microarrays has identified differential gene expression profiles between patients with imatinib sensitivity and resistance.[43] In the future, this technology may predict therapeutic response and thus dictate frontline therapy.

Regardless of the therapeutic strategy, large prospective studies that correlate quantitative RT-PCR data with long-term clinical outcome are anticipated. These investigations

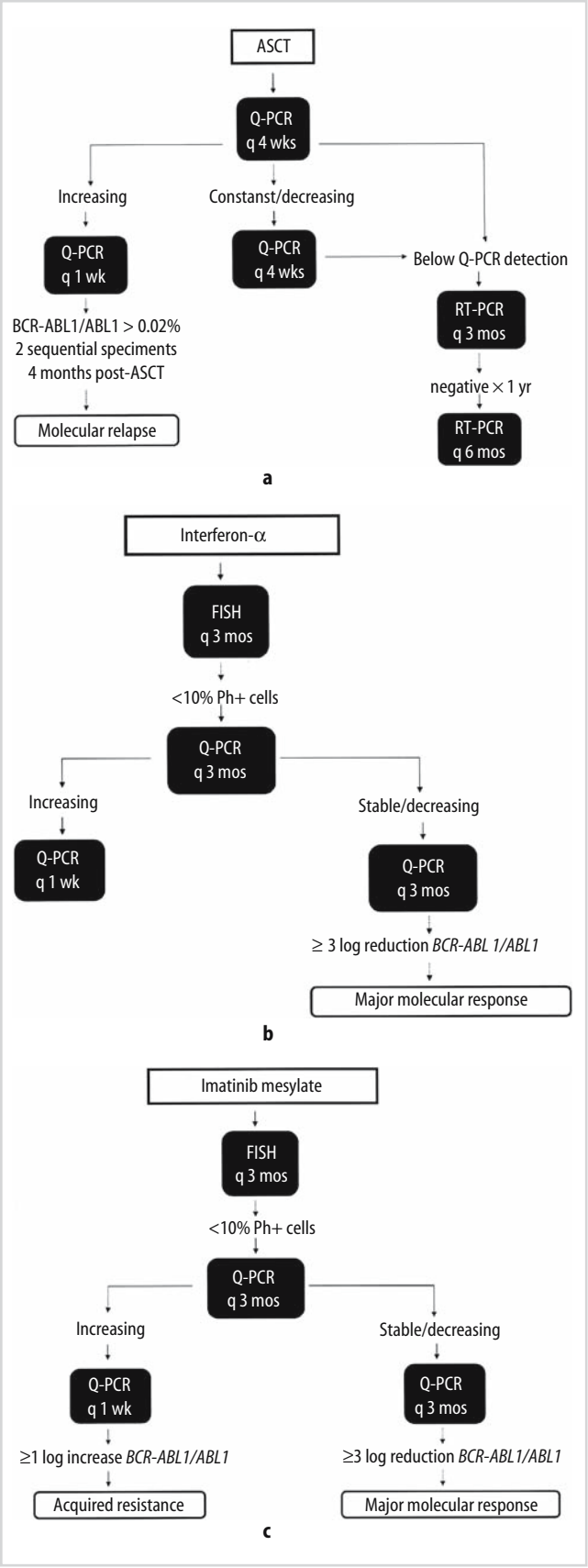


Figure 35-7. Algorithms for the frequency of therapeutic monitoring in CML: (a) after allogenic stem cell transplantation (ASCT), (b) with interferon-α therapy, and (c) with imatinib mesylate therapy. FISH, fluorescence in situ hybridization; mos, months; Ph, Philadelphia chromosome; RT-PCR, qualitative reverse transcription polymerase chain reaction; Q-PCR, quantitative reverse transcription polymerase chain reaction; wk(s), week(s). Recent recommendations propose that FISH is no longer an appropriate method for monitoring patients in complete cytogenetic remission.[78]

will direct the formulation of new guidelines to assist with the future therapeutic monitoring and clinical interpretation of these tests. Until then, a number of algorithms have been proposed for therapeutic monitoring[25,26,44,45] (Figure 35-7).

One final caveat in molecular RT-PCR analysis, whether qualitative or quantitative, is that it cannot assess the clonal evolution of disease. For this assessment, cytogenetic karyotyping is required and must be performed every 6 months throughout the monitoring process regardless of the type of therapy. The emergence of cytogenetically abnormal (but Ph-negative) clones with the use of imatinib, with a reported incidence ranging between 2% and 24%, further underscores the need for periodic conventional cytogenetic analysis.[46]

Laboratory Issues

Several issues arise during the laboratory evaluation of CML. Some have already been discussed in detail above, and so are only briefly noted here. These include the expression of *BCR-ABL1* in normal individuals, false-positive results induced by PCR contamination, particularly using nested PCR techniques, and false-negative results due to variant breakpoints.

A variety of controls (positive, negative, and internal) are used for each test, whether qualitative or quantitative. RNA extracted from cell lines known to be positive and negative for *BCR-ABL1* fusion are commonly used as controls. K562 is a cell line positive for the e14a2 (b3a2) *BCR-ABL1* transcript, KBM-7 is positive for the e13a2 (b3a2) transcript, and SUPB15 is positive for the e1a2 transcript. For clinical testing, a simplex RT-PCR test should be able to detect one K562 cell in a background of 100,000 normal cells (10^{-5}) and one SUPB15 cell in a background of 10,000 normal cells (10^{-4}). Although plasmids with incorporated *BCR-ABL1* sequences also are available, these are rarely used in the clinical laboratory due to the risk of plasmid contamination of specimens or reagents.

In any RT-PCR assay, false-negative results may be due to mRNA degradation, and therefore, an internal control transcript is routinely evaluated to confirm the presence of intact mRNA and absence of RT-PCR inhibitors. In qualitative assays, a variety of transcripts can serve this function, including *ABL1*, *BCR*, *G6PDH*, and *B2M*. In quantitative assays, however, the control transcript also serves as the standard to which the *BCR-ABL1* transcript is

normalized and should have a stable level of expression and approximately equivalent amplification efficiency to *BCR*/*ABL1*. Although *ABL1* is used most widely as a control gene, *BCR* or *GUSB* are equally suitable; others, such as *GAPDH*, are now considered suboptimal.

A silent polymorphism in exon e13/b2 of the M-bcr has been identified in ~30% of individuals with the replacement of thymidine by cytosine at the eighth position of this exon. Although this has no apparent effect on the structure or function of BCR or BCR-ABL1 proteins, it may alter the annealing of PCR primers and probes for real-time RT-PCR.[47]

OTHER MYELOPROLIFERATIVE DISORDERS AND MYELODYSPLASTIC SYNDROMES

The other MPD and MDS (Table 35-1) constitute a broad spectrum of entities previously characterized as mostly lacking specific and distinctive genetic markers analogous to *BCR-ABL1* in CML. Consequently, molecular testing has been much less developed for these diseases, with the exception of chronic eosinophilic leukemia, with the identification of the *FIP1L1-PDGFRA* fusion in many cases.[48] However, based on more recent findings, there is the significant potential that the molecular diagnostic landscape may change for the diagnosis of polycythemia vera (PV) and other MPD.[49,50] Diagnosis has traditionally relied on the integration of morphologic features (from the bone marrow and peripheral blood) with both routine and sophisticated laboratory studies (including complete blood cell counts, cytogenetics, endogenous colony formation, cytokine sensitivity of hematopoietic progenitors, and quantitation of hemoglobin F) as well as clinical manifestations and history.[51–54] However, in early 2005 a number of groups almost simultaneously identified a unifying molecular genetic abnormality in most patients with PV, and up to one half of patients with essential thrombocythemia (ET) and idiopathic myelofibrosis (IMF).[49,50] This abnormality is an acquired mutation (V617F) in *JAK2*, which encodes a key intracellular kinase involved in signaling from multiple hematopoietic growth factor receptors, including the erythropoietin receptor. The mechanism by which one mutation results in three related but distinct MPDs remains to be determined; perhaps additional genetic alterations result in the different hematologic phenotypes. Alternatively, the subsets of ET and IMF with these mutations might reflect PV variants, possibly in early or even spent phase. A variety of technical approaches, including ASO-PCR, have become available to test for this mutation.[55]

Recurrent cytogenetic abnormalities have been described in MDS, such as –5/del(5q), –7/del(7q), trisomy 8, and del(20q); however, these are not currently amenable to routine molecular testing since none of these has been definitively characterized at the molecular level.[56] Some other recurrent, but rare, genetic abnormalities have been characterized at the molecular level in MPD, with two interesting breakpoint clusters at 8p11[57] and 5q33,[58] which generate constitutively active tyrosine kinase fusions (Table 35-3).

Since MPD and MDS are clonal diseases arising from a hematopoietic stem cell, the determination of monoclonality can aid in the establishment of neoplasia (in an appropriate clinical and morphological context). One molecular approach to clonality assessment exploits the physiologic process of X-chromosome inactivation in which one X-chromosome is inactivated in the somatic tissues of all females to compensate for unequal gene dosage compared to males.[73] This inactivation process occurs by methylation of cytosine nucleotides in cytosine-guanine dinucleotide (CpG)-rich regions of DNA. In any cell population, this process yields random inactivation of maternal and paternal alleles, such that the ratio of methylated maternal to methylated paternal alleles approximates 1:1. In a clonal population, one allele predominates and alters the ratio, providing a useful marker for clonality assessment. The human androgen receptor (*HUMARA*) assay is one such DNA-based PCR clonality assay. The human androgen receptor gene contains a trinucleotide repeat region that is polymorphic in more than 95% of individuals and is preceded by a CpG-rich region. Both methylation-specific restriction enzymes and methylation-specific PCR analysis using chemical modification with sodium bisulfite have been used[74,75] (Figure 35-8). Nonrandom X inactivation is conventionally defined by a ratio that exceeds 3:1.

Clonality assessment with *HUMARA* is not without limitations. First, only females are eligible for analysis, limiting its applicability to 50% of the general population. Second, elderly individuals may have nonrandom age-associated X-linked inactivation (skewing) that, in some circumstances, can be distinguished by comparison with a hematopoietic control (e.g., T cells). The phenomenon of skewing may be particularly confounding to the evaluation of MPD and MDS, since these diseases are most commonly encountered in older individuals. Here, a more stringent ratio, exceeding 10:1, may be more appropriate. Despite these limitations, *HUMARA* and other X-inactivation studies (*G6PD* [glucose-6-phosphate dehydrogenase], *PGK* [phosphoglycerate kinase], *HPRT* [hypoxanthine phosphoryl ribosyl transferase]) are useful for documentation of clonality in these and other hematologic disorders.[76]

Prior to the discovery of *JAK2* mutations, quantitation of *PRV1* mRNA was reported to have diagnostic relevance in PV and ET. *PRV1* (polycythemia rubra vera 1) is a gene that encodes a hematopoietic cell surface receptor homologous to the neutrophil alloantigen NB1/CD117. Overexpression of the *PRV1* mRNA has been detected in patients with PV as well as in a subset of patients with ET but not in healthy controls or in those with reactive erythrocytosis or thrombocytosis.[76] Curiously, PRV1 protein levels do not show this discordant expression. Nevertheless, quantitative RT-PCR for *PRV1* may be useful in differentiating PV and ET from their reactive counterparts.

Table 35-3. Recurrent Molecular Genetic Abnormalities Associated with Various Non-CML Myeloproliferative Disorders and Myelodysplastic Syndromes

Target	Examples	MPD/MDS Association
Myeloproliferative Disorders		
JAK2	Point mutation [V617F]	Majority (65%–95%) polycythemia vera, and up to 50% essential thrombocythemia and idiopathic myelofibrosis
Translocations involving 8p11 [*FGFR1*]:	t(6;8)(q27;p11) [*FOP-FGFR1*]	
	t(8;9)(p11;q33) [*CEP110-FGFR1*]	
	t(8;13)(p11;q12) [*ZNF198-FGFR1*]	
	t(8;22)(p11;q22) [*BCR-FGFR1*]	8p11 myeloproliferative syndrome (EMS) associated with T-cell lymphoma[56]
Translocations involving 5q33 [*PDGFRB*]:	t(5;7)(q33;p11) [*HIP1-PDGFRB*]	
	t(5;10)(q33;q21) [*H4-PDGFRB*]	
	t(5;12)(q33;p13) [*TEL-PDGFRB*]	Myeloproliferative/myelodysplastic disorders, in particular CMML with eosinophilia[57]
	t(5;14)(q33;p32) [*CEV14-PDGFRB*]	
	t(5;14)(q33;p24) [*NIN-PDGFRB*]	
	t(5;17)(q33;p13) [*RAB5-PDGFRB*]	
	t(5;17)(q33;p11) [*HCMOGT-PDGFRB*]	
	t(1;5)(q23;q33) [*PDE4DIP-PDGFRB*]	
4q12 [*FIP1L1-PDGFRA*]	del(4q12)	Chronic eosinophilic leukemia/hypereosinophilic syndrome[48] ? Systemic mastocytosis
GATA1	Point mutation	Transient myeloproliferative disorder of Down syndrome[59]†
KIT	Point mutation [D816V]	Mastocytosis[60]†
Myelodysplastic Syndromes		
RAS	Point mutation	CMML, other MDS[61]
NF1, PTPN11	Point mutation	JMML[62,63]
TP53	Point mutation	RAEB, 17p syndrome[64–66]
P15	Methylation	Progression/transformation[67,68]
FMS	Point mutation	Progression/transformation[69]
CBFA2	Mutation, deletion	Therapy-related MDS[70]
Mitochondrial DNA*	Mutation, deletion	RARS[71,72]

* The association of mitochondrial mutations with MDS is controversial.
† Not formally designated a MPD by WHO.
CMML, chronic myelomonocytic leukemia; JMML, juvenile myelomonocytic leukemia; MDS, myelodysplastic syndrome; MPD, myeloproliferative disorder; RAEB, refractory anemia with excess blasts; RARS, refractory anemia with ringed sideroblasts.

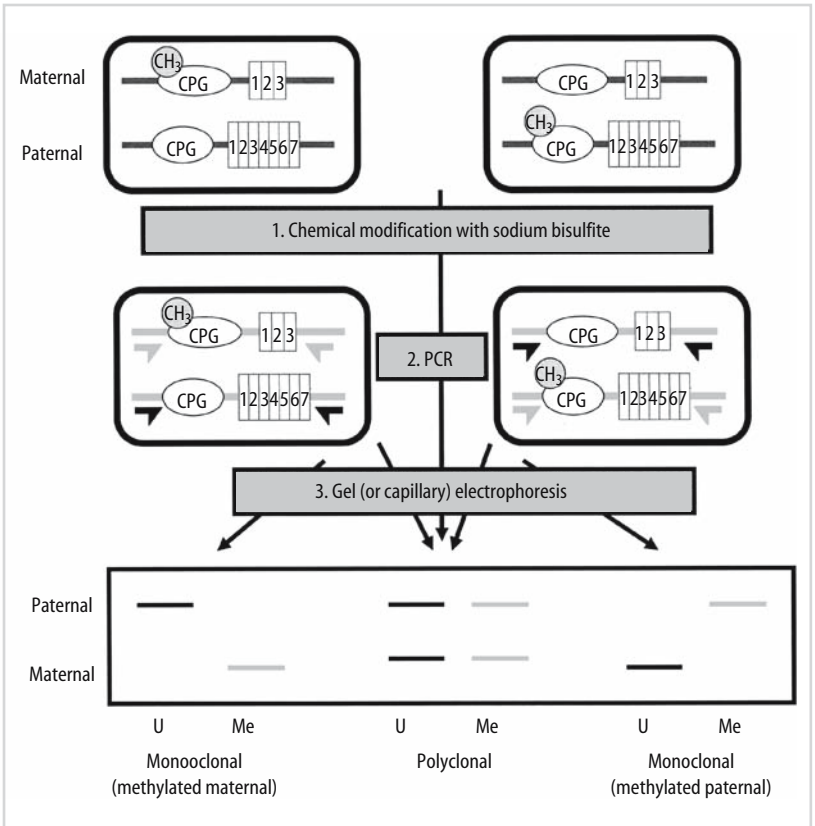


Figure 35-8. Schematic representation of the methylation-specific polymerase chain reaction analysis of the human androgen receptor assay (MSP-HUMARA), used to determine clonality status for female patients. Sodium bisulfite modifies DNA by the permanent conversion of unmethylated (i.e., active) cytosine to uracil. Methylated (i.e., inactive) cytosine residues, located 5′ to guanine residues, are resistant to this modification. PCR amplification with primers specific for either methylated (unmodified) or unmethylated (modified) DNA of the *HUMAR* gene sequence is performed with subsequent visualization of the PCR products by gel or capillary electrophoresis. A methylated maternal to methylated paternal allelic ratio that exceeds either 3:1 or 1:3 is indicative of nonrandom X inactivation. (Used with permission from Vergilio JA and Bagg A. Chronic myeloid leukaemia—Molecular diagnosis and monitoring. In: encyclopedia of Medical Genomics and Proteomics (EMGP). Jürgen Fuchs and Maurizio Podda, (eds.) Marcel Dekker 2005:252–258.)

Although our molecular understanding of non-CML MPD and MDS is far from that of CML, the recent identification of *JAK2* mutations notwithstanding, molecular testing can contribute to laboratory evaluation. With more sophisticated technology (e.g., gene expression profiling) and continued investigation, it is hoped that the molecular pathogenesis of MPD and MDS will be delineated, which will result in improved methods for diagnosis and monitoring of these disorders. Ultimately, such molecular discoveries may result in an evolution from a clinicopathologic-based to a molecular-based classification of these disorders.

References

1. Jaffe ES, Harris NL, Stein H, et al. *World Health Organization Classification of Tumours, Pathology and Genetics, Tumours of Haematopoetic and Lymphoid Tissues.* Lyon, France: IARC Press; 2001.
2. Nowell PC. Progress with chronic myelogenous leukemia: a personal perspective over four decades. *Annu Rev Med.* 2002;53:1–13.
3. Barnes DJ, Melo JV. Cytogenetic and molecular genetic aspects of chronic myeloid leukemia. *Acta Haematol.* 2002;108:180–202.
4. Melo JV. The diversity of BCR-ABL fusion proteins and their relationship to leukemia phenotype. *Blood.* 1996;93:3167–3215.
5. Faderl S, Jeha S, Kantarjian HM. The biology and therapy of adult acute lymphoblastic leukemia. *Cancer.* 2003;98:1337–1354.
6. Pane F, Frigeri F, Sindona M, et al. Neutrophilic-chronic myeloid leukemia: a distinct disease with a specific molecular marker (BCR/ABL with C3/A2 junction). *Blood.* 1996;88:2410–2414.
7. Elliott MA, Hanson CA, Dewald GW, Smoley SA, Lasho TL, Tefferi A. WHO-defined chronic neutrophilic leukemia: a long-term analysis of 12 cases and a critical review of the literature. *Leukemia.* 2005;19:313–317.
8. Shepter PJ, Ben-Yehuda D. Molecular evolution of chronic myeloid leukaemia. *Semin Cancer Biol.* 2001;11:313–322.
9. Calabretta B, Perrotti D. The biology of CML blast crisis. *Blood.* 2004;103:4010–4022.
10. Johansson B, Fioretos T, Mitelman F. Cytogenetic and molecular genetic evolution of chronic myeloid leukemia. *Acta Haematol.* 2002;107:76–94.
11. Huntly BJ, Bench A, Green AR. Double jeopardy from a single translocation: deletions of the derivative chromosome 9 in chronic myeloid leukemia. *Blood.* 2003;102:1160–1168.
12. Kaeda J, Chase A, Goldman JM. Cytogenetic and molecular monitoring of residual disease in chronic myeloid leukemia. *Acta Haematol.* 2002;107:64–75.
13. O'Dwyer ME, Mauro MJ, Druker BJ. Recent advancements in the treatment of chronic myelogenous leukemia. *Annu Rev Med.* 2002;53:369–381.
14. Kantarjian HM, Talpaz M, Cortes J, et al. Quantitative polymerase chain reaction monitoring of BCR-ABL during therapy with imatinib mesylate (STI571; Gleevec) in chronic-phase chronic myelogenous leukemia. *Clin Cancer Res.* 2003;9:160–166.
15. O'Brien SG, Guilhot F, Larson RA, et al. Imatinib compared with interferon and low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukemia. *N Eng J Med.* 2003;348:994–1004.
16. Branford S, Rudzki Z, Harper A, et al. Imatinib produces significantly superior molecular responses compared to interferon alfa plus cytarabine in patients with newly diagnosed chronic myeloid leukemia in chronic phase. *Leukemia.* 2003;17:2401–2409.
17. Hughes TP, Kaeda J, Branford S, et al. Frequency of major molecular responses to imatinib or interferon alfa plus cytarabine in newly diagnosed chronic myeloid leukemia. *N Eng J Med.* 2003;349:1423–1432.
18. Kantarjian H, Talpaz M, O'Brien S, et al. High-dose imatinib mesylate therapy in newly diagnosed Philadelphia chromosome-positive chronic phase chronic myeloid leukemia. *Blood.* 2004;103:2873–2878.
19. Kantarjian HM, O'Brien S, Cortes JE, et al. Complete cytogenetic and molecular responses to interferon-α-based therapy for chronic myelogenous leukemia are associated with excellent long-term prognosis. *Cancer.* 2003;97:1033–1041.
20. Deininger M, Buchdunger E, Druker BJ. The development of imatinib as a therapeutic agent for chronic myeloid leukemia. *Blood.* 2005;105:2640–2653.
21. Gilleece MH, Dazzi F. Donor lymphocyte infusions for patients who relapse after allogeneic stem cell transplantation for chronic myeloid leukaemia. *Leuk Lymphoma.* 2003;44:23–28.
22. Tsao AS, Kantarjian H, Talpaz DM. STI-571 in chronic myelogenous leukemia. *Br J Haematol.* 2002;119:15–24.
23. Shannon KM. Resistance in the land of molecular cancer therapeutics. *Cancer Cell.* 2002;2:99–102.
24. Branford S, Rudzki Z, Parkinson I, et al. Real-time quantitative PCR analysis can be used as a primary screen to identify patients with CML treated with imatinib who have BCR-ABL kinase domain mutations. *Blood.* 2004;104:2926–2932.
25. Wang YL, Bagg A, Pear W, et al. Chronic myelogenous leukemia: laboratory diagnosis and monitoring. *Genes Chromosomes Cancer.* 2001;32:97–111.
26. Bagg A. Chronic myeloid leukemia: a minimalistic view of post-therapeutic monitoring. *J Mol Diag.* 2002;4:1–10.
27. Branford S, Hughes TP, Rudzki Z. Monitoring chronic myeloid leukaemia therapy by real-time quantitative PCR in blood is a reliable alternative to bone marrow cytogenetics. *Br J Haematol.* 1999;107:587–599.
28. Hochhaus A. Detection and quantification of leukemia-specific rearrangements. In: Boultwood J, Fidler C, eds. *Methods in Molecular Medicine.* Totowa, NJ: Humana Press; 2002:67–96.
29. Branford S, Hughes TP, Rudzki Z. Dual transcription of b2a2 and b3a2 BCR-ABL transcripts in chronic myeloid leukaemia is confined to patients with a linked polymorphism within the BCR gene. *Br J Haematol.* 2002;117:875–877.
30. Lichty BD, Keating A, Callum J, et al. Expression of p210 and p190 BCR-ABL due to alternative splicing in chronic myelogenous leukaemia. *Br J Haematol.* 1998;103:711–715.
31. Kreuzer KA, Lass U, Bohn A, et al. LightCycler technology for the quantitation of bcr/abl fusion transcripts. *Cancer Res.* 1999;59:3171–3174.
32. Emig M, Saussele S, Wittor H, et al. Accurate and rapid analysis of residual disease in patients with CML using specific fluorescent hybridization probes for real time quantitative RT-PCR. *Leukemia.* 1999;13:1825–1832.
33. Serrano J, Roman J, Sanchez J, et al. Molecular analysis of lineage-specific chimerism and minimal residual disease by RT-PCR of p210 (BCR-ABL) and p190 (BCR-ABL) after allogeneic bone marrow transplantation for chronic myeloid leukemia: increasing mixed myeloid chimerism and p190 (BCR-ABL) detection precede cytogenetic relapse. *Blood.* 2000;95:2659–2665.
34. Formankova R, Honzatkova L, Moravcova J, et al. Prediction and reversion of post-transplant relapse in patients with chronic myeloid leukemia using mixed chimerism and residual disease detection and adoptive immunotherapy. *Leuk Res.* 2000;24:339–347.
35. Biernaux C, Loos M, Sels A, et al. Detection of major bcr-abl gene expression at a very low level in cells of some healthy individuals. *Blood.* 1995;86:3118–3122.
36. Bose S, Deininger M, Gora-Tybor J, et al. The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals: biologic significance and implications for the assessment of minimal residual disease. *Blood.* 1998;92:3362–3367.
37. Radich JP, Gehley G, Gooley T, et al. Polymerase chain reaction detection of the BCR-ABL fusion transcript after allogeneic marrow

transplantation for chronic myeloid leukemia: results and implications in 346 patients. *Blood.* 1995;85:2632–2638.
38. Radich JP, Gooley T, Bryant E, et al. The significance of bcr-abl molecular detection in chronic myeloid leukemia patients "late," 18 months or more after transplantation. *Blood.* 2001;98:1701–1707.
39. Chomel JC, Brizard F, Veinstein A, et al. Persistence of BCR-ABL genomic rearrangement in chronic myeloid leukemia patients in complete and sustained cytogenetic remission after interferon-alpha therapy or allogeneic bone marrow transplantation. *Blood.* 2000;95:404–408.
40. Lion T. Clinical implications of qualitative and quantitative polymerase chain reaction analysis in the monitoring of patients with chronic myelogenous leukemia: The European Investigators on Chronic Myelogenous Leukemia Group. *Bone Marrow Transplant.* 1994;14:505–509.
41. Wang L, Pearson K, Ferguson JE, et al. The early molecular response to imatinib predicts cytogenetic and clinical outcome in chronic myeloid leukaemia. *Br J Haematol.* 2003;120:990–999.
42. Marin D, Kaeda J, Szydlo R, et al. Monitoring patients in complete cytogenetic remission after treatment of CML in chronic phase with imatinib: patterns of residual leukaemia and prognostic factors for cytogenetic relapse. *Leukemia.* 2005;19:507–512.
43. Tipping AJ, Deininger MW, Goldman JM, et al. Comparative gene expression profile of chronic myeloid leukemia cells innately resistant to imatinib mesylate. *Exp Hematol.* 2003;31:1073–1080.
44. Vergilio J, Bagg A. Chronic myeloid leukemia: Molecular diagnosis and monitoring. In: Fuchs J, Podda M, eds. *Encyclopedia of Medical Genomics and Proteomics.* New York: Marcel Dekker; 2005.
45. Goldman J. Monitoring minimal residual disease in BCR-ABL-positive chronic myeloid leukemia in the imatinib era. *Curr Opin Hematol.* 2005;12:33–39.
46. O'Dwyer ME, Gatter KM, Loriaux M, et al. Demonstration of Philadelphia chromosome negative abnormal clones in patients with chronic myelogenous leukemia during major cytogenetic responses induced by imatinib mesylate. *Leukemia.* 2003;17:481–487.
47. Saussele S, Weisser A, Muller MC, et al. Frequent polymorphism in BCR exon b2 identified in BCR-ABL positive and negative individuals using fluorescent hybridization probes. *Leukemia.* 2000;14:2006–2010.
48. Cools J, Stover EH, Wlodarska I, Marynen P, Gilliland DG. The FIP1L1-PDGFRalpha kinase in hypereosinophilic syndrome and chronic eosinophilic leukemia. *Curr Opin Hematol.* 2004;11:51–57.
49. Goldman JM. A unifying mutation in chronic myeloproliferative disorders. *N Engl J Med.* 2005;352:1744–1746.
50. Tefferi A, Gilliland DG. The JAK2V617F tyrosine kinase mutation in myeloproliferative disorders: status report and immediate implications for disease classification and diagnosis. *Mayo Clin Proc.* 2005;80:947–958.
51. Spivak JL. Polycythemia vera: myths, mechanisms, and management. *Blood.* 2002;100:4272–4290.
52. Harrison CN. Current trends in essential thrombocythemia. *Br J Haematol.* 2002;117:796–808.
53. Steensma DP, Tefferi A. Cytogenetic and molecular genetic aspects of essential thrombocythemia. *Acta Haematol.* 2002;108:55–65.
54. Vergilio J, Bagg A. Myelodysplastic syndromes: contemporary biologic concepts and emerging diagnostic approaches. *Am J Clin Pathol.* 2003;199(suppl):S58–S77.
55. Greiner TC. Diagnostic assays for the JAK2 V617F mutation in chronic myeloproliferative disorders. *Am J Clin Pathol.* 2006;125:651–653.
56. Cherian S, Bagg A. The genetics of the myelodysplastic syndromes: Classical cytogenetics and recent molecular insights. *Hematology.* 2006;11:1–13.
57. Macdonald D, Reiter A, Cross NC. The 8p11 myeloproliferative syndrome: a distinct clinical entity caused by constitutive activation of FGFR1. *Acta Haematol.* 2002;107:101–107.
58. Steer EJ, Cross NC. Myeloproliferative disorders with translocations of chromosome 5q31–35: role of the platelet-derived growth factor receptor Beta. *Acta Haematol.* 2002;107:113–122.
59. Mundschau G, Gurbuxani S, Gamis AS, et al. Mutagenesis of GATA1 is an initiating event in Down syndrome leukemogenesis. *Blood.* 2003;101:4298–4300.
60. Gupta R, Bain BJ, Knight CL. Cytogenetic and molecular genetic abnormalities in systemic mastocytosis. *Acta Haematol.* 2002;107:123–128.
61. Padua RA, West RR. Oncogene mutation and prognosis in the myelodysplastic syndromes. *Br J Haematol.* 2000;111:873–874.
62. Emanuel PD. Juvenile myelomonocytic leukemia. *Curr Hematol Rep.* 2004;3:203–209.
63. Side L, Taylor B, Cayouette M, et al. Homozygous inactivation of the NF1 gene in bone marrow cells from children with neurofibromatosis type 1 and malignant myeloid disorders. *N Eng J Med.* 1997;336:1713–1720.
64. Kita-Sasai Y, Horiike S, Misawa S, et al. International prognostic scoring system and TP53 mutations are independent prognostic indicators for patients with myelodysplastic syndromes. *Br J Haematol.* 2001;115:309–312.
65. Lai JL, Preudhomme C, Roumier C, et al. Myelodysplastic syndromes and acute myeloid leukemia with 17p deletion. An entity characterized by specific dysgranulopoiesis and a high incidence of p53 mutations. *Leukemia.* 1995;9:370–381.
66. Soenen V, Preudhomme C, Roumier C, et al. 17p Deletion in acute myeloid leukemia and myelodysplastic syndromes. Analysis of breakpoints and deleted segments by fluorescence in situ. *Blood.* 1998;91:1008–1015.
67. Quesnel B, Guillerm G, Vereecque R, et al. Methylation of the p15(INK4b) gene in myelodysplastic syndromes is frequent and acquired during disease progression. *Blood.* 1998;91:2985–2990.
68. Tien H-F, Tang J-L, Tsay W, et al. Methylation of the $p15^{INK4B}$ gene in myelodysplastic syndrome: it can be detected early at diagnosis or during disease progression and is highly associated with leukaemic transformation. *Br J Haematol.* 2001;112:148–154.
69. Padua RA, Guinn BA, Al-Sabah AI, et al. RAS, FMS and p53 mutations and poor clinical outcome in myelodysplasias: a 10-year follow-up. *Leukemia.* 1998;12:887–892.
70. Slovak ML, Bedell V, Popplewell L, et al. 21q22 balanced chromosome aberrations in therapy-related hematopoietic disorders: a report from an international workshop. *Genes Chromosomes Cancer.* 2002;33:379–394.
71. Reddy PL, Sheet VT, Dutt D, et al. Increased incidence of mitochondrial cytochrome c-oxidase gene mutations in patients with myelodysplastic syndromes. *Br J Haematol.* 2002;116:564–575.
72. Gattermann N. From sideroblastic anemia to the role of mitochondrial DNA mutations in myelodysplastic syndromes. *Leuk Res.* 2000;24:141–151.
73. Gale RE. Evaluation of clonality in myeloid stem-cell disorders. *Semin Hematol.* 1999;36:361–372.
74. Allen RC, Zoghbi HY, Moseley AB, et al. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet.* 1992;51:1229–1239.
75. Uchida T, Ohashi H, Aoki E, et al. Clonality analysis by methylation-specific PCR for the human androgen-receptor gene (HUMARA-MSP). *Leukemia.* 2000;14:207–212.
76. Briere J, el-Kassar N. Clonality markers in polycythaemia and primary thrombocythaemia. *Baillieres Clin Haematol.* 1998;11:787–801.
77. Klippel S, Strunck E, Busse CE, et al. Biochemical characterization of PRV-1, a novel hematopoietic cell surface receptor, which is overexpressed in polycythemia rubra vera. *Blood.* 2002;100:2441–2448.
78. Hughes T, Deininger M, Hochhaus A, et al. Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting BCR-ABL transcripts and kinase domain mutations and for expressing results. *Blood.* 2006;108:28–37.

Section V

Infectious Diseases

Section Editor: Angela M. Caliendo

Chapter 36

Human Immunodeficiency Virus Type 1

Angela M. Caliendo

Human immunodeficiency virus types 1 and 2 (HIV-1 and HIV-2), the causative agents of the acquired immunodeficiency syndrome (AIDS), are RNA viruses belonging to the genus *Lentivirus* of the family Retroviridae. Like all retroviruses, replication involves reverse transcription of the RNA genome into a double-stranded DNA molecule, with subsequent integration into the host genome. This integrated retroviral DNA is referred to as the provirus. Due to this complex replicative cycle, molecular assays used in the diagnosis and management of HIV-1 infection may target either HIV-1 RNA or proviral DNA.

HIV-1 is a very genetically diverse virus; two distinct genetic groups have been identified, the major (M) and outlier (O) groups. Viruses in the M group are further divided into eight subtypes or clades, designated A through H, based on the sequence diversity within the HIV-1 *gag* and *env* genes. The nucleotide sequence of the *env* gene may differ by as much as 25% among the different clades, while viruses within the same clade generally differ by less than 15%.[1] Group M virus is found worldwide, with clade B predominating in Europe and North America. The genetic diversity of HIV-1 plays an important role in the design and interpretation of viral load and resistance tests.

The uses of molecular tests in the clinical management for persons with HIV-1 infection are many and diverse and include diagnosis of infection, monitoring response to therapy, and assessing the efficacy of a treatment regimen. Here we review the available assays and their clinical utility.

Clinical Utility

HIV-1 Viral Load Testing

The clinical utility of HIV-1 viral load testing, which refers to the quantification of HIV-1 RNA, has been well documented. HIV-1 viral load assays are very useful in determining when to initiate antiretroviral therapy, in monitoring response to therapy, and in predicting time to progression to AIDS. With the availability of commercial assays and their proven clinical utility, viral load testing quickly became the standard of care in the clinical management of persons with HIV-1 infection. Clinical guidelines have been established for the initiation of antiretroviral therapy, and viral load levels are among the parameters used in this decision. Initiation of therapy may be considered in patients with a high plasma viral load even if the patient is without symptoms and the CD4 cell count is above 350 cells/mm^3.[2]

HIV-1 viral load is a strong predictor of the rate of progression to AIDS that is independent of CD4 cell count.[3] A study of homosexual men in which the time of seroconversion was documented showed that an initial viral load of 100,000 equivalents/ml of plasma soon after seroconversion was associated with a greater than ten-fold increase in the risk of developing AIDS. However, persons who had a viral load of <1,000 equivalents/ml did not progress to AIDS during the next 5 years.[3] A study of hemophiliacs found similar results; the level of HIV-1 RNA early during chronic infection was a strong predictor of clinical outcome.[4] The same association between viral load level and progression to AIDS is seen in patients for whom the duration of HIV-1 infection is unknown. Mellors et al. showed that HIV-1 seropositive men with a baseline RNA level of <4,530 molecules/ml had a median time to progression to AIDS of >10 years, compared to 3.5 years with a baseline viral load of >36,270 molecules/ml.[5]

Monitoring response to therapy is another key use of viral load testing. The magnitude of the decrease in viral load in response to therapy is dependent on the specific combinations of antiretroviral drugs used. For example, when zidovudine (AZT) or lamivudine (3TC) are used as monotherapy, the decrease in viral load is in the range of 0.5 to 1.0 $\log_{10}$ copies/ml. But when these two drugs are used in combination, the drop in viral load is 1.5 $\log_{10}$ copies/ml and this decrease in viral load is sustained for a longer period of time.[6] When AZT, 3TC, and the protease inhibitor indinavir are used in combination, the decrease in viral load is 2.0 to 2.8 $\log_{10}$ copies/ml and this regimen

slowed the progression of HIV-1 disease in patients with ≤200 CD4 cells/mm^3.[7] Decreases in viral load of approximately 2 $\log_{10}$ have been reported with other combination regimens.[8–10] Based on these findings, the current standard for treating HIV-1 infected individuals is to use combinations of protease inhibitors or nonnucleoside reverse transcriptase inhibitors with nucleoside analogs.[2] This combination therapy is often referred to as highly active antiretroviral therapy or HAART.

The ultimate goal of therapy is to achieve a viral load below the limit of detection of the assay, as studies have shown suppression of plasma viral load to below 20 copies/ml was associated with a longer response to antiretroviral therapy compared with that achieved when the viral load was suppressed to below 500 copies/ml.[11] However, this may not be possible in all cases, particularly in those patients with very high pretreatment viral load levels or in those who have failed prior therapeutic regimens. Guidelines for the use of HIV-1 RNA levels in clinical practice have been published[12] (AIDSinfo [http://aidsinfo.nih.gov/], International AIDS Society–USA [http://www.iasusa.org]). In general, a plasma HIV-1 RNA level should be measured at baseline, immediately before beginning therapy, and then again at 2 to 8 weeks after the start of therapy to determine the initial response. Testing then is repeated at 3- to 4-month intervals to evaluate continued effectiveness of the regimen. Any increase in viral load should be confirmed with repeat testing, as a variety of clinical illnesses can lead to a transient rise in viral load. Once a significant increase in viral load has been documented, HIV-1 resistance testing should be considered.

HIV-1 viral load testing also has proven useful in the diagnosis of acute HIV-1 infection, although the Food and Drug Administration (FDA) has not approved the assays for this indication. Acute HIV-1 infection, also referred to as acute retroviral syndrome, is defined as the "window period" after exposure to the virus prior to seroconversion, when the ELISA and western blot tests are negative or indeterminate. In this "window period," patients often are symptomatic with a mononucleosis-type syndrome, which may include fever, fatigue, rash, lymphadenopathy, and oral ulcers.[13] During acute HIV-1 infection, the level of RNA is very high, usually 10^5 to 10^7 copies/ml of plasma, making viral load measurement a very useful diagnostic tool. A recent study compared the utility of p24 antigen and viral load testing for identifying patients with acute HIV-1 infection.[14] Viral load testing was found to be more sensitive, while p24 antigen testing was more specific. However, the majority of viral load testing in the study was done using the Versant HIV-1 RNA assay (Bayer Diagnostics Corporation, Tarrytown, NY), a signal amplification method, which is known to have a lower specificity compared to target amplification methods. There were no false positives when the Amplicor reverse transcription–polymerase chain reaction (RT-PCR) assay (Roche Diagnostics, Indianapolis, IN) was used, although the sample size was small. When molecular assays for the diagnosis of HIV-1 infection are used, the possibility of a false-positive result must be considered. Patients must be educated about the limitations of these tests and must give informed consent prior to testing. To minimize the likelihood of reporting a false-positive result, an HIV1/2 ELISA should be obtained at the time of viral load testing, and repeat viral load testing should be done on all positive specimens. It is critical to remember that patients with acute retroviral syndrome should have very high levels of HIV-1 RNA. Requiring laboratory approval for HIV-1 diagnostic testing is a prudent approach to ensure that the patient has signs and symptoms consistent with acute HIV-1 infection prior to testing.

Molecular tests are useful in the diagnosis of acute HIV-1 infection in neonates. Since maternal IgG crosses the placenta, an uninfected infant may be seropositive into the second year of life. Studies have established the utility of both qualitative proviral DNA and viral load assays for the diagnosis of HIV-1 infection in newborns.[15–17] HIV-1 viral load testing appears to be more sensitive than proviral DNA testing for the diagnosis of acute HIV-1 infection in newborns.[16,17] Current guidelines for the diagnosis of neonatal HIV-1 infection recommend that the molecular testing be performed at two separate time points, and that testing cord blood samples be avoided.[18]

HIV-1 Drug Resistance Testing

The replication cycle of HIV-1 is error prone because the reverse transcriptase (RT) does not have proofreading activity. As a result, about one error occurs with each replicative cycle. This error rate coupled with a replication rate that produces about one billion viral particles per day gives rise to viral quasispecies. Therefore, a patient has a high probability of having a virus with a drug-resistance mutation in the quasispecies prior to drug exposure. Treatment with an antiretroviral drug leads to selection pressure that allows a resistant virus to predominate. As a result, monotherapy has been found to be ineffective and the standard of care is to treat patients with a combination of antiretroviral drugs. Currently, there are four general classes of antiretroviral drugs that are used in clinical care, nucleoside reverse transcriptase inhibitors (NRTIs), nonnucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), and a new class of drug, fusion inhibitors. Viral resistance can occur with each of these classes of drug, particularly when viral replication is not maximally suppressed while the patient is on therapy. The current standard of care is to use regimens that contain a combination of drugs, usually a PI or NNRTI with several NRTIs, since resistance is less likely to occur on these complex regimens than on monotherapy.

The clinical utility of resistance testing in the management of HIV-1-infected persons has been studied in several prospective randomized clinical trials. Two early studies focused on the clinical utility of genotypic resistance testing, while more recent studies also have assessed the

role of phenotypic resistance testing. The early clinical trials evaluating resistance testing evaluated the use of genotyping in establishing the treatment regimen in patients who had failed therapy with PIs and NRTIs.[19–21] In the control arm, the salvage regimen was determined based on prior antiretroviral drug use, while patients in the genotype arm received therapy based on the results of a genotyping test. In both studies, patients in the genotype arm showed an improved response to therapy compared to the control group, as measured by either the decrease in viral load or the percentage of patients with an undetectable viral load. To clarify the role of expert advice, the Havana trial evaluated the clinical utility of genotyping and expert advice.[22] As with earlier studies, genotyping was found to improve virologic outcomes. In addition, the use of expert advice showed a virologic benefit in patients with two previous treatment regimen failures. A recent study showed that use of genotypic resistance testing to determine the next regimen following treatment failure is cost-effective.[23] In addition, the use of resistance testing to guide the initial treatment regimen appears to be cost-effective if the prevalence of primary genotypic resistance in the population is at least 4%.[23]

The VIRA3001 study was a prospective randomized trial that compared standard of care to phenotypic drug resistance testing in patients who failed a PI-containing regimen. At week 16, an intention-to-treat analysis showed that patients in the phenotype arm had a greater drop in viral load, and a greater proportion of the patients in the phenotype arm had an undetectable viral load compared to the standard-of-care arm.[24] However, not all clinical trials of resistance testing have shown an improved clinical outcome compared to standard of care. The NARVAL trial compared phenotypic testing or genotypic testing to standard of care for choosing an antiretroviral regimen after treatment failure. As with previous studies, the patients had failed therapy that contained a PI. Overall, the resistance assays did not demonstrate a benefit over standard of care. However in patients with the most limited PI use, a significant benefit was observed in the genotyping arm.[25] The California Collaborative Treatment Group Study 575 (CCTG575) randomized patients who had experienced treatment failure to receive phenotypic testing or standard of care. At both 6 and 12 months there was not a significant difference in either reduction in viral load or in the proportion of patients with a viral load of <400 copies/ml. However, in a subgroup of patients with resistance to three or more PIs at baseline, a statistically greater number of patients achieved a viral load of <400 copies/ml at 6 months in the phenotype arm compared to the standard-of-care arm.[26,27]

An International AIDS Society–USA panel of experts has established clinical guidelines for the use of HIV-1 resistance testing in adults.[28] Either genotypic or phenotypic drug resistance testing is recommended in patients who fail an initial antiretroviral regimen or in those who fail after numerous regimens. Treatment failure usually refers to failure to achieve the desired drop in viral load or rebound in viral load after achieving an initial response. Resistance testing also is recommended in pregnant women to optimize treatment and in hopes of minimizing transmission of HIV-1 infection to the neonate. Resistance testing should be considered prior to initiating therapy in patients with primary HIV-1 infection or in treatment of naive persons with established HIV-1 infection.

Assays and Their Performance

Viral Load Assays

Currently, four HIV-1 viral load test kits have been approved by the FDA for clinical use (Table 36-1). Additional tests are under development and are expected to be available within the next few years. The available assays differ in their limit of quantification and linear range. Both the standard and ultrasensitive Amplicor RT-PCR assays are needed to cover the clinically important range of viral load values. Of note, the Amplicor RT-PCR version 1.0 test has been replaced by the version 1.5 test. The lower limit of quantification is a key characteristic of viral load assays, and how this parameter is defined is equally important. For example, the NucliSens NASBA assay can detect as few as 25 copies/ml when 1 ml of plasma is used, but this level is detected in less than 50% of replicates. The limit of quantification is best defined as the amount of nucleic acid

Table 36-1. Commercially Available HIV-1 Viral Load Assays

Test Kit	Method	Manufacturer	Gene Target	Range* (copies/ml)	Specimen Volume
Amplicor HIV-1 monitor test, standard	RT-PCR	Roche Diagnostics (Indianapolis, IN)	HIV-1 *gag*	400–750,000	200 µl
Amplicor HIV-1 monitor test, ultrasensitive	RT-PCR	Roche Diagnostics (Indianapolis, IN)	HIV-1 *gag*	50–100,000	500 µl
Versant HIV-1 RNA 3.0 assay	Branched DNA	Bayer Diagnostics Corporation, (Tarrytown, NY)	HIV-1 *pol*	75–500,000	1 ml
NucliSens HIV-1 QT assay	NASBA	bioMerieux (Durham, NC)	HIV-1 *gag*	176–3,470,000	1 ml†

*See text for a detailed discussion of the dynamic range of the assays.
†Can accommodate volumes between 50 µl and 2 ml.
RT-PCR, reverse transcription–polymerase chain reaction; NASBA, nucleic acid sequence-based amplification.

that can be verified in 95% of replicate samples. Using this definition, the ultrasensitive Amplicor RT-PCR assay has a sensitivity of 50 copies/ml, followed by the Versant bDNA assay at 75 copies/ml, and the NucliSens HIV-1 QT assay (bioMerieux, Durham, NC) at 176 copies/ml.[29–31] It is unclear whether there is a clinical difference when the viral load is decreased to less than 50 copies/ml compared to reduction below 100 copies/ml.[29]

The ability to detect the various subtypes of HIV-1 RNA also differs among the various viral load assays. Though the majority of HIV-1 infections in North America and Europe are subtype B,[32] infections with non-B subtypes are becoming more common in the United States and certainly represent an important cause of HIV-1 infection globally. The Versant bDNA assay accurately quantifies HIV-1 subtypes A through G, while the Amplicor RT-PCR version 1.0 assay underquantifies HIV-l subtypes A, E, and F by 1–2 $\log_{10}$ copies/ml.[33] This problem has been resolved with the FDA approval of the Amplicor RT-PCR version 1.5 assay, which has modified primer sequences that allow accurate quantification of subtypes A through H.[32–34] For the NucliSens NASBA assay, studies have shown decreased quantification of subtype G.[32,33,35] None of the currently available assays are recommended for quantification of group O viruses. An assay is under development that can detect all types of HIV-1 as well as HIV-2 RNA, but it is not available for clinical use at this time.[36]

The viral load values obtained with the different assays may not be comparable because each uses a different molecular technique and may use a different HIV-l RNA standard. The viral load values obtained with the standard Amplicor RT-PCR (version 1.0) are reported to be about two-fold higher than levels obtained with either version 2.0 of the Versant bDNA assay[37,38] or the NucliSens NASBA assay.[39] The viral load levels obtained with version 3.0 of the Versant bDNA assay and the Amplicor RT-PCR assay (version 1.5) were highly correlated in one study, with differences between the two assays within the coefficients of variation for the assays.[40] The Amplicor version 1.0 and 1.5 assays have been reported to generate comparable values on matched clinical specimens.[41] Although the differences in viral load levels obtained with the different assays are narrowing, it is still best to use the same assay when monitoring patients over time.

Proper collection and processing of blood samples are essential to assure accurate assessment of viral load levels; the key is to minimize RNA degradation. For both the Versant bDNA and Amplicor RT-PCR assays, EDTA is the preferred anticoagulant for blood collection. Blood anticoagulated with acid citrate dextrose (ACD) also is acceptable, but the viral load will be decreased approximately 15% due to the volume of anticoagulant.[42–44] Blood anticoagulated with heparin is unacceptable for the Versant bDNA or Amplicor RT-PCR assays. Due to the extraction method used with the NucliSens NASBA assay, any anticoagulant can be accommodated, though EDTA is most commonly used.[45]

To ensure minimal degradation of RNA, plasma must be separated from blood cells within 4 to 6 hours of collection. Delays in processing may lead to a falsely decreased viral load due to RNA degradation. Plasma specimens can be stored at 4°C for several days without significant degradation of RNA. Moreover, HIV-1 RNA remains stable after three cycles of freezing (–70°C) and thawing.[43] For long-term storage, plasma samples should be frozen at or below –70°C.[44] Vacutainer Plasma Preparation Tubes (PPT) also can be used for the collection of blood specimens for viral load testing. The PPT contain a gel barrier, which, after centrifugation, physically separates plasma from the cellular components.[42] Whole blood collected in PPT can be held at room temperature for as long as 6 hours after collection and shipped as plasma (in the original tube) at ambient temperature or on wet or dry ice without affecting the HIV-1 viral load.[46] Freezing PPTs (after separating the plasma) prior to testing can give higher viral load values compared to those obtained when the plasma is separated and stored at 4°C. For this reason freezing PPTs is not recommended. The PPTs provide a closed sample collection system, which is a safe, convenient, and practical approach to shipping specimens collected at sites remote from the laboratory.

Although measuring viral load in plasma is the standard of care in clinical practice, these assays have been adapted for use with other specimens, most notably serum, dried blood spots, cerebrospinal fluid (CSF), seminal fluid or semen, or cervical secretions. When serum specimens are used, the viral load is decreased approximately 50% compared to plasma.[44] Both blood and plasma dried spots can be used for viral load testing; in fact, viral load levels from dried plasma spots are equivalent to those obtained from fresh frozen plasma specimens. HIV-1 RNA in dried plasma spots remains stable for up to 16 days when stored at 4°C or ambient temperature.[47] Similarly, HIV-1 RNA from dried whole blood spots, corrected for hematocrit, yields viral load results comparable to those obtained from plasma.[48] The measurement of HIV-1 viral load in genital secretions has been done with the Amplicor and NucliSens assays, although this testing is largely reserved for research studies.[44,49] Viral load levels in CSF have been used in the evaluation of patients with AIDS dementia.

Proviral DNA Assays

Qualitative proviral DNA assays are used primarily for the diagnosis of neonatal or acute HIV-1 infections. Currently, there is one commercial test kit available, the Amplicor DNA test; however, it is not FDA approved. The assay uses the same primers as version 1.5 of the Amplicor RT-PCR assay, thus allowing reliable detection of subtypes A through H. Due to the limited availability of commercial proviral DNA assays, many laboratories have developed their own tests using either standard or real-time amplification methods. The performance characteristics of laboratory-developed assays can vary, and performance characteristics

to consider include sensitivity, specificity, reproducibility, and the ability to detect non-B subtypes of HIV-1 proviral DNA.

Drug Resistance Assays

Antiretroviral resistance can be detected using either genotypic or phenotypic assays. HIV-1 genotypic assays identify mutations or changes in the nucleotide sequence known to confer decreased susceptibility to antiretroviral drugs. The effective use of genotypic resistance testing in clinical practice requires an extensive understanding of the genetics of antiretroviral resistance. Phenotype, on the other hand, refers to a viral trait or behavior resulting from the expression of a specific genotype. HIV-1 phenotypic assays measure viral replication in the presence of antiretroviral drugs. Results of phenotypic assays are typically reported as the inhibitory concentration of a drug that reduces in vitro HIV-1 replication by 50% (IC_{50}). The IC_{50} is usually reported as the fold change in IC_{50} relative to a wild-type strain. A "virtual phenotype" also is available commercially for assessing HIV-1 drug resistance. With a virtual phenotype, a phenotypic assay is not performed directly; rather, the information is inferred from the genotypic assay. The results of the genotypic assay are entered into a database containing matching genotypic and phenotypic results from thousands of clinical specimens, and the closest matching phenotypic results are averaged and reported as the virtual phenotype.

There are several methods used for genotypic resistance assays, including automated dideoxynucleotide terminator cycle sequencing, DNA hybridization using high-density microarrays, and reverse hybridization to mutation-specific probes. The initial step in all these genotypic assays is the extraction of HIV-1 RNA from a clinical specimen, which is usually plasma. This is followed by reverse transcription and PCR amplification of the entire protease gene and most of the RT gene (Figure 36-1). The amplified products then are analyzed using one of the methods mentioned above. The vast majority of genotypic resistance testing used for clinical management decisions is done by automated dideoxynucleotide terminator cycle sequencing. These methods are complex, requiring sequence alignment and editing, mutation detection by comparison to a standard wild-type sequence, and interpretation of the significance of the mutations identified. Two FDA-cleared assays are available that include reagents for sequencing and software programs to assist with sequence alignment and interpretation (Trugene HIV-1 Genotyping Kit and OpenGene DNA Sequencing System, Bayer Diagnostics Corporation, Tarrytown, NY; and ViroSeq HIV-1 Genotyping System, Abbott Molecular, Des Plaines, IL). In addition to these commercial assays, several laboratories have developed automated cycle sequencing assays for HIV-1 resistance testing. The currently available commercial assays do not amplify the envelope gene, so mutations that lead to resistance to the fusion inhibitor drugs are not detected. However, this testing is available from reference laboratories.

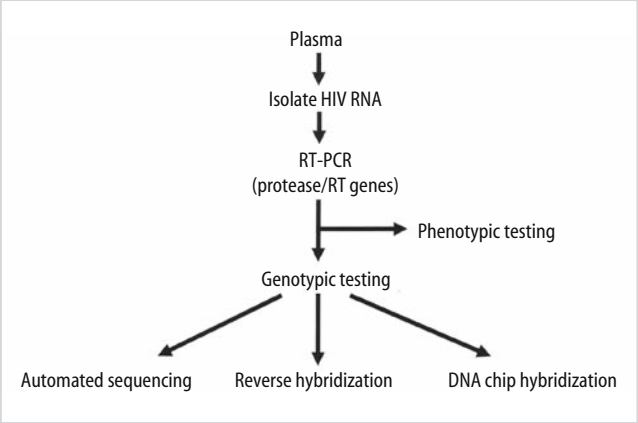


Figure 36-1. Overview of HIV-1 resistance testing.

HIV-1 sequencing also can be done by DNA hybridization of the amplified viral genes to defined oligonucleotide probes using high-density microarrays followed by automated base calling and mutation detection. One such method, GeneChip (Affymetrix, Santa Clara, CA), provides the nucleotide sequence of the entire HIV-1 protease gene and most of the RT gene.[50] Although several studies found good overall concordance (<97%) of sequence data between cycle sequencing and the GeneChip method,[51–53] problems were noted in the ability to detect mixtures of mutant and wild-type virus.[51] However, with continued improvements in DNA chip technology, it is possible that this technology will prove to be successful in the future for detection of resistance mutations.

A reverse hybridization method, the line probe assay (LiPA, manufactured for Bayer Diagnostics Corporation, Tarrytown, NY, by Innogenetics, Gent, Belgium), can rapidly and simultaneously determine the presence of drug resistance mutations in preselected codons rather than the entire gene sequence.[54] The amplified portions of the RT and protease genes are hybridized to nitrocellulose strips containing probes that are complementary to target sequences containing specific mutations. The strips also contain probes complementary to corresponding wild-type target sequences. This method can identify specific mutations but will not provide information about the entire gene sequence. A discordance rate of 8% between LiPA and automated sequencing for the RT gene was noted in one study.[53] Although hybridization failures for some codons exceeded 10% in another study, LiPA appeared to be more sensitive than automated sequencing in detecting minority mutant populations.[55] To date, the LiPA assay has been designed to identify primary resistance mutations only. The nitrocellulose strips will require constant updating as new drugs become available and new mutations are identified. However, because of the relative ease of use, and with ongoing improvements, genotypic methods that rely on

reverse hybridization will likely remain important for clinical and research applications.

There have been few direct comparisons of the various automated sequencing assays used in clinical practice. A study comparing the Trugene and ViroSeq assays showed that both tests identified similar mutations after taking into account differences in drug databases and reference strains.[56] Nineteen of the 21 samples tested were equivalent in the two assays, with an overall concordance of 99% (249 of 252 mutation sites).

Phenotypic assays measure the ability of HIV-1 to grow in the presence of various concentrations of an antiretroviral drug. Results are expressed as the concentration of drug needed to inhibit viral replication by 50% (IC_{50}) or 90% (IC_{90}) compared with a drug-susceptible control virus (Figure 36-2). Initially, phenotypic assays required isolation of infectious HIV-1 from a blood specimen. However, now phenotypic assays are performed using high-throughput automated assays based on recombinant DNA technology. HIV-1 phenotypic resistance testing is available from two different commercial laboratories. For one assay, PhenoSense (ViroLogic, South San Francisco, CA), the protease and RT genes are amplified using RT-PCR and inserted into a modified HIV-1 vector that has a luciferase reporter gene in place of the viral envelope gene. Viral replication, in the presence of various drugs, is measured by quantification of luciferase expression.[57] The reproducibility of the assay is such that increases in IC_{50} of greater than 2.5-fold can be reliably detected in the assay. The second assay, Antivirogram (Virco, Mechelen, Belgium), combines patient and HIV-1 vector sequences using in vitro recombination. Viral replication is measured using a reporter gene system.[58]

There is little data directly addressing specimen collection and processing for HIV-1 resistance testing. These assays are very sensitive to RNA degradation because the methods require the amplification of a large portion of the viral genome (1,200 to 1,600 base pairs). Current recommendations for resistance testing are to follow guidelines established for viral load testing regarding collection, processing, and storage of specimens. The method of RNA extraction can be altered depending on the viral load of the specimen. For specimens with a low viral load (<1,000 to 5,000 copies/ml plasma), the most common approach to improve the yield of RNA is to concentrate virions from a larger specimen volume by high-speed centrifugation (23,000 × g or higher) prior to RNA extraction. Typically a viral load of 1,000 copies/ml or greater is required for reliable results from automated sequencing assays. However, using methods to concentrate virus from a larger volume of plasma, it is possible to obtain sequencing data on specimens with a viral load of <1,000 copies/ml.[56] Considering the cost of these assays and the variability of viral load measurements near the limit of quantification, it may be best to avoid resistance testing until the viral load in the plasma is >1,000 copies/ml. Both the Trugene and ViroSeq assays have been used successfully to genotype non-B subtypes of HIV-1.[59,60]

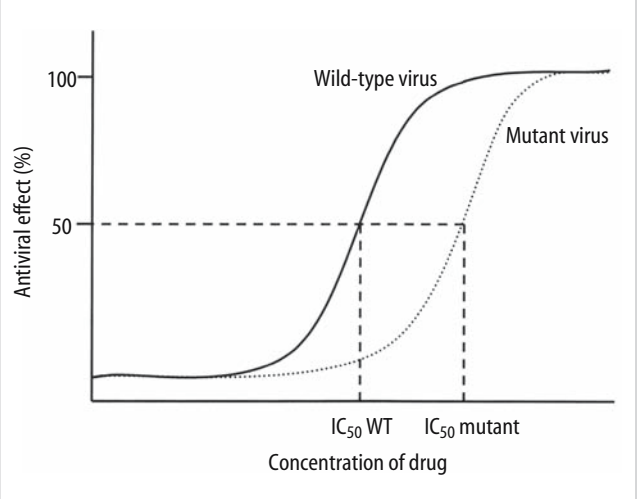


Figure 36-2. HIV-1 phenotypic drug resistance testing results.

Interpretation

Viral Load Assays

HIV-1 viral load assays have become the standard of care for monitoring response to antiretroviral therapy. To effectively use viral load assays in clinical practice there must be an understanding of which changes in viral load represent a clinically important change in viral replication. This requires knowledge of both viral biology and assay performance. The available viral load assays have an intra-assay variability of 0.12 to 0.2 $\log_{10}$ on repeated testing of individual samples,[44,61] with the bDNA assay showing the lowest intra-assay variability. Biologically, HIV-1 RNA levels are fairly stable in individuals who are not receiving therapy; the biological variation is approximately 0.3 $\log_{10}$.[12] Therefore, changes in HIV-1 RNA levels must exceed 0.5 $\log_{10}$ (3-fold) to represent biologically relevant changes in viral replication. For all the viral load assays, the intra-assay variability is even greater near the lower limit of quantification, so for HIV-1 RNA values less than 3 $\log_{10}$ (1,000 copies/ml), it is particularly important not to overinterpret small changes in viral load. Reporting viral load levels as $\log_{10}$-transformed data assists in preventing clinicians from overinterpreting small changes in viral load.

Several clinical illnesses, including herpes simplex virus (HSV) infections, acute infections, and opportunistic infections, as well as vaccinations for influenza, tetanus, or pneumonococcal infection, can lead to transient increases in HIV-1 RNA levels.[62–64] For some individuals these increases in viral load may be quite dramatic, ≥1 $\log_{10}$; however, HIV-1 RNA levels usually return to baseline within a month of the acute event. For this reason it is recommended that viral load measurements be avoided during acute illness or within a month of vaccination.

False-positive results can occur with HIV-1 viral load assays and are attributed to carryover contamination with amplicons, limitations in assay chemistry, or cross-contamination of specimens during specimen processing. Carryover contamination of amplified products is a problem for the Amplicor RT-PCR and NucliSens NASBA assays, as both are target amplification methods. The Amplicor RT-PCR assay has been designed to minimize this problem; specifically, the assay uses dUTP and the DNA repair enzyme uracil-N-glycosylase (UNG) to destroy carryover amplicons before amplification. Carryover contamination will become less of a problem when real-time RT-PCR versions of these assays are used, since the sample is not manipulated after the amplification step. An advantage of the Versant bDNA assay is that carryover contamination does not occur, since it is a signal amplification method. However, the Versant bDNA assay chemistry involves complex hybridization of nucleic acid probes and there can be nonspecific hybridization leading to false-positive results. The assay has a specificity of approximately 98% when testing specimens from HIV-1-negative individuals.[14] Most of the false-positive samples have viral load values of less than 2000 copies/ml. Contamination with HIV-1 RNA during specimen processing can lead to false-positive results with any of the three assays.

Drug Resistance Assays

Interpretation of genotypic resistance testing is very complex and requires a detailed understanding of the genetics of resistance. For many drugs, the mutations associated with resistance have been well characterized when used as monotherapy. However, when drugs are used in combination, as is the standard of care for HIV-1-infected individuals, interactions may occur between mutations, which may increase or decrease drug susceptibilities. These interactions, although complex, must be understood to accurately interpret genotypic results. In general, viral mutations make the virus less fit for replication. The terms "primary" and "secondary" may be used when referring to resistance mutations. Primary mutations are relatively drug specific and may decrease viral susceptibility to the drug. Drug-resistant mutants may have a reduced replication capacity or viral fitness compared to wild-type virus. Secondary mutations alone may have no effect on viral drug susceptibility but may improve viral fitness, allowing a virus with a primary mutation to improve its replicative capacity.[65] A current and comprehensive discussion of the specific mutations associated with each antiretroviral drug and the interactions of mutations is available from a variety of sources,[28] including Los Alamos National Laboratory HIV Databases (http://hiv-web.lanl.gov), International AIDS Society–USA (http://www.iasusa.org), and Stanford University HIV Drug Resistance Database (http://hivdb.stanford.edu).

The proper interpretation of genotypic drug resistance assays involves two independent processes, identification of resistance mutations and interpretation of how these mutations alter viral susceptibility to specific antiretroviral drugs. An error in either process will lead to an inaccurate genotyping interpretation. While establishing appropriate quality control guidelines for the technical aspect of an assay is common practice, the complex interpretation of HIV-1 genotyping assays represents a new challenge for clinical laboratory personnel. Since interpretation of genotypic assays is so complex and critical to patient care, both of the FDA-cleared assays provide software programs that assist in base calling, sequence alignment, and identification of the resistance mutations by comparing the sequence to that of a wild-type virus.

After identifying the resistance mutations, a "rule-based" software program is used to interpret the meaning of the various mutations. For example, with the OpenGene system (Bayer Diagnostics Corporation), the rules are established by an independent panel of experts and updated regularly as new information becomes available. A computer algorithm has been developed based on these rules that accounts for primary and secondary mutations, cross resistance, and interactions of mutations. In addition to listing the mutations identified in the RT and protease genes, an interpretative report is provided that lists each drug and provides a designation of either "no evidence of resistance," "possible resistance," "resistance," or "insufficient evidence."[66] "No evidence of resistance" is used if no known mutations are detected or if reduced susceptibility to a specific drug has not been associated with the identified mutation(s). "Possible resistance" is used when the mutations detected have been associated with diminished virologic response in some but not all patients, or if the mutation has been associated with an intermediate decrease in susceptibility to the drug. "Resistance" refers to mutations that have been associated with a maximum reduction in susceptibility to the drug. If there is inadequate evidence to determine susceptibility, the term "insufficient evidence" is used.[66] A similar approach is used in the ViroSeq assay (Abbott Molecular), though the interpretation of all mutations may not be identical for both systems. These rule-based interpretation systems are essential for providing clinicians with results in a user-friendly format that is easily understood and clinically useful without the need for an extensive knowledge of the genetics of HIV-1 resistance. Ideally, one rule-based system would be used for the interpretation of all genotypic assays. These databases require frequent updates as new data regarding drug resistance become available and new drugs are available for clinical use. Clinically, resistance results must be interpreted in the context of the treatment history of each patient.

Mutations in HIV-1 are reported with a specific nomenclature in which amino acids are reported using single-letter abbreviations. The wild-type amino acid encoded by the nucleotide triplet is followed by the location of the mutation (codon number) and then the mutant amino acid. For example, K103N indicates that the lysine (wild

Table 36-2. Example of Genotypic Resistance Report

Drug	Interpretation
Nucleoside RT Inhibitor Mutations: D67N, K70R, M184V, K219E	
Zidovudine	Resistance
Didanosine	Possible resistance
Zalcitabine	Possible resistance
Lamivudine/emtricitabine	Resistance
Stavudine	Possible resistance
Abacavir	Possible resistance
Tenofovir	No evidence of resistance
Nonnucleoside RT Inhibitor Mutations: K103N	
Nevirapine	Resistance
Delavirdine	Resistance
Efavirenz	Resistance
Protease Inhibitor Mutations: D30N	
Saquinavir	No evidence of resistance
Indinavir	No evidence of resistance
Ritonavir	No evidence of resistance
Nelfinavir	Resistance
Amprenavir/fosamprenavir	No evidence of resistance
Lopinavir and ritonavir	No evidence of resistance
Atazanavir	No evidence of resistance

type) at codon 103 is replaced by an asparagine (mutant). Genotyping reports include a list of the mutations identified as well as the effect of the mutations on antiretroviral drug susceptibility (Table 36-2).

An important issue for the interpretation of HIV-1 phenotypic assays is defining the cutoff change in IC_{50} that is associated with resistance. Initially, this cutoff was established based on the technical performance of the assays, which is the increase in IC_{50} that could be reliably detected when compared to wild-type virus. For the PhenoSense assay, this was established at a 2.5-fold increase in IC_{50}, while for the Antivirogram assay this was set at a 4-fold increase in IC_{50}. More recently there has been a move to define the cutoffs in relationship to the likelihood of a clinical response. Clinically important increases in IC_{50} have been defined for some drugs, and these cutoffs can vary with different drugs. For example, with abacavir, stavudine, and lopinavir the fold change in IC_{50} that is clinically important is 4.5, 1.7, and 10, respectively.[67] However, with NNRTIs, 4- to 10-fold increases in IC_{50} values in therapy-naive subjects were not associated with a poor virologic outcome.[68] It is likely that IC_{50} cutoffs will continue to be modified as more clinical outcomes data become available. Reports of phenotypic assay results include not only the change in IC_{50} value but also an interpretation of whether there is an increase or decrease in susceptibility compared to wild-type virus.

A standard of care has not been established for the preferential use of genotypic or phenotypic resistance testing. Some clinicians prefer phenotypic testing because it is a direct measure of viral susceptibility. Others prefer genotypic testing because the development of a mutation may precede phenotypic expression of resistance. In addition, genotypic testing can be completed in 1 to 2 days, compared to 1 to 2 weeks for phenotypic testing. Another important factor is that phenotypic testing is significantly more expensive than genotypic testing. Due to issues of availability and cost, providers often use genotypic testing routinely and rely on phenotypic testing for patients who have failed multiple regimens and have very complex genotypic results. If both assays are used, it is key to remember that the results of both assays may not agree, as the presence of a resistance mutation does not assure its expression in a phenotypic assay. Depending on the drug regimen, genotypic and phenotypic testing may provide complementary information.[67]

The currently available genotypic and phenotypic methods can detect a mutant virus only if it comprises at least 20% of the total viral population. So resistance testing is best used to predict failure of a regimen rather than ensure success, because a mutant may be present at a concentration below the detection limit of the assay. The one exception is the LiPA assay, which can detect mutants that comprise as little as 4% of the total viral population,[51,54] though this test is not widely used in clinical laboratories. Detection of these minor mutants is important because, in the presence of drug selection pressure, these minority populations will quickly predominate. Likewise, drug selection pressure is required for mutations to persist, and in the absence of drug selection pressure the viral population can revert back to wild type. One study showed that this reversion back to wild type can occur within several weeks of withdrawing the antiretroviral drug.[69] Although the virus may appear susceptible to the specific drug, reintroducing the drug will select for resistance again. For this reason, it is recommended that specimens for resistance testing be obtained while the patient is on antiretroviral therapy.

Laboratory Issues

With the proven clinical utility and the availability of commercial assays, viral load testing for HIV-1 is performed routinely in many clinical laboratories. Three companies have FDA-approved test kits; the appropriate choice of a viral load assay will vary depending on the laboratory. Issues to consider include available space, volume of testing, turnaround time, expertise of technologists, and cost of reportable result. Each of the viral load assays has its strengths and weaknesses. The Versant bDNA assay has the advantages of high throughput and a broad dynamic range, without concern for carryover contamination of amplified product. The System 340 instrument for the Versant bDNA assay allows for a greater level of automation. However, the assay does not have an internal control to compensate for loss of RNA during nucleic acid extraction, and false-positive results have been reported.[14] The Amplicor RT-PCR assay has good specificity, but the dynamic range is limited, requiring the use of both a stan-

dard and an ultrasensitive version of the assay, and adequate space is needed for proper laboratory design to separate the pre- and post-PCR steps of analysis. The Cobas Amplicor instrument, which automates the amplification and detection steps of the test, is available. However, the throughput of the instrument is limited, requiring 6 hours to test 24 specimens. The NASBA assay has a broad dynamic range and can accommodate a wide range of clinical specimen types and specimen volumes, but the assay is slightly less sensitive than the other available assays. The MiniMag and EasyMag are instruments that automate the nucleic acid extraction step of the NucliSens assay and have greatly reduced the hands-on time needed to perform the test.

The initial step in some genotypic and all phenotypic assays involves amplification of viral RNA using RT-PCR technology. As a result, strict precautions must be followed to avoid contamination of specimens with amplified products and to prevent contamination between specimens during processing and amplification. Genetic fingerprinting programs are a very important tool to assist the laboratory in detecting contamination or sample mix-up. Each sequence that is generated is compared to a laboratory-specific database that contains the sequences of other specimens and controls. The database then displays the prior sequences most closely related to the current sample. Due to the variation in the HIV-1 sequence, there should not be an exact or a highly similar match within the database unless it is an earlier specimen from the same patient. If different patients are found to have highly similar sequences, this can be a clue to problems with contamination in the laboratory. This genetic fingerprinting function is included in the Trugene System, allowing all sequenced specimens to be checked against the fingerprinting database. The ViroSeq system uses UTP and uracil N-glycosylase (UNG), which is a very effective method to control for contamination of amplified products. As with viral load testing, extreme care must be taken to avoid cross contamination of specimens during processing.

The College of American Pathologists(CAP) Surveys Program offers proficiency testing for HIV-1 viral load measurements (HIV/HV2 survey) three times per year. The Centers for Disease Control and Prevention (CDC) offers laboratory performance evaluations twice a year (the CDC Model Performance Evaluation).

Phenotypic resistance testing, which requires culture of HIV-1, is not performed in clinical laboratories but is available from two commercial laboratories (Virologic and Virco). Genotypic assays that rely on automated sequencing technology are performed in clinical laboratories; the testing is very complex and challenging for laboratories because quality control is required for both the generation of sequence and the subsequent interpretation of the results. An international HIV-1 genotyping proficiency program (ENVA-3)[70] distributed five plasma samples to 175 laboratories. One sample in the panel contained a 50:50 mixture of wild-type and mutant HIV-1 with mutations in various resistance codons. For this specimen, at least eight of the nine resistance mutations were detected by 55% of the laboratories using ViroSeq, 62% of the laboratories using Trugene, and 33% of the laboratories using laboratory-developed assays. Overall these results are quite disappointing considering that sequencing technology is reported to detect mutants that comprise as little as 20% of the total viral population. In addition to difficulties identifying resistance mutations, there was extensive variation between laboratories in the interpretation of the significance of the mutations.[71] These results clearly point out the need for proficiency testing for both identification and interpretation of resistance mutations.

The interpretation of an assay as complex as HIV-1 genotyping presents a new challenge for clinical laboratories. Many laboratories rely on databases and algorithms provided by commercial companies, since experts in the interpretation of resistance testing may not be available locally. One problem with this approach is that the laboratory has no control over how frequently these databases are updated. In addition, the databases may use different reference strains, base-calling algorithms, or resistance reference files, leading to different interpretations of the same resistance mutations. Standardization of these test components is needed to minimize misinterpretation of genotyping results. FDA clearance has been an important step for genotyping assays, but standardization of interpretive criteria is needed.

Proficiency testing for HIV-1 genotyping is available from the College of American Pathologists (Northfield, IL) Surveys Program (HIV/HV2 surveys), with three challenges per year. To date, this proficiency program has focused on identification of specific resistance mutations, rather than interpretation of these mutations. In the future, proficiency programs will need to include mutation identification as well as interpretation challenges.

An important advance for HIV-1 drug resistance testing has been the establishment of CPT codes for both phenotypic (87903) and genotypic (87901) testing. However, reimbursement rates vary greatly from state to state and may be below the cost of performing the tests. Adequate reimbursement is essential to assure access to resistance testing for all HIV1-infected individuals.

The Future

The use of HAART in combination with HIV-1 viral load and drug resistance assays has revolutionized the management of HIV-1-infected persons. Future advancements for these tests will likely focus on improved assay performance and automation. Real-time RT-PCR methods, which allow the simultaneous amplification and detection of amplified product, offer many improvements over the current assays. The technical time to perform the assay and time to result will be decreased. The risk of contamination due to carry-over of amplified product will be greatly reduced since

there is no postamplification handling of the sample. Analysis is done early in the log phase of amplification, thus decreasing differences in amplification efficiency between samples. Real-time HIV-1 viral load assays are under development that are ultrasensitive (20 to 50 copies/ml) with a 6 $\log_{10}$ linear range. Some of these assays also quantify group O virus.

The currently available genotypic drug resistance assays are very labor-intensive. Future assays will need to be much more automated with a higher throughput. With advances in DNA chip technology, this approach may prove to be applicable to complex genotypic assays. Other needed improvements include standardization of databases and interpretive algorithms, optimization of assay performance to improve detection of minor mutant populations, and establishing appropriate quality control and quality assurance programs. HIV-1 drug resistance assays will need to be very flexible to accommodate new antiretroviral drugs that target other genes such as *gag*, *integrase*, or *envelope*. Finally, the clinical correlation between genotypic and phenotypic assays needs to be better defined.

In the future, the clinical management of HIV-1 infected individuals may include therapeutic drug monitoring, and pharmacogenomic testing to identify individuals with altered metabolism of drugs or those at risk for adverse reactions to antiretroviral drugs. The use of these types of assays may allow for better individualization of antiretroviral therapy and improved clinical outcomes.

References

1. Connor R, Ho D, Kuritzkes DR, Richman D. Human Immunodeficiency Virus. In: *Clinical Virology*, 1st ed. Richman DD, Whitley RJ, and Hayden FG (eds.) Churchill Livingstone, New York; 1997:707–754.
2. Yeni PG, Hammer SM, Hirsch MS, et al. Treatment for adult HIV infection: 2004 recommendations of the International AIDS Society–USA Panel. *JAMA*. 2004;292:251–265.
3. Mellors JW, Kingsley LA, Rinaldo CR Jr, et al. Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. *Ann Intern Med.* 1995;122:573–579.
4. O'Brien TR, Blattner WA, Waters D, et al. Serum HIV-1 RNA levels and time to development of AIDS in the Multicenter Hemophilia Cohort Study. *JAMA*. 1996;276:105–110.
5. Mellors JW, Rinaldo CR Jr, Gupta P, White RM, Todd JA, Kingsley LA. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science.* 1996;272:1167–1170.
6. Eron JJ, Benoit SL, Jemsek J, et al. Treatment with lamivudine, zidovudine, or both in HIV-positive patients with 200 to 500 CD4+ cells per cubic millimeter. *N Eng J Med.* 1995;333:1662–1669.
7. Hammer SM, Squires KE, Hughes MD, et al. A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and CD4 counts of 200 per cubic millimeter or less. *N Eng J Med.* 1997;337:725–739.
8. Cameron DW, Heath-Chiozzi M, Danner S, et al. Randomised placebo-controlled trial of ritonavir in advanced HIV-1 disease. The Advanced HIV Disease Ritonavir Study Group. *Lancet.* 1998;351:543–549.
9. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4+ lymphocytes in HIV-1 infection. *Nature.* 1995;373:123–126.
10. Wei X, Ghosh SK, Taylor ME, et al. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature.* 1995;373:117–122.
11. Raboud JM, Montaner JS, Conway B, et al. Suppression of plasma viral load below 20 copies/ml is required to achieve a long-term response to therapy. *AIDS.* 1998;12:1619–1624.
12. Saag MS, Holodniy M, Kuritzkes DR, et al. HIV viral load markers in clinical practice. *Nature Med.* 1996;2:625–629.
13. Kahn JO, Walker BD. Acute human immunodeficiency virus type 1 infection. *N Eng J Med.* 1998;339:32–39.
14. Daar ES, Little S, Pitt J, et al. Diagnosis of primary HIV-1 infection. Los Angeles County Primary HIV Infection Recruitment Network. *Ann Intern Med.* 2001;134:25–29.
15. Young NL, Shaffer N, Chaowanachan T, et al. Early diagnosis of HIV-1-infected infants in Thailand using RNA and DNA PCR assays sensitive to non-B subtypes. *J Acquir Immune Defic Syndr.* 2000;24:401–407.
16. Cunningham CK, Charbonneau TT, Song K, et al. Comparison of human immunodeficiency virus 1 DNA polymerase chain reaction and qualitative and quantitative RNA polymerase chain reaction in human immunodeficiency virus 1-exposed infants. *Pediatr Infect Dis J.* 1999;18:30–35.
17. Steketee RW, Abrams EJ, Thea DM, et al. Early detection of perinatal human immunodeficiency virus (HIV) type 1 infection using HIV RNA amplification and detection. *J Infect Dis.* 1997;175:707–711.
18. Centers for Disease Control. 1994 Revised classification system for human immunodeficiency virus infection in children less than 13 years of age. *MMWR Morb Mortal Wkly Rep.* 1994;43:(RR-12):1–11.
19. Durant J, Clevenbergh P, Halfon P, et al. Drug-resistance genotyping in HIV-1 therapy: the VIRADAPT randomised controlled trial. *Lancet.* 1999;353:2195–2199.
20. Clevenbergh P, Durant J, Halfon P, et al. Persisting long-term benefit of genotype-guided treatment for HIV-infected patients failing HAART. The VIRADAPT study: week 48 follow-up. *Antivir Ther.* 2000;5:65–70.
21. Baxter JD, Mayers DL, Wentworth DN, et al. A randomized study of antiretroviral management based on plasma genotypic antiretroviral resistance testing in patients failing therapy. *AIDS.* 2000;14:F83–F93.
22. Tural C, Ruiz L, Holtzer C, et al. Clinical utility of HIV-1 genotyping and expert advice—the Havana trial. *AIDS.* 2002;16:209–218.
23. Weinstein MC, Goldie SJ, Losina E, et al. Use of genotypic resistance testing to guide HIV therapy: clinical impact and cost-effectiveness. *Ann Intern Med.* 2001;134:440–450.
24. Cohen CJ, Hunt S, Sension M, et al. A randomized trial assessing the impact of phenotypic resistance testing on antiretroviral therapy. *AIDS.* 2002;16:579–588.
25. Meynard J-L, Vray M, Morand-Joubert L, et al. Phenotypic or genotypic resistance testing for choosing antiretroviral therapy after treatment failure: a randomized trial. *AIDS.* 2002;16:727–736.
26. Miller V. HIV drug resistance: overview of clinical data. *J HIV Ther.* 2001;6:68–72.
27. Haubrich RH, Kemper CA, Hellmann NS, et al. A randomized, prospective study of phenotype susceptibility testing versus standard of care to manage antiretroviral therapy: CCTG 575. *AIDS.* 2005;19:295–302.
28. Hirsch MS, Conway B, D'Aquila RT, et al. Antiretroviral drug resistance testing in adults with human immunodeficiency virus type 1: 2003 recommendations of an International AIDS Society–USA panel. *Clin Infect Dis.* 2003;37:113–128.
29. Erice A, Brambilla D, Bremer J, et al. Performance characteristics of the QUANTIPLEX HIV-1 RNA 3.0 assay for detection and quantitation of human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol.* 2000;38:2837–2845.
30. Sun R, Ku J, Jayakar H, et al. Ultrasensitive reverse transcription-PCR assay for quantitation of human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol.* 1998;36:2964–2969.
31. Witt DJ, Kemper M, Stead A, Ginocchio C, Caliendo AM. Relationship of incremental specimen volumes and enhanced detection of

human immunodeficiency virus type 1 RNA with nucleic acid amplification technology. *J Clin Microbiol.* 2000;38:85–89.
32. Parekh B, Phillips S, Granade TC, Baggs J, Hu DJ, Respess R. Impact of HIV type 1 subtype variation on viral RNA quantitation. *AIDS Res Hum Retroviruses.* 1999;15:133–142.
33. Jagodzinski LL, Wiggins DL, McManis JL, et al. Use of calibrated viral load standards for group M subtypes of human immunodeficiency virus type 1 to assess the performance of viral RNA quantitation tests. *J Clin Microbiol.* 2000;38:1247–1249.
34. Pasquier C, Sandres K, Salama G, Puel J, Izopet J. Using RT-PCR and bDNA assays to measure non-clade B HIV-1 subtype RNA. *J Virol Methods.* 1999;81:123–129.
35. Nkengasong JN, Bile C, Kalou M, et al. Quantification of RNA in HIV Type 1 subtypes D and G by NucliSens and Amplicor Assays in Abidjan, Ivory Coast. *AIDS Res Hum Retroviruses.* 1999;15:495–498.
36. Abravaya K, Esping C, Hoenle R, et al. Performance of a multiplex qualitative PCR LCx assay for detection of human immunodeficiency virus type 1 (HIV-1) group M subtypes, group O, and HIV-2. *J Clin Microbiol.* 2000;38:716–723.
37. Nolte FS, Boysza J, Thurmond C, Clark WS, Lennox JL. Clinical comparison of an enhanced-sensitivity branched-DNA assay and reverse transcription-PCR for quantitation of human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol.* 1998;36:716–720.
38. Segondy M, Izopet J, Pellegrin I, et al. Comparison of the Quantiplex HIV-1 RNA 2.0 assay with the Amplicor HIV-1 monitor 1.0 assay for quantitation of levels of human immunodeficiency virus type 1 RNA in plasma of patients receiving stavudine-didanosine combination therapy. *J Clin Microbiol.* 1998;36:3392–3395.
39. Dyer JR, Pilcher CD, Shepard R, Schock J, Eron JJ, Fiscus SA. Comparison of NucliSens and Roche Monitor assays for quantitation of levels of human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol.* 1999;37:447–449.
40. Elbeik T, Charlebois E, Nassos P, et al. Quantitative and cost comparison of ultrasensitive human immunodeficiency virus type 1 RNA viral load assays: Bayer bDNA quantiplex versions 3.0 and 2.0 and Roche PCR Amplicor Monitor version 1.5. *J Clin Microbiol.* 2000;38:1113–1120.
41. Triques K, Coste J, Perret JL, et al. Efficiencies of four versions of the AMPLICOR HIV-1 MONITOR test for quantification of different subtypes of human immunodeficiency virus type 1. *J Clin Microbiol.* 1999;37:110–116.
42. Ginocchio CC, Wang XP, Kaplan MH, et al. Effects of specimen collection, processing, and storage conditions on stability of human immunodeficiency virus type 1 RNA levels in plasma. *J Clin Microbiol.* 1997;35:2886–2893.
43. Sebire K, McGavin K, Land S, Middleton T, Birch C. Stability of human immunodeficiency virus RNA in blood specimens as measured by a commercial PCR-based assay. *J Clin Microbiol.* 1998;36:493–498.
44. Lew J, Reichelderfer P, Fowler M, et al. Determinations of levels of human immunodeficiency virus type 1 RNA in plasma: reassessment of parameters affecting assay outcome. TUBE Meeting Workshop Attendees. Technology Utilization for HIV-1 Blood Evaluation and Standardization in Pediatrics. *J Clin Microbiol.* 1998;36:1471–1479.
45. Boom R, Sol C, Salimans M, Jansen C, Wertheim-van Dillen PM, van derNoordaa J. Rapid and simple method for purification of nucleic acids. *J Clin Microbiol.* 1990;28:495–503.
46. Holodniy M, Rainen L, Herman S, Yen-Lieberman B. Stability of plasma human immunodeficiency virus load in VACUTAINER PPT plasma preparation tubes during overnight shipment. *J Clin Microbiol.* 2000;38:323–326.
47. Cassol S, Gill MJ, Pilon R, et al. Quantification of human immunodeficiency virus type 1 DNA from dried plasma spots collected on filter paper. *J Clin Microbiol.* 1997;35:2795–2801.
48. Fiscus SA, Brambilla D, Grosso L, Schock J, Cronin M. Quantitation of human immunodeficiency virus type 1 RNA in plasma by using blood dried on filter paper. *J Clin Microbiol.* 1998;36:258–260.
49. Bremer J, Nowicki M, Beckner S, et al. Comparison of two amplification technologies for detection and quantitation of human immunodeficiency virus type 1 RNA in the female genital tract. *J Clin Microbiol.* 2000;38:2665–2669.
50. Kozal MJ, Shah N, Shen N, et al. Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat Med.* 1996;2:753–759.
51. Gunthard HF, Wong JK, Ignacio CC, Havlir DV, Richman DD. Comparative performance of high-density oligonucleotide sequencing and dideoxynucleotide sequencing of HIV type 1 pol from clinical samples. *AIDS Res Hum Retroviruses.* 1998;14:869–876.
52. Hanna GJ, Johnson VA, Kuritzkes DR, et al. Comparison of sequencing by hybridization and cycle sequencing for genotyping of human immunodeficiency virus type 1 reverse transcriptase. *J Clin Microbiol.* 2000;38:2715–2721.
53. Wilson JW, Bean P, Robins T, Graziano F, Persing DH. Comparative evaluation of three human immunodeficiency virus genotyping systems: the HIV-Genotype R method, the HIV PRT GeneChip assay, and the HIV-1 RT line probe assay. *J Clin Microbiol.* 2000;38:3022–3028.
54. Stuyver L, Wyseur A, Rombout A. Line probe assay for rapid detection of drug-selected mutations in the human immunodeficiency virus type 1 reverse transcriptase gene. *Antimicrob Agents Chemother.* 1997;41:284–291.
55. Servais J, Lambert C, Fontaine E, et al. Comparison of DNA sequencing and a line probe assay for detection of human immunodeficiency virus type 1 drug resistance mutations in patients failing highly active antiretroviral therapy. *J Clin Microbiol.* 2001;39:454–459.
56. Erali M, Page S, Reimer LG, Hillyard DR. Human immunodeficiency virus type 1 drug resistance testing: a comparison of three sequence-based methods. *J Clin Microbiol.* 2001;39:2157–2165.
57. Petropoulos CJ, Parkin NT, Limoli KL. A novel phenotypic drug susceptibility assay for human immunodeficiency virus type 1. *Antimicrob Agents Chemother.* 2000;44:920–928.
58. Hertogs K, de Bethune MP, Miler V, et al. A rapid method for simultaneous detection of phenotypic resistance to inhibitors of protease and reverse transcriptase in recombinant human immunodeficiency virus type 1 isolates from patients treated with antiretroviral drugs. *Antimicrob Agents Chemother.* 1998;42:269–276.
59. Mracna M, Becker-Pergola G, Dileanis J, et al. Performance of the Applied Biosystems ViroSeq HIV-1 genotyping system for sequence-based analysis of non-subtype B human immunodeficiency virus type 1 from Uganda. *J Clin Microbiol.* 2001;39:4323–4327.
60. Rozenbaum W, Charpentier C, Delphin N, et al. Comparative performance evaluation of a kit-based HIV-1 genotyping assay and a laboratory developed "home-brew" genotyping assay with samples of varying viral loads and non-B subtypes of group M virus [abstract 171]. *Antiviral Therapy, 5th International Workshop on HIV Drug Resistance & Treatment Strategies.* Scottsdale, AZ; 2001;6:130.
61. Brambilla D, Leung S, Lew J, et al. Absolute copy number and relative change in determinations of human immunodeficiency virus type 1 RNA in plasma: effect of an external standard on kit comparisons. *J Clin Microbiol.* 1998;36:311–314.
62. Donovan RM, Bush CE, Markowitz NP, Baxa DM, Saravolatz LD. Changes in virus load markers during AIDS-associated opportunistic diseases in human immunodeficiency virus-infected persons. *J Infect Dis.* 1996;174:401–403.
63. O'Brien WA, Grovit-Ferbas K, Namazi A, et al. Human immunodeficiency virus-type 1 replication can be increased in peripheral blood of seropositive patients after influenza vaccination. *Blood.* 1995;86:1082–1089.
64. Staprans SI, Hamilton BL, Follansbee SE, et al. Activation of virus replication after vaccination of HIV-1-infected individuals. *J Exp Med.* 1995;182:1727–1737.
65. Hirsch MS, Conway B, D'Aquila RT, et al. Antiretroviral drug resistance testing in adults with HIV infection. *JAMA.* 1998;279:1984–1991.

66. D'Aquila RT. Incorporating antiretroviral resistance testing into clinical practice. Medscape HIV/AIDS CME Circle. Web MD Medscape Health Network. 2002. Available at: http://www.medscape.com.
67. Parkin N, Chappey C, Maroldo L, Bates M, Hellmann NS, Petropoulos CJ. Phenotypic and genotypic HIV-1 drug resistance assays provide complementary information. *J Acquir Immune Defic Syndr.* 2002;31:128–136.
68. Van Houtte M. Update on resistance testing. *J HIV Ther.* 2001;6:61–64.
69. Devereux HL, Youle M, Johnson MA, Loveday C. Rapid decline in detectability of HIV-1 drug resistance mutations after stopping therapy. *AIDS.* 1999;13:F123–F127.
70. Keulen W, Brambilla D, Buimer M, et al. A study on HIV-1 genotyping proficiency in 125 laboratories, using the ENVA-3 panel [abstract 166]. *Antiviral Therapy, 5th International Workshop on HIV Drug Resistance & Treatment Strategies.* Scottsdale, AZ; 2001;6:127.
71. Wensing AM, Keulen W, Buimer M, Brambilla D, Schuurman R, Boucher C. Analysis of the world-wide evaluation study on HIV-1 genotype interpretation: ENVA-3 [abstract 133]. *Antiviral Therapy, 5th International Workshop on HIV Drug Resistance & Treatment Strategies,* Scottsdale, AZ; 2001;6:101.

Chapter 37

Hepatitis B and C Viruses

Frederick S. Nolte

Introduction

Viral hepatitis is believed to have existed in antiquity, with references traced back to the fifth century BC. A new era in viral hepatitis was ushered in by the landmark discovery of Australia antigen, subsequently renamed hepatitis B surface antigen (HbsAg) by Blumberg and coworkers in 1965.[1] What followed was a rapid growth in information about the hepatitis B virus (HBV), development of serologic and molecular tests for HBV, understanding of the natural history and pathogenesis of infection, development and approval of antiviral therapies, and most importantly, the development of effective vaccines for prevention of HBV infection.

Although non-A, non-B hepatitis was thought to have a viral etiology since 1974, the virus eluded investigators for more than a decade. It was discovered in 1989 by a brute force application of molecular cloning techniques through the joint efforts of the Centers for Disease Control and Prevention (CDC) and the Chiron Corporation, and was named the hepatitis C virus (HCV).[2] This discovery led rapidly to the development of serologic screening assays for HCV infection prior to blood donation, which dramatically reduced the incidence of posttransfusion chronic hepatitis. Sequencing of the HCV genome also provided impetus for the development of molecular assays for detection, quantitation, and characterization of HCV. Although there has been explosive growth in information about this medically important virus since its discovery, much remains to be learned about its pathogenesis, treatment, and prevention.

The key characteristics of HBV and HCV are summarized in Table 37-1. Both viruses represent major global public health problems, with an estimated 350 million and 170 million persons chronically infected with HBV and HCV, respectively. Although clinical characteristics and risk factors for infection may give some indication of the specific etiology of viral hepatitis, the diagnosis is laboratory based. Laboratory diagnosis is based on serologic and molecular tests because cell culture techniques for isolation of HBV and HCV are not effective. This chapter reviews the molecular tests that are available to detect, quantitate, and characterize HBV and HCV and how these tests can be used for effective diagnosis and clinical management of patients.

HEPATITIS B VIRUS

The HBV genome is a 3.2 kilobase (kb), relaxed circular, partially double-stranded DNA molecule. It has four partially overlapping open reading frames encoding the viral envelope (pre-S and S), nucleocapsid (precore and core), polymerase, and X proteins. Although HBV is a DNA virus, it replicates by a reverse transcriptase that lacks proofreading activity and, as a result, is prone to errors. The overlapping open reading frames of the genome limit the types of mutations that can be tolerated. However, variations in HBV sequences have been detected in almost all regions of the genome. Consequently, HBV exists as quasispecies, and different patients may be infected with different strains and genotypes.

There are seven phylogenetic genotypes (A through G), most of which have distinct geographic distribution. Genotypes are defined by intergroup divergence of greater than 8% in the complete genome nucleotide sequence. Genotypes A and D are common in the United States and Europe; genotypes B and C are most frequent in China and Southeast Asia. There is insufficient data to determine whether differences in clinical outcome or response to treatment correlate with virus genotype.

Clinical Utility

Serologic assays with high levels of sensitivity, specificity, and reproducibility have been developed to detect HBV antigens and their respective antibodies. This complicated system of serologic markers is used for diagnosis of HBV infection and to define the phase of infection, degree of infectivity, prognosis, and immune status. The presence of

Table 37-1. Characteristics of Hepatitis B Virus and Hepatitis C Virus

Characteristic	Hepatitis B Virus	Hepatitis C Virus
Type of virus	Hepadnavirus	Hepacivirus
Viral genome	3.2 kb, relaxed circular, partially double-stranded DNA	9.5 kb, positive-sense, single-stranded RNA
Routes of transmission	Parenteral Sexual Injection drug use Perinatal	Parenteral Injection drug use Blood products before 1990 Perinatal (infrequent) Sexual (infrequent)
Frequency of acute icteric disease	Common in adults Uncommon in children	Uncommon
Frequency of evolution to chronic disease	Infrequent (<10%) in adults Common in children	Frequent (>80%)
Estimated number of acute infections/yr in the United States	185,000	38,000
Estimated number of chronically infected persons in the United States	1,250,000	2,700,000
Estimated number of chronically infected persons in the world	350,000,000	170,000,000
Treatment	Interferon-α Lamivudine	Interferon-α with ribavirin
Prophylaxis	Recombinant vaccine Hepatitis B immune globulin	None

Source: Adapted from Lauer GM, Walker BD. Hepatitis C virus infection. *N Engl J Med.* 2001;345:41–52.

HBV DNA in the serum is a marker of viral replication in the liver and has replaced hepatitis Be antigen (HBeAg) as the most sensitive marker of viral replication. Efforts are currently under way to develop HBV DNA tests for screening blood donors as an early marker of infection.

The initial evaluation of patients found to have HBsAg in serum should include routine liver tests and a variety of virologic tests including HBV DNA tests.[3] Chronic HBV infection is a disease of variable course, and it is important to establish baseline laboratory values. Monitoring disease activity in chronically HBV-infected patients is best done by measuring aminotransferase (ALT) levels at 6-month intervals. Serial HBV DNA testing is not necessary.

The determination of HBV DNA levels (viral load) is important in the pretreatment evaluation and monitoring of therapeutic response in patients with chronic infection.[3] Currently, therapy for HBV infection is difficult and has limited long-term efficacy. The decision to treat should be based on ALT elevations, presence of HBeAg or high HBV DNA levels or both, presence of moderate disease activity and fibrosis on liver biopsy, and virologic testing to exclude concurrent infections with hepatitis D virus (HDV), HCV, and human immunodeficiency virus (HIV). Currently, the only two approved therapies for HBV infection are interferon-α (IFN) and lamivudine. Several factors predict a favorable response to IFN treatment, the most important being high ALT and low serum HBV DNA levels, which are indirect markers of immune clearance.

Lamivudine is an orally administered nucleoside analog, which is a potent inhibitor of HBV DNA synthesis. A 1-year course has been shown in clinical trials to be as effective as a 16-week course of IFN but is better tolerated by patients.[4,5] The antiviral response appears to be durable, although most treated patients had detectable serum HBV DNA when sensitive tests were used in follow-up studies. A major problem with lamivudine is the development of drug-resistant mutants. Genotypic resistance has been reported in 15% to 25% of patients treated for 1 year, increasing to 49% of patients treated for 3 years. The development of resistance usually is associated with a rebound in viral load. However, in many patients the viral load and ALT level remain lower than pretreatment values because the mutations that occur in the viral polymerase affect viral replication efficiency. The two most common mutations are a methionine-to-valine or -isoleucine substitution at codon 552 (M552V/I) and a leucine-to-methionine substitution at codon 528 (L528M) in the HBV polymerase gene.[6]

Although the ideal goal of therapy is to eradicate the virus, this is not always achievable because of the difficulty of eliminating the covalently closed circular form of the HBV genome from the liver and the existence of extrahepatic reservoirs of HBV. Endpoints of treatment have traditionally been clearance of HBeAg, development of anti-HBe antibodies, and undetectable serum HBV DNA using insensitive hybridization assays with detection limits of approximately 10^6 genome copies/ml. Achieving these endpoints usually is accompanied by remission of liver disease as evidenced by normalization of the ALT level and decreased histologic activity on liver biopsy. The response usually is sustained at long-term follow-up. Nevertheless, most responders continue to have detectable HBV DNA when a sensitive nucleic acid amplification test is used.

Several variations in the nucleotide sequence of HBV have important clinical consequences. An important mutation in the S gene is a glycine-to-arginine substitution at codon 145 (G145R) in the conserved "a" determinant, which causes decreased affinity of the HBsAg for anti-HBs antibodies.[7] HBV with this mutation has been found in children of HBsAg-positive mothers in whom HBV infection develops despite vaccination and an adequate anti-HBs antibody response after vaccination, as well as in liver transplant recipients who have recurrent infection despite administration of HBV immune globulin.[8,9] These immune

escape mutants have raised concern about vaccine efficacy and serologically silent infections. The G145R mutation has been reported in many countries and is responsible for 2% to 40% of vaccine failures. Although there is diminished binding to anti-HBs antibodies, the vast majority of S mutants can be readily detected with the current generation of HBsAg tests. Thus, the initial concern that widespread use of HBV immune globulin and vaccination would result in HBV mutants that would escape detection in the HBsAg test was unfounded.

Core promoter and precore mutants produce virions that do not produce HBeAg. The most common core promoter mutation has a dual change of A to T at nucleotide (nt) 1762 and G to A at nt 1764 that diminishes the amount of mRNA and hence HBeAg secretion.[10] The predominant mutation is a G-to-A change at nt 1896 (A1896), which leads to premature termination of the precore protein at codon 28, thus preventing the production of HBeAg.[11] The A1896 mutation is infrequent in North America and western Europe but is geographically widespread. This geographic variability in frequency is related to the predominant genotypes in a geographic region because the mutation is found only in genotypes B, C, D, and E.

The A1896 mutation was first reported in patients with chronic active hepatitis or fulminant hepatitis. However, the A1896 mutation also can be present in asymptomatic carriers and viruses with this mutation replicate no more efficiently than wild-type HBV. Thus, the pathophysiologic significance of this mutation is unclear.[12] However, the clinical picture of persistent HBV replication and active liver disease in HBeAg-negative patients appears to be increasingly prevalent, and in some regions the A1896-mutant virus may be more prevalent than wild-type virus.

Available Assays

HBV DNA Detection and Quantitation

The commercially available tests for quantitation of HBV DNA in serum and plasma are listed in Table 37-2. These tests employ either signal or target amplification. All of these assays are available as research-use-only kits or analyte-specific reagents, since the Food and Drug Administration (FDA) has not cleared any of these reagents for diagnostic use. In addition, a number of laboratory-developed conventional and real-time polymerase chain reaction (PCR) tests have been described.[4,13–15] All of the quantitative HBV DNA test formats have been used in monitoring the status of HBV infection before and after treatment.

The Hybrid Capture tests (Digene, Gaithersburg, MD) are liquid hybridization assays employing full-length genomic RNA probes and antibodies that bind to DNA-RNA hybrids to quantitate HBV DNA in serum and plasma specimens. The dynamic ranges of the Standard and Ultrasensitive Hybrid Capture II assays are 200,000 to 2 billion genome copies/ml and 5,000 to 60 million genome copies/ml, respectively.

The Versant HBV assay (Bayer Corp, Tarrytown, NY) uses branched DNA (bDNA) technology for HBV DNA quantitation. The virus is concentrated from 1 ml of serum by ultracentrifugation to achieve greater sensitivity. Quantitation of HBV DNA in clinical samples is determined from a standard curve generated with each run using a set of HBV DNA standards included as part of the kit. The Versant HBV assay has a dynamic range of 700,000 to 5 billion copies/ml.

The Amplicor HBV Monitor Test (Roche Diagnostics, Indiannapolis, IN) uses PCR for HBV DNA quantitation[16] and has the greatest sensitivity of all the available assays, detecting as few as 200 HBV genome copies/ml. The PCR primers are biotin labeled and target conserved sequences in the precore/core region of the HBV genome. A known number of copies of a quantitation standard (QS) are added to each specimen aliquot prior to DNA isolation. The QS is a DNA fragment with 5′ and 3′ sequences that are complementary to the HBV PCR primers with a unique internal sequence. After PCR amplification, the HBV and QS amplicons are detected using specific probes for internal sequences of the target and QS in a solid phase assay with an avidin-enzyme conjugate and a colorimetric substrate. HBV DNA in the serum sample is quantitated by reference to the QS signal. The amplification, hybridization,

Table 37-2. Commercially Available HBV DNA Tests

Test Kit	Method	Manufacturer	Gene Target	Dynamic Range*
Standard Hybrid Capture II	Liquid hybridization	Digene (Gaithersburg, MD)	Full-length RNA probe†	2×10^5 to 2×10^9
Ultra-Sensitive Hybrid Capture II	Liquid hybridization	Digene (Gaithersburg, MD)	Full-length RNA probe†	5×10^3 to 6×10^7
Versant HBV	Branched DNA	Bayer Corp (Tarrytown, NY)	Multiple probes	7×10^5 to 5×10^9
Amplicor HBV Monitor	PCR	Roche Diagnostics (Indiannapolis, IN)	Precore/core	2×10^2 to 2×10^5
TaqMan HBV Analyte-Specific Reagent	Real-time PCR	Roche Diagnostics (Indiannapolis, IN)	Precore/core	2×10^2 to 2×10^7

* Genome copies/ml.
† Genotypes A and D.

and detection steps can be performed using the COBAS Amplicor analyzer.[17]

A real-time PCR assay for HBV DNA based on TaqMan chemistry is available from Roche Diagnostics as an analyte-specific reagent (ASR). This assay combines the analytical sensitivity of conventional PCR with the broad dynamic range of the Hybrid Capture and Versant bDNA tests, and simplifies the analytical process relative to the PCR tests by combining the amplification and detection steps.

HBV Genotyping

The only commercially available test for genotyping HBV is a line probe assay (Innogenetics, Ghent, Belgium) for mutations in the HBV polymerase gene associated with resistance to lamivudine.[18] A 341 base pair (bp) PCR product from the polymerase gene is amplified with primers biotinylated at their 5′ ends. The PCR product is denatured and hybridized to probes immobilized on a nitrocellulose strip. The assay uses 38 probes to cover codon positions 528, 552, and 555 of the polymerase gene in 19 different probe lines on a nitrocellulose strip. The hybrids are visualized on the strip after addition of streptavidin-alkaline phosphatase and colorimetric substrate. The mutations are identified by the colored patterns of PCR product hybridization to the probes.

Interpretation of Test Results

The availability of clinical molecular tests for detection and quantitation of HBV DNA in serum has improved our understanding of the clinical manifestations and natural history of HBV infection and facilitated the monitoring of response to therapy. However, the use of increasingly sensitive tests for this purpose has led to new questions and dilemmas. For example, most patients with chronic HBV infection have detectable HBV DNA in serum when sensitive PCR assays are used, even in HBsAg carriers without apparent disease. This leads to the question of what serum level of HBV DNA should be used to determine the need for treatment. Likewise, because of the limitations of currently available therapies, most treatment responders continue to have HBV DNA detectable in serum when the most sensitive assays are used.[19] This raises the questions of when to stop treatment and how virologic treatment response should be defined.

Early studies using insensitive hybridization assays demonstrated that most patients who developed spontaneous or treatment-induced anti-HBeAg seroconversion with undetectable serum HBV DNA have normal ALT levels, reduced histologic activity, decreased risk of hepatic decompensation, and improved survival.[20] The majority, however, have detectable HBV DNA when PCR assays are used. Therefore, it seems that patients with low serum HBV DNA levels may not require treatment. Although there appears to be a level of serum HBV DNA below which hepatitis B is inactive and nonprogressive, this level may range from 10^4 to 10^6 genome copies/ml and may vary with the patient and the assay used to determine the levels. There is poor agreement between HBV DNA results generated with the different assays.[15,21]

Laboratory Issues

The major laboratory issue in HBV DNA quantitation is the widely divergent results obtained with the different assays. To address interassay differences, the World Health Organization (WHO) developed an international standard for HBV DNA, designated 97/746.[22] The WHO standard is a lyophilized plasma specimen that has been analyzed by several laboratories using different nucleic acid tests and has an assigned potency of 10^6 international units (IU)/ml. In the future, all new assays for HBV DNA should be calibrated against the WHO standard and the results reported in IU/ml. In the interim, laboratories should be aware of the substantial interassay differences in quantitative results.

The lack of FDA-cleared test kits for HBV DNA places increased burden on laboratories to verify the test performance characteristics. Unfortunately, there are no formal proficiency surveys for HBV nucleic acid tests, although control and reference materials are available from several companies.

Serum and plasma, with EDTA or citrate dextrose as an anticoagulant, are acceptable specimens for most nucleic acid tests for HBV. In the absence of stability data, samples for nucleic acid testing should be processed to separate blood cells from the plasma or serum within 6 hours of collection and either tested within 24 hours or stored at –70°C. In one study HBV DNA was stable in separated serum samples for at least 5 days when specimens were stored at 4°C.[23]

Future Directions

As more antiviral agents for treatment of HBV are developed, the key surrogate endpoints that will be used are suppression of serum HBV DNA levels and prevention of the development of antiviral resistance. Long-term suppression of viral replication is associated with both biochemical and histological improvement in liver disease. These factors make the development, clinical evaluation, and availability of reliable, standardized quantitative tests for HBV DNA a major priority for the near future. Similarly, increased emphasis will be placed on characterization of drug resistance mutants and HBV genotyping to assist with clinical management of patients as new drugs are developed.

HEPATITIS C VIRUS

HCV is an RNA virus with a positive-sense, single-stranded genome of approximately 9500 nt encoding a single polyprotein of about 3000 amino acids. The long open reading frame is flanked at each end by a short untranslated region (UTR). The genome structure is most similar to viruses of the family Flaviviridae, which includes many of the arthropod-borne viruses. As in other flaviviruses, the three N-terminal proteins of HCV (core, envelope 1, and envelope 2) are probably structural and the four C-terminal proteins (nonstructural 2, 3, 4, and 5) are thought to function in viral replication.

The 5′ UTR is a highly conserved region of 341 nt and has a complex secondary structure. It contains an internal ribosome entry site and presumably is important in the translation of the long open reading frame. The 3′ UTR contains a short region that varies in sequence and length, followed by a polypyrimidine stretch of variable length, and finally a highly conserved sequence of 98 nt, which constitutes the terminus of the genome. The function of the 3′ UTR is not known but is thought to be essential for viral replication.

The envelope 1 (E1) and 2 (E2) regions of HCV are the most variable regions within the genome at both the nucleotide and amino acid levels. Two regions in the E2, called hypervariable regions 1 and 2 (HVR1 and HRV2), show extreme sequence variability, which is thought to result from selective pressure by antiviral antibodies. E2 also contains the binding site for CD81, the putative HCV receptor or coreceptor.

The nonstructural regions 2 (NS2) and 3 (NS3) contain a Zn-dependent autoprotease that cleaves the polyprotein at the NS2-NS3 junction. The aminoterminal portion of the NS3 protein also is a serine protease that cleaves the polyprotein at several sites. The carboxyterminal portion of the NS3 protein has helicase activity, which is important for HCV replication. The NS4A protein is a cofactor for NS3 serine protease. The NS5B region encodes the RNA-dependent RNA polymerase, which replicates the viral genome. A region in NS5A has been linked to response to IFNα and is therefore called the IFNα-sensitivity determining region (ISDR).

HCV Genotypes

The first complete HCV genome sequence was reported by Choo et al. in 1991.[24] As additional genome sequences from isolates from different parts of the world were determined and compared, it was evident that HCV exists as distinct genotypes with as much as 35% sequence diversity over the whole viral genome.[25] Much of the early literature on genotyping is confusing because investigators developed and used their own classification schemes. However, a consensus nomenclature system was developed in 1994. In this system the genotypes are numbered using Arabic numerals in order of their discovery, and the more closely related strains within some types are designated as subtypes with lowercase letters. The complex of genetic variants found within an individual isolate is termed the "quasispecies." The quasispecies result from the accumulation of mutations that occur during viral replication in the host. The terminology and degree of nucleotide similarity that define the relationships of HCV variants are given in Table 37-3.

Table 37-3. Terms Describing Genomic Heterogeneity of HCV

Term	Definition	% Nucleotide Similarity*
Genotype	Heterogeneity among different viruses	66–69
Subtype	Closely related viruses within each genotype	77–80
Quasispecies	Complex of genetic variants within individual viruses	91–99

Source: Adapted from Zein NN. Clinical significance of hepatitis C virus genotypes. *Clin Microbiol Rev.* 2000;13:223–235.
*Full-length genome sequence identity.

Six major HCV genotypes have been identified. Sequence analysis of the E1 region suggested that HCV could be grouped into six major genotypes and 12 subtypes.[26] The same investigators sequenced 573 nt of the core region of the same isolates to confirm this classification scheme.[27] Simmonds et al.[28] also were able to classify HCV isolates into the same six major genotypes and numerous subtypes using sequence analysis of the NS5B region. Analyses of full-length open reading frame sequences have confirmed the original classification scheme based on analyses of individual gene regions.[29]

Genome sequence analysis of HCV isolates from Southeast Asia have led some authors to propose new major genotypes 7, 8, 9, 10, and 11.[30,31] However, other investigators suggested that these variants could be classified within the six major genotypes originally described.[32] Under this scheme, genotype 10 is a divergent subtype of genotype 3, and genotypes 7, 8, 9, and 11 are divergent subtypes of genotype 6.

HCV genotypes 1, 2, and 3 are found throughout the world, but there are clear differences in their distribution.[33] HCV subtypes 1a, 1b, 2a, 2b, 2c, and 3a are responsible for more than 90% of infections in North and South America, Europe, and Japan. In the United States, type 1 accounts for approximately 70% of the infections with equal distribution between subtypes 1a and 1b. In Japan, subtype 1b causes more than 70% of HCV infections. Although subtypes 2a and 2b have wide distributions in North America, Europe, and Japan, subtype 2c is widespread in a region of northern Italy. HCV subtype 3a is common among intravenous drug users in the United States and Europe. Other subtypes of genotype 3 are common in Nepal, Bangladesh, India, and Pakistan.

Genotype 4 is prevalent in North Africa and the Middle East, and genotypes 5 and 6 are limited to South Africa

and Hong Kong, respectively.[28] Subtype 4a constitutes the majority of infections in Egypt, and this and other subtypes of genotype 4 are found in Zaire and Gabon. Subtype 5a is a particular problem in South Africa, where some reports indicate that it is responsible for more than 50% of infections. Subtype 6a infections are common in Hong Kong. The isolates classified as genotypes 7, 8, and 9 by some investigators have been found only in Vietnamese patients.[31] Putative genotypes 10 and 11 have been identified only in patients from Indonesia.[30]

The retrospective nature of most of the studies makes it difficult to determine the role of genotype as a risk factor for disease progression and to separate it from other known risk factors, such as older age at infection, male gender, alcohol consumption, and concurrent viral infection. However, in two prospective studies, viral genotype did not correlate with disease progression.[34,35]

Clinical Utility

HCV RNA Detection and Quantitation

Detection of HCV RNA in serum or plasma by nucleic acid amplification methods is important for confirming the diagnosis of HCV, distinguishing active from resolved infection, assessing the virologic response to therapy, and screening the blood supply. These tests are incorporated into diagnostic algorithms for HCV proposed by the National Institutes of Health,[36] the CDC,[37] European Association of the Study of the Liver,[38] and National Academy of Clinical Biochemistry.[39]

The detection of HCV RNA in the plasma or serum is the earliest marker of infection, appearing 1 to 2 weeks after infection and weeks before the appearance of alterations in liver enzyme levels or anti-HCV antibodies. Approximately 80% of individuals infected with HCV will be chronically infected with the virus. In antibody-positive individuals, HCV RNA tests can distinguish active from resolved infections. In patients with a high pretest probability of infection, a positive serological screening test is usually "confirmed" with a qualitative test for HCV RNA rather than the supplemental recombinant immunoblot assay (RIBA). This strategy is cost-effective and more informative than using the RIBA to confirm positive antibody screening tests in a diagnostic setting. It is important to remember that the RIBA confirms the presence of antibody, whereas an HCV RNA test detects the presence of the virus.

The CDC recently developed guidelines for HCV antibody testing in which the signal to cutoff (s/co) ratio of the serological screening test can be used effectively to determine which screening test results need supplemental testing for confirmation.[40] Screening enzyme-linked immunoassay results with s/co ratios ≥3.8 predict a RIBA-positive result >95% of the time in all populations tested. Consequently, supplemental testing is required only for those screening tests with low s/co ratios. Either a RIBA or an HCV RNA test can be used for confirmation, but all negative HCV RNA tests should be followed by a RIBA since a single negative RNA test does not rule out infection. Confirmation of an HCV antibody screening test is important to ensure that further clinical evaluation is limited to those patients who are truly antibody positive and to reduce psychological stress on patients who test falsely positive for antibody.

HCV RNA testing also is helpful for the diagnosis of infection in infants born to HCV-infected mothers, due to persistence of maternal antibody, and in immunocompromised or debilitated patients who may have blunted serological responses. A qualitative HCV RNA test also should be used for patients suspected of having an acute infection and in patients with hepatitis of no identifiable cause.

Qualitative HCV RNA tests are the most reliable means of identifying patients with active HCV infection. A negative HCV RNA test in a serologically positive individual may indicate that the infection has resolved or that the viremia is intermittent. Up to 15% of chronically infected individuals may have intermittent viremia and, as a result, a single negative HCV RNA determination may not be sufficient to exclude active infection when the index of clinical suspicion is high.[41] In these individuals a second specimen should be collected and tested.

IFN-based regimens, either IFN in combination with ribavirin or IFN monotherapy for those patients who cannot tolerate ribavirin, are the standard treatment for patients with chronic HCV infection. Because these regimens produce sustained virological responses in only about half of the patients treated, are of long duration (24 to 48 weeks), and are associated with significant adverse events, early identification of patients who are likely or unlikely to respond to treatment is desirable.

Early clearance of HCV RNA as assessed with a sensitive assay or a rapid decline in HCV RNA levels during the early treatment period is predictive of sustained virological response among patients receiving either standard or pegylated IFN alone, and pegylated IFN plus ribavirin.[42,43] Patients who have detectable or less than a 2 $\log_{10}$ decrease in viral RNA levels after 12 weeks of treatment have a minimal chance of achieving a sustained virological response (negative predictive value of 98%).

In contrast, early viral kinetics are less predictive of sustained virological response among patients treated with standard IFN and ribavirin because late responses are seen in some patients. Thus, discontinuing therapy in patients based on a positive HCV RNA test early in treatment would deny the treatment to a substantial number of patients who would achieve a sustained response. The virological response to combination therapy with IFN and ribavirin should be assessed after 24 weeks of therapy with a sensitive (≤50 IU/ml) qualitative assay for HCV RNA.[44] Individuals with a positive qualitative RNA test after week 24 of therapy are considered treatment failure and therapy is discontinued. Therapy also can be safely stopped at 24 weeks in

patients infected with HCV genotypes 2 and 3 who have a negative HCV RNA test, because there is no benefit to longer therapeutic regimens. However, an additional 24 weeks of therapy is recommended for patients infected with other genotypes, even with a negative qualitative HCV RNA test. In addition, a qualitative HCV RNA test should be performed 24 weeks after completion of treatment in all patients to establish if a sustained virological response was achieved regardless of the treatment regimen.

The use of anti-HCV antibody tests to screen the blood supply has dramatically reduced the risk of transfusion-associated HCV infection in developed countries. The risk in the United States from blood that is negative for anti-HCV antibodies is less than 1 in 103,000 transfused units.[45] To drive the risk of infection even lower, blood donor pools are currently tested for the presence of HCV RNA.[46] The serologic screening tests for HCV have a 70-day window period of seronegativity, and antigen detection tests are not yet available for blood bank screening. It is estimated that HCV RNA testing will reduce the detection window by 25 days and reduce the number of infectious units from 116 to 32 per year.[47]

An enzyme immunoassay (EIA) for the detection and quantitation of total HCV core antigen in serum has recently been developed and is available as a research-use-only test (Ortho Trak C, Ortho Clinical Diagnostics, Raritan, NJ). This test significantly shortens the serologically silent window period using seroconversion panels, and its performance correlates closely with RNA detection tests in blood donors.[48,49] However, the analytical sensitivity is less than most RNA tests, at approximately 20,000 IU/ml. The analytical sensitivity of the core antigen test is too high to be used in the monitoring of late events during and after treatment.[50] Antigen detection may represent a cost-effective alternative to HCV RNA testing to distinguish active from resolved infections in resource-poor settings.

HCV viral load testing is useful in pretreatment evaluations of patients being considered for therapy, since a viral load of less than 2×10^6 copies/ml (800,000 IU/ml) is one of several predictors of achieving a sustained virological response.[51,52] Other factors associated with achieving a sustained response to therapy include the absence of cirrhosis, age less than 40 years, female gender, white race, and viral genotype other than 1. Viral load testing also can be used in an early assessment of viral kinetics in patients treated with IFN alone or with pegylated IFN plus ribavirin.[42,43] Patients who fail to achieve less than a 2 $\log_{10}$ decline in viral load with treatment have little chance of achieving a sustained virological response to these therapies.

HCV viral load does not predict disease progression and is not associated with severity of liver disease.[53] This is in sharp contrast to HIV-1, in which the viral load is the principal factor determining the rate of disease progression. Monitoring HCV viral load in untreated patients is not warranted and should be discouraged.

HCV Genotyping

Sequence analysis of variable regions of the HCV genome has been used to investigate outbreaks of infection and to study modes of transmission. Two large outbreaks of infection associated with contaminated lots of anti-rhesus D immunoglobulin (anti-D) in Ireland and Germany were investigated using molecular typing.[54,55] In both studies, sequence analysis showed that the virus infecting the women was the same as that found in the implicated batches of anti-D. In another report, sequencing part of the NS3 region provided evidence of patient-to-patient transmission during colonoscopy.[56] Sequence analysis also is becoming a routine part of investigations of HCV infections associated with blood transfusions. In addition, molecular analysis has been used to study vertical and sexual transmission of HCV.[57–60]

Although a number of baseline factors are predictive of response to treatment of chronic hepatitis C infection, HCV genotype is the strongest and most consistent predictor for achieving a sustained virological response. In the large clinical trials of combination therapy with IFN and ribavirin, only 30% of patients infected with genotype 1 had a sustained response compared to 65% of patients infected with genotypes 2 or 3.[51,52] Too few patients infected with genotypes 4, 5, and 6 were included in the clinical trials to adequately assess the likelihood of therapeutic response with these genotypes.

In practice, HCV genotype can be used in tailoring the duration of therapy to individual patients. In an algorithm proposed by Poynard et al.,[44] the five independent predictors of sustained virological response are considered at the end of 24 weeks of combination therapy. The predictors of sustained response are HCV genotype other than 1, viral load of <2,000,000 copies/ml, age <40 years, female gender, and no or only portal fibrosis on liver biopsy. If HCV RNA is still detectable in the serum at the end of 24 weeks of therapy, then the patient is unlikely to benefit from an additional 24 weeks of IFN and ribavirin, and therapy can be stopped. However, if the HCV RNA is undetectable after 24 weeks of therapy and the patient has fewer than four favorable factors, then therapy should be continued for an additional 24 weeks since these patients may obtain benefit from additional therapy. Conversely, therapy can be safely stopped if the patient has four or five favorable factors and is HCV RNA undetectable at 24 weeks. This algorithm ensures that only those patients who may benefit from 48 weeks of therapy have to endure it.

There is considerable interest in identifying isolates within a genotype that exhibit different responses to IFN. For example, Enomoto et al.[61] found that genotype 1b isolates in Japan identical to the HCV-J prototype strain within the carboxyterminal part of the NS5A protein were resistant to IFN. In contrast, patients infected with genotype 1b strains with mutations within this region had a better response to IFN. This region of amino acids 2209 to 2248 in the NS5A protein has been termed the interferon-

sensitivity-determining region (ISDR). The same group recently confirmed and extended their observations. They found that sustained response rates correlated with the number of mutations in the ISDR with a larger group of patients and identified the mutation sites within the ISDR that were significantly associated with sustained response.[62] Although these results have been confirmed for Japanese isolates, in studies of patients from other countries the majority of IFN-sensitive isolates showed no or only a few mutations in the ISDR.[5,63,64] The reasons for the different findings are not clear but may involve racial, virological, or other factors.

Available Assays

A variety of test kits for detection or quantitation of HCV RNA are available commercially (Table 37-4). These assays are based on traditional reverse transcription–polymerase chain reaction (RT-PCR), transcription-mediated amplification (TMA), bDNA, or, most recently, real-time RT-PCR technologies.

Qualitative Detection HCV Assays

There are currently two FDA-cleared qualitative HCV RNA test kits available for diagnostic use, the Amplicor HCV test v2.0 (Roche) and the Versant HCV RNA qualitative test (Bayer).

The Amplicor HCV test v2.0 is based on RT-PCR amplification of a portion of the 5′ UTR and has an analytical sensitivity of 50 IU/ml.[65] The test incorporates an internal control to detect PCR inhibitors and deoxyuridine triphosphate (dUTP) and uracil-N-glycosylase in the reaction mixture to prevent false positives due to amplicon carryover. The test is available in two formats, a manual microwell plate assay and a semiautomated assay designed for the COBAS instrument. The performance characteristics of the Amplicor HCV tests are well established[66] and are the most commonly used qualitative HCV RNA tests in clinical laboratories participating in the proficiency-testing surveys of the College of American Pathologists.

The Versant HCV RNA qualitative test also targets the 5′ UTR but uses TMA to amplify the targeted region. The analytical sensitivity of the Versant HCV RNA assay is 5 IU/ml. The increase in analytical sensitivity over the Amplicor tests is due in part to the larger sample volume (500 µl versus 200 µl) and the use of a specific target capture step to isolate HCV RNA rather than total RNA precipitation. This test also employs an internal control RNA to detect the presence of amplification inhibitors.

The Versant HCV RNA assay can detect residual serum HCV RNA in some patients with no detectable HCV RNA as determined by the Amplicor v2.0 assay at the end of treatment with IFN and subsequently experienced virological relapse.[67,68] However, this difference was not observed with end-of-treatment samples from patients treated with pegylated IFN.[69] The two qualitative HCV RNA assays demonstrated excellent concordance in a study designed to

Table 37-4. Commercially Available HCV RNA Tests

Test Kit	Method	Manufacturer	Lower Limit of Detection (IU/ml)	Dynamic Range	Clinical Application
Qualitative					
Amplicor HCV Test v2.0*	RT-PCR	Roche Diagnostics (Indiannapolis, IN)	50	NA	Diagnose active infection and assess response to therapy
Versant HCV RNA Assay*	TMA	Bayer Corp, (Tarrytown, NY)	5	NA	Diagnose active infection and assess response to therapy
Ampliscreen HCV Test v2.0*	RT-PCR	Roche Diagnostics (Indiannapolis, IN)	<50	NA	Blood screening
Procleix HIV-1/ HCV Assay*	TMA	GenProbe, San (Diego, CA)	<50	NA	Blood screening
Quantitative					
Amplicor HCV Monitor Test v2.0	RT-PCR	Roche Diagnostics (Indiannapolis, IN)	600	2.9 $\log_{10}$	Determine duration of therapy and early prediction of therapy failure
Versant HCV RNA Assay v3.0*	bDNA	Bayer Corp (Tarrytown, NY)	615	4.1 $\log_{10}$	Determine duration of therapy and early prediction of therapy failure
TaqMan HCV Analyte Specific Reagent	Real-time RT-PCR	Roche Diagnostics (Indiannapolis, IN) Abbott Molecular (Des Plaines, IL)	10	6.0 $\log_{10}$	Diagnose active infection, assess response to therapy, determine duration of therapy, and early prediction of therapy failure

*FDA-cleared tests.
NA, not applicable.

compare the performance characteristics with specimens submitted for HCV diagnosis.[70] The difference in analytical sensitivity between the two tests does not result in any meaningful difference in clinical sensitivity when the tests are used diagnostically, because it is rare for patients with chronic hepatitis C infection to present for initial evaluation with viral loads of less than 10^5 copies/ml.

The Procleix HIV-1/HCV test (GenProbe, San Diego, CA) also uses TMA technology and is approved by the FDA for the screening of blood products.[71] The Ampliscreen HCV RNA test v2.0 is an RT-PCR assay (Roche) designed for blood screening that is based on the Amplicor HCV test. Both assays have analytical sensitivities of less than 50 IU/ml.

Quantitative HCV Assays

Currently there are several commercially available and proprietary laboratory-developed methods used to quantitate HCV RNA levels in patients. The Versant HCV RNA Assay v3.0 (Bayer Corp.),[72] which is FDA approved, and the Amplicor HCV Monitor Test v2.0 (Roche Diagnostics)[65] are the most widely used commercially available assays. The HCV Superquant is a proprietary RT-PCR assay developed by National Genetics Institute (Los Angeles, CA), which has been used by many investigators in clinical trials of treatment for HCV.

The Amplicor HCV Monitor test v2.0 is a quantitative RT-PCR assay that amplifies the same target region as the qualitative Amplicor HCV test. The assay uses an internal quantitation standard to calculate the amount of HCV RNA in a sample. The assay has a sensitivity of 600 IU/ml and a 2.9 $\log_{10}$ dynamic range. Specimens with values greater than the upper limit of quantitation (500,000 IU/ml) can be diluted 100-fold and retested as recommended by the manufacturer. The Amplicor HCV Monitor test is available in both manual microwell plate and semiautomated COBAS instrument formats.

The Versant HCV RNA Assay v3.0 is based on bDNA technology. The sensitivity of the Versant HCV RNA Assay v3.0 is 615 IU/ml, with a 4.1 $\log_{10}$ dynamic range. A number of changes to the design of the v3.0 test improved both the sensitivity and specificity over the v2.0 test. These include an increased number of capture probes, improved probe design, the use of nonnatural synthetic nucleotides in detection probes, and redesigned label extenders. The System 340 bDNA analyzer automates all incubations, wash steps, readings, and data analysis. The instrument can process two 96-microwell plates per run.

Overall, the results of the Versant HCV RNA Assay v3.0 and the Amplicor HCV Monitor Test v2.0 have been found to be in substantial agreement.[72,73] The results of the Amplicor Monitor test are reported as IU/ml and those of the Versant assay are reported as either copies/ml or IU/ml. Both tests are free of significant HCV genotype bias and have similar analytical sensitivities. The Versant results tend to be more precise than the Amplicor Monitor results, but the Amplicor Monitor test is better able to discriminate low-positive from negative specimens. The Versant assay has a much greater dynamic range than the Amplicor Monitor test; however, a pretest sample dilution permits quantification of high viral load specimens by the Amplicor Monitor test. The throughput and level of automation are much greater with the Versant assay than with the Amplicor Monitor test.

A number of TaqMan RT-PCR assays for detection and quantitation for HCV RNA have been described.[74] These tests have sensitivities comparable to qualitative tests, have a broad dynamic range, and provide precise quantitation of viral load. These tests also generate results more rapidly than conventional RT-PCR assays and are not prone to amplicon carryover contamination since the amplification and detection steps are combined in a single closed tube. With the development of real-time RT-PCR, clinical laboratories will no longer need to have separate qualitative and quantitative assay formats for HCV RNA to achieve both high sensitivity and broad dynamic range. Roche and Abbott manufacture TaqMan HCV analyte-specific reagents that contain the necessary HCV-specific primers, dual-labeled fluorescent HCV and HCV quantitation standard specific probes, enzyme, and dNTPs for the detection and quantitation of HCV RNA.[75,76]

HCV Genotyping Assays

A variety of laboratory-developed and commercial assays are used for HCV genotyping. The methods include nucleic acid sequencing, reverse hybridization, subtype-specific PCR, DNA fragment length polymorphism, heteroduplex mobility analysis, melting curve analysis and serological genotyping. The FDA has not cleared any of these methods for clinical diagnostic use.

A commercially available reverse hybridization line probe assay is the most commonly used method for genotyping HCV among clinical laboratories participating in the HCV proficiency-testing surveys of the College of American Pathologists. This reverse hybridization assay was developed by Innogenetics to genotype HCV and is now marketed as the Versant HCV Genotype Assay by Bayer. In this line probe assay (LiPA), biotinylated PCR products from the 5′ UTR are hybridized under stringent conditions with 19 type- and subtype-specific oligonucleotide probes attached to a nitrocellulose strip. Hybridized PCR products are detected with a streptavidin-alkaline phosphatase conjugate. The second-generation assay discriminates among genotypes 1a, 1b, 2a/c, 2b, 3a, 3b, 3c, 4a–h, 5a, and 6a.[77] The results from the Versant HCV Genotype LiPA Assay correlate well with results obtained by direct sequencing assays of the 5′ UTR and other genes in published evaluations, but may not distinguish between genotypes 1a and 1b in 5% to 10% of cases, and does not distinguish between genotypes 2a and 2c.[32,59,78,79] The LiPA is the most common method

used in clinical laboratories for HCV genotyping because it can be used with amplicons from both the qualitative and quantitative Amplicor HCV tests, and is easy to perform and interpret. Mixed genotype infections are easily recognized as unusual patterns of hybridization with the typing probes. However, the LiPA requires a considerable amount of amplicon for typing, and the assay may regularly fail when the viral load is less than 10^4 copies/ml.

Sequence analysis of amplified subgenomic sequences is the most definitive way to genotype HCV strains. Genotyping schemes based on sequencing variable genes such as E1, C, and NS5B provide enough resolution to determine types and subtypes.[26–28] The 5′ UTR is too highly conserved to discriminate all subtypes reliably.[59] Genotyping methods targeting highly variable regions have higher failure rates due to primer mismatches and failed amplification reactions. Sequencing reactions can be performed directly on PCR products or on cloned amplicons. Mixed infections with multiple genotypes may be missed with direct sequence analysis. Definitive detection of mixed infections requires analysis of a large number of clones. Cloning may, however, emphasize artifactual nucleotide substitutions introduced by the DNA polymerase during amplification or by selection during the cloning procedure,[80] and is generally not practical for the clinical laboratory.

A standardized direct sequencing system has been recently developed for clinical use by Visible Genetics (Suwanee, GA) and now marketed by Bayer. The Trugene HCV 5′NC genotyping kit targets the 5′ UTR (nt 96 to 282) and employs proprietary single-tube chemistry that is robust and highly sensitive. This method can be used with the 244 bp amplicon generated by either the Roche Amplicor HCV or Amplicor HCV Monitor tests as the sequencing template after a column purification step.[81] The sequencing chemistry produces bidirectional sequences. The software acquires the sequence data in real time, and each pair of forward and reverse sequences is combined. A reference sequence library module contains approximately 200 sequences from the six major genotypes and 24 subtypes of HCV. The software automatically aligns the patient HCV sequence with the reference sequences in the library and reports type, subtype, and closest isolate determinations. The Trugene HCV 5′NC genotyping system is a rapid and reliable method for determining HCV genotypes but, like all approaches targeting the conserved 5′ UTR, cannot reliably distinguish all HCV subtypes.[36,48,81,82]

The practice of using sequence analysis of a single subgenomic region for HCV genotyping has recently been challenged by the description of a naturally occurring intergenotypic recombinant of two HCV genotypes.[83] This virus was found in patients in St Petersburg, Russia, and was assigned to two different genotypes, 2 and 1, by sequence analysis of the 5′ UTR and the NS5B region, respectively. The crossover point for the genome was mapped within the NS2 region. The extent to which such viruses occur in other patient populations is largely unknown.

A DNA EIA (Sorin Biomedica, Saluggia, Italy) for HCV genotyping is based on hybridization of denatured amplicon from the core region to genotype-specific probes that are bound to the wells of a microtiter plate. Mouse monoclonal antibodies to double-stranded DNA are used to detect the hybrids. The results of the DNA EIA were highly concordant with the results of other genotyping methods in two evaluations.[79,81]

A variety of laboratory-developed methods have been used to genotype HCV, including subtype-specific PCR,[25] primer-specific and mispair extension analysis,[84] nested restriction site–specific PCR,[85] restriction fragment length polymorphism,[86] heteroduplex mobility analysis,[87] and melting-curve analysis with fluorescence resonance energy transfer probes.[88,89]

Genotype-specific antibodies directed against the immunodominant epitopes in NS4 have been used to develop serotyping or serological genotyping tests. Two serological genotyping tests are commercially available. An NS4 recombinant immunoblot assay (Chiron Corporation, Emeryville, CA) uses synthetic proteins from the NS4 and core regions to discriminate among HCV genotypes 1, 2, and 3.[90] The competition enzyme-linked immunoassay (Murex Diagnostics Ltd, Dartford, UK) uses eight branched synthetic peptides to detect genotype-specific anti-NS4 antibodies to discriminate among HCV genotypes 1 to 6.[14] The high degree of cross reactivity among genotypes to these synthetic peptides necessitates absorbing the cross-reacting antibodies with an excess of heterologous peptides in solution prior to use. Both of these serologic assays lack sensitivity and specificity as compared with direct sequencing or the Versant HCV Genotype LiPA Assay for HCV genotyping; however serological genotyping is inexpensive, is simple to perform, and lends itself well to large epidemiological studies. It is the only way to determine the genotype of a virus in patients with low-level viremia or who have cleared their infection, as well as for specimens in which the RNA has been destroyed by improper handling.

Interpretation of Test Results

The presence of HCV RNA in serum or plasma defines active infection, and HCV RNA is usually detectable within the first week after exposure. However, a single negative HCV RNA test result does not exclude the possibility of active infection because viremia may be intermittent in some chronically infected patients. An HCV RNA test may be the only evidence of infection in individuals with false-negative antibody tests. False-negative HCV antibody tests can occur in HIV-1 infected individuals, patients undergoing hemodialysis, and patients with HCV-associated essential mixed cryoglobulinemia.[91–93]

In addition to HCV diagnosis, qualitative HCV RNA tests are used to assess virological response to therapy. A negative HCV RNA test at the completion of therapy defines an

end-of-treatment response (ETR), and a negative test 6 months after the completion of therapy defines a sustained virological response (SVR).

HCV RNA quantitation is useful in planning duration of therapy and in predicting the likelihood of response to treatment. Patients with high HCV RNA levels tend to respond less well to IFN and ribavirin, but lengthening the course of therapy from 24 to 48 weeks more than doubles the response rate. Patients with viral loads greater than 800,000 IU/ml are considered to have high HCV RNA levels. HCV viral load does not predict disease progression, is not correlated with disease severity, and, consequently, should not be routinely monitored in untreated patients.

Viral load testing can be used in an early assessment of viral kinetics in patients undergoing treatment with IFN and ribavirin. Patients who fail to achieve at least a 2 $\log_{10}$ decline in viral load after 12 weeks of treatment have little chance of an SVR and therapy should be discontinued.

Viral genotyping, like viral load, helps predict the outcome of therapy and determine its duration. Currently, the only clinically relevant distinction for patients undergoing therapy is between genotype 1 and other genotypes. Patients with a genotype 1 infection require higher doses of ribavirin and treatment for 48 weeks rather than 24 weeks to increase the likelihood of an SVR. There is no proven association between genotype and disease progression or severity, so genotyping should be reserved for those patients being considered for treatment. Routine determination of HCV subtypes (e.g., 1a or 1b), other than for epidemiological purposes, is not warranted.

In summary, a qualitative HCV test is used for the initial evaluation of HCV antibody-positive individuals to distinguish active from resolved infections. Before initiating treatment, HCV viral load and genotype tests are performed to determine the dosing and length of treatment. After 12 weeks of treatment, the HCV viral load is measured to confirm at least a 2 $\log_{10}$ decline in viral load, which will be used to decide whether to continue or terminate treatment. At the end of either 24 or 48 weeks of treatment, a negative qualitative HCV test defines an ETR, and a negative qualitative test 6 months after the end of treatment defines an SVR. Assays with a limit of detection of ≤50 IU/ml should be used to define ETR and SVR to avoid misclassifying therapeutic outcomes. For patients not achieving an SVR, there is little experience with retreatment.

Laboratory Issues

The use of internal controls for a test increases the confidence in negative HCV RNA test results and eliminates concerns that the test is falsely negative because of the presence of amplification inhibitors in the sample or poor recovery of the viral RNA. In a large clinical trial of the Amplicor HCV test v2.0, the failure rate for the internal control was only 1.1%.[66] False-negative results may occur with PCR-based tests if blood samples are collected through central lines due to contamination with heparin. False-negative HCV RNA results also may occur if the specimen was not processed or stored appropriately due to the lability of the viral RNA. Serum or plasma should be separated from the cellular components of blood within 6 hours of collection to avoid significant loss of HCV RNA. Once separated, the sample is stable for 3 days at 4°C. Storage for longer periods should be at −70°C.

False-positive HCV RNA tests can result from contamination of the sample with HCV target RNA or amplicon. The widely used Amplicor HCV test employs the uracil-N-glycolase/dUTP protocol and an equivocal zone for low-level results to limit the number of false-positive test results. Real-time RT-PCR tests for HCV RNA eliminate the risk of amplicon carryover contamination because both amplification and detection are accomplished in a sealed reaction tube.

Much has been written about the lack of agreement and the genotype bias of early commercially available versions of the HCV quantitative assays. The development the WHO First International HCV RNA standard and its acceptance by the manufacturers of these assays as a calibrator was a significant advance in HCV RNA quantitation.[94] As a result, viral load values obtained with current versions of the Amplicor Monitor and Versant quantitative tests are in much better agreement and are largely interchangeable. The remaining disagreement probably results from differences in the dynamic ranges of the assays. The agreement between results for high viral load samples can be improved if the samples are diluted 1 : 100 prior to testing with the Amplicor Monitor assay.

Both the current versions of the Amplicor Monitor and Versant quantitative tests are free of significant genotype bias due to improvements in assay design. The Versant test is more amenable to high-volume testing and is less labor-intensive than the Amplicor Monitor test, while the Amplicor Monitor test may have advantages for laboratories with a lower test volume. Recently a sample preparation instrument, the COBAS Ampliprep, was developed to automate the cumbersome manual sample preparation method used for the Amplicor Monitor test. The automated sample preparation protocol compared favorably to the manual protocol for use with the COBAS Amplicor HCV Monitor test v2.0 in a recent clinical evaluation.[95]

Unfortunately, there is a trade-off between sensitivity and dynamic range for most of the HCV RNA tests currently used in clinical laboratories. As a result, most laboratories have separate qualitative and quantitative tests. The analytical sensitivities of the Amplicor Monitor and Versant quantitative tests are inadequate for assessing ETR or for reliable diagnostic use for some untreated chronically infected patients with low viral loads. Maintaining multiple assay formats for the same virus is costly for the laboratory and often confusing to the ordering physician. The best hope for unifying the HCV RNA assay platforms lies in the real-time PCR methods, which combine the sensitivity required for diagnosis and assessment of treatment

response with the broad dynamic range required for viral load determinations.

Genotyping schemes based on variable subgenomic regions such as E1, core, and NS5B provide enough resolution to reliably determine HCV types and subtypes; however, the 5′ UTR is too conserved for accurate discrimination of all subtypes.[59] Nevertheless, the conserved nature of this region makes it the preferred target for pangenomic HCV RNA detection tests, and sequence analysis of amplicons from these tests is an efficient way to genotype HCV in a clinical laboratory since both tests can be completed with the product from a single amplification reaction. The genotyping results obtained using 5′ UTR amplicons are highly accurate at the genotype level but are not accurate for identifying different subtypes.[48,82,96]

The widespread use of tests not cleared by the FDA for HCV RNA quantitation and genotyping has placed an increased burden on clinical laboratories to verify the performance characteristics of these tests prior to clinical use. When validating HCV tests, laboratories should take advantage of the published evaluations and commercially available panels for HCV RNA quantitation and genotyping to streamline the verification process.

The College of American Pathologists has a well-established proficiency-testing program for laboratories performing tests for detection, quantitation, and characterization of HCV RNA. These surveys have shown a steady improvement in the performance of laboratories over time that probably reflects progress in both the available technologies and laboratory practices.

Future Directions

Although the incidence of new HCV infections in the United States is declining, the number of individuals infected for more than 20 years who are at risk for serious complications is expected to increase until about the year 2015. In the 15 years since the discovery of HCV, the major route of transmission through blood transfusion has all but been eliminated, improvements in therapy have resulted in better response rates, and molecular tests have proven invaluable in the diagnosis and management of patients with HCV. As new therapies evolve and laboratory assays change, the clinical relevance and use of laboratory tests will evolve.

References

1. Blumberg BS, Alter HJ, Visnich S. A "new" antigen in leukemia sera. *JAMA*. 1965;191:541–546.
2. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A non-B viral hepatitis genome. *Science*. 1989;224:359–362.
3. Lok AS, Heathcote EJ, Hoofnagle JH. Management of hepatitis B: 2000—summary of a workshop. *Gastroenterology*. 2001;120:1828–1853.
4. Dienstag JL, Schiff ER, Wright TL, et al. Lamivudine as initial treatment for chronic hepatitis B in the United States. *N Eng J Med*. 1999;341:1256–1263.
5. Duverlie G, Khorsi H, Castelain S, et al. Sequence analysis of the NS5A protein of European hepatitis C virus 1b isolates and relation to interferon sensitivity. *J Gen Virol*. 1998;79:1373–1381.
6. Allen MI, Deslauriers M, Andrews CW, et al. Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. Lamivudine Clinical Investigation Group. *Hepatology*. 1998;27:1670–1677.
7. Carman WF, Zanetti AR, Karayiannis P, et al. Vaccine-induced escape mutant of hepatitis B virus. *Lancet*. 1990;336:325–329.
8. Hsu HY, Chang MH, Ni YH, Lin HH, Wang SM, Chen DS. Surface gene mutants of hepatitis B virus in infants who develop acute or chronic infections despite immunoprophylaxis. *Hepatology*. 1997;26:786–791.
9. McMahon G, Ehrlich PH, Moustafa ZA, et al. Genetic alterations in the gene encoding the major HBsAg: DNA and immunological analysis of recurrent HBsAg derived from monoclonal antibody-treated liver transplant patients. *Hepatology*. 1992;15:757–766.
10. Okamoto H, Tsuda F, Akahane Y, et al. Hepatitis B virus with mutations in the core promoter for an e antigen-negative phenotype in carriers with antibody to e antigen. *J Virol*. 1994;68:8102–8110.
11. Carman WF, Jacyna MR, Hadziyannis S, et al. Mutation preventing formation of hepatitis B e antigen in patients with chronic hepatitis B infection. *Lancet*. 1989;2:588–591.
12. Miyakawa Y, Okamoto H, Mayumi M. The molecular basis of hepatitis B e antigen (HBeAg)-negative infections. *J Viral Hepat*. 1997;4:1–8.
13. Pas SD, Fries E, De Man RA, Osterhaus AD, Niesters HG. Development of a quantitative real-time detection assay for hepatitis B virus DNA and comparison with two commercial assays. *J Clin Microbiol*. 2000;38:2897–2901.
14. Pawlotsky JM, Prescott L, Simmonds P, et al. Serological determination of hepatitis C virus genotype: comparison with a standardized genotyping assay. *J Clin Microbiol*. 1997;35:1734–1739.
15. Zaaijer HL, ter Borg F, Cuypers TM, Hermus MCAH, Lelie PN. Comparison of methods for detection of hepatitis B virus DNA. *J Clin Microbiol*. 1994;32:2088–2091.
16. Gerken G, Gomes J, Lampertico P, et al. Clinical evaluation and applications of the Amplicor HBV Monitor test, a quantitative HBV DNA PCR assay. *J Virol Methods*. 1998;74:155–165.
17. Noborg U, Gusdal A, Pisa EK, Hedrum A, Lindh M. Automated quantitative analysis of hepatitis B virus DNA by using the Cobas Amplicor HBV monitor test. *J Clin Microbiol*. 1999;37:2793–2797.
18. Stuyver L, Van Geyt C, De Gendt S, et al. Line probe assay for monitoring drug resistance in hepatitis B virus-infected patients during antiviral therapy. *J Clin Microbiol*. 2000;38:702–707.
19. Lok AS, Chung HT, Liu VW, Ma OC. Long-term follow-up of chronic hepatitis B patients treated with interferon alfa. *Gastroenterology*. 1993;105:1833–1838.
20. de Jongh FE, Janssen HL, de Man RA, Hop WC, Schalm SW, van Blankenstein M. Survival and prognostic indicators in hepatitis B surface antigen-positive cirrhosis of the liver. *Gastroenterology*. 1992;103:1630–1635.
21. Pawlotsky JM, Bastie A, Hezode C, et al. Routine detection and quantification of hepatitis B virus DNA in clinical laboratories: performance of three commercial assays. *J Virol Methods*. 2000;85:11–21.
22. Saldanha J, Gerlich W, Lelie N, Dawson P, Heermann K, Heath A; WHOCSG. An international collaborative study to establish a World Health Organization international standard for hepatitis B virus DNA nucleic acid amplification techniques. *Vox Sang*. 2001;80:63–71.
23. Krajden M, Comanor L, Rifkin O, Grigoriew A, Minor JM, Kapke GF. Assessment of hepatitis B virus DNA stability in serum by the Chiron Quantiplex branched-DNA assay. *J Clin Microbiol*. 1998;36:382–386.
24. Choo QL, Richman KH, Han JH, et al. Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci USA*. 1991;88:2451–2455.

25. Okamoto H, Kurai K, Okada S, et al. Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *Virology.* 1992;188: 331–341.
26. Bukh J, Purcell RH, Miller RH. At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. *Proc Natl Acad Sci USA*. 1993;90:8234–8238.
27. Bukh J, Purcell RH, Miller RH. Sequence analysis of the core gene of 14 hepatitis C virus genotypes. *Proc Natl Acad Sci USA*. 1994;91: 8239–8243.
28. Simmonds P, Holmes EC, Cha T-A, et al. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *J Gen Virol.* 1993;74:2391–2399.
29. Bukh J, Miller RH, Purcell RH. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin Liver Dis.* 1995;15: 41–63.
30. Tokita H, Okamoto H, Iizuka H, et al. Hepatitis C virus variants from Jakarta, Indonesia classifiable into novel genotypes in the second (2e and 2f), tenth (10a) and eleventh (11a) genetic groups. *J Gen Virol.* 1996;77:293–301.
31. Tokita H, Okamoto H, Tsuda F, et al. Hepatitis C virus variants from Vietnam are classifiable into the seventh, eighth, and ninth major genetic groups. *Proc Natl Acad Sci U S A.* 1994;91:11022–11026.
32. Lau JYN, Davis GL, Prescott LE, et al. and the Hepatitis Intervential Therapy Group. Distribution of hepatitis C virus genotypes determined by line probe assay in patients with chronic hepatitis C seen at tertiary referral centers in the United States. *Ann Intern Med.* 1996;124:868–876.
33. Dusheiko G, Schmilovitz-Weiss H, Brown D, et al. Hepatitis C virus genotypes: an investigation of type-specific differences in geographic origin and disease. *Hepatology.* 1994;19:13–18.
34. Benvegnu L, Pontisso P, Cavalletto D, Noventa F, Chemello L, Alberti A. Lack of correlation between hepatitis C virus genotypes and clinical course of hepatitis C virus-related cirrhosis. *Hepatology.* 1997;25:211–215.
35. Poynard T, Bedossa P, Opolon P. Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups. *Lancet.* 1997;349: 825–832.
36. National Institutes of Health. Management of hepatitis C. *NIH Consens Statement.* 2002;19:1–46.
37. Centers for Disease Control and Prevention. Recommendations for prevention and control of hepatitis C virus (HCV) infection and HCV-related chronic disease. *MMWR Morb Mortal Wkly Rep.* 1998; 47:1–39.
38. EASL International Consensus Conference on Hepatitic C. Consensus statement. *J Hepatol.* 1999;30:956–961.
39. Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS, Seeff LB. Diagnosis and monitoring of hepatic injury. II. Recommendations for use of laboratory tests in screening, diagnosis, and monitoring. *Clin Chem.* 2000;46:2050–2068.
40. Centers for Disease Control and Prevention. Guidelines for laboratory testing and result reporting of antibody to hepatitis C virus. *MMWR Morb Mortal Wkly Rep.*. 2003;52:1–15.
41. Fanning L, Kenny-Walsh E, Levis J, et al. Natural fluctuations of hepatitis C viral load in the homogeneous patient population: a prospective study. *Hepatology.* 2000;31:225–229.
42. Fried MW, Shiffman ML, Reddy KR, et al. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Eng J Med.* 2002; 347:975–982.
43. Lee SS, Heathcote EJ, Reddy KR, et al. Prognostic factors and early predictability of sustained viral response with peginterferon alfa-2a (40KD). *J Hepatol.* 2002;37:500–506.
44. Poynard T, McHutchison JG, Goodman ZD, Ling M-H, Albrecht J for the ALGOVIRC Project Group. Is an "a la carte" combination interferon alfa-2b plus ribavirin regimen possible for the first line treatment in patients with chronic hepatitis C. *Hepatology.* 2000;31: 211–218.
45. Schreiber GB, Busch MP, Kleinman SH, Korelitz JJ. The risk of transfusion-transmitted viral infections. The Retrovirus Epidemiology Donor Study. *N Eng J Med.* 1996;334:1685–1690.
46. Legler TJ, Riggert J, Simson G, et al. Testing of individual blood donations for HCV RNA reduces the residual risk of transfusion-transmitted HCV infection. *Transfusion.* 2000;40:1192–1197.
47. Kolk DP, Dockter J, Linnen J, et al. Significant closure of the human immunodeficiency virus type 1 and hepatitis C virus preseroconversion detection windows with a transcription-mediated-amplification-driven assay. *J Clin Microbiol.* 2002;40:1761–1766.
48. Ansaldi F, Torre F, Bruzzone BM, Picciotto A, Crovari P, Icardi G. Evaluation of a new hepatitis C virus sequencing assay as a routine method for genotyping. *J Med Virol.* 2001;63:17–21.
49. Icardi G, Ansaldi F, Bruzzone BM, et al. Novel approach to reduce the hepatitis C virus (HCV) window period: clinical evaluation of a new enzyme-linked immunosorbent assay for HCV core antigen. *J Clin Microbiol.* 2001;39:3110–3114.
50. Bouvier-Alias M, Patel K, Dahari H, et al. Clinical utility of total HCV core antigen quantification: a new indirect marker of HCV replication. *Hepatology.* 2002;36:211–218.
51. McHutchison JG, Gordon S, Schiff ER, et al. for the Hepatitis Interventional Therapy Group. Interferon alpha-2b alone or in combination of ribavirin as initial treatment for chronic hepatitis C. *N Eng J Med.* 1998;339:1485–1499.
52. Poynard T, Marcellin P, Lee SS, et al. for the International Hepatitis Interventional Therapy Group (IHIT). Randomised trial of interferon α2b plus ribavirin for 48 weeks or for 24 weeks versus interferon α2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. *Lancet.* 1998;352:1426–1432.
53. McCormick SE, Goodman ZD, Maydonovitch CL, Sjogren MH. Evaluation of liver histology, ALT elevation, and HCV RNA titer in patients with chronic hepatitis C. *Am J Gastroenterol.* 1996;91:1516–1522.
54. Hohne M, Schreier E, Roggendorf M. Sequence variability in the envcoding region of hepatitis C virus isolated from patients infected during a single source outbreak. *Arch Virol.* 1994;137:25–34.
55. Power JP, Lawlor E, Davidson F, Holmes EC, Yap PL, Simmonds P. Molecular epidemiology of an outbreak of infection with hepatitis C virus in recipients of anti-D immunoglobulin. *Lancet.* 1995;345:1211–1213.
56. Bronowicki JP, Venard V, Botte C, et al. Patient-to-patient transmission of hepatitis C virus during colonoscopy. *N Eng J Med.* 1997;337:237–240.
57. Aizaki H, Saito A, Kusakawa I, et al. Mother-to-child transmission of a hepatitis C virus variant with an insertional mutation in its hypervariable region. *J Hepatol.* 1996;25:608–613.
58. Chayama K, Kobayashi M, Tsubota A, et al. Molecular analysis of intraspousal transmission of hepatitis C virus. *J Hepatol.* 1995;22: 431–439.
59. Healey CJ, Smith DB, Walker JL, et al. Acute hepatitis C infection after sexual exposure. *Gut.* 1995;36:148–150.
60. Weiner AJ, Thaler MM, Crawford K, et al. A unique, predominant hepatitis C virus variant found in an infant born to a mother with multiple variants. *J Virol.* 1993;67:4365–4368.
61. Enomoto N, Sakuma I, Asahina Y, et al. Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *N Eng J Med.* 1996;334:77–81.
62. Watanabe H, Enomoto N, Nagayama K, et al. Number and position of mutations in the interferon (IFN) sensitivity—determining region of the gene for nonstructural protein 5A correlate with IFN efficacy in hepatitis C virus 1b infection. *J Infect Dis.* 2001;183: 1195–1203.
63. Hofgartner WT, Polyak SJ, Sullivan DG, Carithers RL, Gretch DR. Mutations in the NS5A gene of hepatitis C virus in North American patients infected with HCV genotype 1a or 1b. *J Med Virol.* 1997;53:118–126.
64. Sarrazin C, Berg T, Lee JH, et al. Improved correlation between multiple mutations within the NS5A region and virological response in

European patients chronically infected with hepatitis C virus type 1b undergoing combination therapy. *J Hepatol.* 1999;30:1004–1013.
65. Lee SC, Antony A, Lee N, et al. Improved version 2.0 qualitative and quantitative AMPLICOR reverse transcription-PCR tests for hepatitis C virus RNA: calibration to international units, enhanced genotype reactivity, and performance characteristics. *J Clin Microbiol.* 2000;38:4171–4179.
66. Nolte FS, Fried MW, Shiffman ML, et al. Prospective multicenter clinical evaluation of AMPLICOR and COBAS AMPLICOR hepatitis C virus tests. *J Clin Microbiol.* 2001;39:4005–4012.
67. Comanor L, Anderson F, Ghany M, et al. Transcription-mediated amplification is more sensitive than conventional PCR-based assays for detecting residual serum HCV RNA at end of treatment. *Am J Gastroenterol.* 2001;96:2968–2972.
68. Sarrazin C, Teuber G, Kokka R, Rabenau H, Zeuzem S. Detection of residual hepatitis C virus RNA by transcription-mediated amplification in patients with complete virologic response according to polymerase chain reaction-based assays. *Hepatology.* 2000;32:818–823.
69. Sarrazin C, Hendricks DA, Sedarati F, Zeuzem S. Assessment, by transcription-mediated amplification, of virologic response in patients with chronic hepatitis C virus treated with peginterferon alpha-2a. *J Clin Microbiol.* 2001;39:2850–2855.
70. Krajden M, Ziermann R, Khan A, et al. Qualitative detection of hepatitis C virus RNA: comparison of analytical sensitivity, clinical performance, and workflow of the Cobas Amplicor HCV test version 2.0 and the HCV RNA transcription-mediated amplification qualitative assay. *J Clin Microbiol.* 2002;40:2903–2907.
71. Jackson JB, Smith K, Knott C, et al. Sensitivity of the Procleix HIV-1/HCV assay for detection of human immunodeficiency virus type 1 and hepatitis C virus RNA in a high-risk population. *J Clin Microbiol.* 2002;40:2387–2391.
72. Beld M, Sentjens R, Rebers S, et al. Performance of the New Bayer VERSANT HCV RNA 3.0 assay for quantitation of hepatitis C virus RNA in plasma and serum: conversion to international units and comparison with the Roche COBAS Amplicor HCV Monitor, version 2.0, assay. *J Clin Microbiol.* 2002;40:788–793.
73. Germer JJ, Heimgartner PJ, Ilstrup DM, Harmsen WS, Jenkins GD, Patel R. Comparative evaluation of the Versant HCV RNA 3.0, Quantiplex HCV RNA 2.0, and COBAS Amplicor HCV Monitor version 2.0 Assays for quantification of hepatitis C virus RNA in serum. *J Clin Microbiol.* 2002;40:495–500.
74. Kleiber J, Walter T, Haberhausen G, Tsang S, Babiel R, Rosenstraus M. Performance characteristics of a quantitative, homogeneous TaqMan RT-PCR test for HCV RNA. *J Mol Diagn.* 2000;2:158–166.
75. Barbeau JM, Goforth J, Caliendo AM, Nolte F. Performance characteristics of a quantitative, TaqMan hepatitis C virus RNA analyte specific reagent. *J Clin Microbiol.* 2004:in press.
76. Caliendo AM, Valsamakis A, Zhou Y, et al. Multilaboratory comparison of hepatitis C virus viral load assays. *J Clin Microbiol.* 2006;44:1726–1732.
77. Stuyver L, Wyseur A, van Arnhem W, Hernandez F, Maertens G. Second-generation line probe assay for hepatitis C virus genotyping. *J Clin Microbiol.* 1996;34:2259–2266.
78. Andonov A, Chaudhary RK. Subtyping of hepatitis C virus isolates by a line probe assay using hybridization. *J Clin Microbiol.* 1995;33:254–256.
79. Viazov S, Zibert A, Ramakrishnan K, et al. Typing of hepatitis C virus isolates by DNA enzyme immunoassay. *J Virol Methods.* 1994;48:81–91.
80. Smith DB, McAllister J, Casino C, Simmonds P. Virus "quasispecies": making a mountain out of a molehill? *J Gen Virol.* 1997;78:1511–1519.
81. Ross RS, Viazov SO, Holtzer CD, et al. Genotyping of hepatitis C virus isolates using CLIP sequencing. *J Clin Microbiol.* 2000;38:3581–3584.
82. Halfon P, Trimoulet P, Bourliere M, et al. Hepatitis C virus genotyping based on 5′ noncoding sequency analysis (Trugene). *J Clin Microbiol.* 2001;39:1771–1773.
83. Kalinina O, Norder H, Mukomolov S, Magnius LO. A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg. *J Virol.* 2002;76:4034–4043.
84. Hu YW, Balaskas E, Kessler G, et al. Primer specific and mispair extension analysis (PSMEA) as a simple approach to fast genotyping. *Nucleic Acids Res.* 1998;26:5013–5015.
85. Krekulova L, Rehak V, Wakil AE, Harris E, Riley LW. Nested restriction site-specific PCR to detect and type hepatitis C virus (HCV): a rapid method to distinguish HCV subtype 1b from other genotypes. *J Clin Microbiol.* 2001;39:1774–1780.
86. Davidson F, Simmonds P, Ferguson JC, Jarvis LM, et al. Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5′ non-coding region. *J Gen Virol.* 1995;76:1197–1204.
87. White PA, Zhai X, Carter I, Zhao Y, Rawlinson WD. Simplified hepatitis C virus genotyping by heteroduplex mobility analysis. *J Clin Microbiol.* 2000;38:477–482.
88. Bullock GC, Bruns DE, Haverstick DM. Hepatitis C genotype determination by melting curve analysis with a single set of fluorescence resonance energy transfer probes. *Clin Chem.* 2002;48:2147–2154.
89. Schroter M, Zollner B, Schafer P, et al. Genotyping of hepatitis C virus types 1, 2, 3, and 4 by a one-step LightCycler method using three different pairs of hybridization probes. *J Clin Microbiol.* 2002;40:2046–2050.
90. Dixit V, Quan S, Martin P, et al. Evaluation of a novel serotyping system for hepatitis C virus: strong correlation with standard genotyping methodologies. *J Clin Microbiol.* 1995;33:2978–2983.
91. Agnello V, Chung RT, Kaplan LM. A role for hepatitis C virus infection in type II cryoglobulinemia. *N Eng J Med.* 1992;327:1490–1495.
92. Cribier B, Rey D, Schmitt C, Lang JM, Kirn A, Stoll-Keller F. High hepatitis C viraemia and impaired antibody response in patients coinfected with HIV. *AIDS.* 1995;9:1131–1136.
93. Fabrizi F, Poordad FF, Martin P. Hepatitis C infection and the patient with end-stage renal disease. *Hepatology.* 2002;36:3–10.
94. Saldanha J, Lelie N, Heath A. Establishment of the first international standard for nucleic acid amplification technology (NAT) assays for HCV RNA. WHO Collaborative Study Group. *Vox Sang.* 1999;76:149–158.
95. Stelzl E, Kormann-Klement A, Haas J, et al. Evaluation of an automated sample preparation protocol for quantitative detection of hepatitis C virus RNA. *J Clin Microbiol.* 2002;40:1447–1450.
96. Nolte F, Green AM, Fiebelkorn KR, et al. Clinical evaluation of two methods for genotyping hepatitis C virus based on analysis of the 5′ noncoding region. *J Clin Microbiol.* 2003;41:1558–1564.

Chapter 38

Viral Infections in Transplant Patients

Andrea Ferreira-Gonzalez and Angela M. Caliendo

Introduction

Viruses are particularly problematic pathogens in transplant recipients. Viral infections not only can cause disease but also can enhance susceptibility to opportunistic infections by both causing tissue injury and contributing to systemic immunosuppression. Such infections have been shown to increase the rate of graft rejection and increase the risk of cancer. The outcome of viral infections in the transplant setting is the result of a balance among infection in tissues, host antiviral immune function, and the level of immunosuppression required to maintain graft function. Diagnosis of viral infections in immunocompromised patients remains a challenge due to the need to differentiate asymptomatic infection from clinically relevant disease.

The use of molecular tests has greatly enhanced the diagnosis of viral infections, in part due to the slow turnaround time or inadequate sensitivity of conventional diagnostic tests. Unlike conventional methods, which rely on viral culture or antigen detection, molecular methods detect viral nucleic acid. A number of nucleic acid–based methods have been applied to the diagnosis of viral infection in transplant patients, and the most widely used method is still the polymerase chain reaction (PCR), but other methods are used with increasing frequency. The RNA amplification methods (nucleic acid sequence–based amplification [NASBA] and transcription-mediated amplification [TMA]) and the signal-amplification methods (hybrid capture and branched DNA [bDNA]) also are available for the detection and quantitation of viral pathogens.

Molecular methods often provide increased sensitivity, a more rapid turnaround time, and, for some assays, improved specimen stability and ease of use compared to conventional methods. Even though the increased sensitivity of nucleic acid–based tests has been a major advantage, these tests also have posed some challenges. A large number of viral infections in transplant patients are due to viruses that remain in a latent state after primary infection. Nucleic acid from these viruses can be detected in patients with asymptomatic infection without evidence of clinical disease. Methods to improve clinical specificity of molecular tests have been developed and are discussed here.

One of the biggest challenges facing laboratories and clinicians is the lack of standardized molecular assays. A limited number of commercial test kits are available, and even fewer have been approved or cleared by the Food and Drug Administration (FDA). A large number of molecular assays currently in use have been developed by individual laboratories. The performance characteristics of these assays vary widely due to differences in specimen type, target sequence, limit of detection, and quantitative standard. As a result, values obtained from different laboratory-developed assays may not be comparable, which makes it very difficult to establish interpretive guidelines that can be applied broadly in clinical practice.

This chapter reviews the molecular tests that are available to detect and quantify cytomegalovirus (CMV), BK virus (BKV), hepatitis C virus (HCV), Epstein-Barr virus (EBV), and human herpes virus type 6 (HHV-6) in transplant recipients and discusses their clinical utility.

CMV

CMV remains a major cause of morbidity and mortality in solid organ transplant (SOT) and bone marrow transplant (BMT) recipients.[1] Primary CMV infection occurs during infancy, preschool age, and sexually active years. During primary infection, the virus is disseminated through the body via the bloodstream and the infection can be subclinical with diffuse organ involvement. After resolution of the primary infection, the virus remains latent in a nonreplicating state or with an undetectable level of replication. However, in immunocompromised individuals, the virus can reactivate and result in clinical symptoms. CMV infection of the transplant population occurs by transmission, reactivation, or superinfection with a new strain of CMV. The clinical manifestations of CMV disease include fever, interstitial pneumonia, chorioretinitis, hepatitis,

esophagitis, gastritis, and colitis. Preventive strategies have been developed for CMV disease in this patient population and include the use of CMV-seronegative blood products, CMV-seronegative donor selection whenever possible, and passive immunoprophylaxis with immunoglobulin. In recent years, prophylactic as well as preemptive treatment with antiviral drugs has been used with good success to reduce the risk of developing active CMV disease.[1]

The occurrence and severity of CMV disease in transplant recipients is influenced by the type of organ transplanted, the CMV serostatus of the organ recipient and donor, and the overall level of immunosuppression. Individuals who acquire their primary CMV infection while immunosuppressed are at risk for severe disease. Symptomatic disease can occur in CMV-seropositive recipients regardless of the serologic status of the donor. Furthermore, CMV-seronegative individuals who receive an organ from a CMV-seropositive donor are at higher risk for severe disease. The diagnostic challenge with CMV disease is distinguishing asymptomatic infection from clinically significant disease, because infectious virus, viral antigen, and viral nucleic acid can be detected in patients without clinically significant disease.[2]

CMV is a member of the Herpesviridae family and is a large, enveloped, double-stranded DNA virus. Traditionally, the detection and quantitation of CMV has relied on culture and antigen detection. Urinary shedding of CMV and even seroconversion are common but are not predictive of clinically significant CMV disease. Conventional culture methods are not sensitive enough to detect viremia in immunocompromised patients but can be useful in detecting CMV from respiratory secretions, urine, and tissue samples. However, the turnaround time for conventional culture results is slow, between 2 and 3 weeks, and as a result rarely impacts clinical care. The rapid centrifugation shell vial method has greatly reduced the time for virus detection, with positive results obtained in 24 to 48 hours, although this method lacks sensitivity for detecting viremia. The CMV antigenemia assay is limited to the detection of CMV in white blood cells and is more sensitive than conventional culture or shell vial methods. However, the antigenemia assay is labor-intensive and the sensitivity of the assay varies among different laboratories.[2–4]

Molecular biology assays have had a major impact on the rapid diagnosis of CMV disease.[5–7] The commercially available molecular tests for detection and quantitation of CMV nucleic acid are shown in Table 38-1, two of which have been cleared by the FDA. The Hybrid Capture test (Digene Corp, Gaithersburg, MD) is designed to detect CMV DNA in whole blood, while the NucliSens test (bioMerieux, Durham, NC) utilizes NASBA technology to detect CMV pp67 mRNA in whole blood. The NASBA assay is qualitative, while the Hybrid Capture assay may be either qualitative or quantitative (though it has been cleared by the FDA only as a qualitative test).

The more widely used molecular tests are based on PCR methodology, either laboratory-developed or commercially available test kits. Furthermore, analyte-specific reagents using real-time quantitative PCR (RQ-PCR) technology have recently become available. The laboratory-developed assays utilize both standard PCR and RQ-PCR methods and their performance varies due to differences in specimen type, nucleic acid extraction methods, target gene, primer sequences, quantitation standard, and detection method. These differences become quite important in the comparison of quantitative assays since viral load values may vary when the same sample is tested using different assays. Such assay variation makes comparison of results from different studies difficult, and also impedes establishment of viral load values that correlate with clinical disease. The challenges in assessing clinical performance of molecular tests include developing an understanding of the analytical performance characteristics of the

Table 38-1. Molecular Assays for the Detection of CMV

Assay	Method	Target	Lower Limit (copies/ml)	Upper Limit (copies/ml)	Test Kit	FDA Cleared
Qualitative Methods						
NucliSens CMV (bioMerieux, Durham, NC)	NASBA	pp67 mRNA	Not defined	N/A	Yes	Yes
Hybrid Capture CMV v3.0 (Digene Corp, Gaithersburg, MD)	Hybrid capture	17% of genome	1400	N/A	Yes	Yes
Laboratory-developed assays	PCR/ Real-time PCR	Variable	Variable	N/A	No	No
Quantitative Methods						
Amplicor CMV Monitor (Roche Diagnostics, Indianapolis, IN)	PCR	Pol gene	400	1×10^5	Yes	No
Hybrid Capture CMV v3.0 (Digene Corp, Gaithersburg, MD)	Hybrid capture	17% of genome	1400	600,000	Yes	No
Laboratory-developed assays and analyte specific reagents (ASR)	PCR/ Real-time PCR	Variable	Variable	Variable	No	No

N/A, not applicable.

different assays, determining what each test is measuring, and assessing the clinical implications of the results. For example, PCR assays using whole blood or leukocytes and the Hybrid Capture test detect intracellular DNA,[8] which is present in both latent and actively infected cells. On the other hand, PCR assays using plasma detect DNA in circulation that is the result of active viral replication. Furthermore, the NucliSens assay detects pp67 mRNA, which is expressed at a high level in patients with active CMV disease but is not detectable during latent infection.[9–11]

One of the primary uses of CMV molecular testing is for the diagnosis of active CMV disease, which involves distinguishing patients with asymptomatic infection from patients with clinically significant CMV disease. Early studies evaluating the clinical utility of CMV DNA tests used very sensitive qualitative assays and peripheral blood mononuclear cells (PBMC) as the specimen of choice. As a result, CMV DNA was detected in patients with CMV disease as well as in immunocompromised patients without active disease.[5,12] Several different approaches have been used in an effort to improve the clinical specificity of molecular tests for active CMV disease, including the use of quantitative assays, detecting DNA in plasma rather than whole blood or leukocytes, and detection of mRNA rather than DNA. Several studies have proven the clinical utility of quantitative PCR assays using plasma specimens.[10,13] Quantitative methods may allow the definition of threshold values that distinguish asymptomatic infection from active CMV disease. Detection of CMV DNA in plasma rather than in leukocytes may provide a better correlation with clinical disease because the detection of CMV DNA in plasma correlates with active viral replication due to the spread of virus from leukocytes and tissue to the plasma compartment.[10] The detection of CMV DNA in plasma is an early marker for active CMV disease in renal[7] and liver[14,15] transplant recipients.

Several studies have evaluated the clinical utility of laboratory-developed quantitative PCR assays for the diagnosis of CMV disease in transplant patients. Patients with active CMV disease have consistently higher viral load values than those with asymptomatic infection.[14,16] In light of these findings, cutoff values were proposed as a means to discriminate infection from clinically relevant active disease. Cutoff values associated with active disease vary by specific assay type, as well as by the type of SOT or BMT. Though these tests have proven to be very useful for individual laboratories, it has been difficult to establish broadly applicable, clinically relevant guidelines for the use of laboratory-developed CMV viral load assays.

Recently several studies have evaluated the clinical utility of two commercially available test kits, the Amplicor CMV Monitor (Roche Diagnostics, Indianapolis, IN) and the CMV Hybrid Capture (Digene Corp.). Humar et al.[15] performed a prospective study in which weekly plasma CMV viral load was measured in 97 liver transplant patients during the 12 weeks after transplantation. No preemptive therapy was given; however, CMV-seronegative recipients who received a CMV-seropositive liver received 12 weeks of prophylactic ganciclovir (GCV). Sixty-one of the 97 patients developed CMV infection, defined as a positive laboratory test for CMV. Of these 61 patients, 21 developed documented CMV disease. Clinical manifestations of CMV disease included hepatitis, CMV viral syndrome, and esophagitis or colitis. As shown with early studies, patients with CMV disease had a statistically significant higher peak viral load than those with asymptomatic infection (median peak 55,000 copies/ml vs 1820 copies/ml, respectively). The predictive value of viral load testing also was assessed by evaluating viral load values obtained at least 3 days prior to the development of CMV disease. Using receiver-operator characteristic curves, the authors determined that the optimal viral load for detecting clinically significant CMV disease was between 2000 and 5000 copies/ml of plasma. As the viral load cutoff level for detecting CMV disease was increased, there was a loss in sensitivity but an improvement in specificity.

Two therapeutic approaches are used to prevent CMV disease in transplant patients who are at risk for CMV disease, prophylactic treatment and preemptive treatment. Prophylactic therapy is the treatment of all patients considered to be at risk for CMV disease without the use of molecular testing, without further stratification of risk. Prophylactic therapy is based on the assumption that the entire population is at substantial risk of infection and merits preventive treatment. Thus, the use of prophylactic therapy involves treatment of a large number of patients. Preemptive therapy is used for at-risk patients who are positive for CMV by a laboratory test. Molecular CMV assays are commonly used to make decisions concerning the administration of preemptive therapy.[12,15,17–21] For example, all at-risk patients would be tested for the presence of CMV DNA in their blood, and only those with a positive result would be treated. The use of preemptive therapy focuses preventive treatment on a subgroup of patients determined to be at a higher risk of CMV disease. Thus, the laboratory test is used to determine whether the level of infection makes invasive disease likely. Both these approaches provide therapy prior to the development of symptoms, hence preventing the development of active disease.

A number of studies have evaluated the clinical utility of the NucliSens CMV test (bioMerieux) for initiation of preemptive therapy.[9,18] In a prospective study by Hebart et al.,[22] a total of 33 BMT recipients at risk for CMV disease were monitored once a week for 18 months after transplantation using the NucliSens NASBA assay, a whole-blood laboratory-developed PCR assay, and viral culture. Preemptive antiviral therapy was initiated after two positive PCR results; clinical decisions were not made based on NucliSens results. The NucliSens and the PCR assays had a high level of agreement (87.9% for patients and 85.3% for samples). None of the NASBA-negative patients developed CMV disease. Sixteen of the 18 patients receiving preemptive

GCV treatment also were NASBA positive. Three of the 33 patients at risk for developing CMV disease developed disease, all of whom were PCR and NASBA positive. There was no difference between assays for the time of first positive result. Ongoing studies evaluating the clinical utility of quantitative PCR assays and the NASBA assay will provide further information regarding the clinical utility of these tests for guiding preemptive antiviral therapy.

Several studies have looked at the clinical utility of quantitative molecular tests for monitoring antiviral therapy response in transplant patients.[21,23] A greater than 90% reduction in viral load, as measured by the Hybrid Capture test, has been observed after initiation of therapy for CMV infection.[5] Several studies have reported that viral loads become undetectable several weeks after initiating therapy;[7,20,21] however, the time to clear CMV is related to the viral load. Patients with low pretreatment viral loads will become undetectable for CMV more rapidly than those with high pretreatment viral load values. Failure of viral loads to decline to undetectable levels after initiation of antiviral therapy is a cause of concern because patients with documented GCV resistance have persistent CMV viral loads.[7]

Another application of molecular tests is identifying patients at risk for relapsing CMV infection.[23–25] The optimal duration of antiviral treatment for CMV infection has not been established. Relapsing episodes of CMV disease have been described in 15% to 35% of SOT patients. Although relapsing CMV disease might be due to the evolution of resistant CMV strains, it also may be the result of incomplete suppression of viral replication by antiviral treatment. In a prospective study of 24 SOT patients with CMV infection or disease, patients were treated with a 14-day course of GCV,[24] and viral load values were measured before and after treatment. Eight patients developed relapsing CMV infection and had viral load measurements prior to the initiation of therapy that were statistically higher compared to the nonrelapsing group (80,150 copies/10^6 leukocytes vs 5,500 copies/10^6 leukocytes, respectively, $p = 0.007$). In addition, the relapsing group had detectable viral load at the end of treatment (mean 18,800 copies/10^6 leukocytes), while viral load levels were undetectable in the nonrelapsing group. Similar results have been reported with the NucliSens assay. In one study, all six SOT recipients with detectable RNA after completing a course of therapy required further therapy for their CMV infection.[18]

The rate of decline in CMV load also is predictive of relapsing CMV infections. Fifty-two organ transplant recipients with CMV disease were monitored with the Amplicor CMV Monitor test (Roche Diagnostics), of which 12 developed relapsing CMV disease. The time to clearance of CMV DNA from plasma was 33.8 days in the group with relapsing CMV disease compared to 17.2 days in the group without recurrent disease ($p = 0.002$). The viral load half-life was 8.8 days compared to 3.2 days ($p = 0.001$) in the group with and without recurrent disease, respectively.[25] Sequential monitoring of viral load levels after initiation of therapy may allow the identification of patients at risk of recurrent infection, thus providing the opportunity to intensify therapy and possibly prevent relapse.

Viral load measurements may be useful for individualizing treatment regimen length. Fisher et al.[26] prospectively evaluated and followed 18 SOT patients diagnosed with CMV disease. Plasma samples were obtained for viral load measurements weekly during treatment and monthly for 6 months following the end of therapy. Patients were treated with immunoglobulin for a period of time and with GCV until CMV viral load was undetectable (<100 copies/ml plasma). No patients relapsed with a mean follow-up of 16 months. This study suggests that viral load measurements have clinical utility for determining the length of time for GCV treatment. Further studies with a greater number of patients are needed to determine the utility of viral load testing for determining the duration of antiviral therapy, but this study indicates that discontinuing therapy when virus is still detectable in the plasma is a risk for relapsing infection.

Generally, transplant patients requiring treatment for CMV disease receive a lengthy course of antiviral therapy.[27–30] Such long-term treatment regimens are associated with development of antiviral drug resistance. There are currently three antiviral drugs that are FDA approved for the treatment of systemic CMV disease: GCV, cidofovir (CDV), and foscarnet (phosphonoformic acid [FOS]); of these three, GCV is the most widely used. Development of resistance does not occur very frequently because CMV is a DNA virus that replicates using a DNA polymerase that has proofreading activity, which prevents high mutation rates. Thus, selection of drug-resistant virus from the initial wild-type drug-sensitive population generally occurs only after weeks to months of therapy.

Two different types of testing are used for the detection of drug-resistant virus: phenotypic and genotypic. Phenotypic assays rely on analysis of growth characteristics of viral isolates in the presence of different concentrations of antiviral drugs and therefore require viable virus. The phenotypic assays are labor-intensive, the readout is subjective, and they require at least a month to obtain results, which is generally not therapeutically useful. The lengthy time requirement has led to development of genotypic assays based on detection of nucleotide sequence changes that are known to cause drug resistance. The currently used antiviral drugs target the UL97 gene (GCV) or the DNA polymerase gene (GCV, CDV, FOS) or both. Point mutations or short deletions in these genes confer drug resistance. GCV resistance mutations are most frequently mapped to the UL97 gene within codons 460 and 520 and either point mutations or deletions within the codon range 590 to 607.[28,29] GCV resistance due to mutations in the DNA polymerase gene appears to occur less frequently than those in the UL97 gene.[28,29] The majority of polymerase mutations conferring resistance to GCV also are cross-resistant to CDV,[30] whereas mutations in the UL97 gene confer resist-

ance only to GCV. In general, FOS resistance mutations within the polymerase gene do not cause cross-resistance to either of the other two drugs, although a double deletion of codons 981 and 982 has been reported to confer resistance to all three drugs.[31] At least 23 sites within the polymerase gene have been mapped that confer drug resistance.[32] This covers approximately 2000 nucleotides, a much larger region than that of the UL97 gene. One caveat for genotypic assays is the presence of some normal baseline sequence variability in the UL97 and DNA polymerase genes in drug-sensitive strains. The assumption is made that the same mutation will produce resistance in all genetic backgrounds, which may not prove to be the case as more strains are examined. While phenotypic assays are biologically more relevant for detection of drug resistance because individual strains are tested for replication directly in the presence of antiviral drugs, genotypic results do correlate with the clinical drug resistance and provide rapid results not possible for the presently available phenotypic assays.[33] An advance in genotyping testing is the ability to sequence CMV DNA directly from the plasma specimens in patients failing therapy.

BKV

Since first reported, BKV-associated nephropathy has emerged as an important cause of allograft failure in renal transplant recipients. BKV is ubiquitous, with more than 90% of adults having serologic evidence of past infection. BKV is transmitted at an early age via oral or respiratory routes, and after primary infection remains latent in uroepithelial cells and circulating leukocytes. During immunosuppression, virus reactivation causes cytopathic changes in the uroepithelium, with increased shedding of BKV in the urinary tract. This in turn results in a range of clinical syndromes from viremia and viruria to ureteral ulceration or stenosis, and hemorrhagic cystitis. In renal transplant recipients, persistent active replication of BKV in the renal allograft, called BKV nephropathy, is associated with progressive graft dysfunction and graft loss due to interstitial nephritis. However, BKV nephropathy is rarely diagnosed in transplant recipients of nonrenal organs. BKV nephropathy develops in 1% to 5% of renal transplant recipients, with graft failure in as many as 45% to 67% of affected patients. Risk factors associated with BKV-associated nephritis in renal transplant patients are not well known, but most of the patients with this disorder have received the newer immunosuppressive drugs such as tacrolimus or mycophenolate mofetil.[34,35]

Antiviral treatment is not established for BK nephropathy, and control of antiviral activity is tentatively attained by reduction of immunosuppressive therapy. In patients with diminished renal allograft function and possible BKV infection, a choice must be made between increasing the level of immunosuppression to treat suspected graft rejection or reduction of immunosuppression to allow the immune system to control the BKV infection. Transplant patients with BKV nephropathy who are treated with increased levels of immunosuppressive drugs have a high incidence of graft loss. On the other hand, a reduced level of immunosuppressive therapy may stabilize the graft function but increases the risk of rejection.[36] A recent study evaluated the clinical utility of CDV for the treatment of BKV nephropathy in a limited number of patients. Results were encouraging, with all four patients clearing the viremia without graft loss.[37] Further studies are needed to confirm these findings.

Accurate diagnosis of BKV nephropathy is required to preserve graft function. Progression of BKV nephropathy often occurs without clear clinical signs or symptoms; sometimes the only indication is a rise in creatinine concentration, which occurs over a period of days to weeks before symptoms appear, although not all patients with BKV nephropathy present with an elevated creatinine. Diagnosis of BKV nephropathy is made by histological demonstration of viral infection, which is distinct from signs of rejection. In these patients, histologic examination of allograft biopsy specimens reveals extensive viral cytopathic effect, necrosis of the cells of the tubules and collecting ducts, and varying degrees of interstitial inflammation. Although renal biopsy is useful for the diagnosis of BKV nephropathy, it is an invasive procedure. A noninvasive alternative is to examine the urine for the presence of decoy cells, which contain characteristic intracellular viral inclusion bodies. Identification of decoy cells confirms that there is active replication of BKV, but this finding is not specific for BKV nephropathy.[34–36]

Using a qualitative PCR method, Randhawa et al.[38] demonstrated a positive relationship between the presence of BKV DNA in blood and impairment of renal allograft function, thus indicating that viremic patients are at high risk for progression to BKV-related interstitial nephritis. Moreover, in a recent prospective study, Hirsch et al.[39] have shown that BKV viral load is significantly higher in patients with biopsy-proven BKV nephropathy than in those without BKV nephropathy. A total of 78 renal transplant patients were followed for a median of 85 weeks after transplantation. BKV nephropathy was diagnosed in allograft biopsy specimens from five patients at a median of 28 weeks after transplantation. The diagnosis of infection on the basis of cytologic examination of urine specimens for the presence of decoy cells showed a sensitivity of 100% and specificity of 71% with a positive predictive value of 29% and a negative predictive value of 100%. The diagnosis of infection on the basis of cytologic examination of urine specimens for decoy cells was sensitive but had a low positive predictive value, suggesting that this approach could be used as a screening test. BKV viremia was detected in ten patients at a median of 23 weeks after transplant, with five having histologically proven BKV nephropathy. Interestingly, the mean viral DNA level in plasma was significantly higher in patients with histologically proven BKV nephropathy than in patients without histologic

evidence of nephropathy (2×10^4 copies/ml vs 2×10^3 copies/ml, respectively; $p<0.001$). Furthermore, the plasma viral load increased to 7.7×10^3 copies/ml or more in all patients in whom BKV nephropathy developed, with three of the five patients reaching viral load levels of 1×10^7 copies/ml. For all patients with histologically proven BKV nephropathy, BKV viral load had a sensitivity of 100%, a specificity of 88%, a positive predictive value of 50%, and a negative predictive value of 100%. These findings offer a basis for identifying patients at risk for BKV nephropathy who may be candidates for preemptive reduction in immunosuppressive therapy or initiation of antiviral therapy.

BKV viral load testing of renal transplant patients has shown utility not only for diagnosing BKV nephropathy but also for monitoring the response to antiviral therapy or reduction of immunosuppressive therapy.[40–42] Limaye et al.[40] prospectively collected samples from four renal transplant patients with histologically diagnosed BKV nephropathy (cases) and 16 renal transplant patients who did not develop BKV nephropathy (controls). Serum samples were collected from all patients at the time of transplantation, during clinic visits, and during readmissions to the hospital. The case patients were more likely than the control patients to have BKV detected in serum samples (4 of 4 patients vs 0 of 16 patients). In all four patients, BKV DNA was detected in serum samples before histologic diagnosis of BKV nephropathy. Interestingly, prolonged viremia, in some cases over 50 weeks, was found in three case patients. Furthermore BKV DNA was detected in blood coincident with persistent unexplained inflammatory changes in biopsies weeks to months before characteristic histologic changes consistent with BKV nephropathy became evident. A reduction of the viral load was demonstrated after reduction in immunosuppression or transplant nephrectomy or both.[40]

BKV also is associated with the development of late-onset hemorrhagic cystitis (HC) in allogeneic BMT patients. HC is an important cause of morbidity and occasional mortality in BMT patients. The incidence of HC in this patient population varies from 5% to 35% and is characterized by lower abdominal pain, dysuria, frequent micturition, and hematuria. The manifestations vary from microscopic hematuria to severe hemorrhage, leading to clot retention and renal failure. HC had been ascribed solely to the toxic effects of drugs utilized during BMT, but recently BKV reactivation has been associated with HC. During BMT, intense immunosuppressive therapy leads to increased viral replication, which results in viruria. Since BKV can be detected in urine samples of 77% to 90% of all adult BMT patients, both with and without HC, it is apparent that BKV reactivation alone is not sufficient to cause HC.[43–45] Several studies have tried to identify other factors such as graft-vs-host disease, primary infection, or even specific types of BKV that might be involved in the development of HC, with little success.[34]

One of the biggest challenges in trying to evaluate the clinical utility of molecular tests for the diagnosis and monitoring of BKV-associated disease is the lack of standardized molecular assays. The performance of these assays can vary widely due to differences in specimen type used, methodology, and limit of detection. As a result, values obtained from different laboratory-developed assays may not be comparable, which makes it very difficult to establish interpretive guidelines that can be applied broadly in clinical practice. Further studies are needed to better define the role of BKV infection and the role of molecular qualitative and quantitative assays for diagnosis and monitoring BKV nephropathy.

HCV

End-stage liver disease secondary to HCV infection is the most common indication for liver transplantation in this country, accounting for one third of all liver transplants. A variety of molecular methods for detection and quantitation of HCV RNA genomes are used in the management of patients with HCV infection before and after liver transplantation. Three methods are used to detect and quantify HCV RNA: reverse transcription–PCR (RT-PCR), TMA, and the bDNA technique.[46] Refer to chapter 37 for a detailed discussion of these assays and their performance characteristics.

In patients transplanted for HCV cirrhosis, reinfection and recurrence of HCV disease occurs in most cases.[47–52] Reinfection refers to the presence of HCV in blood, while recurrence of disease refers to histologic evidence of liver damage. Therefore, the diagnosis of recurrent disease requires histologic changes in the presence of persistent HCV viremia. Detectable HCV RNA is present in nearly all HCV-related liver transplant recipients within 1 month after transplant, while histologic recurrence of HCV disease occurs in 50% within 2 years after transplant. The rate of histologic disease development is accelerated in liver transplant patients as compared to the general population, where chronic hepatitis develops in 80% of patients after a decade or more of infection. Moreover, development of cirrhosis also is accelerated in this patient population, with 10% to 20% of patients developing cirrhosis within 5 years of transplant, compared to the general population, in which 20% develop cirrhosis after 2 to 3 decades after infection. Many patients (20% to 40%) will have a benign course after transplant, while a few patients develop more severe disease with rapid progression to fibrosis and cirrhosis.[47–49] Although there is strong data supporting an accelerated disease process in transplanted patients, controversy remains regarding the impact of recurrent HCV on patient and graft survival. These findings underscore the need to identify prognostic factors that may predict the course of disease after transplant.

Prognostic factors for outcome after liver transplantation[53–56] include early recurrence of disease, viral load, and genotype. Early recurrence of HCV disease is associated with increased risk for cirrhosis, as well as decreased graft and patient survival. Liver transplant patients who devel-

oped recurrent HCV disease within 6 months of transplant are two to three times more likely to develop cirrhosis than recipients who developed histologic recurrence after 6 months. Patients who developed HCV recurrence within 1 year of transplantation are more likely to lose their graft.

The role of HCV viral load in predicting poor graft survival is controversial.[53] In an early study, using the National Institute of Diabetes and Digestive and Kidney Diseases Liver Transplantation Database, transplant patients with HCV RNA titers greater than 1×10^6 mEq/ml had worse patient and graft survival rates compared to patients with lower HCV RNA levels.[57] In a more recent study of 79 patients who underwent liver transplantation for chronic HCV infection, neither pretransplant nor posttransplant HCV RNA levels were significantly associated with the occurrence of graft hepatitis. However, there was a trend of more severe recurrent disease in patients with subtype 1b and high viral load soon after transplant.[54]

Another prognostic factor associated with a poor outcome after liver transplant is HCV genotype.[53,55] In nontransplant patients with HCV infection, patients with genotype 1 having a lower rate of sustained virologic response (SVR) to antiviral therapy compared to patients with genotypes 2 or 3 (see chapter 37). The association of genotype and outcome after liver transplantation has been controversial. Earlier studies showed no significant association with genotype 1b and disease severity or graft survival after transplant.[57,58] In contrast, more recent studies have shown that HCV genotype 1b is associated with a higher recurrence of HCV disease 3 years after transplant compared to genotype non-1b.[53] Although there is no uniform agreement regarding the prognostic factors associated with graft and patient survival, HCV RNA testing and genotyping will remain key parameters in future studies.

EBV

EBV-induced posttransplant lymphoproliferative disorders (PTLD) continue to be a rare but severe complication following transplantation. PTLD represents a heterogeneous group of abnormal lymphoid proliferations, generally of B-cell origin, that occur in an environment of ineffective T-cell function. The ineffective T-cell function is the result of pharmacologic immunosuppression after transplant. PTLD represents a spectrum of EBV-related clinical diseases, from benign mononucleosis-like illness to fulminate non-Hodgkin lymphoma. (Refer to chaper 34 for a complete discussion of PTLD.)

The origin of the EBV in PTLD is varied. In SOT patients, the PTLD cells are usually of recipient origin, suggesting that the EBV infection represents reactivation of prior latent virus. Most cases of PTLD in BMT patients involve seropositive donors and recipients, and the PTLD cells are usually of donor origin because the host lymphoid tissue has been eradicated.

EBV-induced PTLD has an incidence of 0.8% to 20% depending on the type of organ transplant, patient age, and the type of immunosuppression.[59] Pretransplantation EBV seronegativity and the development of CMV disease, especially in the donor positive–recipient negative CMV serogroup, are significant risk factors. PTLD occurs more frequently in children than adults, probably because children are more likely to be EBV seronegative (49%) than adults (8%) at the time of transplant. Other risk factors for the development of PTLD include HLA-mismatched T-cell depletion, the use of antilymphocytic antibodies for conditioning, and treatment of graft-vs-host disease.[60–62]

PTLD has a variety of clinical presentations and may be related to viral infection, organ dysfunction, or lymphoma-related symptoms. Early diagnosis of PTLD requires diligent surveillance in appropriate patient populations. The median time to onset of PTLD in SOT patients is 6 months, whereas in BMT patients, signs and symptoms could appear within 70 to 90 days after transplant. Clinical features of PTLD are similar in both SOT and BMT patients, but there appears to be a greater incidence of fulminate, disseminated disease in BMT patients. Given the number of other diagnoses that PTLD can mimic, tissue biopsy is usually required for accurate diagnosis. A variety of special studies have been applied to the diagnosis of PTLD, such as immunophenotyping by flow cytometry, immunohistochemical staining, and molecular tests.[62–64]

Early detection of PTLD may allow for prompt therapy and potentially decreased mortality. Experience with monitoring patients for the development of PTLD is limited; however, several reports suggest that active surveillance for the presence of primary or reactivated EBV infection may be a useful approach to early detection of PTLD.[64–66]

Molecular tests, in particular quantitative PCR assays, are a promising advance that may allow for early diagnosis of PTLD. However, standardized methods and consensus on the appropriate clinical specimens are needed. Since EBV causes a latent infection, PCR assays for the detection of EBV in PBMC will have the same limitations as seen for CMV; that is, EBV DNA will be detected in patients regardless of PTLD status. Quantitative PCR assays have been used to distinguish asymptomatic EBV infection from PTLD. Furthermore, PTLD in SOT patients differs from that in BMT patients with regard to the source of EBV, the time of disease outbreak, and the course of the disease itself. For these reasons, not all patients should be grouped and analyzed together, and broad conclusions from different studies should be drawn cautiously.[64,66,67]

Several studies using quantitative PCR assays on a limited number of patients have shown that high levels of EBV DNA in lymphocytes are associated with the development of PTLD.[68–70] A recent study[68] described the use of a quantitative PCR assay to determine EBV viral load at the time of PTLD diagnosis in a group of allogeneic BMT recipients. The assay quantified between 20 and 10^6 copies of EBV DNA/μg of DNA from peripheral blood lymphocytes (PBL). The EBV viral loads in patients with PTLD were compared with the viral loads of transplant recipients

without PTLD. EBV DNA was detected in 26 of the 59 patients, indicating primary infection or reactivation. The EBV viral load was significantly higher in patients with PTLD compared to those without PTLD. For the nine patients diagnosed with PTLD, the median EBV viral load was 1.6×10^6 copies/μg PBL DNA, compared to 4×10^3 copies/μg PBL DNA in patients without PTLD. Using the different median values in the two patient populations, a threshold value of 10^5 EBV copies/μg PBL DNA for the diagnosis of PTLD was selected, which resulted in both a sensitivity and specificity of approximately 90%. However, in transplant patients with fewer than three major risk factors for PTLD, the positive predictive value of this threshold was low. In addition, there were still a small number of patients with a high viral load that did not develop PTLD during the course of the study.

In an effort to simplify the method needed for EBV detection by PCR, Campe et al.[69] studied the utility of a quantitative PCR assay using sera and PBL in pediatric renal transplant patients. The advantage of using serum is that sample processing does not require the separation and quantitation of PBL from blood. In addition, the authors hypothesized that the detection of EBV DNA in serum would occur only for patients with a high EBV viral load, and might better distinguish patients with PTLD from transplant recipients that have low levels of EBV DNA secondary to immunosuppression, but not PTLD. This prospective study[69] evaluated 25 pediatric renal transplant recipients that were 1 to 8 years after transplant. PBL and serum were collected at each outpatient visit over a 6-month period. A mononucleosis-like syndrome with persistent symptoms of low-grade fever, cervical adenopathy, and hyperplastic adenoids was diagnosed in 11 patients. A total of 38.4% of PBL samples from asymptomatic patients were positive by PCR. PBL samples from two asymptomatic patients were continuously positive for EBV DNA over the period of the study. The mean viral load of positive samples was 480 copies/10,000 PBL. However, 100% of PBL samples from symptomatic patients were positive for EBV DNA with a mean viral load of 1500 copies/10,000 PBL. On the other hand, 14.3% of the serum samples from asymptomatic patients were positive by PCR, with a mean viral load of 13,250 copies/ml serum, while EBV DNA was detected in 81.9% of serum samples of symptomatic patients with a mean viral load of 22,500 copies/ml serum. Detection of EBV DNA in serum always was associated with EBV DNA in the corresponding sample of PBL in symptomatic patients. Thus, positive EBV PCR results in PBL and serum were more likely in symptomatic children. Therefore, patients who have EBV DNA detected simultaneously in PBL and serum may be at higher risk of developing PTLD.

There is some evidence that detection and quantitation of EBV DNA in peripheral blood can be utilized as prognostic markers for development of PTLD, with a correlation between high levels of EBV DNA in the blood and development of PTLD. A study by Kogan-Liberman et al.[70] addressed the value of EBV viral load monitoring and preemptive reduction in immunosuppressive therapy in a group of pediatric liver transplant recipients. The authors prospectively followed monthly EBV viral load in 23 pediatric liver transplant recipients for 18 months after transplantation. A preemptive reduction in immunosuppressive therapy was instituted for significant EBV replication indicated by a positive PCR result. Patients were divided into two groups: group 1 was seropositive for EBV before transplant (13 patients), and group 2 was seronegative for EBV before transplant (10 patients). Nine of the 13 patients in the seropositive group 1 had positive PCR results at a mean time of 22.4 weeks after transplant. All but one of these patients was asymptomatic. In 7 of the 9 PCR-positive patients, preemptive reduction in immunosuppressive therapy was undertaken without development of PTLD or rejection. The two patients in whom immunosuppressive therapy could not be reduced experienced low-grade and medically responsive PTLD. On the other hand, in seronegative group 2 patients, no patient developed a positive EBV viral load, no intervention was necessary, and none of these patients developed PTLD. Since the number of patients enrolled in this study was small, further studies addressing this issue are warranted.

HHV-6

HHV-6 is another herpesvirus of emerging clinical significance. Although overt clinical disease is infrequent in healthy adults, HHV-6 reactivates in immunosuppressed individuals. HHV-6 infection is common in the first 2 years of life, and about 90% of adults are seropositive. HHV-6 can be reactivated following SOT, causing nonspecific febrile illness, bone marrow suppression, interstitial pneumonitis, rash, and hepatitis. The role of HHV-6 in this setting is difficult to determine, since the virus is latent and can be detected in asymptomatic patients, and CMV causes identical syndromes.[71] The reported incidence of HHV-6 infection after transplant varies according to the diagnostic method and ranges from 48% of BMT to 32% of SOT patients.

One of the most interesting features of HHV-6 is its interaction with other viruses. The indirect clinical sequelae of HHV-6 after transplant include altering the net state of immunosuppression with modification of the natural history of CMV disease. Interactions between HHV-6 and CMV have been well documented in liver transplant patients. HHV-6 infection is associated with severe clinical symptoms in SOT patients if concomitant CMV infection occurs. There is evidence that HHV-6 infection in liver or renal transplant patients is an independent risk factor for the development of CMV disease.[72–75]

Laboratory-developed PCR assays have been used to understand the significance of HHV-6 infection in SOT recipients.[74,76,77] One prospective study[74] followed 200 consecutive patients undergoing liver transplantation using a quantitative laboratory-developed PCR assay to detect

HHV-6 DNA in PBL obtained at baseline and weeks 1, 2, 3, 4, 6, 8, 10, and 12, or when clinically indicated. A positive HHV-6 PCR result was obtained in 51.1% of the patients at a median of 27 days after transplant. Peak viral load occurred by day 22 after transplant in 25% of the patients, by day 35 in 50% of the patients, by day 63 in 75% of the patients, and by day 90 in 90% of the patients. The median peak viral load was 200 copies/μg DNA (range 3 to 10,000 copies/μg DNA). To differentiate latent infection from active infection, patients were categorized into two groups, those with peak viral load under 100 copies/μg DNA and those with peak viral load greater than 100 copies/μg DNA. Using this cutoff value, 56 patients had HHV-6 infection. Interestingly, patients on GCV prophylaxis had a lower incidence of HHV-6 infection than those who did not receive prophylaxis (12.9% vs 30.8%, $p = 0.042$). HHV-6-associated disease was documented in only two patients (1%). Both patients had unexplained fever, leukopenia, and thrombocytopenia with negative CMV laboratory results and an HHV-6 viral load greater than 1,000 copies/μg DNA. CMV infection occurred in 40% of the patients, with symptomatic CMV disease in 16.6% of patients. HHV-6 infection was significantly associated with CMV disease. Furthermore, CMV viral load was significantly higher in patients with HHV-6 infection than those without HHV-6 infection (median viral load 1,560 copies/ml (0 to 181,000 copies/ml) vs median viral load undetectable [0 to 14,000 copies/ml], respectively; $p < 0.001$]). Risk factors for development of CMV disease were analyzed in a multivariate logistic regression model, and HHV-6 infection and antilymphocyte globulin administration were the only risk factors identified. In the same study, the association of HHV-6 infection with development of opportunistic infections was evaluated.[74] A total of 74 opportunistic infections, including CMV disease ($n = 32$), invasive fungal infection ($n = 17$), EBV-related PTLD ($n = 6$), disseminated zoster ($n = 11$), and other infections ($n = 8$), were diagnosed. An increase in HHV-6 viral load was significantly associated with the development of opportunistic infections. In addition, HHV-6 infection was associated with rejection in a subgroup of patients whose rejection occurred greater than 30 days after transplantation. Of the 41 patients in this subgroup, the median peak HHV-6 viral load was 40 copies/μg DNA compared to undetectable in those without rejection ($p = 0.004$).

Further studies are needed to better define HHV-6 pathogenesis and its interaction with CMV and other opportunistic pathogens. However, it is clear that molecular methods for the detection and quantitation of HHV-6 will play a crucial role in this area of research, as well as in routine clinical testing in the future.

References

1. Rubin RH. The direct and indirect effects of infection in liver transplantation: pathogenesis, impact, and clinical management. *Curr Clin Top Infect Dis.* 2002;22:125–154.
2. Ljungman P. Beta-herpesvirus challenges in the transplant recipient. *J Infect Dis.* 2002;186(suppl 1):S99–S109.
3. Sissons JG, Carmichael AJ. Clinical aspects and management of cytomegalovirus infection. *J Infect.* 2002;44:78–83.
4. Rubin RH. Cytomegalovirus in solid organ transplantation. *Transpl Infect Dis.* 2001;3(suppl 2):1–5.
5. 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;37:2804–2807.
6. Caliendo AM, Schuurman R, Yen-Lieberman B, et al. Comparison of quantitative and qualitative PCR assays for cytomegalovirus DNA in plasma. *J Clin Microbiol.* 2001;39:1334–1338.
7. Caliendo AM, St George K, Allega J, et al. Distinguishing cytomegalovirus (CMV) infection and disease with CMV nucleic acid assays. *J Clin Microbiol.* 2002;40:1581–1586.
8. Mazzulli T, Drew LW, Yen-Lieberman B, et al. Multicenter comparison of the digene hybrid capture CMV DNA assay (version 2.0), the pp65 antigenemia assay, and cell culture for detection of cytomegalovirus viremia. *J Clin Microbiol.* 1999;37:958–963.
9. Gerna G, Baldanti F, Middeldorp JM, et al. Clinical significance of expression of human cytomegalovirus pp67 late transcript in heart, lung, and bone marrow transplant recipients as determined by nucleic acid sequence-based amplification. *J Clin Microbiol.* 1999;37:902–911.
10. Razonable RR, Brown RA, Wilson J, et al. The clinical use of various blood compartments for cytomegalovirus (CMV) DNA quantitation in transplant recipients with CMV disease. *Transplantation.* 2002;73:968–973.
11. Middeldorp J, Sillekens P, Lunenberg J. Diagnosis of active HCMV infection: the mRNA approach. *Organs and Tissue.* 2000;2:99–107.
12. Weinberg A, Hodges TN, Li S, et al. Comparison of PCR, antigenemia assay, and rapid blood culture for detection and prevention of cytomegalovirus disease after lung transplantation. *J Clin Microbiol.* 2000;38:768–772.
13. Caliendo AM, St George K, Kao SY, et al. Comparison of quantitative cytomegalovirus (CMV) PCR in plasma and CMV antigenemia assay: clinical utility of the prototype AMPLICOR CMV MONITOR test in transplant recipients. *J Clin Microbiol.* 2000;38:2122–2127.
14. Patel R, Smith TF, Espy M, et al. A prospective comparison of molecular diagnostic techniques for the early detection of cytomegalovirus in liver transplant recipients [comment]. *J Infect Dis.* 1995;171:1010–1014.
15. 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;68:1305–1311.
16. Ferreira-Gonzalez A, Fisher RA, Weymouth LA, et al. Clinical utility of a quantitative polymerase chain reaction for diagnosis of cytomegalovirus disease in solid organ transplant patients. *Transplantation.* 1999;68:991–996.
17. Cope AV, Sabin C, Burroughs A, et al. Interrelationships among quantity of human cytomegalovirus (HCMV) DNA in blood, donor-recipient serostatus, and administration of methylprednisolone as risk factors for HCMV disease following liver transplantation. *J Infect Dis.* 1997;176:1484–1490.
18. Oldenburg N, Lam KM, Khan MA, et al. Evaluation of human cytomegalovirus gene expression in thoracic organ transplant recipients using nucleic acid sequence-based amplification. *Transplantation.* 2000;70:1209–1215.
19. Razonable RR, Brown RA, Espy MJ, et al. Comparative quantitation of cytomegalovirus (CMV) DNA in solid organ transplant recipients with CMV infection by using two high-throughput automated systems. *J Clin Microbiol.* 2001;39:4472–4476.
20. Norris S, Kosar Y, Donaldson N, et al. Cytomegalovirus infection after liver transplantation: viral load as a guide to treating clinical infection. *Transplantation.* 2002;74:527–531.

21. Razonable RR, van Cruijsen H, Brown RA, et al. Dynamics of cytomegalovirus replication during preemptive therapy with oral ganciclovir. *J Infect Dis.* 2003;187:1801–1808.
22. Hebart H, Rudolph T, Loeffler J, et al. Evaluation of the NucliSens CMV pp67 assay for detection and monitoring of human cytomegalovirus infection after allogeneic stem cell transplantation. *Bone Marrow Transplant.* 2002;30:181–187.
23. Li H, Dummer JS, Estes WR, et al. Measurement of human cytomegalovirus loads by quantitative real-time PCR for monitoring clinical intervention in transplant recipients. *J Clin Microbiol.* 2003;41:187–191.
24. Sia IG, Wilson JA, Groettum CM, et al. Cytomegalovirus (CMV) DNA load predicts relapsing CMV infection after solid organ transplantation. *J Infect Dis.* 2000;181:717–720.
25. 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;186:829–833.
26. Fisher RA, Saggi BH, Ferreira-Gonzalez A, et al. Prospective validation of quantitative polymerase chain reaction for management of cytomegalovirus disease in solid-organ transplant patients. *Transplantation.* 2002;74:573–576.
27. Chou S. Antiviral drug resistance in human cytomegalovirus. *Transpl Infect Dis.* 1999;1:105–114.
28. Chou S, Erice A, Jordan MC, et al. Analysis of the UL97 phosphotransferase coding sequence in clinical cytomegalovirus isolates and identification of mutations conferring ganciclovir resistance. *J Infect Dis.* 1995;171:576–583.
29. Chou S, Guentzel S, Michels KR, et al. Frequency of UL97 phosphotransferase mutations related to ganciclovir resistance in clinical cytomegalovirus isolates. *J Infect Dis.* 1995;172:239–242.
30. Chou S, Lurain NS, Thompson KD, et al. Viral DNA polymerase mutations associated with drug resistance in human cytomegalovirus. *J Infect Dis.* 2003;188:32–39.
31. Chou S, Miner RC, Drew WL. A deletion mutation in region V of the cytomegalovirus DNA polymerase sequence confers multidrug resistance. *J Infect Dis.* 2000;182:1765–1768.
32. Chou S, Waldemer RH, Senters AE, et al. Cytomegalovirus UL97 phosphotransferase mutations that affect susceptibility to ganciclovir. *J Infect Dis.* 2002;185:162–169.
33. Hamprecht K, Eckle T, Prix L, et al. Ganciclovir-resistant cytomegalovirus disease after allogeneic stem cell transplantation: pitfalls of phenotypic diagnosis by in vitro selection of an UL97 mutant strain. *J Infect Dis.* 2003;187:139–143.
34. Kwak EJ, Vilchez RA, Randhawa P, et al. Pathogenesis and management of polyomavirus infection in transplant recipients. *Clin Infect Dis.* 2002;35:1081–1087.
35. Lin PL, Vats AN, Green M. BK virus infection in renal transplant recipients. *Pediatr Transplant.* 2001;5:398–405.
36. Boubenider S, Hiesse C, Marchand S, et al. Post-transplantation polyomavirus infections. *J Nephrol.* 1999;12:24–29.
37. Vats A, Shapiro R, Singh Randhawa P, et al. Quantitative viral load monitoring and cidofovir therapy for the management of BK virus-associated nephropathy in children and adults. *Transplantation.* 2003;75:105–112.
38. Randhawa PS, Finkelstein S, Scantlebury V, et al. Human polyoma virus-associated interstitial nephritis in the allograft kidney. *Transplantation.* 1999;67:103–109.
39. Hirsch HH, Knowles W, Dickenmann M, et al. Prospective study of polyomavirus type BK replication and nephropathy in renal-transplant recipients [comment]. *N Engl J Med.* 2002;347:488–496.
40. Limaye AP, Jerome KR, Kuhr CS, et al. Quantitation of BK virus load in serum for the diagnosis of BK virus-associated nephropathy in renal transplant recipients [comment]. *J Infect Dis.* 2001;183:1669–1672.
41. Ginevri F, De Santis R, Comoli P, et al. Polyomavirus BK infection in pediatric kidney-allograft recipients: a single-center analysis of incidence, risk factors, and novel therapeutic approaches. *Transplantation.* 2003;75:1266–1270.
42. Leung AY, Chan M, Tang SC, et al. Real-time quantitative analysis of polyoma BK viremia and viruria in renal allograft recipients. *J Virol Methods.* 2002;103:51–56.
43. Azzi A, Cesaro S, Laszlo D, et al. Human polyomavirus BK (BKV) load and haemorrhagic cystitis in bone marrow transplantation patients. *J Clin Virol.* 1999;14:79–86.
44. Iwamoto S, Azuma E, Hori H, et al. BK virus-associated fatal renal failure following late-onset hemorrhagic cystitis in an unrelated bone marrow transplantation. *Pediatr Hematol Oncol.* 2002;19:255–261.
45. Leung AY, Suen CK, Lie AK, et al. Quantification of polyoma BK viruria in hemorrhagic cystitis complicating bone marrow transplantation. *Blood.* 2001;98:1971–1978.
46. Gretch D. Standardization of hepatitis C virus—RNA quantification: advances and unfinished business [comment]. *Hepatology.* 2000; 31:788–789.
47. Chan SE, Rosen HR. Outcome and management of hepatitis C in liver transplant recipients. *Clin Infect Dis.* 2003;37:807–812.
48. Wali MH, Heydtmann M, Harrison RF, et al. Outcome of liver transplantation for patients infected by hepatitis C, including those infected by genotype 4. *Liver Transpl.* 2003;9:796–804.
49. Alonso O, Loinaz C, Abradelo M, et al. Changes in the incidence and severity of recurrent hepatitis C after liver transplantation over 1990–1999. *Transplant Proc.* 2003;35:1836–1847.
50. Berenguer M. Natural history of recurrent hepatitis C. *Liver Transpl.* 2002;8(suppl 1):S14–S18.
51. Firpi RJ, Abdelmalek MF, Soldevila-Pico C, et al. Combination of interferon alfa-2b and ribavirin in liver transplant recipients with histological recurrent hepatitis C. *Liver Transpl.* 2002;8:1000–1006.
52. McCaughan GW, Zekry A. Pathogenesis of hepatitis C virus recurrence in the liver allograft. *Liver Transpl.* 2002;8(suppl 1):S7–S13.
53. Costes V, Durand L, Pageaux GP, et al. Hepatitis C virus genotypes and quantification of serum hepatitis C RNA in liver transplant recipients. Relationship with histologic outcome of recurrent hepatitis C. *Am J Clin Pathol.* 1999;111:252–258.
54. Nuovo GJ, Holly A, Wakely P Jr, et al. Correlation of histology, viral load, and in situ viral detection in hepatic biopsies from patients with liver transplants secondary to hepatitis C infection. *Hum Pathol.* 2002;33:277–284.
55. Gigou M, Roque-Afonso AM, Falissard B, et al. Genetic clustering of hepatitis C virus strains and severity of recurrent hepatitis after liver transplantation. *J Virol.* 2001;75:11292–11297.
56. Feray C, Caccamo L, Alexander GJ, et al. European collaborative study on factors influencing outcome after liver transplantation for hepatitis C. European Concerted Action on Viral Hepatitis (EUROHEP) Group. *Gastroenterology.* 1999;117:619–625.
57. Charlton M, Seaberg E, Wiesner R, et al. Predictors of patient and graft survival following liver transplantation for hepatitis C. *Hepatology.* 1998;28:823–830.
58. Zhou S, Terrault NA, Ferrell L, et al. Severity of liver disease in liver transplantation recipients with hepatitis C virus infection: relationship to genotype and level of viremia. *Hepatology.* 1996;24:1041–1046.
59. Holmes RD, Sokol RJ. Epstein-Barr virus and post-transplant lymphoproliferative disease. *Pediatr Transplant.* 2002;6:456–464.
60. Muti G, Cantoni S, Oreste P, et al. Post-transplant lymphoproliferative disorders: improved outcome after clinico-pathologically tailored treatment. *Haematologica.* 2002;87:67–77.
61. Ho M. Risk factors and pathogenesis of posttransplant lymphoproliferative disorders. *Transplant Proc.* 1995;27(suppl 1):38–40.
62. Ho M, Miller G, Atchison RW, et al. Epstein-Barr virus infections and DNA hybridization studies in posttransplantation lymphoma and lymphoproliferative lesions: the role of primary infection. *J Infect Dis.* 1985;152:876–886.
63. Matsukura T, Yokoi A, Egawa H, et al. Significance of serial real-time PCR monitoring of EBV genome load in living donor liver transplantation. *Clin Transplant.* 2002;16:107–112.

64. Merlino C, Cavallo R, Bergallo M, et al. Epstein Barr viral load monitoring by quantitative PCR in renal transplant patients. *New Microbiol.* 2003;26:141–149.
65. Orentas RJ, Schauer DWJ, Ellis FW, et al. Monitoring and modulation of Epstein-Barr virus loads in pediatric transplant patients. *Pediatr Transplant.* 2003;7:305–314.
66. Limaye AP, Huang ML, Atienza EE, et al. Detection of Epstein-Barr virus DNA in sera from transplant recipients with lymphoproliferative disorders [comment]. *J Clin Microbiol.* 1999;37:1113–1116.
67. Riddler SA, Breinig MC, McKnight JL. Increased levels of circulating Epstein-Barr virus (EBV)-infected lymphocytes and decreased EBV nuclear antigen antibody responses are associated with the development of posttransplant lymphoproliferative disease in solid-organ transplant recipients. *Blood.* 1994;84:972–984.
68. 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 [erratum appears in *J Clin Microbiol.* 2002;40:2316]. *J Clin Microbiol.* 2002;40:351–358.
69. 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 post-transplant lymphoproliferative disease. *Pediatr Transplant.* 2003;7:46–52.
70. 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 [comment]. *J Pediatr Gastroenterol Nutr.* 2001;33:445–449.
71. Griffiths PD, Clark DA, Emery VC. Betaherpesviruses in transplant recipients. *J Antimicrob Chemother.* 2000;45(suppl T3):29–34.
72. Boutolleau D, Fernandez C, Andre E, et al. Human herpesvirus (HHV)-6 and HHV-7: two closely related viruses with different infection profiles in stem cell transplantation recipients. *J Infect Dis.* 2003;187:179–186.
73. Kidd IM, Clark DA, Sabin CA, et al. Prospective study of human betaherpesviruses after renal transplantation: association of human herpesvirus 7 and cytomegalovirus co-infection with cytomegalovirus disease and increased rejection. *Transplantation.* 2000;69:2400–2404.
74. Humar A, Kumar D, Caliendo AM, et al. Clinical impact of human herpesvirus 6 infection after liver transplantation. *Transplantation.* 2002;73:599–604.
75. Gautheret-Dejean A, Manichanh C, Thien-Ah-Koon F, et al. Development of a real-time polymerase chain reaction assay for the diagnosis of human herpesvirus-6 infection and application to bone marrow transplant patients. *J Virol Methods.* 2002;100:27–35.
76. Ljungman P, Wang FZ, Clark DA, et al. High levels of human herpesvirus 6 DNA in peripheral blood leucocytes are correlated to platelet engraftment and disease in allogeneic stem cell transplant patients. *Br J Haematol.* 2000;111:774–781.
77. Nitsche A, Muller CW, Radonic A, et al. Human herpesvirus 6A DNA Is detected frequently in plasma but rarely in peripheral blood leukocytes of patients after bone marrow transplantation. *J Infect Dis.* 2001;183:130–133.

Chapter 39

Viral Infections of the Central Nervous System

Christine C. Ginocchio

Introduction

Viral infections of the central nervous system (CNS) are relatively infrequent and usually result in a benign, self-limiting disease.[1–4] However, in a small percentage of cases, viral infection of the CNS can have extremely serious consequences that result in a spectrum of permanent neurologic damage or death. Viral agents gain access to the CNS by either neuronal or hematogenous spread, and infections can occur at a multitude of sites including the spinal cord, leptomeninges, dorsal nerve roots, nerves, and brain parenchyma. Viral CNS infections are classified clinically as either meningitis or encephalitis, although a close interrelationship occurs between the two disease states.[1–4] Host factors (age, sex, immune status, genetic differences) and viral factors (serotype, receptor preference, cell tropism, viral load) in concert with geographic and seasonal factors contribute to the potential for the development of CNS disease.[1–4]

Epidemiology

Viral meningitis, meningoencephalitis, and encephalitis frequently occur in epidemics and as seasonal outbreaks (late spring through autumn) and account for the majority of viral CNS infections.[1–4] The most common viral agents responsible for CNS disease are listed in Table 39-1. In the United States and in countries that immunize against mumps, enteroviruses account for approximately 80% to 92% of all cases of meningitis in which a causative agent is identified.[3,4] In countries that do not include mumps in their immunization programs, mumps account for as many as 30% of viral meningitis cases.[5] The arthropod-borne viruses (arboviruses) (Table 39-2) account for the majority of the remaining cases in the United States.[6–8]

Encephalitis occurs at a lower frequency, with arboviruses being the leading cause both worldwide and in the United States.[6–8] The flaviviruses and alphaviruses are the most important and are responsible for the majority of mosquito and tick-borne encephalitis that cause epidemic and endemic disease in Asia, Europe, and the Americas.[6–8] Case fatality rates vary greatly, ranging from 5% to 70%. Human immunodeficiency virus (HIV) and rabies (genus *Lyssavirus*) are other important causes of encephalitis worldwide.[1,2] Herpes simplex virus (HSV) CNS infections affect all ages, occur at all times of the year, and have the highest mortality rate in the United States.[1–4] In a small number of cases, childhood viral diseases including rubella, measles, varicella-zoster virus (VZV), mumps, and human herpes virus 6 (HHV-6), can progress to neurologic disease during primary infection or with viral reactivation (VZV, HHV-6). Cytomegalovirus (CMV)[9–11] and JC polyomavirus (JCV)[12] are responsible for encephalitides in immunocompromised patients.

Overview of Diagnostic Testing for Viral CNS Disease

Traditionally, the diagnosis of viral CNS infection was based on laboratory findings in conjunction with patient history, clinical manifestations, and geographic and epidemiologic factors.[1–4,7,8] The diagnosis of viral encephalitis requires supplemental tests, including lumbar puncture, radiographic imaging such as computed tomography (CT) scans and magnetic resonance imaging (MRI) studies, and in some cases brain biopsy.

In general, patients with viral meningitis typically have a cerebrospinal fluid (CSF) pleocytosis with 10 to 500 leukocytes, a slightly elevated protein (<100 md/dL), and a glucose level greater than 40% of a simultaneously drawn serum sample.[1,2] However, there can be a tremendous range of values that often overlap with those indicative of bacterial meningitis. Patients with viral encephalitis or meningitis may or may not exhibit a CSF pleocytosis.[13] Therefore, normal CSF cell counts should not be used as a sole exclusion criterion for both diagnosis and determining which samples should be tested by molecular methods for viral pathogens. CSF glucose levels can be normal or low, as seen

Table 39-1. Primary Viral Agents Causing Central Nervous System Disease

Herpesviruses	Herpes simplex 1 and 2
	Varicella-zoster
	Epstein-Barr
	Cytomegalovirus
	Human herpesvirus 6
Enteroviruses	Poliovirus
	Coxsackievirus
	Echovirus
	Numbered enteroviruses
Arboviruses	See Table 39-2
Childhood illness associated	Measles
	Mumps
	Rubella
Rabies	
JC polyoma virus	
Human immunodeficiency virus type 1	

with bacterial meningitis. Encephalitis can lead to hemorrhagic necrosis with elevated protein levels and the presence of red blood cells.

Cultures for bacterial, fungal, and mycobacterial pathogens can aid in the diagnosis of viral CNS infections but can take several weeks for definitive results. Bacterial cultures may be falsely negative if patients have been treated with antibiotics prior to sample collection. Traditionally, the identification of the specific viral agent relied on viral culture, serologic detection of virus-specific IgM and IgG antibodies, either systemic or intrathecal, or both.[3,4,14–16] However, the ability to isolate the viral agent is highly dependent on the viral species, time of sample collection, sample handling and processing, and treatment of the patient with antiviral agents. In only approximately 10% to 16% of viral meningitis cases is the responsible agent identified using viral culture.[3–5,14–17] Serologic diagnosis can be made in certain cases as early as 5 days after infection with the development of specific IgM antibodies but often takes weeks to demonstrate a diagnostic rise in IgG antibody titers between acute and convalescent serum samples.[3,4,12,14–16,18,19] Some patients with West Nile viral infections have detectable IgM antibodies for up to 500 days after infection, making the distinction between recent and past infection difficult.[19] In addition, immune status of the patient can affect the development of viral-specific antibodies.

Currently, molecular amplification methods that detect viral pathogens in CSF play a critical role in the rapid and accurate diagnosis of viral CNS infections.[14,15,17] This approach has largely abrogated the need for brain biopsy for the diagnosis of encephalitis and, in cases where biopsy is needed, can be used to detect viral pathogens in tissue specimens. In most cases, CSF is easily obtained, especially compared to brain biopsy. CSF should be stored frozen, preferably at −70°C, to maintain the stability of viral nucleic acids. Isolation of the nucleic acids is achieved using a variety of methods (described in chapter 2) that also remove amplification inhibitors and neutralize any DNases or RNases present in the sample. The volume of CSF required can vary significantly and is dependent on the viral target, relative levels of virus present in CSF, and sensitivity of the testing method.

Results can be available within 24 hours and as soon as 2 hours for applications utilizing real-time technologies that incorporate amplification and detection in one step. This is in contrast to viral culture and serology, which can require up to 28 days for a final result. The rapid identification of the correct viral agent directs the selection of appropriate antiviral therapy, decreases the number of patients unnecessarily placed on empiric antibiotic therapy, shortens length of hospitalization, and saves medical costs. Overall, molecular amplification assays are highly sensitive and, depending on the virus and amplification target, can detect as low as one viral particle per reaction. High assay specificity is obtained by gene-specific targeting and often is significantly better than serologic testing, which can demonstrate cross-reactivity among related viral agents. Multiplex assays offer the ver-

Table 39-2. Arthropod-Borne Viruses (Arboviruses) and Endemic Areas

Family	Genus	Virus	Predominant Regions
Togaviridae	*Alphavirus*	Eastern equine	Eastern, southern United States; Canada; Central, South America
		Western equine	Western, central United States; Central, South America
		Venezuelan equine	Central, South America; eastern, western United States
Flaviviridae	*Flavivirus*	St Louis	United States; South America
		Powassan	Canada; United States; Russia
		Tick-borne	Central Europe; Russia
		Japanese	Asia; USSR; India; Sri Lanka
		Murray Valley	United States
		West Nile	North America; Europe; Middle East; Africa; Asia; Australia; Oceania
Bunyaviridae	*Bunyavirus*	California group*	California; upper Midwest; West Virginia; Virginia; Kentucky; Tennessee; North Carolina; Alabama
Reoviridae	*Coltivirus*	Colorado tick fever	Western mountain United States
Rhabdoviruses	*Lyssavirus*	Rabies	Worldwide
		Vesicular stomatitis virus	

*La Crosse, Cache Valley, James-town Canyon, and snowshoe hare.

satility of screening for several pathogens in one test. Quantitative assays are useful for differentiating active from latent infection and for monitoring response to antiviral therapy.

This chapter reviews the advances in the molecular diagnosis of the most common causes of viral meningitis and encephalitis, including the enteroviruses, herpesviruses, and arboviruses. Viral agents affecting persons with immune suppression are briefly discussed, including human immunodeficiency virus type 1 (HIV-1), JC polyomavirus (JCV), and CMV.

ENTEROVIRUSES

Epidemiology and Disease

Enteroviruses are small single-stranded RNA viruses that comprise 64 serotypes distributed into five species of the Picornaviridae family.[20,21] The nonpolio enteroviruses, including the coxsackieviruses, echoviruses, and numbered enteroviruses, are responsible for approximately 50 million infections per year in the United States and possibly more than a billion worldwide.[4,21,22] Enteroviruses cause a diverse array of illnesses in both adults and children, including respiratory, ocular, cardiac, gastrointestinal, and neurologic diseases and skin and oral eruptions.[21,22] In the neonate, enteroviruses can cause a sepsislike picture or meningoencephalitis, which can be severe. Outside of the neonatal period, children under 5 years of age appear to be the most susceptible to infection, partly due to a lack of acquired immunity and poor hygienic habits. Encephalitis is uncommon and meningitis is rarely associated with complicated disease and poor clinical outcome, although it may be more severe in adults than in children. Although enteroviral infections can occur year-round, the majority of the infections occur during the summer through autumn months.

Laboratory Diagnosis

Many enteroviruses can be cultured in human and primate cell lines.[21] No single cell line is optimal for all enteroviral types, and therefore several different susceptible culture cell lines are used. Viral CSF culture has a sensitivity of approximately 65% to 75%, in part because of the lability of the virus, possible low levels in the CSF, and the inability to grow all enteroviral serotypes in tissue culture, including several coxsackievirus A strains that require mouse inoculation for detection. The results of viral culture can take 3 to 8 days and therefore are generally not rapid enough to affect either treatment options or length of hospitalization. Patients are placed on unnecessary antibiotic therapy until bacterial CSF cultures are negative at 48 to 72 hours. A variety of serologic assays can be used to diagnose enteroviral infections, but they can be cross-reactive, nonspecific, and difficult to interpret due to the extended incubation and prodromal periods found with many enteroviral illnesses, and hence are not clinically useful.

Molecular Tests

To increase the sensitivity of enterovirus detection and to reduce the time to results, a variety of molecular tests have been developed.[17,23–35] The methods utilize either reverse transcription–polymerase chain reaction (RT-PCR) amplification combined with enzyme-linked immuno sorbent assay (ELISA),[23–28,31] real-time RT-PCR[30] or nucleic acid sequence–based amplification (NASBA) combined with either electrochemiluminescence (ECL) detection[29,32,34] or molecular beacon technology.[33,35] The assays utilize primer sequences selected from the conserved 5′ nontranslated region of the enteroviral genome. Sequence variations are present in this region, and the scope of detection of the various enteroviral serotypes is dependent on primer selection.[23–35] Since different serotypes circulate in distinct areas of the world, assays should be validated using local clinical isolates and reference strains known to be endemic in the region. Overall, enterovirus molecular tests are highly sensitive, some detecting as low as 0.1 tissue culture infectious dose 50 ($TCID_{50}$) and the majority of enteroviral isolates.[23–35] The tests also are highly specific, showing no cross-reactivity with other viral agents, with the exception of a rare rhinovirus.[30,31,34]

Currently, analyte-specific reagents (ASRs) are available for enteroviral detection from CSF. The Enterovirus Consensus assay (Argene Biosoft, Varilhes, France) utilizes RT-PCR and a stair primer technology that was designed to overcome the problems associated with sequence divergence in the primer binding regions, thus ensuring the detection of all 64 enteroviral serotypes.[31] The assay was shown to be more sensitive than assays using the original Rotbart[23] and Zolli[25] primers, and exhibited cross-reactivity with rhinovirus type 3 only. Studies have determined that the sensitivity of the assay is <0.4 $TCID_{50}$ and can detect approximately 6 RNA copies per input reaction. A NucliSens Basic Kit assay (bioMérieux, Boxtel, the Netherlands) that combines NASBA and ECL detection shows reactivity with all enteroviral isolates tested and a sensitivity in the range of 10 to 100 RNA copies.[29,32,34] Minimal low-level cross-reactivity has been detected for one rhinovirus isolate when tested at high titers. No cross-reactivity was demonstrated for other viral isolates. The assay is 18% to 32% more sensitive than CSF viral culture and 100% specific.[29,32,34] This assay has been modified for rapid real-time detection using molecular beacons and is currently available as an ASR (bioMérieux, Durham, NC).[33,35] This modification has reduced the time to results from 5.5 hours to as little as 3.5 hours, thus permitting multiple test runs per day. An additional enterovirus ASR real-time RT-PCR assay, developed for

use on the SmartCycler and GeneXpert instruments (Cepheid, Sunnyvale, CA), is currently under evaluation.

HERPESVIRUSES

The herpesviruses, including herpes simplex virus types 1 and 2 (HSV-1, HSV-2), VZV, CMV, and HHV-6, cause a broad spectrum of viral CNS disease including meningitis, encephalitis, meningoencephalitis, myelitis, and polyradiculitis.[1,2] Epstein-Barr virus (EBV) DNA has been detected in virtually all AIDS-related cases of primary CNS lymphomas.[36] CMV is discussed briefly under diseases associated with immunosuppression and in detail in chapter 38.

HERPES SIMPLEX

Epidemiology and Disease

HSV-2 is the usual cause of HSV meningitis and accounts for approximately 1% to 5% of all cases of viral meningitis and 4% to 6% of cases of viral encephalitis.[1–4,37,38] The clinical course of the disease usually is self-limiting and generally not associated with permanent neurological damage. HSV-2 also causes Mollaret meningitis, a benign recurrent meningitis.[39] HSV-1 is responsible for the majority of HSV-associated encephalitis and for about 10% to 20% (approximately 1250 cases per year in the United States) of all viral encephalitis cases. Approximately 30% are primary HSV-1 infection, and the remaining cases are due to HSV reactivation. In contrast to HSV meningitis, HSV encephalitis can be a devastating disease, presenting with focal neurologic disease resulting from damage to one or both temporal lobes.[1–4,37,38] Encephalitis occurs in approximately 60% to 75% of babies with disseminated HSV disease, and without appropriate antiviral therapy the mortality rate can be as high as 80%.[40]

Diagnosis

CSF viral cultures are positive only in approximately 15% of the cases of HSV meningitis due to primary infection and, with the exception of neonatal infection, are rarely positive in cases of HSV encephalitis.[1–4,37,38,40] Traditionally, the diagnosis of HSV encephalitis was dependent on brain biopsy for obtaining tissue for viral culture, electron microscopy, immunohistochemical staining, and the demonstration of intrathecal production of HSV-specific antibodies.[1–3,14–16,38,40] Over the last decade, studies comparing HSV CSF PCR with brain biopsy, intrathecal HSV-specific antibody production for the diagnosis of HSV CNS infections, or both demonstrated sensitivities and specificities for PCR ranging from 96% to 98% and 96% to 99%, respectively.[14,40–42] HSV PCR can identify infected persons who fail to either seroconvert in primary infection or demonstrate significant rises in antibody titers after reactivation. Finally, HSV PCR can aid in differentiating HSV antibody increases due to active HSV CNS infection and reactivation in persons with concomitant CNS disease not related to HSV. Another advantage of PCR assays for diagnosis of HSV encephalitis is that the assay sensitivity remains high (95%) up to approximately 1 week after initiating therapy.[42] Since DNA can be slow to clear from the CSF (as many as 21% of CSF specimens can remain PCR positive for more than 15 days after completion of therapy), PCR should not be used as a test of cure.[42]

VARICELLA-ZOSTER

Epidemiology and Disease

The most common manifestation of VZV resulting from primary infection is chickenpox and generally occurs in early childhood.[43] VZV-associated neurologic syndromes include acute cerebella ataxia, diffuse or focal encephalitis, meningitis, transverse myelitis, and Reye syndrome.[1,2,43] Herpes zoster, resulting from the reactivation of latent VZV infection, occurs in approximately 20% of immunocompetent persons, generally 45 years or older, and the incidence is approximately 15% higher in persons with immunosuppression.[43,44] In a small percentage of these cases, severe complications, including meningitis, encephalitis, myelitis, Ramsay Hunt syndrome, Guillian-Barré syndrome, and contralateral hemiplegia can occur.[43–45]

Diagnosis

The recovery of VZV in culture is poor (20% positive), as is serologic diagnosis (48% positive).[46] Antigen detection by immunofluorescence improves detection (82% positive), and PCR is the most sensitive, with a detection rate of approximately 95%.[46] Based on improved detection with PCR assays, VZV accounts for 6% to 30% of all herpesviruses isolated from CSF.[46–54] However, results need to be interpreted cautiously since VZV DNA has been detected in CSF without overt disease, particularly in patients with immunosuppression.

HUMAN HERPES VIRUS 6

Epidemiology and Disease

HHV-6 is generally acquired early in life and is manifested as exanthem subitum, more commonly known as roseola or as a nonexanthemous febrile illness, sometimes accompanied by severe neurologic manifestations, including febrile seizures, meningitis, meningoencephalitis, and encephalitis.[55] Detection of HHV-6 in children with a febrile seizure is important for identifying children with a potential for the recurrence of seizure episodes due to ongoing HHV-6 infection. Immunocompromised patients

can manifest severe postprimary infections, including encephalitis due to HHV-6 reactivation.[55]

Diagnosis

Culture confirmation of HHV-6 infection is not routinely performed in clinical virology laboratories since it requires purification and culture of patient lymphocytes or cocultivation of activated patient lymphocytes and activated human umbilical cord blood lymphocytes.[55] Serologic confirmation of the disease depends on the detection of IgM in primary infection or conversion from either a negative to a positive IgG antibody response or a four-fold or greater rise in IgG antibody titer.[55] Antibody titers can be difficult to interpret for several reasons. IgM antibodies may not develop in some children, can be positive with HHV-6 reactivation, and can remain positive for extended periods of time. Significant rises in HHV-6-specific IgG antibody titers can be found during infections with other herpesviruses. The detection of HHV-6 DNA in the CNS by PCR can support the diagnosis; however, in a percentage of cases, HHV-6 DNA can be found in normal brain tissue and in the CSF of children without evidence of CNS disease.[55] Quantitative methods may be more accurate for correlating the presence of HHV-6 with active disease.[56,57]

Molecular Tests for Herpesviruses

Comparisons of the sensitivity of HSV PCR to viral culture, antibody detection, and direct immunofluorescence assays have clearly established the utility of PCR as a first-line diagnostic test in the clinical laboratory.[5,14,15,17] Most PCR assays are applicable to the majority of conventional or real-time PCR instruments. The assays for herpesviruses use a variety of primer sets targeting a number of genes including: DNA polymerase gene (HSV-1, HSV-2, CMV, EBV, VZV, HHV-6); UL42, glycoprotein B (gB), glycoprotein D (gD), or thymidine kinase (TK) genes (HSV); genes 28 and 29 (VZV); and major capsid protein, U89/U90 (HHV-6). Identification of the specific herpesvirus is achieved using either traditional methods, including the use of species-specific primers, hybridization with species-specific probes, agarose gel electrophoresis, restriction enzyme analysis, and ELISA-based colorimetric detection methods,[46–49,51,52] or fluorescent probe technologies that permit real-time detection.[50,53,54] Currently, no test kits are approved or cleared by Food and Drug Administration (FDA) for the detection of herpesviruses; however, several commercial kits or ASRs are available for the detection of herpesviruses in CSF. Roche Diagnostics (Indianapolis, IN) has developed ASRs for HSV-1, HSV-2, VZV, and EBV identification using the LightCycler instrument.[50,53,54] A single-tube amplification and detection step allows the identification of both HSV-1 and HSV-2 and the differentiation of subtype by melting-curve analysis.[50] The manufacturer's claim of sensitivity is approximately 12.5 copies or 2.5 to 6.3 genome equivalents per reaction. Artus Biotech (Artus GmbH, Hamburg, Germany) has real-time PCR assays (Real/Art LC PCR) for the detection of HSV-1, HSV-2, EBV, and VZV, for individual analytes (EBV, VZV) or as multiplex tests (HSV-1 and HSV-2, VZV and EBV, and HSV-1, HSV-2, VZV, and EBV). Each kit includes all PCR reagents, primers, probes, internal amplification control, and, for HSV, external calibration controls for quantitation. According to the manufacturer, the detection limits (95% probability) of the assays are 1 copy/mL for HSV, 0.7 copies/mL for VZV, and 0.2 copies/mL for EBV.

Multiplex assays that can screen simultaneously for the six major herpesviruses (HSV-1, HSV-2, CMV, VZV, EBV, HHV-6) offer some distinct advantages, including the detection of coinfections in the CNS and the identification of a herpesvirus that may not have been considered in the original diagnosis.[58] Several laboratory-developed assays have been validated for this purpose using either traditional or real-time PCR methods,[14,47–54] and currently two commercial ASR multiplex systems that use standard platform PCR and screen for all six herpesvirus are available.[48] The Herpes Mplex (Prodesse Inc, Waukesha, WI) includes primers, probes, internal and external controls, and all PCR reagents, with the exception of amplification enzyme. The Herpes Generic Consensus kit (Argene Biosoft, Varilhes, France) uses stair primers to compensate for sequence divergence in the primer binding regions of the six herpesviruses.[48] After PCR amplification and generic colorimetric detection of the herpesvirus, the Herpes Identification Hybridomawell kit (Argene Biosoft, Varilhes, France) is used to identify the specific herpesvirus present. All primers, probes, an internal control, external controls, and all PCR reagents, with the exception of amplification enzyme, are provided. Assay sensitivity varies between 5 and 50 copies per PCR reaction, with EBV being the least sensitive. Another Argene Biosoft kit (Herpes Simplex Consensus 1/2) is available to screen for just HSV-1 and HSV-2.

ARBOVIRUSES

Epidemiology and Disease

Alphaviruses, flaviviruses, and bunyaviruses are responsible for the majority of the arboviral encephalopathies (Table 39-2) and are found throughout the world.[1,2,6] The encephalitides are zoonotic, and transmission of the viruses occurs between susceptible amplifier vertebrate hosts, such as birds, squirrels, and chipmunks, via blood-feeding arthropods, including mosquitoes, psychodids, ticks, and ceratopogonids. Peak incidence occurs in summer and continues through the fall months, when arthropods are most active. In warmer regions cases can occur into the winter months. Humans, horses, and domestic animals can develop a clinical illness. The majority of human arboviral infections are either asymptomatic or manifest as a flulike self-limiting illness. Occasionally viruses enter the brain by mechanisms not clearly

understood, and productive infection of brain cells occurs, resulting in encephalitis, often with a fatal outcome or permanent neurologic damage.[1,2,6]

The mosquito-borne viruses were first discovered in the 1930s.[6] Until 1999 and the emergence of West Nile virus (WNV) in Queens, New York,[59] the most important cause of epidemic and endemic encephalitis in the United States was St Louis encephalitis virus (SLE), which is found throughout the lower 48 states. Confirmed cases of SLE average approximately 193 per year, and 4437 cases have been reported to the Centers for Disease Control and Prevention (CDC; Atlanta, GA), since 1964.[7] Less than 1% of the cases are clinically apparent, with a case fatality rate of 5% to 25% for symptomatic persons.[7] Since the initial outbreak in 1999, WNV has migrated across the United States, and by August 2005 all but three states (Alaska, Hawaii, and Washington) have reported bird, mosquito, vertebrate, or human cases.[7,8] WNV is responsible for neuroinvasive disease (encephalitis and meningitis) and what is termed West Nile fever, which is typically less severe and does not show signs of neuroinvasion.[59] As of August 2005, 17,039 WNV cases with 674 deaths have been reported to the CDC.[7,8] Approximately one in 150 WNV cases progresses to meningitis or encephalitis, the most common neurological manifestation. Case fatality rates remain constant at approximately 12%, with advanced age being the most important risk factor for death.[6,7,8] In addition, WNV has been transmitted to persons receiving either transfusions or transplanted organs from asymptomatic donors infected with WNV.[8,59,60] La Crosse virus (LAC) is found in several midwestern and mid-Atlantic states, and an average of 70 cases are reported per year, usually in children under the age of 16. Since 1964, eastern equine encephalitis (EEE) virus and western equine encephalitis (WEE) virus have been responsible for 200 and 639 cases, respectively.[7] In the northern United States, Powassan virus is a minor cause of encephalitis and is transmitted by ticks.[7]

Japanese encephalitis (JE) virus, found mainly in Asia and the Pacific, is the leading cause of encephalitis worldwide, with more than 45,000 cases reported annually.[1,2,6] Venezuelan encephalitis virus remains the third most common cause of equine encephalitides and the most common cause of epidemic encephalitis in Central and South America.[1,2,6]

Diagnosis

By the time persons present with encephalitic symptoms, the viremic phase of the disease often has ceased and infected persons have developed serum and intrathecal IgM or IgG antibodies or both.[6,59] Since IgM antibodies do not cross the blood-brain barrier, the identification of virus-specific IgM antibodies in the CSF is strongly suggestive of CNS infection. A variety of assays including hemagglutination inhibition, complement fixation, plaque reduction neutralization test (PRNT), IgM antibody capture ELISA (MAC-ELISA) and IgG antibody ELISA, indirect immunofluorescence (IFA), and antigen capture ELISA have been useful in the identification of arboviral infections.[6,59,61,62] Confirmation of the virus-specific IgM antibodies or the demonstration of a four-fold or greater rise in neutralizing IgG antibodies in either CSF or serum is achieved by performing PRNT with multiple arboviruses. Older serologic assays were technically demanding, lacked reproducibility, and were often poor measures of early IgM production. Recently developed MAC-ELISAs are more reliable, rapid and reproducible.[59,61,62] In persons with immunosuppressive disorders or who are very early in the course of the infection, antibody titers can be negative and the diagnosis is dependent on isolation of the virus. However, due to the transient nature of the viremic stage and probably low levels of virus in the CSF and serum, virus isolation using cell culture generally has been unsuccessful and not practical for most clinical laboratories.

Although serological confirmation of disease remains the gold standard, molecular tests are important for identifying early infection prior to seroconversion, infected transplant tissues and blood,[59,60] and infection in immunocompromised patients, and for diagnosis confirmation.[59,63–66] A positive result with a nucleic acid amplification assay is diagnostic of an arboviral infection; however, a negative result does not preclude the possibility of an arboviral infection. For example, up to 55% of CSF samples and only approximately 10% of serum samples are positive in patients with serologically confirmed cases of WNV disease.[59] The molecular detection of arboviral nucleic acids is standard for vector-borne disease surveillance studies and control programs[59,63,64,66,67] and for monitoring both the blood supply and transplantation tissues.[59,60]

Molecular Tests

Several assays using traditional RT-PCR, nested RT-PCR, Taqman-based RT-PCR, and NASBA combined with either ECL or molecular beacon detection have been developed for the identification of arboviral infections.[59,63–67] Assay methods are highly sensitive, specific, and rapid, with results available in as little as 2 hours for real-time RT-PCR and NASBA molecular beacon assays. Primers usually target the NS1, NS3, NS5, and 3′ NC regions of the genomes of the flaviviruses. Primer probe sets may be virus specific,[63,64,66] or universal primers can be used for the detection of all flaviviruses.[65] Studies by Lanciotti et al., Division of Vector-Borne Infectious Diseases, CDC (Fort Collins, Colorado), have clearly defined the status of molecular testing for arboviral infections in both humans and surveillance populations, including mosquito pools, dead birds, horses, and other vertebrates.[63,64,66] The results of Lanciotti and colleagues' evaluations of laboratory-developed RT-PCR, Taqman RT-PCR, and NASBA-based assays for the detection of WNV, EEE, SLE, and WEE viruses are summarized in

Table 39-3. Sensitivities of the RT-PCR, Taqman, NASBA ECL, and NASBA Molecular Beacon Assays for the Detection of WN, SLE, EEE, and WEE Viral Culture Isolates (PFU detected*)

Assay	WNV	SLE	EEE	WEE
RT-PCR	1.00	0.15	0.77	35
Taqman	0.10	0.15	0.77	0.35
NASBA-ECL	0.01	0.15	0.77	3.5
NASBA-MB	0.10	0.15	ND	ND

Sources: References 62 and 64.
*Number of plaque-forming units (PFU) detected by assay.
NASBA, nucleic acid sequence–based amplification; ECL, electrochemiluminescence; MB, molecular beacon; WNV, West Nile virus; SLE, St Louis encephalitis virus; EEE, eastern equine encephalitis virus; WEE, western equine encephalitis virus; ND, not done; PFU, plaque-forming units.

Tables 39-3 and 39-4. The Taqman and NASBA assays are currently used by the US Public Health Laboratories for both human diagnosis and surveillance studies. ASR for performing real-time RT-PCR (Real/Art WNV RT-PCR) are available from Artus Biotech with a manufacturer's claim of detection sensitivity of 5 to 20 nucleic acid copies per reaction. Prodesse has two ASR reagents available, one using standard RT-PCR and one using real-time RT-PCR, with limits of detection of 50 and 500 copies/mL, respectively, as per the manufacturer. Full assessment of the Artus Biotech and Prodesse assays is currently limited by the lack of sufficient clinical data.

VIRAL AGENTS ASSOCIATED WITH IMMUNOSUPPRESSION

CYTOMEGALOVIRUS

Epidemiology and Disease

CMV disease may be due to primary infection but generally is related to CMV reactivation in association with progressive immune deficiency.[9–11] Neurological syndromes associated with CMV include peripheral neuropathy, ventriculoencephalitis, myelitis/polyradiculopathy, and diffuse micronodular encephalitis with dementia.[9–11]

Diagnosis

The presumptive diagnosis of CMV CNS disease has been based primarily on clinical presentation, serologic testing, neuroradiologic studies including CT and MRI, CSF chemistries, and the magnitude and types of CSF pleocytosis.[9–11] CSF viral culture is positive in only approximately 50% of patients with CMV CNS disease, and culture is relatively slow, requiring between 1 and 4 weeks for detection.[9–11]

Molecular Tests

Various molecular methods have been used to more accurately and rapidly detect CMV in patients with active CMV CNS disease.[68] Several studies have determined that the detection of CMV DNA by PCR in CSF of aquired immunodeficency syndrome (AIDS) patients was highly sensitive (100% negative predictive value for CNS disease) but did not always correlate with active CMV CNS disease (68.4% positive predictive value for CNS disease).[68] Sensitive qualitative PCR may detect nonreplicating virus, which could clearly limit the use of these tests in the differentiation of asymptomatic infection from active CNS disease. Quantitative PCR assays with established cut-off values for differentiating infection from active disease may provide an alternative superior to qualitative PCR. For example, CSF from AIDS patients with autopsy-proven CMV encephalitis had a median value of 3333 CMV genomes/10^5 cells compared to a median value of 125/10^5 cells for AIDS patients with neurologic symptoms not related to CMV, and a median value of 19/10^5 cells for HIV-1 seronegative controls.[68] High levels of cellular CMV correlated with marked histopathologic changes in AIDS patients.

Table 39-4. Sensitivities of Vero Cell Culture, Taqman Assay, NASBA-ECL Assay, and RT-PCR Assay for the Detection of EEE, WEE, and WNV from Mosquito Pools and WNV from Human CSF Specimens

Mosquito Pools

	Number Tested	Viral Culture Pos	Taqman Pos/Equiv	NASBA-ECL Pos/Equiv	RT-PCR Pos/Equiv
EEE	20	4	9	7	4
WEE	20	0	3	2	3
WNV	68	32	31/2	34	24/7

CSF Samples

	Number Tested	Serology Pos	Taqman Pos/Equiv	NASBA ECL Pos/Equiv	RT-PCR Pos/Equiv
WNV	10	10	4/3	7/0	0/0
WNV	10	0	0	0	0

Sources: References 62 and 64.
NASBA, nucleic acid sequence-based amplification; ECL, electrochemiluminescence; WNV, West Nile virus; SLE, St Louis Encephalitis virus; EEE, eastern equine encephalitis virus; WEE, western equine encephalitis virus; ND, not done; Pos, positive; Equiv, equivocal.

Additional studies are necessary to determine the level of CMV DNA that would be considered indicative of active CNS disease.

Several PCR-based commercial test kits, including the Amplicor Cytomegalovirus test (Roche Molecular Systems, Branchburg, NJ), PrimeCapture CMV DNA Detection Plate System (Synthetic Genetics, San Diego, CA), Herpes Mplex (Prodesse, Inc), and Herpes Generic Consensus kit (Argene, Inc), have been used to detect human CMV in CSF.[64] In addition, the Amplicor CMV Monitor test, the Digene CMV Hybrid Capture assay (Digene Corp, Gaithersburg, MD), and quantitative CMV real-time DNA-based PCR assays developed for use with the LightCycler instrument are available to quantify CMV DNA from a variety of clinical specimens.

A different approach, used in both transplantation recipients and HIV-1-infected persons for diagnosing active disease versus asymptomatic CMV infection, is based on detection of CMV mRNA transcripts (immediate-early and early, UL83, pp150, pp67).[68,69] The NucliSens pp67 assay (bioMérieux, Boxtel, the Netherlands) detects the presence of mRNA encoding for the phosphorylated matrix tegument protein pp67, one of the most abundant late gene transcripts that is detectable when viral replication is occurring. The detection of pp67 mRNA in CSF was found to be the most accurate in diagnosing active CMV CNS disease (100% positive predictive value [PPV], 97% negative predictive value [NPV]), when compared to viral culture (100% PPV, 82.7% NPV) and qualitative DNA PCR (68.4 % PPV, 100% NPV).[69] In addition to diagnostic utility, quantitative PCR and the NucliSens pp67 assay have been effective for monitoring response to therapy for CMV infections.[68]

HUMAN JC POLYOMAVIRUS

Epidemiology and Disease

Primary infection with human JCV generally occurs early in life, is usually asymptomatic, and results in a latent infection of both renal tissues and B lymphocytes.[12] Reactivation occurs during cell-mediated immune deficiency and can lead to the rapidly evolving severe demyelinating disease of the CNS, progressive multifocal leukoencephalopathy (PML). This disease is primarily seen in patients with AIDS and, prior to the era of highly active antiretroviral therapy (HAART), was a serious cause of infection in 2% to 10% of AIDS patients.

Diagnosis

The diagnosis of PML can be difficult, was generally presumptive, and was based on imaging techniques (CT and MRI) in combination with serologic studies, virus isolation by cell culture, and electron microscopy.[12] Often confirmation required histopathological analysis of brain biopsy materials.

Molecular Tests

Several PCR-based strategies to detect JCV in CSF have been established for the diagnosis of PML and to monitor patients with AIDS who achieved a PML remission as a result of HAART.[12,70–72] Testing has included unique or sequential CSF samples, with assays developed using single or nested PCR combined with conventional or real-time detection technologies. Primers have targeted different noncoding and coding (VP1, VP2, small T-antigen, large T-antigen) regions, often with variable results, presumably due to the large amount of sequence variation within the various subtypes of JCV.[12] Using a variety of PCR assays, the rate of detection of JCV in CSF from patients with PML ranged from 30% to 89.5%.[70–72] Specificity of PCR for identifying viral isolates of JCV was 100%, while specificity for the diagnosis of confirmed PML was generally about 95%.[70–72] In summary, CSF examination for JCV DNA was shown to be useful for confirming the diagnosis of PML. However, a negative test does not preclude the possibility of PML and brain biopsy is necessary to confirm such cases.

HUMAN IMMUNODEFICIENCY VIRUS

Early in the course of infection, HIV-1 enters the CNS.[73,74] A complex interaction between the virus and host immune responses leads to neurological damage that is manifested by a variety of syndromes including meningitis, encephalitis, peripheral neuropathies, and AIDS dementia complex (ADC).[73,74]

Molecular Tests

High HIV-1 RNA levels in CSF correlate with an increased incidence of ADC and with the presence of cerebral atrophy.[73] Most persons with high CSF viral load have concomitant high plasma viral load. However, there are patients on HAART with either very-low-level or undetectable plasma viral load (<50 copies/ml) who have signs of neurologic disease and high CSF viral load (>50,000 copies/ml). Studies have indicated that suppression of plasma viral load does not always correspond with HIV-1 suppression in other body compartments, and the monitoring of such compartments may be indicated in certain clinical situations. Therefore, when neurological symptoms are evident, it is important to measure HIV-1 levels in the CSF. Detectable viral load in the CSF can affect the selection of HAART, since not all drugs penetrate the CSF to the same degree. In addition, due to the large overlap of neurological symptoms associated with the various AIDS-related neurological disorders, such as CMV

encephalopathies and PML, CSF viral load can aid in the differential diagnosis of ADC.[69] The measurement of CSF viral load can be achieved using a variety of commercially available methods described in chapter 36.

Conclusions and Future Considerations

Currently, the application of molecular testing for the detection of clinically relevant viral CNS infections is considered the standard of care. For certain viral pathogens, such as HSV, molecular tests are considered the new gold standard, whereas for arboviral infections, serology remains the gold standard for human disease, yet molecular assays are useful for mosquito, bird, and vertebrate surveillance studies. Molecular tests are rapid, accurate and significantly impact patient management and outcome. The selection of the best method or target must be carefully considered, as variations in both can affect the overall performance of the assays. In certain cases, where latent viruses may be detected, quantitative assays or assays that target mRNA rather than DNA may provide more useful clinical correlation for CNS disease. Laboratories must thoroughly evaluate and validate the tests, continue to monitor technical performance, and maintain strict environmental precautions to prevent cross-contamination of samples. Standardization of test methods and cooperative participation in proficiency programs will improve performance and correlation of intra- and interlaboratory results.

References

1. Cassady KA, Whitley RJ. Pathogenesis and pathophysiology of viral central nervous system diseases. In: Scheld WM, Whitey RJ, Durack DT, eds. *Infections of the Central Nervous System.* Philadelphia, Lippincott-Raven; 1997:7–22.
2. Cassady KA, Whitley RJ. Viral central nervous system infections. In: Richman DD, Whitley RJ, Hayden FG, eds. *Clinical Virology.* Washington, DC. ASM Press; 2002:27–44.
3. Hammer SM, Connolly KJ. Viral aseptic meningitis in the United States: clinical features, viral etiologies, and differential diagnosis. *Curr Clin Top Infect Dis.* 1992;12:1–25.
4. Rotbart HA. Viral meningitis. *Semin Neurol.* 2000;20:277–292.
5. Hosoya M, Honzumi K, Sato M, et al. Application of PCR for various neurotropic viruses on the diagnosis of viral meningitis. *J Clin Virol.* 1998;11:117–124.
6. Gubler DJ, Roehrig JT. Arboviruses (*Togaviridae* and *Flaviviridae*). In: Collier L, ed. *Topley and Wilson's Microbiology and Microbial Infections.* ASM Press, London; 1996.
7. Centers for Disease Control and Prevention. Division of Vector-Borne Infectious Diseases. Arboviral encephalitides. August 2005. Available at: http://www.cdc.gov/ncidod/dvbid/index.htm.
8. Centers for Disease Control and Prevention. Division of Vector-Borne Infectious Diseases. West Nile virus. August 2005. Available at: http://www.cdc.gov/ncidod/dvbid/westnile/index.htm.
9. Hibberd PL, Snydman DR. Cytomegalovirus infection in organ transplant recipients. *Infect Dis Clin North Am.* 1995;9:863–877.
10. Ives DV. Cytomegalovirus disease in AIDS. *AIDS.* 1997;11:1791–1797.
11. Cinque P, Marenzi R, Ceresa D. Cytomegalovirus infections of the nervous system. *Intervirology.* 1997;40:85–97.
12. Safak M, Khalili K. An overview: human polyomavirus JC virus and its associated disorders. *J Neurovirol.* 2003;9(suppl 1):3–9.
13. Landry ML. Frequency of normal cerebrospinal fluid protein level and leukocyte count in enterovirus meningitis. *J Clin Virol.* 2005;32:73–74.
14. Cinque P, Cleator GM, Weber T, et al. The role of laboratory investigation in the diagnosis and management of patients with suspected herpes simplex encephalitis: a consensus report. *J Neurol Neurosurg Psychiatry* 1996;61:339–345.
15. Thomson R, Bertram H. Laboratory diagnosis of central nervous system infections. *Infect Dis Clin North Am.* 2001;15:1047–1071.
16. Reiber H, Lange P. Quantitation of virus specific antibodies in cerebrospinal fluid and serum; sensitive and specific detection of antibody synthesis in the brain. *Clin Chem.* 1991;37:1153–1160.
17. Cinque P, Bossolasco S, Lundkvist A. Molecular analysis of cerebrospinal fluid in viral diseases of the central nervous system. *J Clin Virol.* 2003;26:1–28.
18. Johnson A, Martin D, Karabatsos N, et al. Detection of anti-arboviral immunoglobulin G by using a monoclonal antibody-based capture enzyme-linked immunosorbent assay. *J Clin Microbiol.* 2000;38:1827–1831.
19. Roehrig JT, Nash D, Maldin B, et al. Persistence of virus-reactive serum immunoglobulin M antibody in confirmed West Nile virus encephalitis cases. *Emerg Infect Dis.* 2003;9:376–379.
20. van Regenmortel MHG, Fauquet CM, Bishop EB, et al. *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses.* New York: Academic Press; 2000.
21. Pallansch MA, Roos RP. Enteroviruses: polioviruses, coxsackieviruses, echoviruses, and newer enteroviruses. In: Knipe DM, Howley PM, Griffin DE, et al., eds. *Fields Virology.* Philadelphia: Lippincott Williams and Wilkins; 2001:723–775.
22. Rotbart HA. Enteroviral infections of the central nervous system. *Clin Infect Dis.* 1995;20:971–981.
23. Rotbart HA. 1990. Enzymatic RNA amplification of the enteroviruses. *J Clin Microbiol.* 1990;28:438–442.
24. Rotbart HA. Diagnosis of enteroviral meningitis with the polymerase chain reaction. *J Pediatr.* 1990;117:85–89.
25. Zoll GJ, Melchers JG, Kopecka H, et al. General primer-mediated polymerase chain reaction detection of enteroviruses; application for diagnostic routine and persistent infections. *J Clin Microbiol.* 1992;30:160–165.
26. Oberste MS, Maher K, Flemister MR, et al. Comparison of classic and molecular approaches for the identification of untypeable enteroviruses. *J Clin Microbiol.* 2000;38:1170–1174.
27. Ramers C, Billman G, Hartin M, et al. Impact of a diagnostic cerebrospinal fluid enterovirus polymerase chain reaction test on patient management. *JAMA.* 2000;283:2680–2685.
28. Stellrecht KA, Harding I, Hussain FM, et al. A one-step RT-PCR assay using an enzyme-linked detection system for the diagnosis of enterovirus meningitis. *J Clin Virol.* 2000;17:143–149.
29. Fox JD, Han S, Samuelson A, et al. Development and evaluation of nucleic acid sequence based amplification (NASBA) for diagnosis of enterovirus infections using the NucliSens Basic Kit. *J Clin Virol.* 2002;24:117–130.
30. Verstrepen WA, Bruynseels P, Mertens AH. Evaluation of a rapid real-time RT-PCR assay for detection of enterovirus RNA in cerebrospinal fluid specimens. *J Clin Virol.* 2002;25:S39-S43.
31. Bourlet T, Caro V, Minjolle S, et al. New PCR test that recognizes all human prototypes if enterovirus: application for clinical diagnosis. *J Clin Microbiol.* 2003;41:1750–1752.
32. Landry ML, Garner R, Ferguson D. Rapid enterovirus RNA detection in clinical specimens by using nucleic acid sequence-based amplification. *J Clin Microbiol.* 2003;41:346–350.
33. Sillekens P, Foolen H, Overdijk M, et al. Development of a real-time NASBA application for the diagnosis of enteroviral meningitis

[abstract T4]. University of South Florida Clinical Virology Symposium, Pan American Society for Clinical Virology, Clearwater, Florida; April, 2003.
34. Ginocchio CC, Zhang F, Malhotra A, et al. Development, technical performance, and clinical evaluation of a NucliSens Basic Kit application for detection of enterovirus RNA in cerebrospinal fluid. *J Clin Microbiol.* 2005;43:2616–2623.
35. Landry ML, Garner R, Ferguson D. Real-time Rapid nucleic acid sequence-based amplification using molecular beacons for detection of enterovirus RNA in clinical specimens. *J Clin Microbiol.* 2005;43:3136–3139.
36. International Agency for Research on Cancer. Epstein-Barr virus and Kaposi's sarcoma herpesvirus/human herpes virus 8. *IARC Monogr Eval Carcinog Risk Chem Hum.* 1997;70:101–373.
37. Jackson AC. Acute viral infections [review]. *Curr Opin Neurol.* 1995;8:170–174.
38. Whitley RJ. Herpes simplex viruses. In: Knipe DM, Howley PM, Griffin DE, et al., eds. *Fields Virology.* Philadelphia: Lippincott Williams and Wilkins; 2001:2461–2509.
39. Tedder DG, Ashley R, Tyler R, et al. Herpes simplex virus infection as a cause of benign recurrent lymphocytic meningitis. *Ann Intern Med.* 1994;121:334–338.
40. Kimberlin DW, Whitley RJ. Neonatal herpes: what have we learned? *Semin Pediatr Infect Dis.* 2005;16:7–16.
41. Sauerbrei A, Eichorn U, Hottenrott G, et al. Virologic diagnosis of herpes simplex encephalitis. *J Clin Virol.* 2000;17:31–36.
42. Lakeman F, Whitley R, and the National Institute of Allergy and Infectious Diseases Collaborative Antiviral Study Group. Diagnosis of herpes simplex encephalitis: application of polymerase chain reaction to cerebrospinal fluid from brain-biopsied patients and correlation with disease. *J Infect Dis.* 1995;171:857–863.
43. Arvin AM. Varicella zoster virus. In: Knipe DM, Howley PM, Griffin DE, et al., eds. *Fields Virology.* Philadelphia: Lippincott Williams and Wilkins; 2001:2731–2768.
44. Gnann J, Whitley R. Herpes zoster. *N Engl J Med.* 2002;347:340–346.
45. Gilden DH, Kleinschmidt-DeMaster BK, LaGuardia JJ, et al. Neurologic complications of the reactivation of varicella-zoster virus. *N Engl J Med.* 2000;342:635–645.
46. Sauerbrei A, Eichorn U, Schacke M, et al. Laboratory diagnosis of herpes zoster. *J Clin Virol.* 1999;14:31–36.
47. Rowley A, Whitley R, Lakeman F, et al. Rapid detection of herpes simplex virus DNA in cerebrospinal fluid of patients with herpes simplex encephalitis. *Lancet.* 1990;335:440–441.
48. Minjolle S, Michelet C, Jusselin I, et al. Amplification of the six major human herpesviruses from cerebrospinal fluid in a single PCR. *J Clin Microbiol.* 1999;37:950–953.
49. Read SJ, Kurtz JB. Laboratory diagnosis of common viral infections of the central nervous system by using a single multiplex PCR screening assay. *J Clin Microbiol.* 1999;37:1352–1355.
50. Espy M, Uhl J, Mitchell P, et al. Diagnosis of herpes simplex virus infections in the clinical laboratory by LightCycler PCR. *J Clin Microbiol.* 2000;38:759–799.
51. Johnson G, Nelson S, Petric M, et al. Comprehensive PCR-based assay for detection and species identification of human herpesviruses. *J Clin Microbiol.* 2000;38:3274–3279.
52. Quereda C, Corral I, Laguna F, et al. Diagnostic utility of a multiplex herpesvirus PCR assay performed with cerebrospinal fluid from human immunodeficiency virus-infected patients with neurological disorders. *J Clin Microbiol.* 2000;38:3061–3067.
53. Read S, Mitchell J, Fink C. LightCycler multiplex PCR for the laboratory diagnosis of common viral infections of the central nervous system. *J Clin Microbiol.* 2001;39:3056–3059.
54. Weidmann M, Meyer-Konig U, Hufert F. Rapid detection of herpes simplex virus and varicella-zoster virus infections by real-time PCR. *J Clin Microbiol.* 2003;41:1565–1568.
55. Braun DK, Dominguez G, Pellett PE. Human herpesvirus 6. *Clin Microbiol Rev.* 1997;10:521–567.
56. Secchiero P, Zella D, Crowley R, et al. Quantitative PCR for human herpesviruses 6 and 7. *J Clin Microbiol.* 1995;33:2124–2130.
57. Van den Bosch G, Locatelli G, Geerts L, et al. Development of reverse transcriptase PCR assays for detection of active human herpesvirus 6 infection. *J Clin Microbiol.* 2001;39:2308–2310.
58. Tang Y, Espy M, Persing D, et al. Molecular evidence and clinical significance of herpesvirus coinfection in the central nervous system. *J Clin Microbiol.* 1997;35:2869–2872.
59. Campbell GL, Marfin AA, Lanciotti RS, et al. West Nile virus. *Lancet Infect Dis.* 2002;2:519–529.
60. Harrington T, Kuehnert MJ, Kamel H, et al. West Nile virus infection transmitted by blood transfusion. *Transfusion.* 2003;43:1018–1022.
61. Johnson A, Martin D, Karabatsos N, et al. Detection of anti-arboviral immunoglobulin G by using a monoclonal antibody-based capture enzyme-linked immunosorbent assay. *J Clin Microbiol.* 2000;38:1827–1831.
62. Martin D, Muth D, Brown T, et al. Standardization of immunoglobulin M capture enzyme-linked immunosorbent assay for routine diagnosis of arboviral infections. *J Clin Microbiol.* 2000;38:1823–1826.
63. Lanciotti R, Kerst A, Nasci R, et al. Rapid detection of West Nile virus from human clinical specimens, field-collected mosquitoes, and avian samples by a TaqMan reverse transcriptase-PCR assay. *J Clin Microbiol.* 2000;38:4066–4071.
64. Lanciotti R, Kerst A. Nucleic acid sequence-based amplification assays for rapid detection of West Nile and St. Louis Encephalitis viruses. *J Clin Microbiol.* 2001;39:4506–4513.
65. Scaramozzino N, Crance J, Jouan A, et al. Comparison of flavivirus universal primer pairs and development of a rapid, highly sensitive hemi-nested reverse transcription-PCR assay for detection of flaviviruses targeted to a conserved region of the NS5 gene sequences. *J Clin Microbiol.* 2001;39:1922–1927.
66. Lambert A, Martin D, Lanciotti R. Detection of North American eastern and western equine encephalitis viruses by nucleic acid amplification assays. *J Clin Microbiol.* 2003;41:379–385.
67. White D, Kramer L, Backenson P, et al. Mosquito surveillance and polymerase chain reaction detection of West Nile virus, New York State. *Emerg Infect Dis.* 2001;7:643–649.
68. Ginocchio CC. Laboratory diagnosis of human cytomegalovirus (HCMV) central nervous system disease in AIDS patients. *Int J Antimicrob Agents.* 2000;16:447–453.
69. Zhang F, Tetali S, Wang XP, et al. Detection of human cytomegalovirus pp67 late gene transcripts in cerebrospinal fluid of human immunodeficiency virus type 1 infected patients by nucleic acid sequence based amplification. *J Clin Microbiol.* 2000;38:1920–1925.
70. Hammarin A, Bogdanovic G, Svenhem V, et al. Analysis of PCR as a tool for detection of JC virus DNA in cerebrospinal fluid for diagnosis of progressive multifocal leukoencephalopathy. *J Clin Microbiol.* 1996;34:2929–2932.
71. De Viedma D, Alonson R, Miralles P, et al. Dual qualitative-quantitative nested PCR for detection of JC virus in cerebrospinal fluid: high potential for evaluation and monitoring of progressive multifocal leukoencephalopathy in AIDS patients receiving highly active antiretroviral therapy. *J Clin Microbiol.* 1999;37:724–728.
72. Whiley D, Mackay I, Sloots T. Detection and differentiation of Human polyomaviruses JC and BK by LightCycler PCR. *J Clin Microbiol.* 2001;39:4357–4361.
73. Albright AV, Soldan SS, Gonzalez-Scarano F. Pathogenesis of human immunodeficiency virus-induced neurological disease. *J Neurovirol.* 2003;9:222–227.
74. McArthur JC, Haughey N, Gartner S, et al. Human immunodeficiency virus-associated dementia: An evolving disease. *J Neurovirol.* 2003;9:205–221.

Chapter 40

Sexually Transmitted Diseases

Jeanne A. Jordan

CHLAMYDIA TRACHOMATIS AND NEISSERIA GONORRHEAE

Clinical Utility

Chlamydia trachomatis (CT) and *Neisseria gonorrheae* (GC) are presented together, not because of their similarities in disease presentation, but because of the current trend in screening samples for both simultaneously. Historically, these organisms were identified using very different laboratory methods: CT by tissue culture and GC by growth in specialized bacterial medium. However, over the past decade a revolutionary change has taken place in the approach used to detect these two sexually transmitted infections (STIs). In many instances, molecular testing, either nucleic acid hybridization or nucleic acid amplification, has replaced culture and immunoassays.

Chlamydia Trachomatis

Chlamydia trachomatis is a gram-negative, obligate intracellular bacterium of global public health significance. Infection with CT is associated with three different disease presentations; trachoma (serovars A, B, B_1, and C), genital infection (serovars D-K), and lymphogranuloma venereum (LV) (serovars L_1, L_2, and L_3).[1] This chapter focuses on the diagnosis of those serovars causing genital infections.

Genital CT infections represent a major reproductive health problem. The World Health Organization (WHO) estimates that annually there are 89 million new cases of genital CT globally, with more than 4 million cases occurring annually in North America. In the developed world, CT is the likely cause of secondary infertility in females, due to its association with cervicitis, endometritis, and urethritis. Between 10% and 40% of CT infections in females will result in pelvic inflammatory disease (PID), whose sequelae include ectopic pregnancy, tubal factor infertility, and chronic pelvic pain.[2] The risk of developing any one of these sequelae increases with the number of PID episodes that an individual experiences. Additionally, pregnant women infected with CT are at risk of transmitting the infection to their newborn infant in the form of conjunctivitis or pneumonia. In males, genital CT infection is associated with nongonococcal urethritis (NGU), epididymitis, urethritis, prostatitis, proctitis, and Reiter syndrome.[1,3]

The individuals at greatest risk of genital CT infections are adolescents and young adults, ages 14 to 24 years, which is similar to the age of prevalence for human papillomavirus (HPV). Intervention based on selective screening for genital CT infection has reduced its prevalence, as well as the incidence of PID and ectopic pregnancy. These improved outcomes have occurred with the implementation of nucleic acid amplification testing (NAT), use of less-invasive and better-tolerated collection techniques, and development of more stable specimen transport systems. Molecular testing for sexually transmitted diseases (STDs) is an advantageous approach, especially in the high prevalence or high-risk patient populations.

CT infection often is silent, with up to 70% of infected females and 50% of infected males being asymptomatic; therefore, a significant number of individuals do not seek treatment and are at risk of developing complications from the infection. Mass screening programs in Europe have led to a significant reduction in CT prevalence. However, to be successful, these programs must be practical and acceptable to the targeted patient populations. Acceptable screening programs might include the use of self-collected vaginal specimens or first-void urines that could be stably shipped at ambient temperatures. This strategy has the advantage of limiting the number of physician office visits for individuals, thereby increasing compliance.

Neisseria Gonorrhoeae

Neisseria gonorrhoeae is a gram-positive, oxidase-positive, intracellular diplococcus. Microscopically, the microorganism has a characteristic kidney or coffee bean appearance.

Table 40-1. Summary of Selected NAT Platforms for *Chlamydia trachomatis* and *Neisseria Gonorrheae*

Manufacturer	Assay Method	Sample Preparation	Amplification Strategy	Detection Strategy	Contamination Controls	IAC
Becton Dickinson	SDA	Crude cell lysate	Target	Fluorescence	Closed System Postamplification Bleach	Yes
Digene Corporation	hc2	Hybrid capture	Signal	Chemiluminescence	No recommendations	No
Gen-Probe, Inc	TMA	Target capture	Target	Chemiluminescence Dual Kinetics Assay	Postamplification Bleach	No
Roche Diagnostics	PCR	Crude cell lysate	Target	Colorimetric	Preamplification dUTP, UNG	Yes

SDA, strand displacement amplification; IAC, internal amplification control; NASBA, nucleic acid sequence–based amplification; hc2, Hybrid Capture 2; TMA, transcription-mediated amplification; PCR, polymerase chain reaction; NAT, nucleic acid amplification testing.

GC is a fastidious organism, being highly susceptible to temperature extremes and desiccation, which results in less-than-satisfactory isolation by culture, especially when off-site specimen transportation is required before culture.[4]

The fastidious nature of GC has led to a trend from culture toward molecular testing.[5] Interestingly, experiences in GC testing have led some investigators to conclude that NAT for GC does not appreciably increase the number of positive specimens compared to culture methods, if optimal transport conditions have been maintained. However, NAT allows for the use of alternative, less-invasive specimens such as urine and self-collected vaginal swabs for GC detection. This strategy avoids the need to collect urethral or endocervical specimens, which should increase compliance with testing.

Like CT, the incidence of GC is highest among young adults under the age of 24 years. GC remains second only to CT infection in the number of cases of a reportable STD. Worldwide, approximately 62 million new cases of GC occur annually. GC-associated urogenital tract infections can lack symptoms altogether, or present with mild to severe symptoms. In men, symptoms include acute urethritis and discharge, which if left untreated can lead to epididymitis, prostatitis, or urethral strictures. In women, symptoms can include endocervical inflammation and discharge, which if not treated can lead to abscess formation, salpingitis, or PID. Disseminated GC infection, although not common, can occur in a small percentage of infected individuals and is usually has a poor outcome.

Available Tests

Prior to the introduction of molecular testing, CT and GC were detected from clinical specimens primarily by culture. For CT, more rapid tests were developed to facilitate shorter turnaround times and included enzyme immunoassays (EIA) to detect either the LPS or MOMP antigen and direct fluorescence assays (DFA) that used genus-specific or species-specific monoclonal antibodies.[1] For GC, direct gram stain and culture have been used. Guidelines for CT and GC testing now include molecular testing as a recommended test method.

More molecular testing platforms for detection of CT and GC have been cleared or approved by the Food and Drug Administration (FDA) than for any other infectious pathogen. These platforms include both nonamplified (e.g., the PACE DNA probe assay, GenProbe, San Diego, CA) and amplified testing methods (Table 40-1). The nonamplified probe-based assays are not discussed further in this chapter.

The popularity of using NAT for CT and GC can be partially explained by market pressures; the sheer volume of STD testing performed annually drives commercial interests. In addition, there are a wide variety of validated specimen types that have been approved for use with these testing platforms, which help to facilitate compliance in high-risk groups for this ever-growing public health concern.

Table 40-1 lists four of the most commonly used, commercially available NAT platforms for CT and GC; three use target amplification and one uses signal amplification methods.[6–10] The table summarizes the test methods, purification strategies, amplification and detection schemes, and types of contamination controls and internal amplification controls included in each kit. The sample preparation protocols vary widely for these kits and range from using crude cell lysates to using purified nucleic acid as the input specimen. The different sample types used and the extent of sample purification affect the rates of inhibition and thus the need to include an internal amplification control.

Interpretation of Test Results

Tables 40-2 through 40-4 describe in detail the three most commonly used NAT platforms for CT and GC, specifically, ProbeTec ET (Becton, Dickinson and Co, Franklin Lakes, NJ), Aptima Combo 2 (Gen-Probe, Inc, San Diego, CA), and Amplicor CT/NG Assays (Roche Diagnostics, Indianapolis, IN). The tables provide the gene targeted for CT and GC detection, the recommended clinical specimens, optimal transport and storage conditions, processing deadlines, test interpretations, workflow issues, and other features of these test methods. Many investigators have published performance data using these platforms.[6,7,9–16]

Table 40-2. BD ProbeTec™ ET CT/GC SDA Assays: Parameters and Test Interpretation for CT/GC Detection

Target	CT: Cryptic plasmid GC: *Piv*$_{Ng}$ gene			
Specimens	Female and male urine Endocervical and urethral swabs			
Specimen handling	Urines: Adding a UPP to urine sample at specimen collection site permits 15°C to 27°C transport; process within 2 days of collection. Unpreserved urines should be transported at 2°C to 8°C; process within 7 days of collection. See package insert for all approved variations for urine specimens. Swabs: Transport swabs at 2°C to 27°C; process within 4 days.			
Test interpretations				
IAC	CT/GC MOTA Score	AC MOTA Score	Result	Interpretation
No	<2,000	NA	Negative	CT and/or GC not detected
	2,000 to 9,999	NA	Low positive	CT and/or GC likely, supplemental testing
	≥10,000	NA	Positive	CT and/or GC detected
Yes	<2,000	<1,000	Indeterminate	Inhibition
	<2,000	≥1,000	Negative	CT and/or GC not detected
	2,000 to 9,999	Any	Low positive	CT and/or GC likely, supplemental testing
	≥10,000	Any	Positive	CT and/or GC detected
Workflow issues	Individual amplification and detection reactions required for CT and GC, or CT, GC, and IAC.			
Other features	CT: IAC consists of 1,000 copies of a linearized GC DNA containing plasmid. GC: Cross-reactivity to some nongonococcal *Neisseria* species.			

Complete package insert is available at: http://www.bd.com.
IAC, Internal amplification control; MOTA, method other than acceleration; NA, not applicable; UPP, urine processing pouch.

Table 40-3. Gen-Probe Aptima Combo 2 Assays: Parameters and Test Interpretation for CT/GC Detection Using TMA and DKA Technologies

Target	23S rRNA/16S rRNA			
Specimens	Female and male urine Endocervical and male urethral swabs			
Specimen handling	Urines and Swabs: Transport at 2°C to 30°C and process within 60 days of collection. Transport both specimen types in the proper collection devices.			
Test interpretation Signals in relative light units (RLU)				
		Negative	Equivocal	Positive
For CT	CT only	1 to <25	25 to <100	100 to <3000
	CT and GC	1 to <85	85 to <250	250 to <3000
For GC	CT indeterminate	1 to <85	85 to <3000	NA
	Rapid "flasher" kinetics Slower "glower" kinetics	Negative	Equivocal	Positive
	CT only	1 to <60	60 to <150	150 to <4500
	CT and GC	1 to <85	85 to <250	250 to <4500
	CT indeterminate	1 to <85	85 to <4500	NA
Workflow issues	Simultaneous amplification and detection of CT and GC. Options include manual assay or Tecan-assisted assay.			
Other features	No known cross-reactivity with nongonococcal *Neisseria* species.			

Complete package insert is available at: http://www.gen-probe.com.
TMA, transcription-mediated amplification; DKA, dual kinetic assay; NA, not applicable.

Table 40-4. Roche Amplicor and Cobas Amplicor CT/NG PCR Assay: Parameters and Test Interpretation for CT/NG Detection

Targets	Cryptic plasmid/*M·Ngo* PII (*Cytosine DNA methyltransferase*)				
Specimens	Female and male urine (CT only)				
	Endocervical and male urethral swabs (CT/NG)				
Specimen handling	Transport swabs and urines at 2° to 8°C				
	Store swabs and urines at 2° to 8°C; process within 7 days of collection; –20°C: process within 30 to 60 days of collection.				
Test interpretation					
For CT	IC	A_{660}	CT Result	IC Result	Interpretation
	Without	<0.2	NA	NA	Negative
		≥0.2 to <2.0	NA	NA	Equivocal
		≥2.0	NA	NA	Positive
	With	<0.2	Negative	Positive	CT not detected
		<0.2	Negative	Negative	Inhibition
		≥0.2 to <2.0	Equivocal	Any	Equivocal
For NG	IC	A_{660}	NG Result	IC Result	Interpretation
	Without	<0.2	NA	NA	Negative
		≥0.2 to <3.5	NA	NA	Equivocal
		≥2.0	NA	NA	Positive
	With	<0.2	Negative	Positive	GC not detected
		<0.2	Negative	Negative	Inhibition
		≥0.2 to <3.5	Equivocal	Any	Equivocal
		≥2.0	Positive	Any	GC detected
Workflow issues	Simultaneous amplification of CT/NG/IC with separate detection reactions.				
Other issues	IC is a plasmid-containing CT primer binding sites and randomized internal sequence.				
	Cross-reactivity with certain nongonococcal *Neisseria* species.				

PCR, polymerase chain reaction; A_{660}, $Absorbance_{660}$; IC, internal control.
More information is available at: http://www.roche-diagnostics.us.

Laboratory Issues

The following laboratory issues are noteworthy when NAT platforms for CT and GC testing are selected.[17–21]

Inhibitors and Internal Controls

Amplification inhibition is common within urogenital specimens, with consequent negative effects on test results.[22] The percentage of specimens containing amplification inhibitors ranges from 1% to 5% for urines and as much as 20% for cervical swabs. Initial studies found that inhibition could be reduced or eliminated if specimens were first refrigerated overnight or frozen and thawed before testing, pointing to the labile nature of some inhibitors. However, other inhibitors are quite stable and thus more difficult to neutralize. Some of the commercially available platforms include an internal amplification control (IAC) to identify specimens containing inhibitors. Alternatively, inhibitors can be removed if nucleic acid purification steps are included in the sample preparation protocol. For testing platforms that include an IAC, the testing algorithm states that when the internal control fails, the results of a specimen without detectable CT and/or GC cannot be reported as negative due to the likelihood that amplification inhibitors are present in the specimen.

Cross-Reactivity with Nongonnococcal Neisseria Species

The specificity for some GC NAT platforms has been problematic due to cross-reactivity with certain nongonococcal species of *Neisseria*.[23,24] This problem is thought to arise from the intraspecies and interspecies genetic recombination that occurs between *Neisseria* species, which can result in false-positive NAT results with certain of the commensal *Neisseria* species.[16] Of eight nongonococcal species tested, five species demonstrated cross-reactivity: *N. cinerea*, *N. flavescens*, *N. lactamica*, *N. subflava*, and *N. sicca*. False-positive results were seen with *N. flavescens*, *N. lactamica*, *N. subflava*, and *N. cinerea* using the BD ProbeTec assay, and with *N. flavescens*, *N. lactamica*, and *N. sicca* using the Roche Amplicor assay. Since *N. cinerea*, *N. lactamica*, *N. subflava*, and *N. sicca* isolates have been recovered from genital mucosa, it is possible to generate false-positive results from genital specimens as well as pharyngeal specimens. Therefore, confirmatory testing using a different gene target may be useful.

Confirmatory Testing

When a testing platform is being selected, the prevalence of infection in the population being tested is of significance. The analytical performance of any test is depend-

ent on the prevalence of infection, with the risk of generating false-positive results being inversely related to the prevalence. These issues have fueled the debate over the need for confirmatory testing for CT and GC. Both are reportable infections, with the potential for psychosocial and/or medico-legal consequences for a false-positive or a true-positive result.[17,18,25] FDA-cleared confirmatory tests for CT and GC are now available.

The most recent CDC Guidelines for STD testing have addressed the issue of confirmatory testing, providing a variety of options.[26] These guidelines are especially critical for confirmatory testing in low-prevalence populations where the positive predictive value (PPV) of the assay is known to be less than 90%. In fact, the manufacturers' package inserts illustrate this point, with a PPV of 32% to 60% for any of these tests when the GC prevalence is 1%. This also is true for CT testing, with a PPV range from 56% and 74% when the prevalence is 5%. The most recent options recommended by the CDC for confirmatory testing are summarized as follows:

- Test another specimen with a different assay that has a different target.
- Test the original specimen with a different assay that has a different target.
- Test the original specimen using the original test, incorporating a blocking antibody or competitive probe.
- Test the original specimen again using the original test.

To this end, many laboratories have developed algorithms to incorporate confirmatory testing in their workflow. Some laboratories establish an equivocal range for specimens with an initial low-level positive or high-level negative result. Although laboratory personnel feel more confident in reporting positive results that are confirmed by repeat testing, this has cost implications. Questions remain about the proper reporting strategy for specimens that do not reconfirm as positive on repeat testing.

Amplification Contamination Control

The advent of NAT, with its exquisite sensitivity, has given birth to a whole new mind-set for cleanliness in the molecular testing laboratory. No longer is disinfection of the benchtop after a day's work adequate. New standards strive to "nuclease" the laboratory environment to remove or prevent amplicon contamination. Strategies includes daily cleaning of laboratory surfaces with a bleach solution, and frequently replacing disposable gloves and gowns while working in the pre- and postamplification area(s). A regular schedule for performing wipe testing to monitor for amplicon contamination should be enforced. For many of the current NAT platforms, pre- and postamplification steps should be performed, if possible, in separate rooms with positive and negative airflow, respectively. In the future, much of the monumental efforts currently taken to minimize amplicon contamination may be reduced or eliminated altogether with the introduction of fully automated, closed real-time NAT systems that incorporate sample preparation, nucleic acid amplification, and detection in a single closed reaction.

Future Directions

Testing Noninvasive Specimen Types

One of the more recent changes to STD testing has been the use of noninvasively collected samples, including self-collected vaginal swabs and first-void urines.[20,27–32] The greatest impact of this change will be seen in the young adult patient population, age range from 14 to 24 years, where the prevalence of CT and GC infections is highest and the willingness to undergo a pelvic examination or urethral swab collection is lowest.

Out-of-the-Vial Testing

Another concept currently being marketed is out-of-the-vial testing for STDs.[11] This testing approach would use the remainder of liquid-based cytology fluid collected for Pap testing not only for HPV testing but also for CT and GC testing. Many critical issues must be addressed before this approach can be implemented; these include technical issues (cross contamination), ethical issues (waiting to perform the infectious disease testing until after the cytology screening is completed), and legal issues (cancer diagnosis takes precedence over STD diagnosis). Successful implementation of this strategy will require the cooperation of many interested groups. The ability to use a prealiquot of the sample from the liquid-based cytology specimen for STD testing prior to the Pap analysis may provide a workable solution.

Automated NAT

Fully automated instrumentation (including specimen processing, amplification, detection, and reporting) will facilitate testing. Gen-Probe launched the Tigris, a fully automated high-throughput system, while Cepheid has introduced the GeneXpert System. Other manufacturers have similar systems in development.

TRICHOMONAS VAGINALIS

Clinical Utility

Trichomonas vaginalis (TV) is a flagellated protozoan and the only species within its genus that can infect squamous epithelial cells of the human urogenital tract.[33] TV infection is considered to be a nonulcerative STD but is associated with severe local inflammation. In women, symptoms may include vulvar irritation and vaginal discharge, which

appears frothy, mucopurulent, and yellow-green in color. During a TV infection, the vaginal pH is often abnormally elevated (pH > 4.5). Complications of TV infection in untreated women include endometritis, infertility, and cervical erosion. In men infected with TV, symptoms may include profuse purulent urethritis and a form of NGU, with complications including chronic prostatitis, urethral strictures, epididymitis, and/or infertility if the infection is untreated.[34,35]

Detection of TV infections in males has received much less attention than detection in females. However, with the advent of NAT, this understudied infection in men has become more appreciated, and thus research studies have provided important information on its prevalence, clinical symptoms, sequelae, and the most appropriate specimen to collect for proper diagnosis. TV has a greater role in NGU than previously thought.[36] As in females, direct microscopic examination of urethral discharge in males has poor sensitivity for detection of TV.

Worldwide, TV infection accounts for more than 167 million to 200 million cases annually. WHO estimates that TV infection accounts for close to half of all treatable STIs worldwide. In the United States, the annual incidence of TV infection is estimated between 5 million and 8 million cases. However, because TV infection is not a reportable disease, this number may be an underestimate. Unlike CT and GC, where prevalence is higher in adolescents and young adults, TV is more equally distributed among all age groups.

The detection of TV in an individual is considered by healthcare providers to be a red flag for high-risk sexual behavior and is frequently present along with other STIs in the same individual. Diagnosing TV is difficult, because 50% to 70% of all infected individuals are asymptomatic. Without a sensitive assay, infected individuals left untreated continue to act as a reservoir for ongoing disease transmission.

TV infection is associated with two important sequelae: (1) an increased risk of acquiring human immunodeficiency virus (HIV) and (2) an increased risk of perinatal morbidity and mortality.[37–40] HIV transmission is enhanced by the local inflammation reaction within the genital tract that is present with TV infections. In women, TV infection is strongly associated with an abnormal vaginal ecology. Harboring TV may contribute to the change in vaginal flora, which is associated with decreased lactic acid production and subsequent increase in vaginal pH. Lactic acid production and the normally low pH of the genital tract environment help to inactivate HIV. Therefore, a change in the vaginal environment to a less-hostile environment promotes an increased survival of HIV. If this hypothesis is correct, then controlling TV infections could lower HIV acquisition.

The increased risk of perinatal morbidity and mortality with TV infection is associated with premature rupture of membranes (PROM), preterm delivery, and low-birth-weight infants in pregnant women infected with TV.[41] Although controversial, these associations suggest a need for increased efforts to detect and treat this infection in pregnant women.[42]

Screening for TV infection is usually done only in public health STD clinics and some obstetrical practices. Successful control of this STI would be aided greatly by performing contact follow-ups with sexual partners and using a more sensitive diagnostic assay. Improved detection of TV by using NAT, as was the case for CT, would be predicted to reduce the incidence of TV infection and assist in reducing HIV transmission and possibly even poor pregnancy outcomes.[43]

Table 40-5. Summary of the Commonly Used *Trichomonas Vaginalis* Specific Targets for NAT

Primer Pair	DNA Target	Amplicon (bp)
TVK3, TVK7	Repetitive DNA	312
TVK3, TVK4	Repetitive DNA	350
BTUB2, BTUB9	*Beta Tubulin*	112
TV1, TV2	*18S rRNA* gene	312
TV-E650	E650	650
TVA5, TVA6	*Ferredoxin* gene	102

Available Tests

Historically, the most common means of diagnosing TV in urogenital discharge was direct microscopic examination, broth culturing, or both. Direct microscopic examination of genital discharge material on a slide is certainly the most rapid and inexpensive method to use, but it lacks adequate sensitivity, which is reported to be approximately 40% to 70%.[44] The low level of sensitivity with microscopic examination may be due to the rapid loss of the characteristic protozoan motility once TV has been removed from a 37°C environment, and to the fact that nonmotile TV organisms are difficult to differentiate from leukocytes, being similar in size.

Currently, broth culturing is the gold standard for detection of TV.[45] Successful growth in culture can be achieved with as few as 300 to 500 TV organisms per milliliter of vaginal fluid, but this approach requires 2 to 7 days of incubation and daily microscopic examination. Culture methods have sensitivities that range between 50% and 80% but require specialized medium such as Diamond's broth, Tricosel medium, or the In-Pouch system. These specialized media may not be available in the physician's office. In addition, some TV isolates do not grow in culture due to strain requirements, low numbers of organisms, or damaged/nonviable organisms.

These limitations have led some laboratories to develop NAT for TV DNA. Although none of the NAT are commercially available, numerous assays have been described that are specific for TV.[46–53] A range of testing platforms has been developed, including conventional and real-time NAT assays. The most commonly cited TV targets used for NAT are listed in Table 40-5. The repetitive DNA sequences have a higher copy number per organism and therefore may result in a more sensitive assay.

Interpretation of Test Results

Currently, no FDA-approved test kits or analyte-specific reagents for TV are commercially available. Therefore, the individual laboratories performing NAT for TV DNA need to establish the interpretation criteria. The following issues should be included in the development and validation of NAT for TV:

- Rule out cross-reactivity of primers and probes with non-TV targets that could be present within the specimens.
- Validate the different specimen types to be tested, and the handling conditions to be used.
- Establish a protocol for specimen preparation.
- Establish the limit of detection for TV DNA.
- Establish analytical performance of the assay including sensitivity and specificity, and positive and negative predictive values.
- If applicable, include the rationale for establishing positive and negative cutoff values and whether an equivocal zone is used.
- Establish a protocol for repeat testing and reporting of indeterminate results.
- Address resolution of culture-negative, NAT-positive discrepancies.
- Assess the presence of interfering substances by including an internal amplification control.
- Ensure specimen adequacy by testing for an endogenous housekeeping gene such as beta globin.
- Develop an alternative proficiency testing program for TV NAT because an external program is not available.

Laboratory Issues

Target Selection

Validation of the primers and probes is necessary for laboratory-developed NAT to rule out cross-reactivity with other *Trichomonas* species, such as *T. tena*, as well as to other organisms that may be present within the urogenital tract specimen, including human genomic DNA. The TV-specific primer pairs and probes described in the literature recognize different gene targets that vary in their copy number. In general, choosing a gene target that is present in multiple copies per genome is usually preferable to choosing a single-copy gene target.

Internal Control

Both vaginal swabs and urine specimens can contain inhibitors that interfere with the nucleic acid amplification reaction. For interpreting negative results, it is important to include an internal control in the assay, especially if a crude cell lysate is used as the input material for testing. More details on inhibitors can be found within the CT and GC section of this chapter.

Specimen Collection

The optimal specimen to collect differs for women and men. For women, the specimen of choice is vaginal discharge collected using a Dacron-tipped, plastic-shaft swab, while for men, the sediment from the first-void urine is superior to a urethral swab specimen. For increased analytical sensitivity, only the first 20 ml to 30 ml of the first-void urine should be collected.

Specimen Transport

TV expresses and secretes numerous proteases. Therefore, for greatest target stability, specimens should be transported at temperatures below room temperature (≤2°C to 8°C).

Future Directions

Multiplex Testing for CT, GC, and TV

Multiplex testing for CT and GC has been available for years. Recently, there has been increased attention given to including TV as a target. With the prevalence of TV being greater than that of CT and GC combined in many regions, along with the association of TV infection with increased risk of HIV acquisition, inclusion of TV in a multiplex STD panel is a reasonable idea. Depending on the prevalence of other STIs within a patient population or geographic area, additional targets may be considered applicable for a multiplex STD panel.

Greater Use of Noninvasive Specimens

As with CT and GC detection, collection of first-void urines or self-collected vaginal swab-based specimens may improve patient compliance for TV testing.[32] This approach may in turn help to identify the asymptomatic, infected individuals and reduce transmission to sexual partners.

HUMAN PAPILLOMAVIRUS

Clinical Utility

Human papillomaviruses (HPV) are small, nonenveloped, double-stranded, circular DNA (~8 kilobases [kb]) viruses with a very limited cell tropism. To date, more than 100 different types of HPV have been identified.[54] HPV can infect and replicate in the nuclei of only certain squamous epithelial cells, which are classified as either cutaneous or mucosal in nature. In general, the cutaneous types of HPV

infect keratinizing epithelium, while mucosal types infect nonkeratinizing epithelium. This section focuses on the approximately 30 types of anogenital HPV, which can be spread through sexual contact.

The sexually transmitted HPV types can be further classified by their risk of disease progression. Infection with a low-risk HPV is commonly associated with condyloma accuminata or genital warts, and is considered benign or low-risk for progression to malignancy.[55] In contrast, infection with a high-risk HPV can be associated with cervical cancer in women and penile cancer in men.[56–58] High-risk HPV types play a major role in the pathogenesis of epithelial cell cancers of the anogenital tract, as they induce epithelial cell proliferation.[59,60]

HPV is one of the most common sexually transmitted infections. It is estimated that more than 5 million new cases are contracted annually in the United States, with more than 20 million men and women currently infected. The prevalence of HPV is highest among sexually active young women ages 15 through 25 years.

While most men and women infected with anogenital HPV types will not develop cancer, a subset will; a process that may take decades to occur.[61,62] Certain cofactors may be important in disease progression as well, including the individual's immune status, certain HLA types, smoking, nutritional status, or possibly coinfection with HIV-1, herpes simplex virus, or *Chlamydia trachomatis*.[63–67] Women ages 30 years or older who are persistently infected within the cervix with a high-risk HPV type are at an increased risk for developing cervical cancer. Thus, HPV acquisition during adolescence may have long-term consequences for health later in adulthood.[68,69]

Worldwide, cervical cancer is the second most common female cancer, with approximately 493,000 women developing cervical cancer each year. In the United States, approximately 13,000 new cases of cervical cancer are diagnosed annually. Cervical cancer is a highly preventable disease if detected and treated early. Infection with high-risk HPV types is associated with the appearance of clinical lesions. In fact, HPV now is considered to be a necessary cause of invasive cervical cancer.[70] Abnormal cell growth or cervical dysplasia is an early manifestation of HPV infection. The characteristic cellular changes caused by HPV infection can be detected on Pap smear evaluations or in cervical biopsies. The morphologic changes of dysplasia are classified as atypical squamous cells of undetermined significance (ASCUS), low-grade squamous intraepithelial lesion (LSIL), or high-grade squamous intraepithelial lesion (HSIL).[71,72]

Annually, approximately 2 million Pap tests performed in the United States are interpreted as an ASCUS-grade result. ASCUS rates typically fall within the range of 5% to 10% but can be as high as 20% in younger women. Current recommendations include reflex testing (triage) for high-risk HPV for patients with an ASCUS-grade Pap smear result; patients testing negative for HPV DNA can be followed according to routine practice, while those testing positive for high-risk HPV should be referred to colposcopy, as progression to high-grade disease is probable.[73,74] Patients with abnormal Pap smears whose specimens contain high-risk HPV DNA require immediate medical attention.[75–78]

In 2003, the FDA approved the Digene Hybrid Capture 2 (hc2) assay for use with a Pap test for adjunct screening of women ages 30 years and older for high-risk HPV infection. This recommendation was based on the theory that in this age group, the combined negative predictive value of Pap and hc2 would improve the sensitivity of cervical screening. Women who have both a normal Pap test and a negative high-risk HPV screen can be screened once every 3 years instead of annually.

Available Tests

Molecular testing platforms have been particularly helpful in detecting and classifying the HPV DNA within clinical specimens, as viral culturing is not easily accomplished and cytological examinations cannot differentiate between low-risk and high-risk types of HPV. Several nucleic acid–based tests have been developed for the detection and typing of HPV strains. These tests include signal amplification assays, polymerase chain reaction (PCR), and in situ hybridization.

Digene Hybrid Capture 2 Assay

Currently, the Hybrid Capture 2 (hc2) HPV Assay (Digene Corp, Gaithersburg, MD) is the only FDA-approved test kit for detecting low-risk or high-risk HPV types from cervical specimens.[79] This test is a signal amplification assay, with separate tests for high-risk and low-risk HPV types. The indicated uses for both the high-risk and the low-risk HPV hc2 assay include aiding in the diagnosis of sexually transmitted HPV infections. In contrast, the indicated uses for the high-risk assay include (1) screening specimens from patients with an ASCUS-grade Pap test result to determine the need for colposcopy[73,78] and (2) as a primary screening tool in conjunction with a Pap test for women 30 years and older.[76] Recommended specimen types for this assay include cervical swabs and liquid-based cytology specimens. The recommended collection devices and transport media include cervical brushes and cervical biopsies transported in specimen transport medium (STM), or a cervical broom transported in Cytyc PreservCyt (PC) solution.

The principle of the hc2 HPV assay is as follows: Target DNA, liberated from the virus, is denatured and allowed to hybridize with a specific cocktail of RNA probes. The high-risk cocktail contains 13 RNA probes recognizing HPV types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68. The low-risk cocktail contains five RNA probes recognizing HPV types 6, 11, 42, 43, and 44. The resulting DNA-RNA hybrids are captured onto the surface of a microtiter well coated with polyclonal antibodies specific for RNA-DNA hybrids. Alkaline phosphatase (AP) conjugated antibodies

recognizing the immobilized hybrids bind several molecules per target to provide signal amplification. Finally, the chemiluminescent substrate CDP-Star with Emerald II is added to the wells and cleaved by the AP enzyme. The resulting light emission is measured as relative light units (RLUs) using a luminometer.

Interpretation of Test Results

Prior to interpreting the patient specimen results of the hc2 HPV test, assay calibration verification must meet the following criteria:

- Mean value of the negative control must be <250 RLU with a %CV ≤ 25% for at least 2 of the 3 values.
- %CV of the calibrator mean value must be <15% for at least 2 of the 3 values.
- Ratio of the calibrator mean over the negative control mean must be ≥2.0.

If the above criteria are met, then the cutoff value can be calculated for determining positive specimens. The calibrator mean value serves as the cutoff value. A ratio of specimen RLU/cutoff value (CO) is generated, and, according to the package insert:

- Specimens with RLU/CO ratios ≥1.0 are interpreted as positive.
- Specimens with RLU/CO ratios <1.0 are interpreted as negative, or nondetected for the 13 high-risk HPV types tested.
- When testing PC solution specimens, if the RLU/CO ratio of a specimen is ≥1.0 and <2.5, the specimen must be retested. If the initial retest result is positive (≥1.0 RLU/CO), the specimen can be reported as positive. If the first retest is negative (<1.0), then a second retest should be performed. The result of the second retest is considered the final result and is to be reported.

Laboratory Issues

Specimen Expiration Date

Routinely, the liquid-based cytological evaluation (Pap test) is performed before the HPV DNA testing is done. Therefore, the DNA testing laboratory must carefully monitor the specimen-collection and testing dates to assure that the specimens have not exceeded their expiration date as stated in the package insert for the hc2 HPV assay.

Low-Volume Specimens Received for HPV DNA Testing

Inadequately collected specimens may contain fewer cells than normal, which results in a greater volume of the specimen being used for the Pap test, leaving fewer than normal cell numbers for the HPV DNA testing. Lower than usual cell numbers may still occur even if the specimen contains the required 4.0 ml volume. Technologists need to be aware of this issue when visually inspecting the pelleted material during HPV sample preparation to ensure quality results.

Assay Reproducibility

As with any test, reproducibility is most problematic at the assay's limit of detection. Therefore, specimens containing a limited number of HPV also will be the least reproducible specimens. In contrast, specimens containing a high level of HPV will have the greatest reproducibility. It is important to remember that the positive predictive value of any assay will decrease when the test is used in populations with low prevalence.

False-Positive Results

False-positive results are known to occur with the hc2 HPV assay due to the following conditions:

- Contamination of specimen with exogenous alkaline phosphatase.
- Inadequate denaturation of initial specimen.
- Inadequate washing during the assay.
- Cross-reactivity with HPV types 40, 53 and 66.
- Cross-reactivity with low-risk HPV types. Hybridization exists between low-risk HPV types 6 and 42 with the high-risk probe set. Specimens containing high levels (≥4 ng/ml) of HPV 6 or 42 DNA may be falsely positive in the hc2 high-risk HPV assay.

HPV PCR

There are numerous PCR assays in the literature describing HPV DNA amplification.[80–82] A common strategy utilizes the L1 consensus primer pair MY09 and MY11 to amplify HPV DNA, followed by hybridization with specific probes for amplicon detection and confirmation. Recent modifications to the MY09/11 assay have been described using the PGMY primers, which amplify a smaller region internal to the MY09/11 sequence and have improved detection of HPV DNA in genital samples.[83] HPV genotyping can be performed using line blot or line probe assays or DNA sequencing.[84] Currently there are no FDA-cleared PCR assays for HPV detection or typing. However, outside of the United States, Roche Diagnostics offers the Amplicor HPV test, which detects 13 high-risk HPV types, and the Linear Array HPV Genotyping test, which detects 37 high- and low-risk HVP types.

Inform HPV Assay

The Inform HPV assay (Ventana Medical Systems, Inc, Tucson, AZ) is an in situ hybridization (ISH) test used in conjunction with the company's Benchmark IHC and ISH staining system.[85] The Inform HPV assay is not an FDA-approved test kit. The automated ISH technology is a slide-based HPV assay that provides HPV DNA detection in the context of cervical cell morphology. The assay does not detect cell-free HPV DNA. The company offers both high-risk and low-risk HPV genomic fluorescently labeled probe cocktails for ISH testing.

Invader HPV Reagents

Third Wave Technologies (Madison, WI) offers analyte-specific reagents (ASR) for detecting the DNA of 13 HPV types, including types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68, along with the human histone gene. The HPV ASR are packaged as four separate mixtures based on phylogenic relatedness, and include A5/A6, A7, A9 and human histone. Analytical and performance characteristics are established by individual laboratories.

Summary

In summary, the limits of detection for the various assays are between 50 and 100 genome copies for PCR, approximately 500 to 1000 genome copies for signal amplification assays, and approximately 100,000 genome copies for DNA in situ hybridization assays.

Future Directions

A number of possible applications are on the horizon for HPV testing.

- Several manufacturers are marketing out-of-the-vial testing for HPV detection from liquid-based cytology specimens. More comments can be found on this topic in the CT/GC section of this chapter.[86]
- There is potential prognostic value in determining quantitative HPV viral load measurements. Further studies will be needed before any conclusions can be made on this topic including addressing variations in the cell numbers present within specimens and the viral copy number present per cell.
- In light of the recent findings of an elevated risk of cervical precancer and cancer in women with HPV type 16 or 18 compared to the other high-risk HPV types, HPV viral typing may prove informative in differentiating between viral persistence and reinfection.[87]
- Virus persistence and disease progression will be assessed relative to HPV E6 and E7 mRNA gene expression.
- The level of CpG methylation of high-risk HPV DNA may be explored.

References

1. Moss T, ed. *International Handbook of Chlamydia.* Exeter, UK: Polestar Wheatons Ltd; 2001.
2. Scholes D, Stergachis A, Heidrich FE, et al. Prevention of pelvic inflammatory disease by screening for cervical chlamydial infection. *N Engl J Med.* 1996;334:1362–1366.
3. Burstein GR, Zenilman JM. Nongonococcal urethritis—a new paradigm. *Clin Infect Dis.* 1999;28(suppl 1):S66–S73.
4. Judson FN. Gonorrhea. *Med Clin North Am.* 1990;74:1353–1366.
5. Koumans EH, Johnson RE, Knapp JS, et al. Laboratory testing for *Neisseria gonorrhoeae* by recently introduced nonculture tests: a performance review with clinical and public health considerations. *Clin Infect Dis.* 1998;27:1171–1180.
6. Chernesky MA, Martin DH, Hook EW, et al. Ability of new APTIMA CT and APTIMA GC assays to detect *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in male urine and urethral swabs. *J Clin Microbiol.* 2005;43:127–131.
7. Gaydos CA, Quinn TC, Willis D, et al. Performance of the APTIMA combo 2 assay for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in female urine and endocervical swab specimens. *J Clin Microbiol.* 2003;41:304–309.
8. Modarress KJ, Cullen AP, Jaffurs WJ Sr, et al. Detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in swab specimens by the Hybrid Capture II and PACE 2 nucleic acid probe tests. *Sex Transm Dis.* 1999;26:303–308.
9. Van der Pol B. COBAS Amplicor: an automated PCR system for detection of *C. trachomatis* and *N. gonorrhoeae*. *Expert Rev Mol Diagn.* 2002;2:379–389.
10. Van Der Pol B, Ferrero DV, Buck-Barrington L, et al. Multicenter evaluation of the BDProbeTec ET System for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in urine specimens, female endocervical swabs, and male urethral swabs. *J Clin Microbiol.* 2001;39:1008–1016.
11. Bianchi A, Moret F, Desrues JM, et al. PreservCyt transport medium used for the ThinPrep Pap test is a suitable medium for detection of *Chlamydia trachomatis* by the COBAS Amplicor CT/NG test: results of a preliminary study and future implications. *J Clin Microbiol.* 2002;40:1749–1754.
12. Chong S, Jang D, Song X, et al. Specimen processing and concentration of *Chlamydia trachomatis* added can influence false-negative rates in the LCx assay but not in the APTIMA Combo 2 assay when testing for inhibitors. *J Clin Microbiol.* 2003;41:778–782.
13. Diemert DJ, Libman MD, Lebel P. Confirmation by 16S rRNA PCR of the COBAS AMPLICOR CT/NG test for diagnosis of *Neisseria gonorrhoeae* infection in a low-prevalence population. *J Clin Microbiol.* 2002;40:4056–4059.
14. Farrell DJ. Evaluation of AMPLICOR *Neisseria gonorrhoeae* PCR using *cppB* nested PCR and 16S rRNA PCR. *J Clin Microbiol.* 1999;37:386–390.
15. Livengood CH 3rd, Wrenn JW. Evaluation of COBAS AMPLICOR (Roche): accuracy in detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* by coamplification of endocervical specimens. *J Clin Microbiol.* 2001;39:2928–2932.
16. Palmer HM, Mallinson H, Wood RL, et al. Evaluation of the specificities of five DNA amplification methods for the detection of *Neisseria gonorrhoeae. J Clin Microbiol.* 2003;41:835–837.
17. McAdam AJ. Discrepant analysis: how can we test a test? *J Clin Microbiol.* 2000;38:2027–2029.
18. Schachter J, Hook EW, Martin DH, et al. Confirming positive results of nucleic acid amplification tests (NAATs) for *Chlamydia trachomatis:* All NAATs are not created equal. *J Clin Microbiol.* 2005;43:1372–1373.
19. Schachter J, Hook EW 3rd, McCormack WM, et al. Ability of the Digene hybrid capture II test to identify *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in cervical specimens. *J Clin Microbiol.* 1999;37:3668–3671.

20. Stary A. Correct samples for diagnostic tests in sexually transmitted diseases: which sample for which test? *FEMS Immunol Med Microbiol.* 1999;24:455–459.
21. Watson EJ, Templeton A, Russell I, et al. The accuracy and efficacy of screening tests for *Chlamydia trachomatis*: a systematic review. *J Med Microbiol.* 2002;51:1021–1031.
22. Rosenstraus M, Wang Z, Chang SY, et al. An internal control for routine diagnostic PCR: design, properties, and effect on clinical performance. *J Clin Microbiol.* 1998;36:191–197.
23. Hagblom P, Korch C, Jonsson AB, et al. Intragenic variation by site-specific recombination in the cryptic plasmid of *Neisseria gonorrhoeae*. *J Bacteriol.* 1986;167:231–237.
24. Miyada CG, Born TL. A DNA sequence for the discrimination of *Neisseria gonorrhoeae* from other Neisseria species. *Mol Cell Probes.* 1991;5:327–335.
25. Van Der Pol B, Martin DH, Schachter J, et al. Enhancing the specificity of the COBAS AMPLICOR CT/NG test for *Neisseria gonorrhoeae* by retesting specimens with equivocal results. *J Clin Microbiol.* 2001;39:3092–3098.
26. Centers for Disease Control. Sexually transmitted diseases treatment guidelines 2002. *MMWR Recomm Rep.* 2002;51:1–78.
27. Gaydos CA, Crotchfelt KA, Shah N, et al. Evaluation of dry and wet transported intravaginal swabs in detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infections in female soldiers by PCR. *J Clin Microbiol.* 2002;40:758–761.
28. Hardy PH, Hardy JB, Nell EE, et al. Prevalence of six sexually transmitted disease agents among pregnant inner-city adolescents and pregnancy outcome. *Lancet.* 1984;2:333–337.
29. Mahony JB, Jang D, Chong S, et al. Detection of *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, *Ureaplasma urealyticum*, and *Mycoplasma genitalium* in first-void urine specimens by multiplex polymerase chain reaction. *Mol Diagn.* 1997;2:161–168.
30. Ostergaard L, Moller JK, Andersen B, et al. Diagnosis of urogenital *Chlamydia trachomatis* infection in women based on mailed samples obtained at home: multipractice comparative study. *BMJ.* 1996;313:1186–1189.
31. Sugunendran H, Birley HD, Mallinson H, et al. Comparison of urine, first and second endourethral swabs for PCR based detection of genital *Chlamydia trachomatis* infection in male patients. *Sex Transm Infect.* 2001;77:423–426.
32. Wiesenfeld HC, Lowry DL, Heine RP, et al. Self-collection of vaginal swabs for the detection of chlamydia, gonorrhea, and trichomoniasis: opportunity to encourage sexually transmitted disease testing among adolescents. *Sex Transm Dis.* 2001;28:321–325.
33. Petrin D, Delgaty K, Bhatt R, et al. Clinical and microbiological aspects of *Trichomonas vaginalis*. *Clin Microbiol Rev.* 1998;11:300–317.
34. Graves A, Gardner WA Jr. Pathogenicity of *Trichomonas vaginalis*. *Clin Obstet Gynecol.* 1993;36:145–152.
35. Schwebke JR. Update of trichomoniasis. *Sex Transm Infect.* 2002;78:378–379.
36. Krieger JN. Trichomoniasis in men: old issues and new data. *Sex Transm Dis.* 1995;22:83–96.
37. Laga M, Manoka A, Kivuvu M, et al. Non-ulcerative sexually transmitted diseases as risk factors for HIV-1 transmission in women: results from a cohort study. *AIDS.* 1993;7:95–102.
38. Moodley P, Connolly C, Sturm AW. Interrelationships among human immunodeficiency virus type 1 infection, bacterial vaginosis, trichomoniasis, and the presence of yeasts. *J Infect Dis.* 2002;185:69–73.
39. Rendon-Maldonado J, Espinosa-Cantellano M, Soler C, et al. Trichomonas vaginalis: in vitro attachment and internalization of HIV-1 and HIV-1-infected lymphocytes. *J Eukaryot Microbiol.* 2003;50:43–48.
40. Sorvillo F, Kerndt P. *Trichomonas vaginalis* and amplification of HIV-1 transmission. *Lancet.* 1998;351:213–214.
41. Cotch MF, Pastorek JG 2nd, Nugent RP, et al. *Trichomonas vaginalis* associated with low birth weight and preterm delivery. The Vaginal Infections and Prematurity Study Group. *Sex Transm Dis.* 1997;24:353–360.
42. Klebanoff MA, Carey JC, Hauth JC, et al. Failure of metronidazole to prevent preterm delivery among pregnant women with asymptomatic Trichomonas vaginalis infection. *N Engl J Med.* 2001;345:487–493.
43. Sutton MY, Sternberg M, Nsuami M, et al. Trichomoniasis in pregnant human immunodeficiency virus-infected and human immunodeficiency virus-uninfected Congolese women: prevalence, risk factors, and association with low birth weight. *Am J Obstet Gynecol.* 1999;181:656–662.
44. McCann JS. Comparison of direct microscopy and culture in the diagnosis of trichomoniasis. *Br J Vener Dis.* 1974;50:450–452.
45. Garber GE, Sibau L, Ma R, et al. Cell culture compared with broth for detection of *Trichomonas vaginalis*. *J Clin Microbiol.* 1987;25:1275–1279.
46. Caliendo AM, Jordan JA, Green AM, et al. Real-time PCR improves detection of Trichomonas vaginalis infection compared with culture using self-collected vaginal swabs. *Infect Dis Obstet Gynecol.* 2005;13:145–150.
47. Heine RP, Wiesenfeld HC, Sweet RL, et al. Polymerase chain reaction analysis of distal vaginal specimens: a less invasive strategy for detection of Trichomonas vaginalis. *Clin Infect Dis.* 1997;24:985–987.
48. Jordan JA, Lowery D, Trucco M. TaqMan-based detection of *Trichomonas vaginalis* DNA from female genital specimens. *J Clin Microbiol.* 2001;39:3819–3822.
49. Kengne P, Veas N, Vidal J, et al. *Trichomonas vaginalis*: repeated DNA target for highly sensitive and specific polymerase chain reaction diagnosis. *Cell Mol Biol.* 1994;40:819–831.
50. Madico G, Quinn TC, Rompalo A, et al. Diagnosis of *Trichomonas vaginalis* infection by PCR using vaginal swab samples. *J Clin Microbiol.* 1998;36:3205–3210.
51. Mayta H, Gilman RH, Calderon MM, et al. 18S ribosomal DNA-based PCR for diagnosis of *Trichomonas vaginalis*. *J Clin Microbiol.* 2000;38:2683–2687.
52. Paces J, Urbankova V, Urbanek P. Cloning and characterization of a repetitive DNA sequence specific for Trichomonas vaginalis. *Mol Biochem Parasitol.* 1992;54:247–255.
53. Riley DE, Roberts MC, Takayama T, et al. Development of a polymerase chain reaction-based diagnosis of *Trichomonas vaginalis*. *J Clin Microbiol.* 1992;30:465–472.
54. Reid R, Greenberg M, Jenson AB, et al. Sexually transmitted papillomaviral infections. I. The anatomic distribution and pathologic grade of neoplastic lesions associated with different viral types. *Am J Obstet Gynecol.* 1987;156:212–222.
55. Gissmann L, Wolnik L, Ikenberg H, et al. Human papillomavirus types 6 and 11 DNA sequences in genital and laryngeal papillomas and in some cervical cancers. *Proc Natl Acad Sci USA.* 1983;80:560–563.
56. Bosch FX, Manos MM, Munoz N, et al. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International biological study on cervical cancer (IBSCC) Study Group. *J Natl Cancer Inst.* 1995;87:796–802.
57. McCance DJ, Walker PG, Dyson JL, et al. Presence of human papillomavirus DNA sequences in cervical intraepithelial neoplasia. *Br Med J (Clin Res Ed).* 1983;287:784–788.
58. Zbar AP, Fenger C, Efron J, et al. The pathology and molecular biology of anal intraepithelial neoplasia: comparisons with cervical and vulvar intraepithelial carcinoma. *Int J Colorectal Dis.* 2002;17:203–215.
59. Crum CP, Ikenberg H, Richart RM, et al. Human papillomavirus type 16 and early cervical neoplasia. *N Engl J Med.* 1984;310:880–883.
60. Herrero R. Epidemiology of cervical cancer. *J Natl Cancer Inst Monogr.* 1996;21:1–6.
61. Ellerbrock TV, Chiasson MA, Bush TJ, et al. Incidence of cervical squamous intraepithelial lesions in HIV-infected women. *JAMA.* 2000;283:1031–1037.
62. Svare EI, Kjaer SK, Worm AM, et al. Risk factors for genital HPV DNA in men resemble those found in women: a study of male attendees at a Danish STD clinic. *Sex Transm Infect.* 2002;78:215–218.

63. Behbakht K, Friedman J, Heimler I, et al. Role of the vaginal microbiological ecosystem and cytokine profile in the promotion of cervical dysplasia: a case-control study. *Infect Dis Obstet Gynecol.* 2002;10:181–186.
64. Chin-Hong PV, Palefsky JM. Natural history and clinical management of anal human papillomavirus disease in men and women infected with human immunodeficiency virus. *Clin Infect Dis.* 2002;35:1127–1134.
65. Goedert JJ, Cote TR, Virgo P, et al. Spectrum of AIDS-associated malignant disorders. *Lancet.* 1998;351:1833–1839.
66. Smith JS, Herrero R, Bosetti C, et al. Herpes simplex virus-2 as a human papillomavirus cofactor in the etiology of invasive cervical cancer. *J Natl Cancer Inst.* 2002;94:1604–1613.
67. Wallin KL, Wiklund F, Luostarinen T, et al. A population-based prospective study of *Chlamydia trachomatis* infection and cervical carcinoma. *Int J Cancer.* 2002;101:371–374.
68. Cothran MM, White JP. Adolescent behavior and sexually transmitted diseases: the dilemma of human papillomavirus. *Health Care Women Int.* 2002;23:306–319.
69. Ressel GW. CDC releases 2002 guidelines for treating STDs: part II. Human papillomavirus and hepatitis. *Am Fam Physician.* 2002;66: 1996, 1999.
70. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol.* 1999;189:12–19.
71. Ronnett BM, Manos MM, Ransley JE, et al. Atypical glandular cells of undetermined significance (AGUS): cytopathologic features, histopathologic results, and human papillomavirus DNA detection. *Hum Pathol.* 1999;30:816–825.
72. Schenck U, Herbert A, Solomon D, et al. Terminology. International Academy of Cytology Task Force summary. Diagnostic cytology towards the 21st century: an international expert conference and tutorial. *Acta Cytol.* 1998;42:5–15.
73. Cox JT, Lorincz AT, Schiffman MH, et al. Human papillomavirus testing by hybrid capture appears to be useful in triaging women with a cytologic diagnosis of atypical squamous cells of undetermined significance. *Am J Obstet Gynecol.* 1995;172:946–954.
74. Solomon D, Schiffman M, Tarone R. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. *J Natl Cancer Inst.* 2001;93:293–299.
75. Berkova Z, Kaufmann RH, Unger ER, et al. The effect of time interval between referral and colposcopy on detection of human papillomavirus DNA and on outcome of biopsy. *Am J Obstet Gynecol.* 2003;188:932–937.
76. Kulasingam SL, Hughes JP, Kiviat NB, et al. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: comparison of sensitivity, specificity, and frequency of referral. *JAMA.* 2002;288:1749–1757.
77. Nobbenhuis MA, Walboomers JM, Helmerhorst TJ, et al. Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. *Lancet.* 1999;354: 20–25.
78. Wang SS, Walker JL, Schiffman M, et al. Evaluating the risk of cervical precancer with a combination of cytologic, virologic, and visual methods. *Cancer Epidemiol Biomarkers Prev.* 2005;14:2665–2668.
79. Peyton CL, Schiffman M, Lorincz AT, et al. Comparison of PCR- and hybrid capture-based human papillomavirus detection systems using multiple cervical specimen collection strategies. *J Clin Microbiol.* 1998;36:3248–3254.
80. Gravitt PE, Peyton CL, Alessi TQ, et al. Improved amplification of genital human papillomaviruses. *J Clin Microbiol.* 2000;38:357–361.
81. Kornegay JR, Shepard AP, Hankins C, et al. Nonisotopic detection of human papillomavirus DNA in clinical specimens using a consensus PCR and a generic probe mix in an enzyme-linked immunosorbent assay format. *J Clin Microbiol.* 2001;39:3530–3536.
82. Manos MM, Ting Y, Wright DK, et al. The use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. *Cancer Cells.* 1989;7:209–214.
83. Coutlee F, Gravitt P, Kornegay J, et al. Use of PGMY primers in L1 consensus PCR improves detection of human papillomavirus DNA in genital samples. *J Clin Microbiol.* 2002;40:902–907.
84. van Doorn LJ, Quint W, Kleter B, et al. Genotyping of human papillomavirus in liquid cytology cervical specimens by the PGMY line blot assay and the SPF(10) line probe assay. *J Clin Microbiol.* 2002;40:979–983.
85. Menezes G, Euscher E, Schwartz B, et al. Utility of the in situ detection of HPV in Pap smears diagnosed as within normal limits. *Acta Cytol.* 2001;45:919–926.
86. Castle PE, Solomon D, Hildesheim A, et al. Stability of archived liquid-based cervical cytologic specimens. *Cancer.* 2003;99:89–96.
87. Khan MJ, Castle PE, Lorincz AT, et al. The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. *J Nat Cancer Inst.* 2005;97:1072–1079.

Chapter 41

Respiratory Pathogens

Julie D. Fox and Peter A. Tilley

Introduction

Respiratory tract infections are among the most common presenting complaints of patients in both hospital and community settings. They are a considerable burden in terms of both patient morbidity and public health interventions. Laboratory diagnosis of respiratory tract infections should provide guidance in therapy and prognosis, as well as useful epidemiological information reflecting trends in the community. Understanding and monitoring such trends facilitates early recognition of new infectious agents in a population. A summary of the common viruses and bacteria causing respiratory tract infections and their clinical relevance is given in Tables 41-1 and 41-2, respectively.

Even with a significant clinical effort and analysis of multiple specimens, current laboratory methods fail to diagnose approximately half of lower respiratory tract infections. In fact, laboratory diagnosis of community-acquired pneumonia (CAP) is so poor that current clinical practice guidelines do not recommend testing for all but the most severely affected patients and advise use of empiric therapy.[1] This pragmatic approach fails to address issues of antimicrobial overuse and resistance, public health surveillance, and advancement of medical knowledge.

Many "atypical" bacteria are known to cause severe respiratory symptoms, but lack of good diagnostic procedures has hampered the measurement of the real impact of such infections in the community. Despite vaccination policies, *Bordetella pertussis* infection remains relatively common in children and adults and is associated with chronic cough in adults.[2] *Mycoplasma pneumoniae*, *Legionella pneumophila,* and *Chlamydophilia* (previously *Chlamydia*) *pneumoniae* are all recognized causes of lower respiratory tract infections, but again, their impact has not been studied in detail. In addition, despite the well-recognized association of viral infections with upper and lower respiratory tract infections, the current diagnostic virology procedures do not provide an answer rapidly enough to prevent inappropriate antibiotic use or to consider use of antiviral therapy.[3,4]

Molecular techniques have the potential to enhance our diagnostic approaches to respiratory pathogen identification and enable more detailed analysis of outbreaks. Use of nucleic acid detection methods has demonstrated that some organisms are more common and important causes of respiratory infection and disease than previously appreciated. Molecular methods are applied to the diagnostic detection and analysis of the viral and some of the atypical bacterial causes of respiratory infections.

Clinical Utility

Limitations of Conventional Diagnostic Procedures

Conventional diagnostic techniques (culture, antigen, and antibody detection) have been widely utilized for the diagnosis of individual infections and for the identification of respiratory outbreak pathogens with varying success. Culture-based methods are the mainstay for the diagnosis of more typical bacterial infections, such as *Streptococcus pneumoniae* and *Staphylococcus aureus,* and provide isolates for antimicrobial susceptibility testing. For many agents of CAP, however, culture methods have significant drawbacks. In particular, such diagnostic approaches have very low sensitivity for atypical bacteria, due to the fastidious nature of the organisms (e.g., *M. pneumoniae, C. pneumoniae, Bordetella pertussis*), and are too slow to influence patient management. In other cases, culture methods are hazardous and require enhanced containment laboratories (e.g., *Chlamydophilia psittaci, Coxiella burnetii, Francisella tularensis, Yersinia pestis*), which is a costly and not widely available option.

Cell culture for respiratory viruses is cumbersome, expensive, and available only in major medical centers. For many viral infections, suitable culture techniques and antibodies for isolate identification are not available. Thus, infections

Table 41-1. Respiratory Viruses in Acute and Public Health Settings

Viruses	Acute Infection	Public Health Significance and Community Impact	Conventional Diagnostics	Molecular Diagnostics and Typing
Influenza viruses (A and B)	Mild to severe URT and LRT infection	Responsible for outbreaks in the community and in immunocompromised individuals	DFA or culture can be utilized, but sensitivity depends on sample quality	Utilized for subtype and strain identification, particularly when culture is inefficient or impractical
Parainfluenza viruses (HPIV1-4)	Mild to severe URT and LRT infection	Impact of parainfluenza 4 not clear; others associated with outbreaks in the community and in immunocompromised individuals	DFA or culture can be utilized, but sensitivity depends on sample quality; HPIV4 is not usually identified	Some outbreaks investigated by sequencing to confirm relationship among viruses
Coronaviruses (and SARS-CoV)	Thought to be mild for most coronaviruses; SARS-CoV has a high morbidity and mortality	SARS-CoV spread by close contact and responsible for outbreaks in multiple countries; other coronaviruses not well studied	Not readily available and lack sensitivity; antibody responses slow to develop	Used to confirm SARS-CoV as a new introduction to man; sequencing used to differentiate among coronaviruses
Rhinoviruses	Usually mild in healthy individuals; exacerbations in asthmatic individuals reported	Common cause of mild URT symptoms with significant economic impact	Not readily available and lack sensitivity	Of only academic interest to date but potentially useful for studying the impact of this common virus in the community
Enteroviruses	Recognized as a cause of wide-ranging respiratory symptoms	Not well established for respiratory symptoms but recognized for other clinical manifestations	Only a proportion is culturable; DFA not established for all serotypes	Range of methods used to mirror serotyping procedures; molecular methods likely to replace typing methods that depend on culture
Respiratory syncytial virus (RSV)	Cause of mild to severe infection, particularly in infants and elderly	Responsible for outbreaks in the community and in immunocompromised individuals	DFA or culture can be utilized, but sensitivity depends on sample quality	Differentiated into subtypes RSVA and B based on sequence differences, but clinical relevance unproven
Metapneumovirus	Studies to date indicate clinical presentation similar to that of RSV	Common infection; likely responsible for outbreaks in the community and in immunocompromised individuals; common copathogen	Culture thought to be insensitive; DFA not yet available	Differentiated into two lineages based on sequence differences, but clinical relevance unproven
Respiratory adenoviruses	Mild to severe respiratory infection (URT and LRT)	Responsible for outbreaks in the community and in immunocompromised individuals	DFA or culture can be utilized, but sensitivity depends on sample quality	Differentiated into types and subtypes with recognized clinical relevance; persistence may create problems in interpretation

DFA, direct fluorescent antibody; HPIV, human parainfluenza virus; LRT, lower respiratory tract; RSV, respiratory syncytial virus; SARS-CoV, severe acute respiratory syndrome coronavirus; URT, upper respiratory tract.

Table 41-2. Respiratory Bacteria in Acute and Public Health Settings

Organism	Acute Infection	Public Health Significance and Community Impact	Conventional Diagnostics	Molecular Diagnostics and Typing
Bordetella pertussis	Mild to severe infection, especially in infants	Large community outbreaks are common	Culture and DFA have low sensitivity; latter also compromised by artifacts	PCR superior to culture but requires standardization; gradual evolution of strains reported
Legionella species	Mild to fatal infection	Frequent outbreaks, water or soil related	Culture performs well, but is slow; Urine antigen detection useful for *L. pneumophila* serogroup 1 in high-prevalence areas	PCR similar to culture for *L. pneumophila*; likely better for non-*pneumophila* species; typing may be useful for outbreaks, but available data are limited
Mycoplasma pneumonia	Common cause of CAP, usually mild	Endemic with occasional epidemics	IgM test useful for primary infection	PCR useful for rapid, sensitive detection
Chlamydophila pneumoniae	Common cause of CAP requiring hospitalization	Community outbreaks described	Limited; serology by MIF most accurate, but slow	PCR useful for rapid, sensitive detection; role of *C. pneumoniae* prolonged infection or colonization needs to be determined

DFA, direct fluorescent antibody; CAP, community acquired pneumonia; MIF, microimmunofluorescence; PCR, polymerase chain reaction.

with parainfluenza virus type 4, human coronaviruses, rhinoviruses, and some enteroviruses would not ordinarily be identified without RNA detection methods. The impact of such infections is only just being realized, and there is probably an underestimation of their clinical importance, particularly for immunocompromised individuals, the elderly, or those with underlying conditions such as asthma.[5,6]

Human metapneumovirus has been confirmed as an important cause of severe lower respiratory tract infection. The virus has been circulating for more than 50 years, and studies using molecular assays have confirmed its wide distribution,[7,8] but many laboratories have not been successful in isolating this virus. The recent identification of the agent causing severe acute respiratory syndrome (SARS), known as SARS-CoV, has illustrated the need for expansion of diagnostic testing to encompass new emerging viruses and the limitations of conventional virological laboratory approaches to respiratory pathogen diagnosis.[9,10]

Even if it is possible to culture viruses efficiently, isolation and confirmation of the cause of a cytopathic effect can take days to weeks, depending on the pathogen. Waiting for a culture-positive result can take many days, during which time the patient may be inappropriately treated and, if hospitalized, infection control measures may not be initiated. Also, the use of primary primate cells in culture (which gives the best yields of influenza and parainfluenza viruses) is unlikely to be sustainable in the long term.

When available, monoclonal antibodies are useful in direct virus-specific antigen detection tests, and these can be used for rapid diagnosis. Many laboratories are able to provide diagnostic testing for influenza, parainfluenza (types 1–3), respiratory syncytial virus (RSV), and respiratory adenoviruses. A respiratory specimen containing cells is necessary for sensitive detection of viruses by immunofluorescence or other antigen-detection methods. Good-quality diagnostic samples (often lavage or aspirate samples) can usually be obtained from young, immunocompetent, hospitalized individuals, but in other circumstances the ideal sample may not be available. Delays in transportation may reduce specimen quality and compromise assay results. The most-difficult (and least-efficient) specimens for diagnostic testing are swab samples containing minimal cellular material taken from largely asymptomatic individuals in the community. Smears from these samples can be difficult to interpret in a direct antigen test and thus culture of the sample is usually required for pathogen identification. Bacterial antigen detection by a direct fluorescent antibody (DFA) test is similarly compromised by poor sensitivity and has the added concern of artifacts leading to false-positive results, particularly for *B. pertussis* and *L. pneumophila*.

Serological assays of an antibody response to infection are available for some respiratory pathogens. Although useful for retrospective evidence of infection in a community, the results are not timely enough for patient management. For some cell-associated or intracellular bacteria and viruses, antibody responses develop slowly, if they develop at all, and convalescent sera taken many weeks after disease onset are required to make a definitive diagnosis (e.g., for *C. pneumoniae*, *Legionella* species, RSV). For other infections such as influenza, antibody responses are brisk, but frequent reinfection reduces the IgM response and convalescent sera are required to demonstrate changing titers.

Application of Molecular Assays to Respiratory Pathogen Diagnosis

The limitations of conventional testing are well recognized for viral[11,12] and bacterial[13,14] pathogens, and some of the

Table 41-3. Comparison of Nucleic Acid and Culture/Antigen Detection Methods for Respiratory Pathogen Diagnosis

	Nucleic Acid Methods	Culture/Antigen Methods
Cost	Tests tend to be expensive (but getting cheaper)	Relatively inexpensive in laboratories already set up for these procedures, but "real" cost of maintaining cultures often underestimated
Speed	Rapid diagnostic methods	Speed very variable depending on the pathogen and method used
Infrastructure	Specialized laboratory set-up required	Specialized laboratory set-up required
Spectrum	Specific sequence information required for design; triage of testing can be difficult; generic primers may be used to identify novel pathogens	"Catch-all" approach, which may be advantageous when novel pathogens need to be identified
Sensitivity	Exquisite sensitivity but can be prone to cross-contamination problems	Generally less sensitive than nucleic acid detection methods
Specificity	Careful handling required to avoid contamination (common problem); primer/probe design crucial	Careful handling required to avoid contamination, but less-common problem than for molecular methods; DFA subject to over interpretation
Strain typing	Most definitive method	Limited serotyping (e.g., influenza, *Legionella*)
Automation	Automated extraction equipment becoming available; automated detection commonplace	Difficult to automate
Safety	Inactivated before analysis, but antimicrobial sensitivity information requires knowledge of genotypic mutation	Isolates useful for antimicrobial sensitivity testing and phenotyping, but specialized safety requirements needed for culture of category 3/4 pathogens
Quality assurance	Proficiency and validation of methods not well established	Culture depends critically on cell line or medium quality; maintaining quality can be difficult

key differences between nucleic acid and culture or direct antigen testing methods are summarized in Table 41-3. Utilization of rapid viral diagnostic procedures, such as may be provided by molecular amplification methods, could help to reduce the emergence of antibiotic-resistant bacteria. One study demonstrated that a 52% reduction in antibiotic use was possible using molecular methods for viral diagnosis.[15]

Despite the obvious economic burden of CAP, we do not have accurate data on how many of these infections are viral in origin or caused by the atypical bacteria for which routine diagnostic testing is not readily available. The identification of emerging human viral infections (such as H5N1 influenza and SARS-CoV) has heightened awareness of the gaps in respiratory pathogen diagnosis. Nucleic acid detection assays are likely to be utilized more widely to identify novel emerging pathogens that could result in worldwide outbreaks. Molecular amplification assays are being used successfully for the identification of organisms associated with pneumonia.[16]

The potential for nosocomial spread of respiratory pathogens is well recognized for patients admitted to the hospital with RSV, parainfluenza, or other infections, serving as a reservoir for transmission to vulnerable patients and leading to possible outbreaks.[17] Previously, available diagnostic methods were not sensitive enough to identify sources of outbreaks, but the advent of molecular amplification methods has allowed even environmental sampling to be helpful in confirming outbreak sources and linking clustered cases of infection.[18]

Molecular Tests for Respiratory Pathogens

Samples and Nucleic Acid Extraction

Samples used for detection of respiratory pathogens include swabs (usually nasopharyngeal or throat), aspirates (nasopharyngeal or tracheal), sputum (usually from individuals presenting with pneumonia), or bronchoalveolar lavage specimens. For infections involving the entire respiratory tract, nasopharyngeal specimens are practical for diagnosis. For other infections, which are more focal,

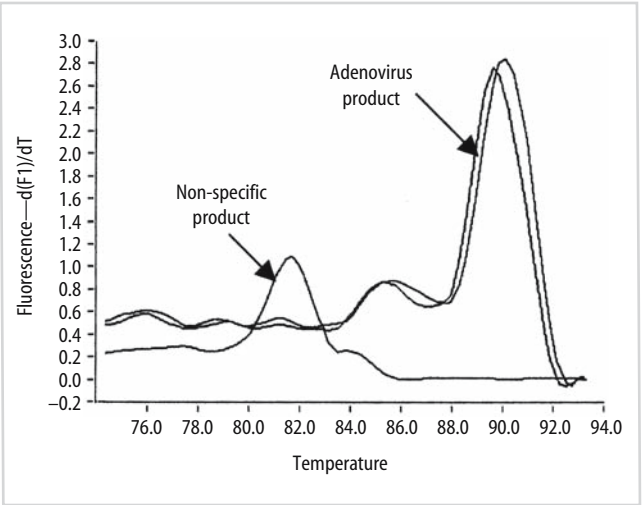


Figure 41-1. Melting-curve analysis of PCR products using the intercalating dye SYBR Green differentiates between specific and nonspecific products without further manipulation of the PCR products.

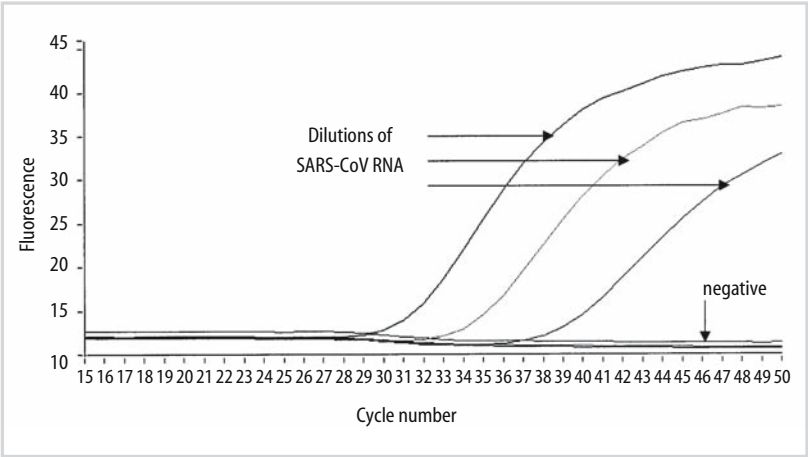


Figure 41-2. Detection of RT-PCR products for severe acute respiratory syndrome coronavirus (SARS-CoV) by a real-time PCR assay utilizing a hydrolysis (TaqMan) probe. The fluorescence signal is produced by measurement of the amount of fluorescent dye released each cycle.

lower respiratory tract specimens are required (e.g., for *Legionella*).

Sample preparation is a critical step for the detection and analysis of organisms. Numerous methods, from simple boiling to sophisticated automated protocols, are available for disruption of the organism and purification of the nucleic acids. Many studies have demonstrated inhibitors in respiratory specimens, making some form of extraction (with or without freezing) necessary to avoid frequent false negative results. Commercial kits for preparation of samples are available and should reduce interlaboratory variation in results. Simultaneous extraction of RNA and DNA facilitates assays for both viruses and bacteria.

Diagnostic Detection of Nucleic Acid

Molecular techniques for the detection and analysis of pathogens associated with respiratory infection provide specific diagnoses for individual cases and for outbreaks. Currently, FDA-cleared molecular tests are not available for the detection of respiratory pathogens, with the exception of *Mycobacterium tuberculosis* (see chapter 43). Molecular tests are performed using either validated laboratory-developed procedures or commercial testing reagents. Published diagnostic methods for detection of respiratory pathogen DNA or RNA directly from clinical specimens utilize target amplification procedures such as polymerase chain reaction (PCR) or nucleic acid sequence-based amplification (NASBA). Although direct detection methods based on nucleic acid hybridization would be theoretically possible, the amount of target nucleic acid in specimens may be minimal and such methods would lack sensitivity compared to amplification methods, unless the organism was propagated before analysis. Thus, the molecular amplification procedures reported for direct detection of respiratory pathogens in clinical samples include PCR (e.g., Reference 19 and Figure 41-1), reverse transcription-polymerase chain reaction (RT-PCR) (e.g., Reference 20 and Figures 41-2 and 41-3), and NASBA (e.g., References 21–23 and Figures 41-4 and 41-5). Target nucleic acid for amplification assays usually is a pathogen-specific gene or genes from the pathogen genomic DNA or RNA, but some

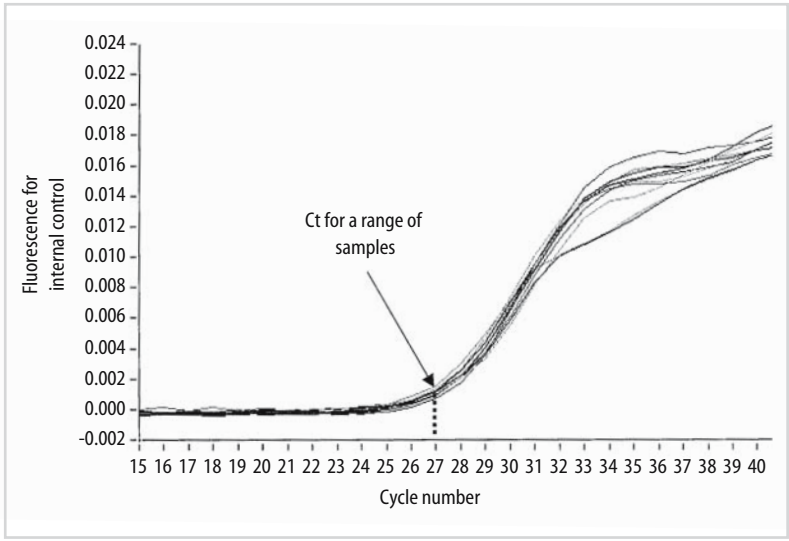


Figure 41-3. Analysis of internal control reactions for clinical samples in a real-time RT-PCR assay ensures that nucleic acid extraction of the sample was efficient and no residual inhibitors are present.

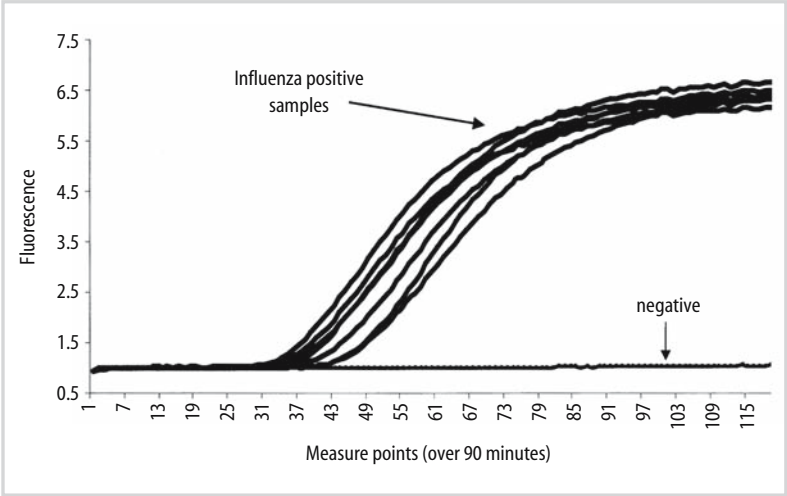


Figure 41-4. Real-time NASBA assay for influenza with detection of amplified products using a molecular beacon probe. Differentiation of positive from negative results is straightforward in this example.

assays have utilized bacterial ribosomal RNA (rRNA; e.g., Reference 22). For cellular samples tested for respiratory pathogens, targeting messenger RNA (mRNA) or genomic antisense RNA may enhance diagnostic sensitivity.

A variety of formats have been utilized for the detection of amplified products. Procedures that separate target amplification from the detection phase (agarose gel analysis or endpoint hybridization) are well established[24,25] and may allow multiple targets to be analyzed in a single reaction, providing added typing information.

For ease of use and incorporation into diagnostic laboratories, most laboratory-developed assays for detection of respiratory pathogens utilize real-time amplification methods in which the amplification and detection steps are combined. Some methods use intercalating dyes with the analysis of PCR product melting temperatures (e.g., as described previously[26] and illustrated in Figure 41-1), whereas others use fluorogenic primers or probes (e.g., TaqMan, hybridization format, and molecular beacons[19,21–23,27]) to ensure the specificity of the reaction. Figure 41-2 illustrates a real-time RT-PCR assay for SARS-CoV using probe-specific detection of amplified products (one fluorescence measurement per cycle). The range of real-time PCR or RT-PCR and NASBA methodologies used for respiratory targets is diverse given the fact that assays are laboratory developed.[28,29]

The difficulty with diagnosis of respiratory infections is the wide range of pathogens with similar presentations. The nucleic acid technologies utilized currently in the majority of diagnostic laboratories are real-time PCR single-target assays. In some cases, generic primers can be designed to pick up several related pathogens. Such generic primers may be based on conserved protein coding sequences (such as those essential for enzyme function) or noncoding regions for which variation is limited because of the need for maintenance of secondary structure. The use of primer sets to pick up genera or even families of organisms has shown promise in limited studies, including analysis of *Legionella* and *Mycobacteria*. Generic assays have the ability to detect many related organisms and may be used to characterize previously undescribed species in respiratory infection.

Limited multiplex procedures have been reported for detection of related organisms using real-time PCR or NASBA procedures, but such assays are difficult to set up and control. For many respiratory pathogens, there is sufficient variation that multiplex approaches have been developed to detect, for example, all possible respiratory

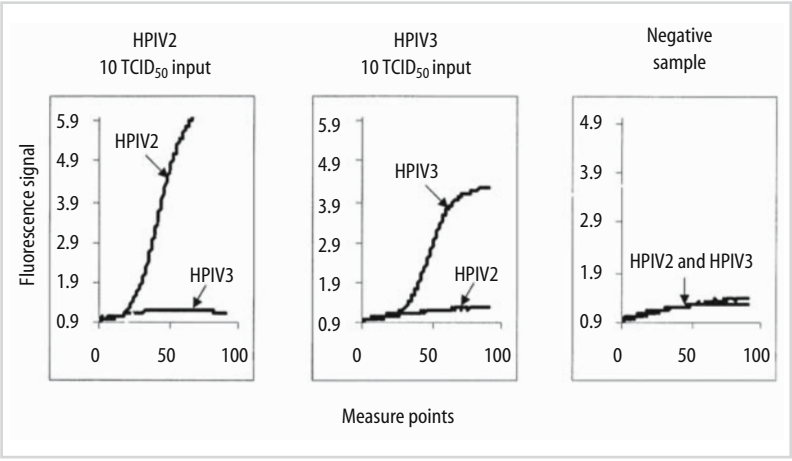


Figure 41-5. Multiplex real-time NASBA assay for human parainfluenza virus (HPIV) with simultaneous detection and differentiation of two target types using two sets of primers and specific molecular beacons labeled with FAM and ROX, respectively. $TCID_{50}$, 50% Tissue Culture Infecting Dose.

adenovirus,[19] influenza,[20] or parainfluenza[21] types. A simple, dual-labeled multiplex NASBA assay is shown in Figure 41-5 that uses separate primer sets and specific molecular beacon probes for parainfluenza types 2 and 3 (HPIV2 and HPIV3) in the same reaction mix. Each probe is labeled with a different fluorophore, allowing detection and differentiation of both viruses in a single reaction. An ambitious multiplex nested RT-PCR procedure with gel analysis of the amplicons detects influenza A, B, and C viruses, RSV (A and B subtypes), and adenoviruses in a single assay.[30] The procedure, while complex to set up and validate, was reported to have good specificity and better sensitivity than antigen/culture procedures.

Interpretation and validation of a negative result are important parts of diagnostic tests based on nucleic acid amplification. Some assays incorporate an internal control system to distinguish true-negative from false-negative results. The internal control may amplify with the pathogen-specific primers but result in an amplicon with a different size or internal sequence from the pathogen amplicon. Alternatively, the internal control may be an external sequence spiked into the reaction (heterologous control) and amplified with a primer set different from the pathogen primers. In the example shown in Figure 41-3, amplification of the RNA heterologous control is consistent across many clinical samples and ensures that there are no gross inhibitors present in the reactions. For cellular samples obtained for detection of respiratory pathogens, the internal control reaction can utilize human DNA, rRNA, or mRNA detection. Such approaches have the added value of assessing for sample collection and integrity, as well as amplification inhibitors.

The relative merits of commercial versus laboratory-developed tests depend on the laboratory facilities, the technical expertise available, and the clinical need for expanded diagnosis. Commercial testing reagents provide quality controls and procedure standardization that facilitate clinical studies.[24,32] Many companies are focusing on providing analyte-specific reagents (ASRs) for respiratory pathogen assays. ASRs will provide the laboratory with quality-controlled primers and probes, while allowing them the flexibility to test for currently known circulating pathogens or according to a local testing algorithm. Microarray-based detection of multiplexed PCR products also has been reported.[31]

Microbial Typing and Respiratory Outbreak Investigation

Classically, typing of bacteria or viruses has used serological techniques that rely on antibody-antigen interactions. One benefit of approaches based on DNA or RNA detection is the more-detailed, quantitative assessment of the relationship between organisms, providing valuable data relevant to outbreak investigations and community health. Variation at the nucleic acid sequence level is not necessarily reflected in altered protein sequence or function; thus, additional sequence variation information may not correlate with conventional typing methods.

Restriction fragment length polymorphism (RFLP) analysis by pulsed field gel electrophoresis (PFGE), either with or without blot hybridization, has been utilized for analysis of complex DNA genomes from a variety of respiratory pathogens. RFLP analysis also has been applied to PCR or RT-PCR products from respiratory bacteria and viruses. In general, such methods can provide resolution down to the subtype level and have proven useful in outbreak investigation, as illustrated in Figure 41-6 for *B. pertussis* isolates. The difficulty with gel-based typing assays, such as PFGE, is standardizing results and sharing data between laboratories. Amplified fragment length polymorphism (AFLP) analysis represents an alternative method with better discriminatory power and portability, but this approach has not been used extensively for respiratory isolates to date.

For respiratory viruses, other methods have been used for typing, including heteroduplex mobility assay (HMA), single-strand conformation polymorphism (SSCP), and RFLP analysis of amplified PCR products. In general, HMA is considered technically complex but has the capacity to distinguish viral quasispecies with >3% nucleotide

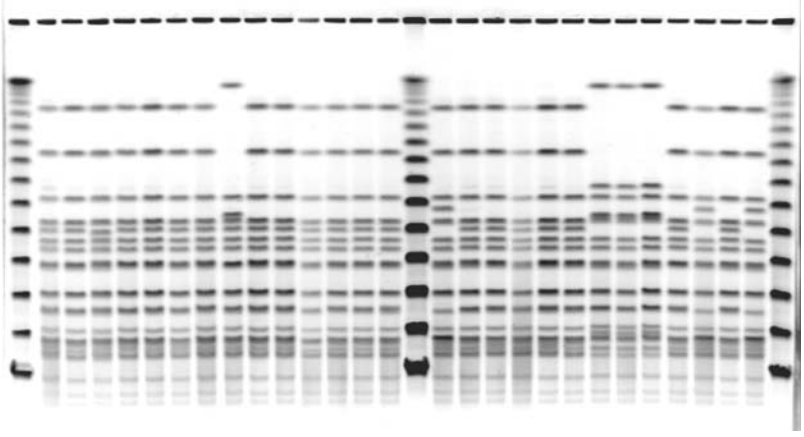


Figure 41-6. Pulsed field gel electrophoresis for analysis of *Bordetella pertussis* isolates. (Figure kindly provided by Dr, M. Peppler, University of Alberta.)

differences. SSCP and RFLP, while technically easier, generally can resolve viruses only to the subtype level, and RFLP has the added constraint of assessing only sequence differences in restriction sites.

The use of sequencing to assess the relationship among viruses is well established, and molecular phylogenetic knowledge is expanding, allowing modeling of viral populations and prediction of new outbreaks.[33] The level of resolution using primary sequence is at one genome, and point mutations can be identified. Sometimes this provides more information than originally sought and creates problems in interpretation; while in other circumstances even small sequence variations can confer important changes in viral transmissibility and disease outcome. Identification of emerging viruses, which may have been recently introduced into the human population (e.g., SARS-CoV and influenza A H5N1 types), is critical to public health. Analysis of such novel viruses has relied heavily on sequencing of isolates or amplicons.[9,10,34]

Human influenza A viruses are associated with enhanced morbidity and mortality compared with influenza B or influenza C viruses. Differences in pathogenicity for subtypes of influenza A also have been reported; for example, H3N2 is associated with more severe infection than H1N1. Such types and subtypes of influenza

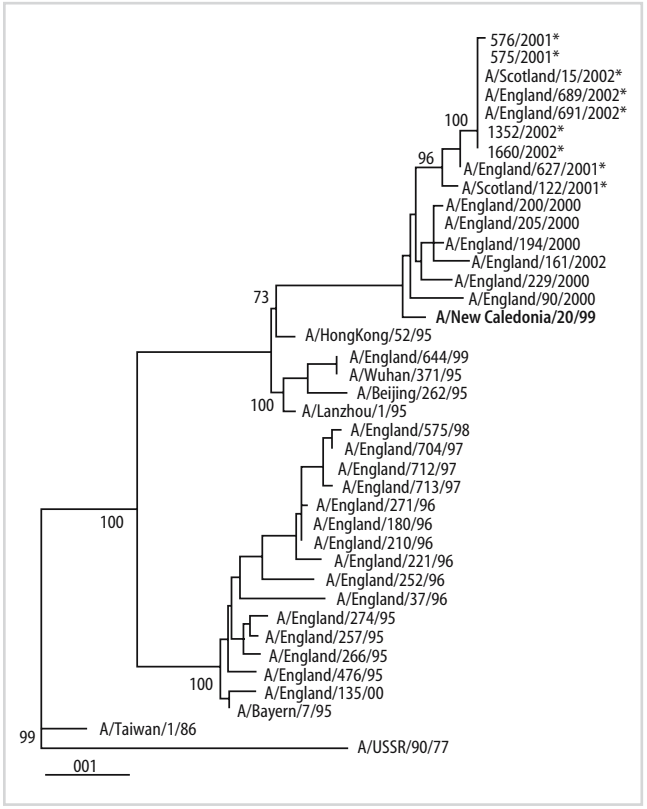


Figure 41-7. Sequence comparison of hemagglutinin (HA) for H1N1 and H1N2 viruses circulating in the United Kingdom in 2001 to 2002. The H1N1 vaccine strain is in bold typeface. H1N2 strains are indicated with asterisks. (Reproduced from Ellis JS, Alvarez-Aguero A, Gregory V, et al. Influenza AH1N2 viruses, United Kingdom, 2001–02 influenza season. *Emerg Infect Dis.* 2003;9:304–310.)

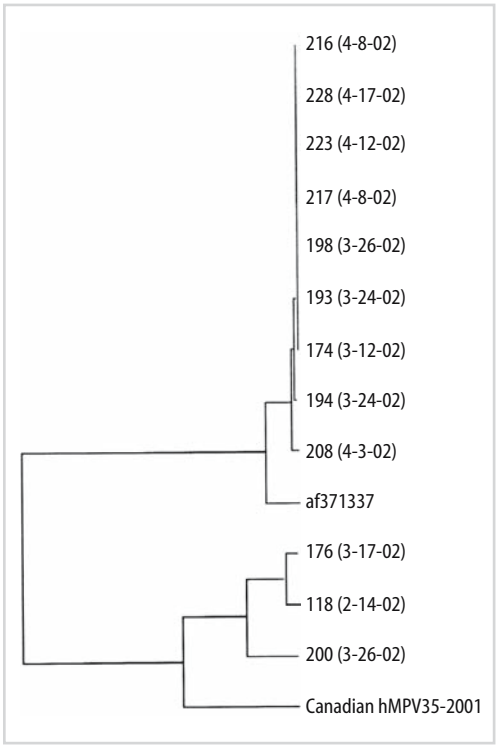


Figure 41-8. Phylogenetic tree illustrating the relationships among fusion gene sequences for 12 human metapneumovirus strains (from 2002). af371337 is the GenBank accession number for the prototype human metapneumovirus strain, and HMPV 35 is a Canadian strain isolated in 2001. (Reprinted from Boivin G, De Serres G, Cote S, et al. Human metapneumovirus infections in hospitalized children. *Emerg Infect Dis.* 2003;9:634–640.)

A can circulate independently, and their identification is important for assessment of current vaccine efficacy. Reassortment of the two predominant influenza subtypes infecting humans in recent times has been reported, and analysis of main hemagglutinin (HA) types has been undertaken by a range of molecular and nonmolecular methods.[35,36] Detailed sequence comparison for HA of H1N1 and H1N2 viruses circulating in the United Kingdom in 2001–2002 (Figure 41-7 and Reference 37) illustrates the utility of sequence analysis in understanding viral divergence and relationship to current vaccine use. Reports of avian H5N1 viruses infecting humans (1997–1998 and 2003–2004[34,38]) emphasize the need for detailed surveillance of influenza viruses and vigilance in identification of emerging viruses of importance to public health. Detailed analysis of avian H5N1 viruses that have infected humans to date have confirmed that all genes are of avian origin and are associated with minimal or very inefficient human-to-human spread. The potential for reassorted viruses that could more easily spread among humans is clear, and molecular methods are now an important part of influenza surveillance.

Recent studies of human metapneumovirus have identified two main lineages, with sequence diversity within each group (Figure 41-8[40]), thus displaying a similar pattern to RSV isolates that are classified into two major

groups, A and B.[39] Further studies will confirm whether this distinction is associated with differences in virulence.

Sequence analysis for typing of bacteria has been slower to develop than that for viruses but has been utilized for investigation of some atypical bacteria associated with outbreaks of respiratory infection (e.g., *Legionella*[41]). Due to the problem of recombination, characterization of a single bacterial gene often does not reflect the organism as a whole. Multilocus sequence typing is a strategy that addresses this and appears useful for analysis of *B. pertussis*.[42]

Microarray hybridization methods have been used to identify and differentiate related pathogens.[31] Such an approach was useful in first identifying the agent of SARS as a coronavirus.[43] Molecular methods provide additional information about the virulence and type of infectious organism, as illustrated by recent experience with SARS-CoV and influenza types.

Interpretation of Test Results

Molecular tests have advantages over conventional procedures, but the sensitivity of molecular amplification methods can lead to problems with interpretation of results. For many organisms, a gold standard method is not available that accurately reflects the enhanced sensitivity of molecular methods, as has been seen with PCR testing for *B. pertussis* or *C. pneumoniae* in clinical samples. Studies confirm that PCR tests are very sensitive, and PCR-positive individuals may be culture negative or asymptomatic, so that results must always be interpreted in the clinical context.

Inhibitors of amplification are common in respiratory specimens, so a negative result must be interpreted in the context of the nucleic acid extraction method and the control results to monitor for nucleic acid degradation and amplification inhibition. When assays for the detection of respiratory pathogens are designed, primers and probes should not cross-react with normal respiratory flora or other respiratory pathogens.

Laboratory Issues

The triage of molecular testing for respiratory infection diagnosis is difficult. Currently, a single respiratory pathogen test detects only one or a few related pathogens. Also, bacterial testing and viral testing are not combined. Thus, many molecular tests must be used to screen for all appropriate pathogens, which increases testing costs. Thus, a laboratory that embarks on using molecular methods for the diagnosis of respiratory infections may require a complex testing algorithm. One approach is use a multiplex amplification procedure to identify multiple pathogens in a single assay, with certain assays now commercially available.[24,32] Unfortunately, such tests tend to be expensive and, if developed by the laboratory, are very difficult to control and ensure equal sensitivity and specificity for all pathogens. Thus, despite the potential for replacement of many culture and antigen procedures with nucleic acid amplification assays, such a molecular diagnostic revolution has not yet happened. The exception is for new pathogens when nucleic acid amplification and detection methods are clearly far superior to alternatives (e.g., metapneumovirus, SARS-CoV) or for testing of samples that are suboptimal for routine procedures (e.g., in surveillance situations).

Future Directions

Respiratory infections are currently underdiagnosed, despite the fact that accurate pathogen identification is important to ensure appropriate patient management and monitor infectious trends in the community. The major stumbling blocks in the diagnosis and investigation of respiratory infections are the complexity of testing algorithms and the number of potential targets that cause both upper and lower respiratory tract symptoms. Real-time PCR methods have vastly improved the sensitivity for detection and recognition of some difficult-to-culture organisms, and will likely become standard practice in the clinical laboratory in the next few years. There is, however, a limit to how many organisms can be "multiplexed" in a single test.

Microarray hybridization of randomly amplified PCR products from respiratory cultures and clinical samples has shown some success.[31] If the promise of early experiments is maintained when applied to large-scale clinical studies, this could answer some of the technical problems surrounding the use of multiplex systems. Microarray hybridization, while not currently as convenient as real-time PCR detection methods, potentially has the benefit of being able to resolve complex product mixtures and provide clinically valuable information.

The use of molecular methods for typing and outbreak investigation of respiratory pathogens of public health importance is well established and is likely to expand. Future directions will incorporate the use of microarray systems for respiratory pathogen detection and analysis to allow crossing of the barriers between conventional virology and bacteriology (and mycology/parasitology). Once microarray systems have been developed and validated, the costs of this enhanced technology may be reduced and justifiable.

Identification of novel viruses, which have presumably only recently been introduced to humans, has reinforced the need for careful surveillance of emerging respiratory pathogens and institution of appropriate infection control measures. Lessons should be learned from the continuous sensitive surveillance and typing of organisms such as influenza and *B. pertussis* to direct the use and efficacy of available vaccines.

Molecular techniques developed for detection and analysis of the microbes responsible for respiratory

infections will be vital to our understanding of pathogenic mechanisms, appropriate management, and prevention of outbreaks in the future. Gel-based typing procedures (PFGE, HMA, SSCP) slowly will be replaced by sequence-based alternatives (e.g., multilocus sequence typing or MLST), which are more amenable to standardization and sharing of data among laboratories.

References

1. Mandell LA, Marrie TJ, Grossman RF, et al. Canadian guidelines for the initial management of community-acquired pneumonia: an evidence-based update by the Canadian Infectious Diseases Society and the Canadian Thoracic Society. The Canadian Community-Acquired Pneumonia Working Group. *Clin Infect Dis.* 2000;31:383–421.
2. Birkebaek NH. Bordetella pertussis in the aetiology of chronic cough in adults. Diagnostic methods and clinic. *Dan Med Bull.* 2001;48: 77–80.
3. Barenfanger J, Drake C, Leon N, et al. Clinical and financial benefits of rapid detection of respiratory viruses: An outcomes study. *J Clin Microbiol.* 2000;38:2824–2828.
4. Zambon MC, Stockton JD, Clewley JP, et al. Contribution of influenza and respiratory syncytial virus to community cases of influenza-like illness: an observational study. *Lancet.* 2001;358:1410–1416.
5. Atmar RL, Guy E, Guntupalli KK, et al. Respiratory tract viral infections in inner-city asthmatic adults. *Arch Intern Med.* 1998;158:2453–2459.
6. van Elden LJ, van Kraaij MG, Nijhuis M, et al. Polymerase chain reaction is more sensitive than viral culture and antigen testing for the detection of respiratory viruses in adults with hematological cancer and pneumonia. *Clin Infect Dis.* 2002;34:177–183.
7. Falsey AR, Erdman D, Anderson LJ, et al. Human metapneumovirus infections in young and elderly adults. *J Infect Dis.* 2003;187:785–790.
8. Stockton J, Stephenson I, Fleming D, et al. Human metapneumovirus as a cause of community-acquired respiratory illness. *Emerg Infect Dis.* 2002;8:897–901.
9. Poutanen SM, Low DE, Henry B, et al. Identification of severe acute respiratory syndrome in Canada. *N Engl J Med.* 2003;348:1967–1976.
10. Drosten C, Gunther S, Preiser W, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med.* 2003;348:1995–2005.
11. Mitchell S, O'Neill HJ, Ong GM, et al. Clinical assessment of a generic DNA amplification assay for the identification of respiratory adenovirus infections. *J Clin Virol.* 2003;26:331–338.
12. Ong GM, Wyatt DE, O'Neill HJ, et al. A comparison of nested polymerase chain reaction and immunofluorescence for the diagnosis of respiratory infections in children with bronchiolitis, and the implications for a cohorting strategy. *J Hosp Infect.* 2001;49: 122–128.
13. Tilley PA, Kanchana MV, Knight I, et al. Detection of Bordetella pertussis in a clinical laboratory by culture, polymerase chain reaction, and direct fluorescent antibody staining; accuracy, and cost. *Diagn Microbiol Infect Dis.* 2000;37:17–23.
14. Qasem JA, Khan ZU, Shiji G, et al. Polymerase chain reaction as a sensitive and rapid method for specific detection of Mycoplasma pneumoniae in clinical samples. *Microbiol Res.* 2002;157:77–82.
15. Woo PCY, Chiu SS, Seto W-H, et al. Cost-effectiveness of rapid diagnosis of viral respiratory tract infections in pediatric patients. *J Clin Microbiol.* 1997;35:1579–1581.
16. Murdoch DR. Nucleic acid amplification tests for the diagnosis of pneumonia. *Clin Infect Dis.* 2003;36:1162–1170.
17. Zambon M, Bull T, Sadler CJ, et al. Molecular epidemiology of two consecutive outbreaks of parainfluenza 3 in a bone marrow transplant unit. *J Clin Microbiol.* 1998;36:2289–2293.
18. Echavarria M, Kolavic SA, Cersovsky S, et al. Detection of adenoviruses (AdV) in culture-negative environmental samples by PCR during an AdV-associated respiratory disease outbreak. *J Clin Microbiol.* 2000;38:2982–2984.
19. Gu Z, Belzer SW, Gibson CS, et al. Multiplexed, real-time PCR for quantitative detection of human adenovirus. *J Clin Microbiol.* 2003;41:4636–4641.
20. Smith AB, Mock V, Melear R, et al. Rapid detection of influenza A and B viruses in clinical specimens by Light Cycler real time RT-PCR. *J Clin Virol.* 2003;28:51–58.
21. Hibbitts S, Rahman A, John R, et al. Development and evaluation of NucliSens basic kit NASBA for diagnosis of parainfluenza virus infection with "end-point" and "real-time" detection. *J Virol Methods.* 2003;108:145–155.
22. Loens K, Ieven M, Ursi D, et al. Detection of Mycoplasma pneumoniae by real-time nucleic acid sequence-based amplification. *J Clin Microbiol.* 2003;41:4448–4450.
23. Fox JD, Han S, Samuelson A, et al. Development and evaluation of nucleic acid sequence based amplification (NASBA) for diagnosis of enterovirus infections using the NucliSens Basic Kit. *J Clin Virol.* 2002;24:117–130.
24. Fan J, Henrickson KJ, Savatski LL, et al. Rapid simultaneous diagnosis of infections with respiratory syncytial viruses A and B, influenza viruses A and B, and human parainfluenza virus types 1, 2, and 3 by multiplex quantitative reverse transcription-polymerase chain reaction-enzyme hybridization assay (Hexaplex). *Clin Infect Dis.* 1998;26:1397–1402.
25. Stockton J, Ellis JS, Saville M, et al. Multiplex PCR for typing and subtyping influenza and respiratory syncytial viruses. *J Clin Microbiol.* 1998;36:2990–2995.
26. Boivin G, Cote S, Dery P, et al. Multiplex real-time PCR assay for detection of influenza and human respiratory syncytial viruses. *J Clin Microbiol.* 2004;42:45–51.
27. Templeton KE, Scheltinga SA, Sillekens P, et al. Development and clinical evaluation of an internally controlled, single-tube multiplex real-time PCR assay for detection of Legionella pneumophila and other Legionella species. *J Clin Microbiol.* 2003;41:4016–4021.
28. Mackay IM, Arden KE, Nitsche A. Real-time PCR in virology. *Nucleic Acids Res.* 2002;30:1292–1305.
29. Chan AB, Fox JD. NASBA and other transcription-based amplification methods for research and diagnostic microbiology. *Rev Med Microbiol.* 1999;10:185–196.
30. Coiras MT, Perez-Brena P, Garcia ML, et al. Simultaneous detection of influenza A, B, and C viruses, respiratory syncytial virus, and adenoviruses in clinical samples by multiplex reverse transcription nested-PCR assay. *J Med Virol.* 2003;69:132–144.
31. Wang D, Coscoy L, Zylberberg M, et al. Microarray-based detection and genotyping of viral pathogens. *Proc Natl Acad Sci U S A.* 2002;99:15687–15692.
32. Hindiyeh M, Hillyard DR, Carroll KC, et al. Evaluation of the Prodesse Hexaplex multiplex PCR assay for direct detection of seven respiratory viruses in clinical specimens. *Am J Clin Pathol.* 2001;116: 218–224.
33. McCormack GP, Clewley JP. The application of molecular phylogenetics to the analysis of viral genome diversity and evolution. *Rev Med Virol.* 2002;12:221–238.
34. Subbarao K, Shaw MW. Molecular aspects of avian influenza (H5N1) viruses isolated from humans. *Rev Med Virol.* 2000;10:337–348.
35. Ellis JS, Zambon MC. Molecular diagnosis of influenza. *Rev Med Virol.* 2002;12:375–389.
36. Poddar SK. Influenza virus types and subtypes detection by single step single tube multiplex reverse transcription-polymerase chain reaction (RT-PCR) and agarose gel electrophoresis. *J Virol Methods.* 2002;99:63–70.
37. Ellis JS, Alvarez-Aguero A, Gregory V, et al. Influenza AH1N2 viruses, United Kingdom, 2001–02 influenza season. *Emerg Infect Dis.* 2003;9: 304–310.

38. Shaw M, Cooper L, Xu X, et al. Molecular changes associated with the transmission of avian influenza A H5N1 and H9N2 viruses to humans. *J Med Virol.* 2002;66:107–114.
39. Peret TC, Hall CB, Schnabel KC, et al. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J Gen Virol.* 1998;79:2221–2229.
40. Boivin G, De Serres G, Cote S, et al. Human metapneumovirus infections in hospitalized children. *Emerg Infect Dis.* 2003;9:634–640.
41. Cloud JL, Carroll KC, Pixton P, et al. Detection of Legionella species in respiratory specimens using PCR with sequencing confirmation. *J Clin Microbiol.* 2000;38:1709–1712.
42. van Loo IH, Heuvelman KJ, King AJ, et al. Multilocus sequence typing of Bordetella pertussis based on surface protein genes. *J Clin Microbiol.* 2002;40:1994–2001.
43. Rota PA, Oberste MS, Monroe SS, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science.* 2003;300:1394–1399.

Chapter 42

Bacterial Pathogens

Ruth Ann Luna and James Versalovic

Introduction

Bacterial infections represent important diseases worldwide despite decades of antibiotic therapy. Diverse microbial pathogens continue to rapidly evolve and present challenges for medical practice that will require ongoing refinements in laboratory-based diagnostic strategies. Since the 1970s, the steady parade of bacterial pathogen discoveries such as *Legionella pneumophila*, *Helicobacter pylori*, and *Bartonella henselae* have highlighted the ongoing importance of bacterial evolution in human infectious diseases. Established bacterial pathogens such as *Streptococcus pyogenes* and *Mycobacterium tuberculosis* have reemerged during the past two decades. Drug-resistant pathogens including multidrug-resistant organisms spread to different geographic areas, ignoring regional boundaries with the assistance of global immigration and travel. Advances in medicine including oncology and transplantation have resulted in greater numbers of immunocompromised patients with increased risks for invasive bacterial infections.

During the 20th century, clinical microbiology developed into an important part of clinical laboratory services and modern laboratory medicine. Since the advent of the Gram stain in the late 1800s, microbiologists visualized bacterial morphology. Bacteriologic culture on plated or in liquid media usually represented the gold standard for the diagnosis of bacterial infections. Continual development of new media formulations resulted in the improved ability to cultivate diverse bacterial pathogens. Challenges with difficult-to-culture or unculturable pathogens were partly addressed with the development of culture-independent serologic and antigen-detection methods.

Despite the established utility of different methods in clinical microbiology, the advent of molecular microbiology methods has resulted in many improvements in the ability to diagnose and monitor bacterial infections. During the past two decades, molecular methods and applications have gained acceptance by the clinical microbiology community. The successful integration of molecular methods with other laboratory approaches in clinical microbiology requires careful evaluation of existing techniques and clinical relevance.

Bacterial Identification and Culture Confirmation

Conventional clinical microbiology depends on microbial culture and subsequent biochemical testing for accurate identification of bacterial pathogens. DNA probe–based solution hybridization with chemiluminescent signal detection represents a widely accepted strategy for bacterial identification using molecular methods. DNA:RNA hybridization assays, especially those using synthetic oligonucleotide probes targeting ribosomal RNA (rRNA), have been used widely for the identification of a number of medically important microorganisms either following culture, as confirmation, or directly in specimens, as detection. The AccuProbe technology (Gen-Probe Inc, San Diego, CA) is the most widely used method and is based on acridinium ester-labeled DNA probes that are complementary to species-specific 16S rRNA sequences.[1] The AccuProbe method includes DNA:RNA hybridization by the hybridization protection assay (HPA) in a liquid format and chemiluminescence for rapid signal detection. These assays are limited by the availability of commercial DNA probes, and each assay enables identification of a single pathogen or closely related pathogens.

Nucleic acid genotyping and sequencing methods have provided alternative strategies for bacterial identification (Figure 42-1). Sequence and chromosomal DNA profile approaches facilitate identification that is database driven and not limited to particular pathogens for individual assays. Novel pathogens also may be detected by these methods, as sequences or DNA profiles may not match entries in databases and represent unique patterns.

The characterization and comparisons of small and large subunit rRNA sequences (5S, 16S, and 23S) for bacterial identification and phylogenetic studies[2] have

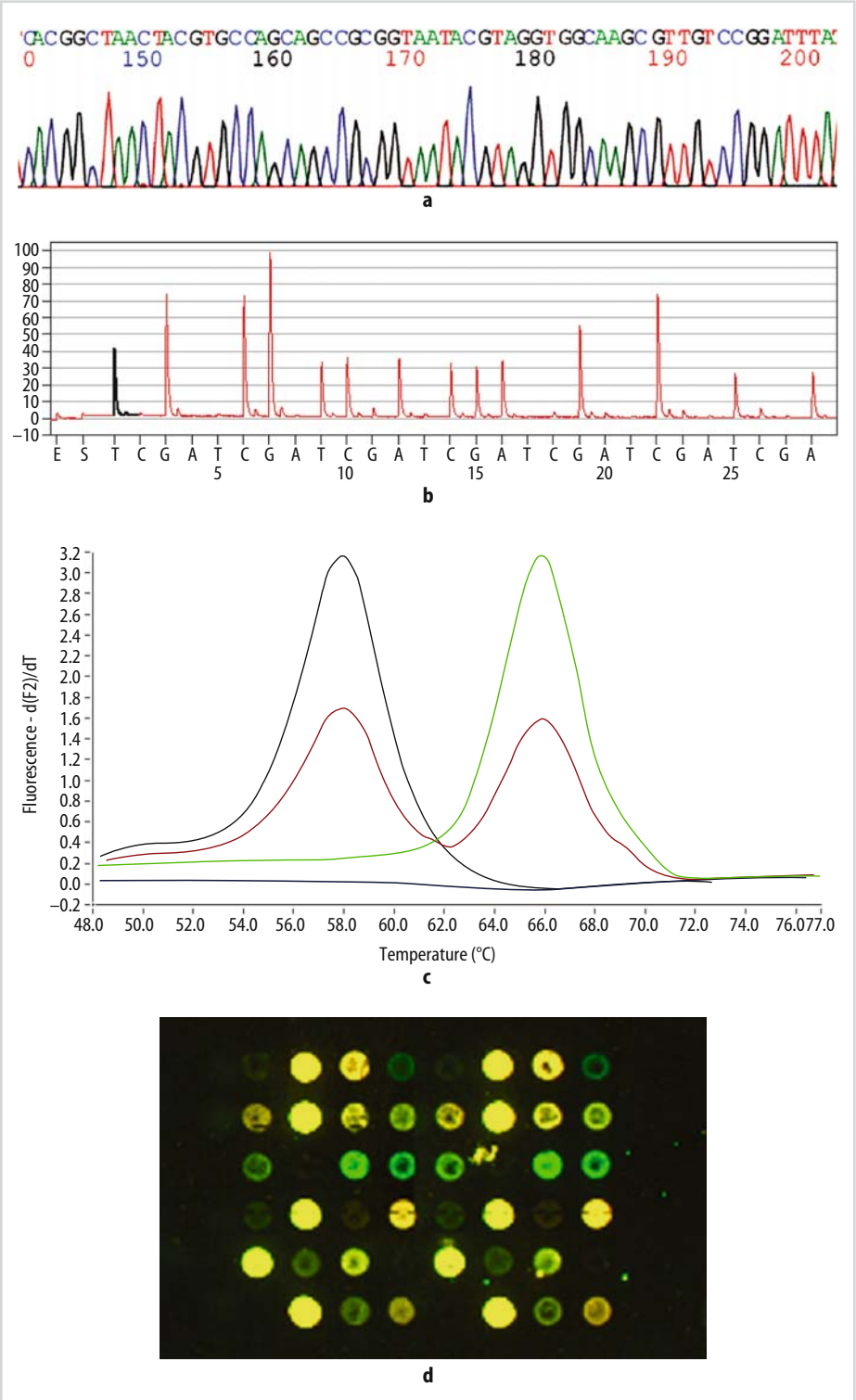


Figure 42-1. Genotyping strategies. (a) Chromatogram generated by DNA sequencing and fluorescent signal detection. (b) Pyrogram generated by pyrosequencing and bioluminescence detection. (c) Melting-curve analyses for characterization of amplicon sequences generated by real-time PCR. (d) Microarray data generated by gene expression profiling of two cell populations.

facilitated research and development of diagnostic applications. Bacterial rRNA sequences have been compiled into separate sequence databases (e.g., Ribosomal Database Project II) and combined with software tools for bacterial identification and phylogenetic studies (e.g., http://rdp.cme.msu.edu).[3,4] DNA sequencing of all or part of the 16S and 23S rRNA genes has been useful for identification of diverse bacterial pathogens[5] and has resulted in the development of commercial sequencing platforms (e.g., MicroSeq, Applied Biosystems, Foster City, CA) for microbial identification. Genotypic comparisons based on sequencing of the first 527 base pairs (bp) of the 16S rRNA gene resulted in successful genus- and species-level identification (100% and 93.1%, respectively) of unusual aerobic pathogenic gram-negative bacilli.[6] The juxtaposition of highly conserved and variable sequences

within the 16S or 23S rRNA genes[7,8] has culminated in subsequence-based strategies for cost-effective and streamlined identification. Pyrosequencing of selected variable regions, specifically the variable 1 and 3 (V1 and V3) regions, within the 16S rRNA gene provides successful identification of staphylococcal and streptococcal species with 50 bp or less of sequence information.[9] The combination of broad-range polymerase chain reaction (PCR) with cloning and pyrosequencing of V1 and V3 target sequences in the 16S rRNA gene enabled the detection of multiple bacterial contaminants in molecular biology reagents.[10] Targeted genotyping approaches offer the possibility of practical molecular strategies for sequence-based bacterial identification. Finally, future molecular identification strategies may include microarray and microfluidics chromosomal DNA profiling methods that will scan multiple genes and intergenic sequences in parallel.

BLOODSTREAM AND INVASIVE INFECTIONS

The incidence of sepsis and the number of sepsis-related deaths have consistently increased from 1979 to 2000,[11] making sepsis one of the top ten causes of death in the United States. The annualized incidence of sepsis increased 8.7 percent, from 83 cases per 100,000 population in 1979 to 240 cases per 100,000 population in 2000.[11] The diagnosis of bacteremia in cases of sepsis requires the growth of pathogenic bacteria in liquid media, optimally with continuous monitoring by automated detection systems. Blood samples from infected individuals typically contain low levels of organisms (1–10 organisms/ml in adults) so that incubations with continuous monitoring usually require a minimum of 8 to 12 hours prior to detection.

Broad-range PCR has been used for the evaluation of positive blood culture specimens in the clinical microbiology laboratory. Peripheral blood specimens are typically added to bacteriologic liquid media in a glass bottle format, and the blood:media mixtures are monitored continuously in automated blood culture systems. Upon signal (e.g., fluorescence) detection, Gram stain visualization of positive blood cultures and subculturing on plated media are performed for identification. Alternatively, molecular methods may be used for the direct identification of bacterial pathogens in positive blood cultures prior to subculturing specimens on plated media (Table 42-1).

PCR amplification of 16s rRNA gene sequences combined with amplicon sequencing effectively identifies bacteria in blood culture specimens and correlates completely with Gram stain or BACTEC 9240 instrument (Becton Dickinson, Franklin Lakes, NJ) true-positive and true-negative results. Interestingly, the blood culture instruments yielded a 1.3% false-positive signal rate in this study, and all instrument false positives (negative Gram stain and cultures) lacked evidence of bacterial DNA.[12] Several isolates of gram-negative bacilli, including *Acinetobacter* and *Bordetella* organisms, that were difficult to analyze by conventional methods were identified only by DNA sequencing of the 16S rRNA gene using MicroSeq[6] (Applied Biosystems).

Alternative strategies support the potential utility of short subsequences of the 16S rRNA gene for rapid bacterial identification by pyrosequencing.[13] Successful identification of 28 *S. aureus* and 18 streptococcal isolates was achieved by interrogation of 10 nucleotides in the V1 region of 16S rDNA[9]. Pyrosequencing data from the V1 and V3 regions were combined (approximately 30 bp of sequence data) to generate species identification of related enterobacteria. Alternatively, bacterial pathogens may be specifically identified in positive blood cultures by fluorescence in situ hybridization (FISH).[14]

Staphylococcus aureus and coagulase-negative staphylococci (CoNS) represent the most common bacterial isolates from human blood cultures in the United States. In

Table 42-1. Principal Bacterial Pathogens and Molecular Targets

Pathogen	Specimen(s)	Gene (detection)	Gene (drug resistance)
B. henselae	Lymph node aspirates and biopsies	*ribC*	
B. parapertussis	Nasopharyngeal secretions, nasal swabs, sputa, throat swabs	IS*1001*	
B. pertussis	Nasopharyngeal secretions, nasal swabs, sputa, throat swabs	IS*481*	
Brucella species	Blood cultures, tissues, abscess material, bone marrows	IS*711*	
C. pneumoniae	Bronchoscopic washes, nasopharyngeal secretions, sputa, throat swabs	*pmp4*, Pst-1 fragment	
CoNS	Blood cultures, nasal swabs, wound swabs	16S rRNA	*mecA*
EHEC	Stool cultures	*stx*	
ETEC	Stool cultures	LT, ST genes	
H. pylori	Gastric biopsies	*ureC*	*rdxA*, 23S rRNA
L. pneumophila	Bronchoscopic washes, lung biopsies, sputa	*mip*	
M. pneumoniae	Bronchoscopic washes, CSF, lung biopsies, nasopharyngeal secretions, sputa, throat swabs	P1 gene, ATPase Operon gene	
N. meningitidis	Blood cultures, CSF, throat swabs	*ctrA*, *porA*	
S. aureus	Blood cultures, nasal swabs, wound swabs	*Sa442*	*mecA*
S. pneumoniae	Blood cultures, CSF, nasal swabs	*ply*, *lytA*	
S. pyogenes	Throat swabs	*ptsI*	
T. whippelii	CSF, intestinal biopsies, lymph node biopsies, synovial fluids	16S rRNA	

CSF, cerebrospinal fluid; CoNS, coagulase-negative staphylococci; EHEC, enterohemorrhagic *E. coli*; ETEC, enterotoxigenic *E. coli*.

addition to species identification, the detection of methicillin-resistant *S. aureus* (MRSA) is important for patient management (Table 42-1). The presence of the *mecA* gene effectively confers methicillin resistance in staphylococci and is a convenient molecular target for PCR-based drug resist ance detection. Real-time PCR can rapidly identify MRSA within 4 hours (DNA preparation, amplification, and detection) by the concurrent application of species-specific and *mecA*-specific DNA probes in parallel or multiplex testing strategies. Real-time PCR strategies have successfully identified MRSA among a panel of bacterial isolates from clinical specimens,[15,16] using both species-specific and *mecA*-specific primer-probe combinations. Investigators have demonstrated that real-time PCR enables rapid MRSA detection directly from smear-positive blood culture bottles, including the BACTEC[17] (Becton Dickinson) and BacT/ALERT[18] (bioMérieux, Hazelwood, MO) systems. The *mecA* gene was detected in 108 of 109 known MRSA isolates, which highlights the potential utility of combining real-time species identification with *mecA* gene detection.[15] Based on the methods already published, appropriate therapies could be instituted 24 to 36 hours earlier if smear-positive blood culture bottles were directly assessed for MRSA. The presence of MRSA dictates the antibiotic treatment strategies, as glycopeptides (e.g., vancomycin) are commonly used in this setting. Differences of 24 to 36 hours may significantly improve patient management by reducing the use of inappropriate antibiotics and diminishing the spread of resistant organisms.

CoNS represent a diverse group of more than 30 different species. The identification of particular species may have clinical implications such as the determination of the presence of *Staphylococcus epidermidis* from a sterile site. Identification of CoNS represents a challenge for development of multiplex real-time PCR methods and melting-curve analyses. A creative primer-biprobe combination resulted in the identification of 15 different CoNS species with only three different real-time PCR reactions.[19]

Streptococcus agalactiae (group B streptococci) remains the leading cause of neonatal sepsis and meningitis and is frequently transmitted as a perinatal infection. Rapid and accurate diagnosis using molecular methods (Table 42-1) facilitates timely management with antimicrobial therapy to prevent perinatal transmission.[20] Although *S. agalactiae* can be successfully cultured from maternal anal or vaginal specimens using enrichment broth culture techniques, the 24- to 48-hour delay in culture detection reduces the utility of group B streptococcal cultures at the time of delivery. For this reason, empiric ampicillin therapy has gained acceptance in obstetrics practice. Real-time PCR assays for rapid detection of *S. agalactiae* directly in vaginal specimens have been developed, targeting species-specific sequences in the *cfb*, or CAMP factor, gene.[21] As bacterial concentrations have not been correlated with different risks of transmission, qualitative real-time PCR detection is sufficient. The Food and Drug Administration (FDA) approved a real-time PCR assay (i.e., IDI-Strep B assay, Becton, Dickinson, and Company, Franklin Lakes, NJ) for detection of *S. agalactiae* directly in anal or vaginal specimens. PCR assays also will be useful for direct detection of invasive group B streptococci in peripheral blood in cases of neonatal sepsis.

CENTRAL NERVOUS SYSTEM INFECTIONS

Rapid diagnosis of central nervous system (CNS) infections may have a dramatic impact on patient outcomes. The most important causes of acute bacterial meningitis are *Neisseria meningitidis* and *Streptococcus pneumoniae*. Childhood vaccination strategies have diminished the risk of meningitis due to *Haemophilus influenzae* serotype b in North America. If acute bacterial meningitis is suspected, cerebrospinal fluid (CSF) specimens are submitted for Gram stain and culture. Immediate Gram stain evaluation is useful, but has limited sensitivity because 10^5 organisms/ml for are required visualization. Laboratory evaluation requires waiting 24 to 48 hours for culture results and a definitive diagnosis.

Although meningococci can be visualized by microscopy and may be cultured in the clinical microbiology laboratory, these organisms present diagnostic challenges with serious implications for patients. PCR testing for detection of *N. meningitidis* infection has been proposed as a strategy for the rapid diagnosis of meningococcal meningitis[22] (Table 42-1) and has been embraced in the United Kingdom and Australia as an important element of laboratory testing. Real-time PCR has been used for direct detection of meningococcal DNA and genogrouping of meningococci for epidemiological purposes in Sweden.[23] *N. meningitidis* DNA was amplified directly from CSF or peripheral blood specimens using the 16S rRNA gene for species detection and the *porA* gene for genogrouping.

Broad-spectrum molecular approaches have been advocated for molecular screening of CSF specimens independent of Gram stain and culture methods (Table 42-1). The conserved 16S rRNA gene contains species-specific target sequences for *N. meningitidis* and has been used in molecular screening strategies. Broad-range PCR based on the 16S rRNA gene has been used successfully to evaluate CSF specimens directly for the presence of bacterial DNA, with a sensitivity and specificity of 100% and 98.2%, respectively.[24] An alternative approach for molecular screening is the multiplex strategy, whereby different genes are used for detection of each species. Corless et al.[25] demonstrated the simultaneous detection of *Neisseria meningitidis*, *Haemophilus influenzae*, and *Streptococcus pneumoniae* by real-time PCR with different genetic targets. A 5′ nuclease (TaqMan) multiplex approach was used to directly amplify bacterial DNA from CSF and peripheral blood specimens. The respective sensitivities of real-time PCR for detection of *N. meningitidis*, *H. influenzae*, and *S. pneumoniae* were 88.4%, 100%, and 91.8%, respectively. More than 4000 CSF specimens were screened in this study,[25] and the improvement in the meningococcal detection rate was 2.9%. That is, 87 additional cases of

meningococcal meningitis were identified by PCR alone. Refinements in primer selection targeting less variable sequences in the capsular transport (*ctrA*) gene of *N. meningitidis* accounted for improvements in the sensitivity of the real-time PCR assay.[25] Endpoint PCR detection of *N. meningitidis* increased the laboratory confirmation of clinically suspected cases by 36% when compared to culture and antigen detection.[22]

Molecular methods may yield benefits for the diagnosis of other CNS bacterial infections. *Mycoplasma pneumoniae* is an important cause of atypical pneumonia and is difficult to culture, in contrast to the genital mycoplasmas. In addition to its role in respiratory infections, *M. pneumoniae* may cause meningoencephalitis or transverse myelitis in pediatric patients. PCR tests for the detection of *M. pneumoniae* have been useful for the provision of timely and accurate diagnoses of *M. pneumoniae* CNS infections in children.[26]

GASTROINTESTINAL INFECTIONS

Enteric bacterial pathogens cause well-recognized gastric and intestinal infections. Pathogens such as *Campylobacter*, *Salmonella*, and *Shigella* represent important causes of gastroenteritis and bacillary dysentery and may be cultured from freshly collected stool specimens in infected patients. Successful culture and biochemical identification of enteric pathogens from fecal specimens provide cost-effective diagnoses. However, specific diagnostic challenges that require subspecies or toxigenic strain identification highlight the limitations of current approaches (Table 42-1).

Intestinal *E. coli* is recognized as a commensal organism and pathogen of the human intestine, so the ability to distinguish nonpathogenic and pathogenic isolates is important. Specific pathogenic *E. coli* strains or clones have been associated with human enteric infections. Enterohemorrhagic *E. coli* (EHEC) causes hemorrhagic colitis and hemolytic uremic syndrome. Candidate EHEC isolates can be serotyped as O157:H7 by latex agglutination or tested for Shiga-like toxin production by antigen detection. The limitation of serotyping is highlighted by the fact that more than 20% of EHEC isolates lack the O157:H7 serotype. The Shiga-like toxin genes may be directly detected by DNA amplification and may reduce the turnaround time for diagnosis of this potentially fatal infection.[27] Enterotoxigenic *E. coli* (ETEC) is the most important cause of traveler's diarrhea and, with *Campylobacter jejuni*, remains one of the two most common causes of nonviral gastroenteritis. Culture methods cannot distinguish toxigenic from nontoxigenic strains except by antigen detection. Molecular methods that enable the specific detection of either the heat-labile (LT) or heat-stable (ST) toxin genes provide options for direct ETEC detection in stool isolates.[28]

Tropheryma whippelii is the etiologic agent of Whipple's disease, a rare multisystemic disease primarily involving the small intestine and characterized by inclusion-rich macrophages in the lamina propria. Patients with Whipple's disease may present with diarrhea, weight loss, arthralgias, abdominal pain, and dementia. As up to 15% of cases may not involve the intestinal tract, sampling of extraintestinal sites (e.g., CSF, synovial fluid) may be important for molecular diagnosis of this disorder. Intestinal biopsy specimens, lymph node biopsy specimens, and synovial fluid may be tested directly by PCR amplification of *Tropheryma*-specific DNA.[29,30] As this organism is generally nonculturable, bacterial DNA detection is very useful for the establishment of a specific diagnosis. Importantly, this organism often persists in asymptomatic individuals and its DNA may be detected in healthy individuals using fecal or oral specimens.[30] Colonization of the intestinal tract may be difficult to distinguish from infection, and thus the molecular data should be interpreted in the context of the intestinal histology.

In addition to intestinal infections, *Helicobacter pylori* and other gastric *Helicobacters* may cause chronic gastritis and peptic ulcer disease.[31] Long-term infections with *H. pylori* have been associated with gastric adenocarcinomas and mucosa-associated lymphoid tissue (MALT) lymphomas. Serologic tests are useful for IgG screening, and fecal antigen testing is effective for direct detection. The bacteria reside adjacent to the gastric mucosa and require biopsy sampling by endoscopy for successful culture. In addition to histology and rapid urease testing, endpoint PCR provided strategies for successful detection of *H. pylori* and macrolide resistance mutations in gastric biopsies[32,33] (Table 42-1). Endpoint PCR studies using the 16S rRNA gene for genus-specific PCR detection yielded novel findings such as the detection of *Helicobacter cinaedi* in gastric biopsy specimens obtained from patients with gastritis.[34] More recently, real-time PCR assays have been used for detection and molecular resistance testing of *H. pylori* in gastric biopsies.[35] He et al.[36] described real-time quantitative PCR of *H. pylori* with the *ureC* gene as the target in an assay spanning 6 logs of DNA concentrations. Interestingly, 24 of 27 specimens that were negative by culture were positive by real-time PCR, although contamination could not be excluded.[36]

RESPIRATORY INFECTIONS

Bacterial pathogens represent important causes of respiratory tract infections. Pneumonias caused by various pathogens have overlapping features and are not clinically distinguishable with respect to etiologic agent. Diagnosis of upper or lower respiratory tract infections may require sampling and culture of different sites. Respiratory specimens include throat or nasopharyngeal swabs, sputa, tracheal aspirates, and bronchoscope-assisted collections of fluid or tissue of the lower respiratory tract. In addition to effective differentiation among pathogenic and commensal organisms, colonization must be distinguished from infection. Among bacterial pathogens that cause respiratory infections, several organisms continue to be impractical to culture, and indirect serologic tests may be used.

The diagnosis of *Mycoplasma pneumoniae* as a cause of atypical pneumonia requires IgM serologic testing or molecular methods for direct detection. The utility of PCR has been demonstrated for the detection of *M. pneumoniae* DNA in throat swabs, sputa, and lower respiratory specimens[37,38] (Table 42-1). The risk for pneumonia is especially prominent among children 5 to 15 years of age, in whom 30% of infections progress to pneumonia if untreated.[39] Molecular methods have altered our view of *M. pneumoniae* as a cause of both atypical and acute bacterial pneumonia. PCR methods for detection have been recommended alone and in combination with IgM serologic tests for the diagnosis of respiratory infections caused by *M. pneumoniae*.[38,40] Sensitivities and specificities of PCR tests have ranged from 78% to 92% and 92% to 100%, respectively.[40] In one study, the positive predictive value of endpoint PCR was 100% and contrasted sharply with the relatively low positive predictive value (50%) of IgM detection by the indirect immunofluorescence assay (IFA).[41] Real-time detection of *M. pneumoniae* DNA enhanced molecular diagnostic strategies by improving speed and test handling.[42] Multiplex quantitative real-time PCR has been employed for the simultaneous assessment of atypical pneumonia caused by *Chlamydia pneumoniae*, *Legionella pneumophila*, or *M. pneumoniae*.[43] A comparison of results obtained by real-time PCR with 73 specimens yielded an overall 98.3% agreement with conventional PCR.[43]

Pertussis or whooping cough caused by *Bordetella pertussis* or *Bordetella parapertussis* is an important disease in children and adults. As *B. pertussis* is a fastidious pathogen and requires specialized media for culture, PCR amplification has been adopted as a strategy for diagnosis (Table 42-1). Direct fluorescence assay (DFA) and microbiologic culture have reduced sensitivities when compared to PCR assays in multiple studies. Endpoint and real-time PCR methods are commonly used for the detection of *B. pertussis* or *B. parapertussis* DNA in nasopharyngeal swabs[44–47] from patients. Documented sensitivities for endpoint PCR (95.0%) and real-time PCR (100%) compared favorably with microbiologic culture (11.6% and 36.0%) and DFA tests (11.4%).[44,45] Insertion sequences IS*481* and IS*1001* have been useful for the detection of *B. pertussis* and *B. parapertussis*, respectively. In addition to IS*481*, the pertussis toxin promoter contains species-specific target sequences for PCR detection. A systematic comparison of real-time PCR, endpoint PCR, culture, and DFA studies indicated that real-time PCR based on insertion sequence detection had the highest sensitivity and best overall performance.[47] Careful PCR assay selection and quality assessment procedures must be emphasized, as a published investigation presented evidence for substantial over- and misdiagnosis of pertussis cases in two disease outbreaks in New York.[48]

The detection of *Streptococcus pyogenes* (group A streptococci) in throat swabs represents an important strategy for the diagnosis of streptococcal pharyngitis. The typical strategy includes rapid antigen testing with culture confirmation of all specimens negative by antigen testing. This strategy is currently recommended by the American Academy of Pediatrics.[49] DNA probe hybridization methods for the detection of *S. pyogenes* DNA have been proposed as a stand-alone approach instead of the established two-tiered testing procedure.[50] Uhl et al.[51] recommended real-time PCR methods for detection of *S. pyogenes* DNA in throat swabs as a stand-alone test. Real-time PCR matched culture, outperformed rapid antigen testing, and could be performed more rapidly than culture for confirmatory testing. The high prevalence of this disease and potential impact for patient management supports the consideration of molecular testing for the diagnosis of streptococcal pharyngitis (Table 42-1).

Molecular Resistance Testing

Molecular resistance testing refers to the direct detection of genes or mutations conferring antimicrobial resistance. The detection of the *mecA* gene was previously discussed as a strategy for identifying MRSA isolates (see "Bloodstream and Invasive Infections"). Gene detection is relatively straightforward and akin to qualitative bacterial pathogen detection. In contrast, mutation detection requires different methods designed to identify the presence of nucleotide substitutions, deletions, or insertions (Figure 42-1). Parallel identification and detection of drug resistance mutations in bacteria may be complicated by the presence of many different mutations in single or several genes. The ability to detect multiple mutations in parallel may require microarray or high-throughput sequencing strategies in the future.

Detection of the *vanA* and *vanB* glycopeptide resistance genes in enterococci represents a useful molecular strategy for surveillance testing of vancomycin-resistant enterococci (VRE). Like *mecA* detection in MRSA, the presence of *vanA* or *vanB* is sufficient for the reporting of VRE isolates in patients. Endpoint multiplex PCR has been used to detect VRE directly in surveillance specimens and compared favorably with culture and biochemical testing.[52] The sensitivity and specificity of multiplex PCR were 98.4% and 96.1%, respectively.[52] Use of multiplex PCR methods resulted in a turnaround time reduction of approximately 48 hours and resulted in cost savings of approximately 40%.[52] Reductions in turnaround times and labor costs may be further reduced by the implementation of real-time PCR detection of *vanA* or *vanB* target sequences.[53] Cost savings in laboratory testing for *vanA* and *vanB* detection were calculated over time and, when combined with training costs, generated a payback period of 3 years for laboratory costs alone. Further cost benefits for the healthcare system caused by more timely adjustments in patient management and reductions in spread of resistant organisms due to changes in isolation practices were not computed.

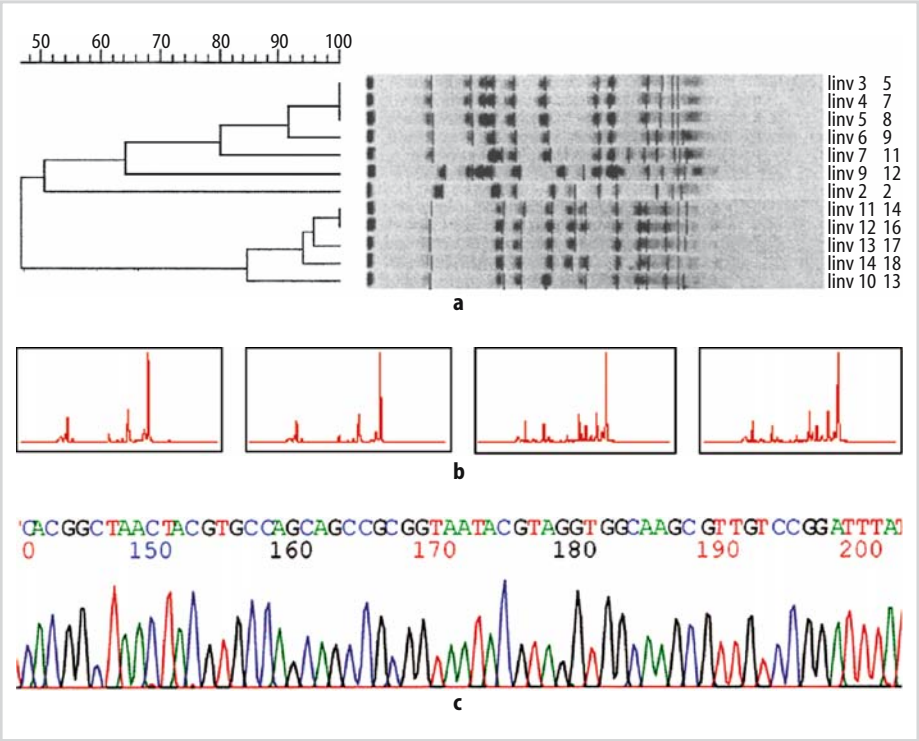


Figure 42-2. DNA typing of bacterial pathogens. (a) Gel image of pulsed field gel electrophoresis (PFGE) data. (b) Peak profiles of DNA fragments generated by repetitive extragenic palindramic PCR (rep-PCR) and microfluidics fragment resolution. (c) One of multiple chromatograms generated by multilocus sequence typing (MLST).

Detection of particular nucleotide substitutions associated with drug resistance may be practical if small sets of mutations are responsible for the vast majority of resistance in particular organisms. Since two point mutations in the 23S rRNA gene account for greater than 90% of macrolide resistance in *Helicobacter pylori*,[54,55] convenient polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP) methods[33,54] and reverse hybridization[32] assays have been developed for direct mutation detection (Table 42-1). Gastric biopsy specimens may be obtained directly from rapid urease tests and used for molecular resistance testing several weeks following tissue collection.[33]

Molecular Epidemiology

Nosocomial, or hospital-acquired, infections represent the most common complication affecting hospitalized patients and comprise an important category of adverse events in hospitals.[56] Nosocomial infection surveillance activities and epidemiologic studies in hospitals are essential aspects of infection control efforts. Classical epidemiology, usually performed by infection control personnel, includes traditional case-based outbreak investigations that correlate microbiologic identification and antimicrobial susceptibility testing patterns with reviews of case clusters. Molecular epidemiology requires the correlation of classical epidemiology with molecular data for delineation of outbreak-related clones. Molecular epidemiology refers to clonal investigations of bacterial pathogens by genotypic methods and includes methods such as DNA fingerprinting, DNA profiling, and DNA typing (Figure 42-2 and Table 42-2). Molecular methods yield patterns of different-sized DNA fragments or aggregate sequence information that may be useful for epidemiologic or clonal investigations. Studies have demonstrated the importance of real-time molecular typing strategies for the control of nosocomial infections. Hacek et al. demonstrated hospital savings of

Table 42-2. DNA Typing Strategies

Method	Instrument(s)	Time
AFLP	Thermal cycler, electrophoresis equipment	1–2 days
MLST	DNA sequencer	1–2 days
PFGE	PFGE apparatus	2–3 days
RAPD (AP-PCR)	Thermal cycler, electrophoresis equipment	1 day
rep-PCR	Thermal cycler, electrophoresis equipment, microfluidics device	4 hours–1 day
Ribotyping	Electrophoresis/blotting system	1 day
Spoligotyping	Thermal cycler, miniblotter apparatus	2 days

AFLP, amplified fragment length polymorphism; MLST, multilocus sequence typing; PFGE, pulsed field gel electrophoresis; RAPD, random amplification polymorphic DNA; AP-PCR, arbitrarily primed PCR; rep-PCR, repetitive extragenic palindromic PCR.

greater than $4.3 million during a 2-year period following implementation of routine molecular epidemiologic investigations.[57]

The established gold standard of molecular epidemiology is pulsed field gel electrophoresis (PFGE; Table 42-2), a method that depends on the electrophoretic separation of large DNA fragments created by digestion of bacterial DNA with rare-cutting restriction enzymes.[58,59] Cultivation of bacteria is required prior to initiation of typing studies, and the PFGE studies require approximately 2 days for completion. Several multicenter studies have been published that document the relative abilities of laboratories to perform PFGE and participate in interlaboratory comparisons.[60–62] In one study,[60] four of 12 laboratories failed to generate interpretable PFGE data with *S. aureus*, which highlights the formidable technical challenges of this test. Centralized computer studies of data obtained from four selected laboratories generated similarities of 85% with a group of identical *S. aureus* isolates.[60] Improved harmonization of PFGE protocols recently was achieved by focusing on several critical parameters, but only seven of ten laboratories provided data of sufficient quality for interlaboratory comparisons.[62] Although PFGE has evolved as the historical standard for molecular typing of bacterial pathogens, data comparability and reproducibility represent ongoing challenges due to lack of standardization.

Chromosomal or plasmid RFLP studies, based on conventional electrophoretic separation of DNA fragments, have been less useful for typing bacterial pathogens. Plasmids represent nonessential extrachromosomal DNA molecules that often are absent in bacterial pathogens and limit unique identification with this method. Chromosomal RFLP approaches produce highly complex patterns that are difficult to interpret and limit interlaboratory comparisons.

Ribotyping is a variant of chromosomal RFLP methods and is performed by digestion of chromosomal DNA and hybridization with rRNA operon probes.[63] Generally, ribotyping is less discriminatory compared with other methods due to the limited complexity of ribotyping patterns. A commercial ribotyping method, the Riboprinter Microbial Characterization System (DuPont Qualicon, Wilmington, DE), is available, but the reduced discrimination with this assay continues to be a disadvantage relative to other typing methods such as PFGE and DNA fingerprinting by PCR.[64,65]

DNA typing methods using PCR include amplified fragment length polymorphism (AFLP),[66,67] arbitrarily primed PCR (AP-PCR) or random amplification polymorphic DNA (RAPD) studies,[68,69] multilocus sequence typing (MLST),[70] repetitive extragenic palindromic-PCR (rep-PCR),[71,72] and spoligotyping.[73,74] PCR methods facilitate DNA typing by providing nucleic acid amplification for specimens with limited amounts of target DNA. AFLP yields complex DNA profiles with relatively high discriminatory capabilities when testing various bacterial pathogens.[67] AP-PCR and RAPD methods are similar to AFLP but utilize short random oligonucleotide primers and fortuitously anneal to multicopy chromosomal DNA sequences. Multiple amplicons generated by PCR amplification can be separated by conventional agarose gel electrophoresis. Rep-PCR yields different-sized DNA fragments that are separated by various distances between interspersed repetitive DNA elements.[72] Amplicons generated by rep-PCR may be separated by agarose gel electrophoresis or in microfluidic channels with laser-based fragment detection (DiversiLab, Bacterial Barcodes, Houston, TX).[75] Rep-PCR studies suggest that particular MRSA clones are more frequently associated with disease outbreaks.[76] Development of rep-PCR DNA profile databases can be combined with digitized microfluidic channel data to facilitate Internet pattern matching with sophisticated clustering algorithms. MLST has been proposed as a next-generation molecular typing strategy[70] and, in selected studies, MLST matches PFGE for relative discriminatory abilities.[77] High-throughput sequencing capabilities are required for MLST to be practically useful in a diagnostic laboratory setting. As diagnostic laboratories migrate toward more strain- or subspecies-level identification, molecular typing methods may have a prominent role in the diagnosis of bacterial infections.

Summary and Future Trends

Molecular methods have matured and gained acceptance as an important component of laboratory testing for the diagnosis of infectious diseases caused by bacterial pathogens. The initial hype has been supplanted by practical and selective data-driven molecular strategies for diagnosis and patient monitoring. Rapid advances in microbial genomics and bioinformatics have created many opportunities for the development of molecular assays with improved specificity over earlier assays. As the entire genomes of many important bacterial pathogens have been sequenced, it is possible to use multistrain genomic sequencing and comparative genomics for identification of individual pathogens. The greatly increased quantities of genomic data will be translated into the selection of optimal genetic targets for species-specific and strain-specific detection by DNA amplification methods. The development of comprehensive gene and mutation databases will drive the development of new tests for specific virulence genes, toxin genes, and drug resistance genotypes. The next generation of molecular technologies will include developments in high-throughput sequencing, custom microarray applications, and microfluidics or chromatography chromosomal scanning approaches. The rapid change in testing methods will challenge laboratories to maintain leading-edge technical proficiencies and provide clinically relevant information in clinically useful reporting formats. Demands for quality control and quality assessment will require new approaches to keep pace with the rapid evolution of molecular diagnostics and its applications in medicine.

References

1. Arnold LJ Jr, Hammond PW, Wiese WA, et al. Assay formats involving acridinium-ester-labeled DNA probes. *Clin Chem.* 1989;35:1588–1594.
2. Olsen GJ, Woese CR. Ribosomal RNA: a key to phylogeny. *FASEB J.* 1993;7:113–123.
3. Cole JR, Chai B, Marsh TL, et al. The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.* 2003;31: 442–443.
4. Devulder G, Perriere G, Baty F, et al. BIBI, a Bioinformatics Bacterial Identification Tool. *J Clin Microbiol.* 2003;41:1785–1787.
5. Kolbert CP, Persing DH. Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. *Curr Opin Microbiol.* 1999;2: 299–305.
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;36:3674–3679.
7. Gray MW, Sankoff D, Cedergren RJ. On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Res.* 1984;12:5837–5852.
8. Van de PY, Chapelle S, De Wachter R. A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Res.* 1996;24:3381–3391.
9. Jonasson J, Olofsson M, Monstein HJ. Classification, identification and subtyping of bacteria based on pyrosequencing and signature matching of 16S rDNA fragments. *APMIS.* 2002;110:263–272.
10. Grahn N, Olofsson M, Ellnebo-Svedlund K, et al. Identification of mixed bacterial DNA contamination in broad-range PCR amplification of 16S rDNA V1 and V3 variable regions by pyrosequencing of cloned amplicons. *FEMS Microbiol Lett.* 2003;219:87–91.
11. Martin GS, Mannino DM, Eaton S, et al. The epidemiology of sepsis in the United States from 1979 through 2000. *N Engl J Med.* 2003; 348:1546–1554.
12. Qian Q, Tang YW, Kolbert CP, et al. Direct identification of bacteria from positive blood cultures by amplification and sequencing of the 16S rRNA gene: evaluation of BACTEC 9240 instrument true-positive and false-positive results. *J Clin Microbiol.* 2001;39:3578–3582.
13. Ronaghi M, Karamohamed S, Pettersson B, et al. Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem.* 1996;242:84–89.
14. 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;40:247–251.
15. Grisold AJ, Leitner E, Muhlbauer G, et al. Detection of methicillin-resistant *Staphylococcus aureus* and simultaneous confirmation by automated nucleic acid extraction and real-time PCR. *J Clin Microbiol.* 2002;40:2392–2397.
16. Elsayed S, Chow BL, Hamilton NL, et al. Development and validation of a molecular beacon probe-based real-time polymerase chain reaction assay for rapid detection of methicillin resistance in *Staphylococcus aureus*. *Arch Pathol Lab Med.* 2003;127:845–849.
17. Tan TY, Corden S, Barnes R, et al. Rapid identification of methicillin-resistant *Staphylococcus aureus* from positive blood cultures by real-time fluorescence PCR. *J Clin Microbiol.* 2001;39:4529–4531.
18. Shrestha NK, Tuohy MJ, Hall GS, et al. Rapid identification of *Staphylococcus aureus* and the *mecA* gene from BacT/ALERT blood culture bottles by using the LightCycler system. *J Clin Microbiol.* 2002;40: 2659–2661.
19. Edwards KJ, Kaufmann ME, Saunders NA. Rapid and accurate identification of coagulase-negative staphylococci by real-time PCR. *J Clin Microbiol.* 2001;39:3047–3051.
20. Noya FJ, Baker CJ. Prevention of group B streptococcal infection. *Infect Dis Clin North Am.* 1992;6:41–55.
21. Ke D, Menard C, Picard FJ, et al. Development of conventional and real-time PCR assays for the rapid detection of group B streptococci. *Clin Chem.* 2000;46:324–331.
22. Pollard AJ, Probe G, Trombley C, et al. Evaluation of a diagnostic polymerase chain reaction assay for *Neisseria meningitidis* in North America and field experience during an outbreak. *Arch Pathol Lab Med.* 2002;126:1209–1215.
23. Molling P, Jacobsson S, Backman A, et al. Direct and rapid identification and genogrouping of meningococci and *porA* amplification by LightCycler PCR. *J Clin Microbiol.* 2002;40:4531–4535.
24. Saravolatz LD, Manzor O, VanderVelde N, et al. Broad-range bacterial polymerase chain reaction for early detection of bacterial meningitis. *Clin Infect Dis.* 2003;36:40–45.
25. Corless CE, Guiver M, Borrow R, et al. Simultaneous detection of *Neisseria meningitidis*, *Haemophilus influenzae*, and *Streptococcus pneumoniae* in suspected cases of meningitis and septicemia using real-time PCR. *J Clin Microbiol.* 2001;39:1553–1558.
26. Bitnun A, Ford-Jones EL, Petric M, et al. Acute childhood encephalitis and *Mycoplasma pneumoniae*. *Clin Infect Dis.* 2001;32:1674–1684.
27. Wang G, Clark CG, Rodgers FG. Detection in *Escherichia coli* of the genes encoding the major virulence factors, the genes defining the O157:H7 serotype, and components of the type 2 Shiga toxin family by multiplex PCR. *J Clin Microbiol.* 2002;40:3613–3619.
28. Tsen HY, Jian LZ. Development and use of a multiplex PCR system for the rapid screening of heat labile toxin I, heat stable toxin II and Shiga-like toxin I and II genes of *Escherichia coli* in water. *J Appl Microbiol.* 1998;84:585–592.
29. Fenollar F, Fournier PE, Raoult D, et al. Quantitative detection of *Tropheryma whippelii* DNA by real-time PCR. *J Clin Microbiol.* 2002;40:1119–1120.
30. Dutly F, Altwegg M. Whipple's disease and "*Tropheryma whippelii*." *Clin Microbiol Rev.* 2001;14:561–583.
31. Versalovic J, Fox JG. *Helicobacter*. In: Murray P, Baron EJ, Jorgensen JH, et al., eds. *Manual of Clinical Microbiology.* Washington, DC: ASM Press; 2003:915–928.
32. van Doorn LJ, Henskens Y, Nouhan N, et al. The efficacy of laboratory diagnosis of *Helicobacter pylori* infections in gastric biopsy specimens is related to bacterial density and *vacA*, *cagA*, and *iceA* genotypes. *J Clin Microbiol.* 2000;38:13–17.
33. Pena JA, Fox JG, Ferraro MJ, et al. Molecular resistance testing of *Helicobacter pylori* in gastric biopsies. *Arch Pathol Lab Med.* 2001;125:493–497.
34. Pena JA, McNeil K, Fox JG, et al. Molecular evidence of *Helicobacter cinaedi* organisms in human gastric biopsy specimens. *J Clin Microbiol.* 2002;40:1511–1513.
35. Chisholm SA, Owen RJ, Teare EL, et al. PCR-based diagnosis of *Helicobacter pylori* infection and real-time determination of clarithromycin resistance directly from human gastric biopsy samples. *J Clin Microbiol.* 2001;39:1217–1220.
36. He Q, Wang JP, Osato M, et al. Real-time quantitative PCR for detection of *Helicobacter pylori*. *J Clin Microbiol.* 2002;40:3720–3728.
37. Buck GE, Eid NS. Diagnosis of *Mycoplasma pneumoniae* pneumonia in pediatric patients by polymerase chain reaction (PCR). *Pediatr Pulmonol.* 1995;20:297–300.
38. Daxboeck F, Krause R, Wenisch C. Laboratory diagnosis of *Mycoplasma pneumoniae* infection. *Clin Microbiol Infect.* 2003;9:263–273.
39. Taylor-Robinson D. Infections due to species of *Mycoplasma* and *Ureaplasma*: an update. *Clin Infect Dis.* 1996;23:671–682.
40. Ferwerda A, Moll HA, de Groot R. Respiratory tract infections by *Mycoplasma pneumoniae* in children: a review of diagnostic and therapeutic measures. *Eur J Pediatr.* 2001;160:483–491.
41. Dorigo-Zetsma JW, Zaat SA, Wertheim-van Dillen PM, et al. Comparison of PCR, culture, and serological tests for diagnosis of *Mycoplasma pneumoniae* respiratory tract infection in children. *J Clin Microbiol.* 1999;37:14–17.
42. Hardegger D, Nadal D, Bossart W, et al. Rapid detection of *Mycoplasma pneumoniae* in clinical samples by real-time PCR. *J Microbiol Methods.* 2000;41:45–51.

43. Welti M, Jaton K, Altwegg M, et al. Development of a multiplex real-time quantitative PCR assay to detect *Chlamydia pneumoniae*, *Legionella pneumophila* and *Mycoplasma pneumoniae* in respiratory tract secretions. *Diagn Microbiol Infect Dis.* 2003;45:85–95.
44. Tilley PA, Kanchana MV, Knight I, et al. Detection of *Bordetella pertussis* in a clinical laboratory by culture, polymerase chain reaction, and direct fluorescent antibody staining; accuracy, and cost. *Diagn Microbiol Infect Dis.* 2000;37:17–23.
45. Chan EL, Antonishyn N, McDonald R, et al. The use of TaqMan PCR assay for detection of *Bordetella pertussis* infection from clinical specimens. *Arch Pathol Lab Med.* 2002;126:173–176.
46. Qin X, Turgeon DK, Ingersoll BP, et al. *Bordetella pertussis* PCR: simultaneous targeting of signature sequences. *Diagn Microbiol Infect Dis.* 2002;43:269–275.
47. Sloan LM, Hopkins MK, Mitchell PS, et al. Multiplex LightCycler PCR assay for detection and differentiation of *Bordetella pertussis* and *Bordetella parapertussis* in nasopharyngeal specimens. *J Clin Microbiol.* 2002;40:96–100.
48. Lievano FA, Reynolds MA, Waring AL, et al. Issues associated with and recommendations for using PCR to detect outbreaks of pertussis. *J Clin Microbiol.* 2002;40:2801–2805.
49. Committee on Infectious Diseases AAP. Group A streptococcal infections. In: Pickering LK, Peter G, Baker C, et al., eds. *Red Book 2000: Report of the Committee on Infectious Diseases.* Elk Grove Village, IL: American Academy of Pediatrics; 2000:526–536.
50. Chapin KC, Blake P, Wilson CD. Performance characteristics and utilization of rapid antigen test, DNA probe, and culture for detection of group A streptococci in an acute care clinic. *J Clin Microbiol.* 2002;40:4207–4210.
51. Uhl JR, Adamson SC, Vetter EA, et al. Comparison of LightCycler PCR, rapid antigen immunoassay, and culture for detection of group A streptococci from throat swabs. *J Clin Microbiol.* 2003;41(1):242–249.
52. Petrich A, Luinstra K, Page B, et al. Effect of routine use of a multiplex PCR for detection of *vanA*- and *vanB*- mediated enterococcal resistance on accuracy, costs and earlier reporting. *Diagn Microbiol Infect Dis.* 2001;41:215–220.
53. Palladino S, Kay ID, Costa AM, et al. Real-time PCR for the rapid detection of *vanA* and *vanB* genes. *Diagn Microbiol Infect Dis.* 2003;45:81–84.
54. Versalovic J, Shortridge D, Kibler K, et al. Mutations in 23S rRNA are associated with clarithromycin resistance in *Helicobacter pylori*. *Antimicrob Agents Chemother.* 1996;40:477–480.
55. Versalovic J, Fox JG. *Helicobacter*. In: Murray P, Baron EJ, Jorgensen JH, et al., eds. *Manual of Clinical Microbiology.* Washington, DC: ASM Press; 2003:915–928.
56. Burke JP. Patient safety: infection control—a problem for patient safety. *N Engl J Med.* 2003;348:651–656.
57. Hacek DM, Suriano T, Noskin GA, et al. Medical and economic benefit of a comprehensive infection control program that includes routine determination of microbial clonality. *Am J Clin Pathol.* 1999;111:647–654.
58. Goering RV. Molecular epidemiology of nosocomial infection: analysis of chromosomal restriction fragment patterns by pulsed-field gel electrophoresis. *Infect Control Hosp Epidemiol.* 1993;14:595–600.
59. Tenover FC, Arbeit RD, Goering RV, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol.* 1995;33:2233–2239.
60. van Belkum A, van Leeuwen W, Kaufmann ME, et al. Assessment of resolution and intercenter reproducibility of results of genotyping *Staphylococcus aureus* by pulsed-field gel electrophoresis of *Sma*I macrorestriction fragments: a multicenter study. *J Clin Microbiol.* 1998;36:1653–1659.
61. Deplano A, Schuermans A, Van Eldere J, et al. Multicenter evaluation of epidemiological typing of methicillin-resistant *Staphylococcus aureus* strains by repetitive-element PCR analysis. The European Study Group on Epidemiological Markers of the ESCMID. *J Clin Microbiol.* 2000;38:3527–3533.
62. Murchan S, Kaufmann ME, Deplano A, et al. Harmonization of pulsed-field gel electrophoresis protocols for epidemiological typing of strains of methicillin-resistant *Staphylococcus aureus*: a single approach developed by consensus in 10 European laboratories and its application for tracing the spread of related strains. *J Clin Microbiol.* 2003;41:1574–1585.
63. Grimont F, Grimont PA. Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. *Ann Inst Pasteur Microbiol.* 1986;137B:165–175.
64. Price CS, Huynh H, Paule S, et al. Comparison of an automated ribotyping system to restriction endonuclease analysis and pulsed-field gel electrophoresis for differentiating vancomycin-resistant *Enterococcus faecium* isolates. *J Clin Microbiol.* 2002;40:1858–1861.
65. Aarnisalo K, Autio T, Sjoberg AM, et al. Typing of *Listeria monocytogenes* isolates originating from the food processing industry with automated ribotyping and pulsed-field gel electrophoresis. *J Food Prot.* 2003;66:249–255.
66. D'Agata EM, Gerrits MM, Tang YW, et al. Comparison of pulsed-field gel electrophoresis and amplified fragment-length polymorphism for epidemiological investigations of common nosocomial pathogens. *Infect Control Hosp Epidemiol.* 2001;22:550–554.
67. Fry NK, Bangsborg JM, Bergmans A, et al. Designation of the European Working Group on Legionella Infection (EWGLI) amplified fragment length polymorphism types of *Legionella pneumophila* serogroup 1 and results of intercentre proficiency testing using a standard protocol. *Eur J Clin Microbiol Infect Dis.* 2002;21:722–728.
68. Williams JG, Kubelik AR, Livak KJ, et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 1990;18:6531–6535.
69. Welsh J, McClelland M. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* 1990;18:7213–7218.
70. Clarke SC. Nucleotide sequence-based typing of bacteria and the impact of automation. *Bioessays.* 2002;24:858–862.
71. Versalovic J, Koeuth T, Lupski JR. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res.* 1991;19:6823–6831.
72. Versalovic J, Schneider M, de Bruijn FJ, Lupski JR. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol Cell Biol.* 1994;5:25–40.
73. Warren RM, Streicher EM, Sampson SL, et al. Microevolution of the direct repeat region of *Mycobacterium tuberculosis*: implications for interpretation of spoligotyping data. *J Clin Microbiol.* 2002;40:4457–4465.
74. van der Zanden AG, Kremer K, Schouls LM, et al. Improvement of differentiation and interpretability of spoligotyping for *Mycobacterium tuberculosis* complex isolates by introduction of new spacer oligonucleotides. *J Clin Microbiol.* 2002;40:4628–4639.
75. Healy M, Huong J, Bittner T, et al. Microbial DNA typing by automated rep-PCR. *J Clin Microbiol.* 2005;43:199–207.
76. van der ZA, Verbakel H, van Zon JC, et al. Molecular genotyping of *Staphylococcus aureus* strains: comparison of repetitive element sequence-based PCR with various typing methods and isolation of a novel epidemicity marker. *J Clin Microbiol.* 1999;37:342–349.
77. Peacock SJ, de Silva GD, Justice A, et al. Comparison of multilocus sequence typing and pulsed-field gel electrophoresis as tools for typing *Staphylococcus aureus* isolates in a microepidemiological setting. *J Clin Microbiol.* 2002;40:3764–3770.

Chapter 43

Mycobacterial Infections

Betty A. Forbes

Clinical Utility

Microbiology laboratories have developed and introduced clinical molecular assays for mycobacteria during the last several years because of the need for a reliable and rapid means of diagnosing tuberculosis (TB) for public health and therapeutic reasons. Molecular tests are used for identification of mycobacteria directly in clinical specimens, for mycobacterial identification, and for determination of drug susceptibilities.

Direct Detection

The limitations of routine acid-fast bacilli (AFB) smear and culture are well known. Although rapid, a sputum AFB smear has a reported sensitivity range of 22% to 78%. In recent years, the specificity of AFB smears has been a problem if specimens are obtained from individuals with a high incidence of disease due to nontuberculous mycobacteria (NTM), such as HIV-infected individuals. Similarly, culture of sputum is positive in the majority of patients with pulmonary TB if multiple specimens are obtained;[1] however, almost 30% of patients reported to have TB, including 22% of patients with pulmonary TB, are not culture confirmed.[2] Moreover, due to the slow growth rate of mycobacteria, cultures usually take at least a week, and in some instances as long as 6 to 8 weeks, to become positive.

Initially, amplification techniques were expected to dramatically reduce turnaround time and to have better sensitivity and specificity compared to conventional AFB smear and culture methods. However, despite numerous and significant advances, molecular methods have not replaced AFB smears and cultures for the diagnosis of mycobacterial infections from respiratory or extrapulmonary sites. Conventional tests remain a primary means of establishing a diagnosis as well as for drug susceptibility testing for optimal therapy.

Respiratory Specimens

The appropriate clinical use of nucleic acid amplification (NAA) assays has been difficult to delineate. (Only commercially available NAA assays are discussed here; however, similar issues also are pertinent to laboratory-developed amplification assays.)

Commercially available NAA assays were initially evaluated as screening tests (i.e., all specimens were tested regardless of the suspected risks for TB). When testing smear-positive specimens, these assays performed well (sensitivity 95% to 96%, specificity, 99% to 100%), but with smear-negative specimens, the sensitivity was much lower (48% to 53%).[3] Based on the results of these early trials, the Food and Drug Administration (FDA) limited the use of two commercially available NAA kits, the Amplified Mycobacterium Tuberculosis Direct (AMTD) test (Gen-Probe, Inc, San Diego, CA) and the Amplicor Mycobacterium tuberculosis test (Roche Diagnostic Systems, Inc, Indianapolis, IN) to AFB smear-positive respiratory specimens only from patients who had not received antituberculosis drugs for 7 or more days or had not been treated for TB within the last 12 months.

The utility of NAA assays for the detection of the 20% to 30% of culture-negative cases was difficult to determine from these early studies because culture growth was used as the gold-standard method.[4] In 1998, a clinical trial was performed to evaluate a reformulated AMTD test (AMTD 2, Gen-Probe, Inc, San Diego, CA) compared to a physician estimation of the probability that a patient had TB, with review by a physician panel of experts for cases that were not proven TB.[5] The enhanced AMTD2 test was performed on respiratory specimens from 339 patients, and results were compared with culture and clinical diagnosis results. The AMTD2 test had an overall sensitivity of 85.9%, a specificity of 97.8%, a positive predictive value (PPV) of 91.0%, and a negative predictive value (NPV) of 96.6%. The FDA approved the new version AMTD2 test in 1999 for testing respiratory specimens regardless of the AFB smear results.

Table 43-1. Evaluation of the Enhanced AMTD2 Test Based on Comprehensive Clinical Diagnosis

Clinical Suspicion of TB	AFB Smear: PPV and NPV (%)	AMTD Test: Sensitivity, Specificity, PPV, NPV (%)
Low (<10%)	36, 96	83, 97, 59, 99
Intermediate (10–80%)	30, 71	75, 100, 100, 91
High (>80%)	94, 37	87, 100, 100, 91

Source: Data from Reference 6.
AMTD2, Amplified Mycobacterium Tuberculosis Direct test (Gen-Probe, Inc, San Diego, CA); AFB, acid-fast bacilli; NPV, negative predictive value; PPV, positive predictive value; TB, tuberculosis.

Subsequent clinical trials of the enhanced AMTD2 test demonstrate the importance of using not only mycobacterial culture results but also multiple other parameters including clinical signs and symptoms, response to therapy, and other laboratory results to interpret the results of NAA tests.[6–9] Data from these studies support the use of the enhanced AMTD2 test as a rapid method for the diagnosis of pulmonary TB in patients for whom there is a moderate to high suspicion of TB, regardless of the AFB smear result.

In another study,[6] enrolling physicians were asked to quantify their degree of clinical suspicion for TB using a scale from 0% to 100%; subjects were broken down into three major groups. In addition, a conservative consensus standard for the diagnosis of pulmonary TB was established and an independent expert panel reviewed all cases with a clinical suspicion of less than 80% and only one culture or no cultures positive for *M. tuberculosis* (Table 43-1). NAA tests were most useful for cases with a higher clinical suspicion.

As more of these studies are published, optimal use of NAA tests in conjunction with clinical information and other test methods will become manifest, allowing NAA tests to be utilized in a cost-effective manner with positive impact on patient management. Clearly, based on recent studies, decisions about when and how to use NAA tests for TB diagnosis should be individualized according to the clinical setting and NAA results interpreted within the context of the clinical suspicion for TB and on the basis of laboratory performance. To help laboratories and clinicians in this regard, guidelines (Table 43-2) for the use of NAA tests in the diagnosis of TB were set forth by the Centers for Disease Control and Prevention (CDC).[10,11]

Nonrespiratory Specimens

Both laboratory-developed and commercially available NAA assays have been used to test nonrespiratory specimens, although no commercially available test kit is approved by the Food and Drug Administration (FDA) for this purpose. Because clinical diagnosis often is uncertain, and AFB smear and culture lack sensitivity, the use of NAA tests is particularly attractive for suspected cases of extrapulmonary TB.

In particular, NAA assays have been evaluated in patients suspected of having tuberculous meningitis.[12–14] Signs and symptoms of this disease are nonspecific. AFB smears in patients with TB meningitis are positive in only 10% of cases. Isolation of *M. tuberculosis* is positive in only approximately 50% of adults[15] and is insufficiently timely to aid clinical judgment with respect to treatment. Although promising, results of these NAA studies have varied with respect to sensitivity, while specificity was greater than 98% in most instances.

Similar results with other extrapulmonary specimens have been obtained. In many studies, the performance of NAA tests with extrapulmonary specimens has been similar to their performance with respiratory specimens,[7,16,17] while in other studies, the sensitivity was quite low. Based on studies published to date, it is clear that more studies must be performed to establish optimal

Table 43-2. Guidelines for the Use of NAA Tests on Respiratory Specimens for TB

AFB Smear	Specimen No.	NAA Result	Action
Positive	1	Positive	Presumed to have TB*
Positive	1	Negative	Test for inhibitors. If inhibitors are detected, NAA is no diagnostic help. If no inhibitors are detected, repeat NAA on no more than 2 additional specimens
	2, 3 (repeats)	Negative	Presumed to have NTM
		Positive	Clinician must rely on clinical judgment
Negative	1	Positive	Repeat NAA
	2, 3 (repeats)	Positive	Presumed to have TB
		Negative	Clinician must rely on clinical judgment
Negative	1	Negative	Repeat NAA
	2, 3 (repeats)	Negative	Presumed not to be infectious†
		Positive	Clinician must rely on clinical judgment

Source: Data from Reference 7.
* Unless there is concern about NTM, the NAA test adds little to the diagnostic workup.
† Clinician must rely on clinical judgment regarding the need for antituberculous therapy and further diagnostic workup.
NAA, nucleic acid amplification; TB, tuberculosis; NTM, nontuberculous mycobacteria.

Table 43-3. Commercially Available, Molecular-Based Amplification Assays for the Diagnosis and Treatment of Mycobacterial Infections

Test (Manufacturer)	Method	Applications/Comments
Amplified *Mycobacterium tuberculosis* Direct test (Gen-Probe, Inc, San Diego, CA)	TMA	FDA approved for testing respiratory specimens Target: rRNA Direct detection of *M. tuberculosis* in broth cultures
Amplicor *Mycobacterium tuberculosis* test (Roche Diagnostics, Indianapolis, IN)	PCR	FDA approved for testing AFB smear-positive respiratory specimens only Automated version: COBAS Amplicor MTB Target: 16S rDNA Direct detection of *M. tuberculosis* in broth cultures
BDProbe Tec (Becton Dickinson Diagnostic Instrument Systems, Sparks, MD)	SDA	Fully automated Targets: IS6110 and 16S rDNA
LCx *M. tuberculosis* test (Abbott Laboratories, Abbott Park, IL)	LCR	Semiautomated
INNO-LiPA Mycobacteria v2 (Innogenetics, Gent, Belgium)	PCR with reverse hybridization	Identifies genus *Mycobacteria* and 16 mycobacterial species Target: 16S-23S ribosomal spacer gene region
INNO-LiPA Rif. TB (Innogenetics, Gent, Belgium)	PCR with reverse hybridization	Detects *M. tuberculosis* complex plus rifampin resistance Target: region of the *rpoB* gene
GenoType Mycobacteria (Hain, Diagnostica, Nehren, Germany)	PCR with reverse hybridization	Identifies 13 mycobacterial species and a supplemental kit with 16 additional species Target: 23S rRNA gene

FDA, Food and Drug Administration; LCR, ligase chain reaction; PCR, polymerase chain reaction; rRNA, ribosomal RNA; SDA, strand displacement amplification; TMA, transcription-mediated amplification.

sample volume, nucleic acid extraction and amplification procedures, and uniform criteria for interpretation of results for each specimen type, including formalin-fixed, paraffin-embedded tissues. However, based on current data, a negative NAA result does not rule out extrapulmonary TB, particularly if the AFB smear for the specimen is negative.

NAA tests appear to be useful for early identification of *M. tuberculosis* complex for all specimen types grown in liquid cultures, except blood.[18,19] Sensitivity and specificity of NAA assays using this approach are both greater than 98%. In addition, the INNO-LiPA Mycobacteria assay (Innogenetics NV, Ghent, Belgium) (Table 43-3) successfully identifies mycobacteria directly from aliquots of culture medium from the MB/BacT ALERT 3D System (MB/BacT) (Organon Teknika, Boxtel, the Netherlands).[20]

NAA tests also have been evaluated for their usefulness in monitoring therapeutic efficacy. Although beneficial for initial diagnosis, most NAA assays appear to be unsuitable for monitoring treatment of TB patients since nucleic acid targets persist long after AFB smears and cultures become negative.[21,22] However, detection of *M. tuberculosis* messenger RNA (mRNA) may prove to be a potentially useful method for monitoring therapeutic efficacy,[23] although more studies are required to determine its clinical utility.

Identification

Mycobacterial isolates have traditionally been identified to the species level based on phenotypic and biochemical tests. These methods are slow and cumbersome and may fail to obtain an identification, with results varying among different isolates of the same species. Therefore, laboratories are increasingly using molecular methods for identification. Rapid identification of *M. tuberculosis* is of paramount importance for therapeutic and public health reasons. Rapid detection and identification of NTM, for which there are currently at least 100 species, is becoming more important since NTM lung disease is caused by many NTM species with a range of clinical presentations.[24] Toward this goal, a variety of molecular approaches have been published for identification of not only *M. tuberculosis* complex organisms, but NTM as well (e.g., References 25 and 26). Finally, two commercially available systems using amplification of a region of the 16S-23S ribosomal RNA (rRNA) gene spacer region, or the 23S rRNA gene with subsequent hybridization to a membrane strip containing probes to the most commonly isolated mycobacterial species, have been used to identify mycobacteria.[27]

Susceptibility Testing

To successfully control the spread of TB, cases must be detected and treated in a timely manner. At best, conventional susceptibility testing methods for *M. tuberculosis* are available within 7 to 14 days, after a culture result is positive. Thus, molecular methods that can rapidly detect drug resistance are attractive. Mutations have been delineated that are responsible for resistance to the primary drugs used to treat TB, including rifampin (RIF), isoniazid (INH), ethambutol (EMB), streptomycin (STR), and

Table 43-4. Examples of Laboratory-Developed Methods for Mycobacterial Infections

Application	Methods	Comments
Direct detection	PCR: single or multiplex	Targets include *IS6110*, *MPB64* and protein antigen B
	Real-time PCR	
	Peptide nucleic acids	
Identification	PCR-RFLP with agarose gel or CE	Targets include *hsp65*, 16S rDNA, *recA*, *rpoB*, *dnaJ,* and 32-kDa protein
	SSCP	
	High-density DNA probe arrays	Numerous techniques and targets for identification of *M. tuberculosis* complex
Drug susceptibility	PCR-RFLP	To date, DNA sequencing is a more successful method for detecting RIF resistance
	Heteroduplex analysis	
	RT-PCR	
	Real-time PCR	
	Sequencing	
	SSCP	
	High-density DNA probe arrays	
	Oligonucleotide arrays	
Epidemiology	*IS6110* RFLP	*IS6110* RFLP has high discriminatory power
	Secondary markers	
	Polymorphic guanine-cytosine-rich repetitive RFLP typing	
	Spoligotyping (PCR-based)	Spoligotyping is useful when discriminating isolates of *M. tuberculosis* with few *IS6110* bands, and is economical, easy to perform, and rapid
	Mixed-linker PCR	
	Variable number tandem repeat analysis	
	Mycobacterial interspersed repetitive units analysis	

CE, capillary electrophoresis; PCR, polymerase chain reaction; PCR-RFLP, PCR–restriction fragment length polymorphism; RIF, rifampin; SSCP, single-strand conformation polymorphism.

pyazinamide (PZA). Because RIF resistance is an excellent marker for multidrug-resistant TB and 95% of all RIF-resistant strains have mutations localized in an 81 base pair (bp) region of the bacterial RNA polymerase gene, *rpoB*, which encodes the active site of the enzyme,[28] numerous molecular strategies have been developed to successfully detect RIF resistance. In contrast, research has shown that more than a single gene mutation is frequently responsible for resistance to INH, EMB, and STR. Although numerous molecular formats have been used to detect RIF drug resistance, the complexity of the drug resistance mechanisms for other agents combined with the complexity of the technology has hampered their broader acceptance in the clinical laboratory setting. Nevertheless, relevant mutations for resistance to most of the primary antituberculosis drugs have been described; most commonly, conventional or real-time polymerase chain reaction (PCR) in conjunction with identification of specific mutations within the amplicon using a variety of methods are employed to detect drug resistance mutations.

Epidemiology

Molecular epidemiological methods have made significant contributions to our understanding of the pathogenesis and transmission of TB within populations. For example, molecular fingerprinting techniques have provided epidemiological evidence of exogenous reinfection as well as quantification of the level of infectiousness among AFB smear-negative patients.[29] Of great significance has been the use of DNA fingerprinting of *M. tuberculosis* isolates to determine the occurrence of laboratory cross-contamination of cultures. Laboratory cross-contamination represents a significant problem that can result in unnecessary treatment and drug toxicity for a patient. Molecular typing also has provided insight into the pathogenesis of cavitary and noncavitary disease caused by NTM.[24]

Available Tests

There are hundreds of publications in which molecular methods, using a variety of formats, have been used to directly detect mycobacteria in clinical specimens, identify mycobacteria, and detect drug resistance mutations in *M. tuberculosis*. Examples of NAA methods used for mycobacteria are provided in Tables 43-3 and 43-4.

Interpretation of Results

As with any laboratory test, prior to the interpretation of results of molecular tests on patient specimens or AFB clinical isolates, results of controls included with each run must be interpreted first. Selection and use of appropriate controls is an essential feature of any NAA test. Regardless of the format, positive and negative controls must be run in

parallel with patient samples. If the test is a laboratory-developed assay to directly detect *M. tuberculosis*, the concentration of the positive control should be near the lower limit of detection of the assay. If the run is large, multiple negative controls should be included to monitor for cross-contamination. As with any other laboratory test, if these controls do not perform as expected, the run must be repeated. Optimally, an internal inhibition control with upstream and downstream primer recognition sequences should be added to each patient sample reaction to monitor for inhibition. If an internal inhibition control is not included (some commercially available assays do not include an inhibition control, e.g., the AMTD test), a second reaction should be performed with addition and amplification of a known target nucleic acid. If inhibition is detected, a test result cannot be reported. For assays using gel or capillary electrophoresis (CE), molecular-weight markers are used.

Inherent problems and limitations associated with NAA tests are false-positive results due to cross-contamination, as well as false-negative results due to inhibition or inadequate sample collection, transport, and processing. In addition, sampling error due to low numbers of mycobacteria (paucibacillary) or inadequate sample volumes also can lead to false-negative test results. Results should be interpreted in the context of the patient's history, physical examination, and clinical course. Thus, it is imperative that clinicians as well as clinical microbiologists have a thorough understanding of the advantages and limitations of the particular NAA assay used for patient testing as well as the specific microbiology and pathogenesis of the identified mycobacterium.

Sequenced-based or PCR-RFLP methods for identification of NTM provide a rapid tool for identification of this vast group of organisms. However, paradoxically, the identification of mycobacteria has become even more complex with the introduction of these techniques, since these methods have uncovered greater complexity of mycobacterial species. In a study of 72 mycobacterial isolates that did not belong to any officially recognized species, only three cases had agreement of species identification by conventional tests, high-performance liquid chromatography, and genetic sequencing.[30] Further complicating the interpretation of molecular identification results were findings that specific species identification was not the rule using analysis of the 16S rRNA gene of patient strains.[31] Based on these results, identification of mycobacteria will mandate attention to quality control of available databases used for mycobacterial identification, coupled with the realization that the taxonomy of the genus *Mycobacterium* is far from elucidated. Thus, interpretation of identification results using molecular-based methods will require constant changes and updates to available databases.[30,32]

In conclusion, molecular-based diagnostic tests should be interpreted within the context of clinical information and test performance characteristics. Of great importance is strict adherence to guidelines for method validation that include the determination of clinical utility for accurate interpretation of results.

Laboratory Issues

There are numerous laboratory issues that pertain to the development, introduction, and performance of any clinical molecular test. Unfortunately, only guidelines, not universal standards, currently exist for the validation and subsequent quality control and assurance of laboratory-developed tests.[33,34] For example, all NAA methods require adequate and appropriately designed space, as well as other measures, is to minimize cross-contamination of samples.[35,36] This section addresses those laboratory issues that are specific to the diagnosis of mycobacterial infections by molecular assays.

False-positive amplification results can occur from carryover of amplicon while setting up and performing the molecular assay, as well as from cross-contamination of clinical specimens during processing for AFB smear and culture.[37] A review of 14 studies revealed a median false-positive rate of AFB culture of 3.1%.[38] Of 236 patients reported with false-positive cultures, 67% of patients were subsequently treated, and some had toxicity from therapy as well as unnecessary hospitalizations, tests, and contact investigations. Thus, performance of NAA tests to directly detect and identify *M. tuberculosis*, with their inherent potential for false-positive results, underscores the critical need for adherence to strict laboratory technique, inclusion of appropriate negative controls, and careful interpretation of results within the context of the clinical presentation.

As is the case with molecular testing for other infectious agents, the lack of a perfect gold standard complicates the interpretation of a positive amplification result with a negative culture. Likewise, false-negative amplification results can occur because of inhibition, sampling error, and inadequate sample preparation. The theoretical detection of one AFB has not generally been achieved in reality, which is a particular issue for the diagnosis of *M. tuberculosis* infection where there are often few organisms present in the clinical sample. This coupled with the tendency of mycobacteria to clump and thereby cause an uneven distribution of organisms in a sample can result in duplicate tests with discrepant results that are difficult to interpret.

Owing to the nature of NAA methods, quality control is essential for these procedures. A quality control program should consist of an internal quality control program, as well as participation in an external quality control program such as that offered by the CDC. The critical need for strict adherence to quality control measures was underscored when 20 unknowns were sent to 30 laboratories performing NAA tests; only five laboratories correctly identified the presence or absence of *M. tuberculosis* DNA in all samples.[39]

Specific to molecular-based identification methods, laboratories performing these techniques must be aware of the multiple problems with present sequence repositories such as base errors, ambiguous base designation, and incomplete sequences.[31] It is becoming apparent that visual inspection is becoming more cumbersome and challenging due to the growing number of *hsp65* alleles described in the literature.[32] Another issue slowly arising from PCR-RFLP analysis for identification is the lack of standardization for electrophoresis conditions, which makes comparison of data from different laboratories challenging. For example, difficulties in PCR-RFLP interpretation stemming from similarities in a number of band sizes needed to discriminate species via agarose gel electrophoresis could be alleviated using 10% polyacrylamide gel electrophoresis.[40]

As previously mentioned, the complexity of drug resistance in *M. tuberculosis* has hindered the utility of molecular tests. However, this limitation may be ultimately overcome by development of other molecular approaches (see "Future Applications"). Another caveat to testing is that the presence of a resistance gene does not always imply expression of that gene and phenotypic resistance.

Cost-effectiveness is another laboratory issue associated with the performance of NAA tests for the diagnosis of TB and identification of mycobacteria. Since NAA tests cannot currently replace conventional methods for the diagnosis and management of TB, a NAA test is an additional test with associated costs. Although cost savings may be realized because of possibly preventing more invasive and costly diagnostic procedures (e.g., bronchoalveolar lavage), limiting unnecessary or potentially toxic empiric antituberculous therapy, shortening hospital stays in costly isolation rooms, and limiting transmission, this type of outcomes research has yet to be systematically done. Our understanding of the natural history of infection caused by *M. tuberculosis* will continue to evolve as different applications are explored and evaluated in the clinical setting. Many questions must be answered to fully exploit and utilize the potential of molecular tests for the diagnosis of mycobacterial infections. Toward this end, continued objective evaluation of the analytic and clinical performance of molecular tests, and the impact on patient outcomes, is imperative.

Future Applications

Continued advances in nucleic acid-based technologies will serve only to enhance capabilities for the diagnosis of mycobacterial infections. Efforts employing a variety of strategies are already well under way.

Miscellaneous Methods

Different strategies and formats to allow for easier and more rapid means by which to detect and identify mycobacteria are continually being published. To illustrate, DNA extracted directly from slides of AFB smear-negative specimens was tested for *M. tuberculosis* by PCR and sequence analysis for RIF resistance; this technique was 100% sensitive and specific for the detection of *M. tuberculosis* and RIF resistance.[41]

Another recent approach is the use of fluorescently labeled peptide nucleic acids (PNAs) to directly detect *M. tuberculosis* microscopically.[42] PNAs are DNA-like molecules in which the sugar-phosphate backbone is replaced with a peptidelike structure that can hybridize to specific DNA sequences with specific base pairing. PNAs can be labeled with a fluorescent dye to allow for visualization of PNA binding to specific DNA sequences present on a slide. Of significance, the PNA can easily pass through an intact cell wall and bind specifically to intracellular nucleic acid sequences. This strategy for the direct microscopic identification of *M. tuberculosis* and NTM appears to hold some promise. Finally, new rapid phenotypic methods, such as the luciferase phage assay or the PhaB assays[43] (Biotec Laboratories, Ipswich, UK), avoid some pitfalls associated with genotypic methods for detecting drug resistance in *M. tuberculosis*. New strategies will continue to be developed and evaluated in direct detection and identification of mycobacteria as well as drug susceptibility testing.

Automation and Miniaturization

The development of real-time PCR assays and other amplification formats in which single or multiple nucleic acid targets can be amplified and analyzed in a single closed tube in minutes rather than hours is particularly suited for subsequent automation. Instruments that automate the extraction of either RNA or DNA also have been introduced, allowing for the performance of a greater number of molecular-based assays while also providing a more consistent quality of nucleic acid for analysis.

In addition to automation of NAA assays for the diagnosis of mycobacterial infections, there will be miniaturization of analytical devices by micromachining technology that will perform PCR in microreactors consisting of either silicon or silicon and glass microchips.[44] Advances in this area will have a major impact on the ability to diagnose TB in the field and might prove to be particularly attractive for use in underdeveloped countries where the prevalence of TB is high.

Microarrays

High-density oligonucleotide arrays can rapidly examine large numbers of DNA sequences with a single hybridization step. DNA microarrays can be employed in two modes: (1) DNA-based comparison of the genomic content of different strains and (2) RNA-based monitoring of gene

expression. Preliminary work has begun in the simultaneous detection, species identification, drug resistance profiling, and strain genotyping of mycobacteria involved in human disease.[45,46] Utilizing a high-density array, simultaneous species identification of 121 isolates was achieved, as well as detection of RIF resistance in 41 isolates of *M. tuberculosis*.[46] Similarly, high-density probe arrays for species identification and RIF resistance using sequences from the 16S rDNA gene and the *rpoB* gene, respectively, were used to evaluate 51 *M. tuberculosis* RIF resistance-causing rpoB mutations; arrays also included 2.2 kilobases (kb) of the *M. tuberculosis* wild-type *katG* gene, previously shown to confer INH resistance.[47] The total process from culture, including sample preparation and amplification, took less than 4 hours manually. As *M. tuberculosis* drug resistance determinants are gradually delineated, this type of platform could be expanded, even with the addition of epidemiological markers.

Clinical Knowledge

In addition to further technological innovation, epidemiologic and operational research is needed to develop algorithms for optimal use of molecular-based assays for patient management, particularly in light of test expense. In developed countries, TB remains a low-prevalence disease, and the most rapid way to improve test performance is to select patients for testing in whom results will have the greatest predictive value. Toward this end, there is a need for well-designed, patient-based clinical trials with NAA tests. Numerous questions need to be answered regarding the use of these tests in previously treated individuals, children, and other subpopulations, in paucibacillary forms of TB, as well as regarding how to assess the infectiousness of individual patients, when to isolate a patient, and when to begin contact investigations.[3] Additional questions remain as to whether NAAs can ultimately be used to distinguish latent from active *M. tuberculosis* infections and monitor response to chemotherapy. As the natural history of mycobacterial infections becomes better understood in conjunction with the advancement of molecular test methods, significant potential for rapid, sensitive, specific, and cost-effective diagnosis of mycobacterial infections will become a reality.

References

1. Bradley SP, Reed SL, Cantanzaro A. Clinical efficacy of the amplified *Mycobacterium tuberculosis* direct test for the diagnosis of pulmonary tuberculosis. *Am J Resp Crit Care Med.* 1996;153:1606–1610.
2. Nardell E, Bautolte S, Shampra P, et al. Tuberculosis without culture confirmation (abstract). *Am Rev Respir Dis.* 1995;151:A336.
3. American Thoracic Society Workshop: rapid diagnostic tests for tuberculosis. What is the appropriate use? *Am J Respir Crit Care Med.* 1997;155:1804–1814.
4. Barnes PF. Rapid diagnostic tests for tuberculosis. *Am J Respir Crit Care Med.* 1997;155:1497–1498.
5. Jonas V, Acedo M, Clarridge JE, et al. A multi-center evaluation of MTD and culture compared to clinical diagnosis [abstract]. In: *Abstracts of the 98th General Meeting of the American Society for Microbiology*; May 17–21, 1998; Atlanta, GA. Abstract L-31.
6. Cantanzaro A, Perry S, Clarridge JE, et al. The role of clinical suspicion in evaluating a new diagnostic test for active tuberculosis. *JAMA.* 2000;283:639–645.
7. O'Sullivan CE, Miller DR, Schneider PS, et al. Evaluation of Gen-Probe Amplified *Mycobacterium tuberculosis* Direct Test by using respiratory and nonrespiratory specimens in a tertiary care center laboratory. *J Clin Microbiol.* 2002;40:1723–1727.
8. Bergmann JS, Yuoh G, Fish G, et al. Clinical evaluation of the enhanced Gen-Probe Amplified *Mycobacterium tuberculosis* Direct Test for rapid diagnosis of tuberculosis in prison inmates. *J Clin Microbiol.* 1999;37:1419–1425.
9. Peirsimoni C, Callegaro A, Scarparo C, et al. Comparative evaluation of the new Gen-Probe *Mycobacterium tuberculosis* Direct Test and the semiautomated Abbott LCx *Mycobacterium tuberculosis* assay for the direct detection of *Mycobacterium tuberculosis* complex in respiratory and extrapulmonary specimens. *J Clin Microbiol.* 1998;36:3601–3604.
10. Centers for Disease Control and Prevention. Nucleic acid amplification tests for the diagnosis of tuberculosis. *MMWR Morb Mortal Wkly Rep.* 1996;45:950–952.
11. Centers for Disease Control and Prevention. Update: nucleic acid amplification tests for tuberculosis. *MMWR Morb Mortal Wkly Rep.* 2000;49:593–594.
12. Baker CA, Cartwright CP, Williams DN, et al. Early detection of central nervous system tuberculosis with the Gen-Probe nucleic acid amplification assay: utility in an inner city hospital. *Clin Infect Dis.* 2002;35:339–342.
13. Bonington A, Strang JIG, Klapper PE, et al. Use of Roche AMPLICOR *Mycobacterium tuberculosis* PCR in early diagnosis of tuberculous meningitis. *J Clin Microbiol.* 1998;36:1251–1254.
14. Lang AM, Feris-Iglesias J, Pena C, et al. Clinical evaluation of the Gen-Probe Amplified Direct Test for detection of *Mycobacterium tuberculosis* complex organisms in cerebrospinal fluid. *J Clin Microbiol.* 1998;36:2191–2194.
15. Roos KL. *Mycobacterium tuberculosis* meningitis and other etiologies of the aseptic meningitis syndrome. *Semin Neurol.* 2000;20:329–335.
16. Peirsimoni C, Scarparo C, Piccoli P, et al. Performance assessment of two commercial amplification assays for direct detection of *Mycobacterium tuberculosis* complex from respiratory and extrapulmonary specimens. *J Clin Microbiol.* 2002;40:4138–4142.
17. Scarparo C, Piccoli P, Rigon A, et al. Comparison of enhanced *Mycobacterium tuberculosis* Amplified Direct Test with COBAS AMPLICOR *Mycobacterium tuberculosis* assay for direct detection of *Mycobacterium tuberculosis* complex in respiratory and extrapulmonary specimens. *J Clin Microbiol.* 2000;38:1559–1562.
18. Forbes BA, Hicks KE. Ability of PCR assay to identify *Mycobacterium tuberculosis* in BACTEC 12B vials. *J Clin Microbiol.* 1994;32:1725–1728.
19. Katila ML, Katila P, Erkinjuntti-Pekkanen R. Accelerated detection of mycobacteria with MGIT 960 and COBAS AMPLICOR systems. *J Clin Microbiol.* 2000;38:960–964.
20. Scarparo C, Piccoli P, Rigon A, et al. Direct identification of mycobacteria from MB/BacT Alert 3D bottles: comparative evaluation of two commercial probe assays. *J Clin Microbiol.* 2001;39:3222–3227.
21. Bennedsen J, Thosen V, Pfyffer GE, et al. Utility of PCR in diagnosing pulmonary tuberculosis. *J Clin Microbiol.* 1996;34:1407–1411.
22. Desjardin LE, Chen Y, Perkins MD, et al. Comparison of the ABI 7700 System (TaqMan) and competitive PCR for quantitation of IS6110 DNA in sputum during treatment of tuberculosis. *J Clin Microbiol.* 1998;36:1964–1968.

23. Hellyer TJ, Desjardin LE, Hehman GL, et al. Quantitative analysis of mRNA as a marker for viability of *Mycobacterium tuberculosis*. *J Clin Microbiol.* 1999;37:290–295.
24. Griffith DE, Brown-Elliott BA, Wallace RJ. Diagnosing nontuberculous mycobacterial lung disease. *Infect Dis Clin N Am.* 2002;16: 235–249.
25. Kirschner P, Springer B, Vogel U, et al. Genotypic identification of mycobacteria by nucleic acid sequence determination: report of a two year experience in a clinical laboratory. *J Clin Microbiol.* 1993;31:2882–2889.
26. Brunello F, Ligozzi M, Cristelli E, et al. Identification of 54 mycobacterial species by PCR-restriction fragment length polymorphism analysis of the hsp65 gene. *J Clin Microbiol.* 2001;39:2799–2806.
27. Padilla E, Gonzalez V, Manterola JM, et al. Comparative evaluation of the new version of the INNO-LiPA mycobacteria and GenoType *Mycobacterium* assays for identification of *Mycobacterium* species from MB/BacT liquid cultures artificially inoculated with Mycobacterial strains. *J Clin Microbiol.* 2004;42:3083–3088.
28. Ramaswamy S, Musser JM. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber Lung Dis.* 1998;79:3–29.
29. Burgos MV. Molecular epidemiology of tuberculosis. *Eur Respir J.* 2002;20(suppl 36):54s–65s.
30. Tortolli E, Bartoloni A, Böttger EC, et al. Burden of unidentifiable mycobacteria in a reference laboratory. *J Clin Microbiol.* 2001;39: 4058–4065.
31. Turenne CY, Tschetter L, Wolfe J, et al. Necessity of quality-controlled 16S rRNA gene sequence databases: identifying nontuberculous *Mycobacterium* species. *J Clin Microbiol.* 2001;39:3637–3648.
32. da Silva Rocha A, Barreto AM, Campos CED, et al. Novel allelic variants of mycobacteria isolated in Brazil as determined by PCR-restriction enzyme analysis of hsp65. *J Clin Microbiol.* 2002;40: 4191–4196.
33. National Committee for Clinical Laboratory Standards. *Molecular Diagnostic Methods for Infectious Diseases: Approved Guidelines [MM3-A].* Wayne, PA: National Committee for Clinical Laboratory Standards; 1993.
34. National Committee for Clinical Laboratory Standards. *Quantitative Molecular Methods for Infectious Disease: Proposed Guidelines.* Wayne, PA: National Committee for Clinical Laboratory Standards; 2001. NCCLS Document MM6-P.
35. Lisby G. Application of nucleic acid amplification in clinical microbiology. *Mol Biotechnol.* 1999;12:75–99.
36. Wolk D, Mitchell S, Patel R. Principles of molecular biology testing methods. *Infect Dis Clin N Am.* 2001;15:1157–1204.
37. Chang CL, Kim HH, Son HC, et al. False-positive growth of *Mycobacterium tuberculosis* attributable to laboratory contamination confirmed by restriction fragment length polymorphism analysis. *Int J Tuberc Lung Dis.* 2001;5:861–867.
38. Burman WJ, Reves RR. Review of false-positive cultures for *Mycobacterium tuberculosis* and recommendations for avoiding unnecessary treatment. *Clin Infect Dis.* 2000;31:1390–1395.
39. Noordhoek GT, vanEmbden JDA, Kolk AH. Reliability of nucleic acid amplification for detection of Mycobacterium tuberculosis: an international collaborative quality control study among 30 laboratories. *J Clin Microbiol.* 1996;34:2522–2525.
40. Brunello F, Ligozzi M, Cristelli E, et al. Identification of 54 mycobacterial species by PCR-restriction fragment length polymorphism analysis of the hsp65 gene. *J Clin Microbiol.* 2001;39:2799–2806.
41. Patnaik M, Liegmann K, Peter JB. Rapid detection of smear-negative *Mycobacterium tuberculosis* by PCR and sequencing for rifampin resistance with DNA extracted directly from slides. *J Clin Microbiol.* 2001;39:51–52.
42. Stender H, Lund K, Petersen KH, et al. Fluorescence in situ hybridization assay using peptide nucleic acid probes for differentiation between tuberculous and nontuberculous *Mycobacterium* species in smears of *Mycobacterium* cultures. *J Clin Microbiol.* 1999;37:2760–2765.
43. Riska PF, Su Y, Bardarov S, et al. Rapid film-based determination of antibiotic susceptibilities of *Mycobacterium tuberculosis* strains by using a luciferase reporter phage and the Bronx box. *J Clin Microbiol.* 1999;37:1144–1149.
44. Park DJ, Drobniewski FA, Meyer A, et al. Use of a phage-based assay for phenotypic detection of mycobacteria directly from sputum. *J Clin Microbiol.* 2003;41:680–688.
45. Schneegass I, Köhler JM. Flow-through polymerase chain reactions in chip thermocyclers. *J Biotech.* 2001;82:101–121.
46. Gingeras TR, Ghandour G, Wang E, et al. Simultaneous genotyping and species identification using hybridisation pattern recognition analysis of generic *Mycobacterium* arrays. *Genome Res.* 1998;8:435–448.
47. Troesch A, Nguyen H, Miyada CG, et al. *Mycobacterium* species identification and rifampin resistance testing with high-density DNA probe arrays. *J Clin Microbiol.* 1999;37:49–55.

Section VI

Identity Testing

Section Editor: Vivianna M. Van Deerlin

Chapter 44

Forensic DNA Typing

Victor W. Weedn

Historical Overview

The beginning of the forensic DNA typing revolution was marked by the 1985 publication of a landmark article by Sir Alec Jeffreys of Leicester, England, in which he coined the term "DNA fingerprint" and suggested the potential application of DNA fingerprinting in forensic investigations.[1,2] Using restriction fragment length polymorphism (RFLP) analysis and DNA probes complementary to polymorphic regions of the genome called "minisatellites," he described how bar code–like patterns were produced. The autoradiographic patterns seemed to be different in every person tested; hence the term "DNA fingerprint" likened this molecular typing to an individual's unique digital fingerprint. In the same year, Jeffreys applied his revolutionary technique to resolve a paternity dispute and to help solve a double rape-homicide in England.

Queen v Pitchfork

The first criminal investigation using DNA typing was in a double rape-homicide (of Linda Mann in 1983 and of Dawn Ashworth in 1986) in the English countryside. Richard Buckland, a person of low intelligence and sexual fetishes, became the focus of early suspicion and was charged but then exonerated by the new Jeffreys DNA tests. Males in the community between 13 and 30 years of age were asked to volunteer blood samples for DNA testing. There were no matches despite 4500 "bloodings." However, police discovered that a man named Ian Kelly had substituted his blood for Colin Pitchfork's. Pitchfork was subsequently DNA matched and then convicted of both homicides.[3]

Within a year of this DNA typing discovery (1986), the first widespread commercial applications of identity testing based on molecular techniques came into existence when two commercial laboratories opened in the United States to offer DNA typing services primarily for the purpose of parentage testing: Forensic Science Associates (Redmond, California) and Lifecodes (Valhalla, New York). Shortly thereafter, a company called Cellmark Diagnostics became operational in England and in Germantown, Maryland. While the commercial applications drove the growth of DNA typing, the fact that the earliest testing was performed by commercial laboratories with proprietary databases of allele frequencies was initially a problem that delayed the use of this new information for legal forensic work until statistical data necessary for interpretation of results was made public.

Pennsylvania v Pestinikas

The first use of DNA typing in the United States was in a 1986 nursing home negligent homicide case. Forensic Science Associates performed DNA tests to prove that organs in the autopsy had not been switched as was alleged by one expert. The DNA in this case had become highly degraded, averaging fragments of approximately 100 base pairs (bp).

Florida v Andrews

The first U.S. criminal conviction based on DNA tying was of a serial rapist, Tommy Lee Andrews (1987). A series of breaking and entering women's homes and rapes began in 1986 in Orlando, Florida. A stakeout resulted in an arrest, and Lifecodes matched the suspect's DNA to vaginal swabs of two of the rape victims.

Forensic Identification Methods

The Federal Bureau of Investigation (FBI) began performing DNA typing casework in December 1988. A few months later, in March 1989, the state of Virginia became the first state with an operational crime laboratory. Both laboratories used RFLP analysis similar to that described by Jeffreys.

Most forensic evidence links an item to the crime scene but does not identify the perpetrator. Other than imaging data, only fingerprint friction ridge analysis and DNA typing permit the positive identification of their source. Although the individuality of fingerprint friction ridge patterns was described centuries ago, fingerprints came into use for forensic purposes only in the latter part of the 19th century and were admitted into U.S. courts as evidence in the 1930s. The FBI owes its origins to the need for a centralized database of fingerprints. Today, millions of fingerprints are filed for criminal and civil purposes. In a similar way, DNA typing provides a biological fingerprint of an individual, and the government now also maintains databases of DNA profiles. Thus, DNA not only confirms a detective's hunch but also provides a powerful investigatory tool to identify suspects.

Serology tests were the forerunners to forensic DNA identity testing of biological materials and included primarily ABO blood group typing and serum protein isoenzyme testing. Unfortunately, these tests did not have sufficient discriminatory power to unequivocally identify an individual. Additionally, these traditional forensic tests had technical drawbacks, such as requiring a large quantity of biological specimen, rapid degradation of protein targets, and limited types of acceptable specimens. For example, full serologic testing could be performed on blood, but not semen, saliva, and other tissues amenable to DNA testing. With the discovery of the DNA fingerprint and technological advances in molecular biology such as the Southern blot and the polymerase chain reaction (PCR), it became feasible to target the genomic source of the variations in the protein marker systems, that is, the differences or polymorphisms in an individual's DNA, rather than the protein product itself. Advantages of DNA typing over serologic methods include being more informative, species specific, tissue independent, sensitive, and resistant to degradation.[4]

In 2004, as least 150 government forensic laboratories and approximately a half dozen private laboratories were performing some level of DNA examination. The Bureau of Justice Statistics (BJS) conducted a survey of publicly operated crime laboratories performing DNA examinations.[5] In 2000, laboratories received 31,000 DNA investigative case submissions (typically with several samples each) and analyzed 25,000 of the cases; 81% ended the year with a backlog of 16,000 cases. These numbers represented a marked increase from previous years. In addition, the laboratories received nearly 177,000 convicted offender database samples, resulting in a median of 176 cases received per laboratory. With a median staff size of six full-time DNA analysts, 45% of the laboratories outsourced some of their workload to private laboratories in attempts to meet workload requirements. The significant impact of DNA testing on forensic sciences is evidenced by the 2003 *Advancing Justice Through DNA Technology* initiative proposed by US President George W. Bush, which called for a total of more than $1 billion funding over 5 years to the exclusion of all other forensic science programs of the National Institute of Justice (NIJ).[6]

Although identity testing by DNA typing has its origins in forensics and parentage testing, DNA typing has many other applications in clinical medicine and pathology that are described in other chapters of this book. This chapter focuses on the forensic applications of DNA typing, beginning with an overview of the methods and genetic systems used for DNA typing, followed by a discussion of technical and laboratory testing issues.

Forensic Testing and Sample Considerations

Forensic tests differ from clinical tests in several respects. The samples to be tested vary substantially. An analyst may routinely see cigarette butts but then must be prepared to face a partially eaten orange for the first time. The sample may be of trace amounts, permitting testing only once, although routine practice is to attempt to save a portion of the sample for potential testing by the defense. The evidentiary sample is usually neither fresh nor pristine. For example, semen samples are generally admixed with vaginal cells and microbial flora in a rape swab (Figure 44-1), spit on a sidewalk has been exposed to the sun and rain, and blood on the floor may be that of both the victim and the assailant. Furthermore, the testing is performed with an eye to court challenge. Thus, the forensic scientist will use only well-validated protocols, will document all aspects of the laboratory processing, and will stand ready to defend the science, the procedures, and the testing against any legal attack. Even the purpose of the testing is usually different from that of clinical testing. The forensic scientist most often answers the question "Who did this come from?" rather than, as in the clinical laboratory, "How much is here?" Moreover, the clinical laboratory staff often is not

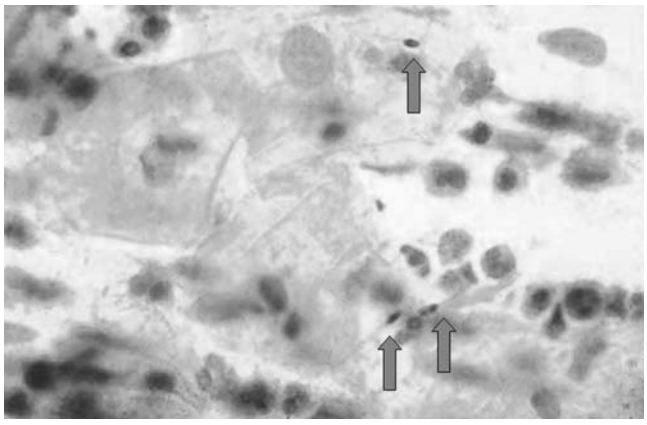


Figure 44-1. The most common DNA evidence in U.S. crime laboratories is a vaginal swab from a rape kit. This photomicrograph is a stained vaginal smear from a rape kit. The arrows indicate spermatozoa. In addition to the DNA from the male contributor, there is also DNA from the female epithelial squamous and white cells, as well as that of the microbial flora.

familiar with the regulations, standards, and quality assurance practices of forensic laboratories. Thus, clinical laboratories, though capable, are often not well suited to address forensic issues.

In the United States, rape kits dominate the evidential submissions to the forensic DNA laboratory (Figure 44-1). A typical rape kit collection will include vaginal, anal, and oral swabs, a blood collection tube, pubic combings, and exemplars of pubic and scalp hairs. If a condom is used by a rapist and later found, it may yield semen from the male perpetrator on the inside and vaginal epithelial cells and white blood cells from the female victim on the outside. Blood and other specimens from homicides are the next most common evidentiary specimens submitted to forensic laboratories in the United States. The FBI reports that biological evidence is collected in 60% of violent crimes.

A sample of blood from windows broken to gain entry into a house in cases of burglary are not uncommon. In the United Kingdom, testing is most frequently done for property crimes; in contrast, this testing is not performed in the United States, where fewer than 20% of property crimes are solved. However, this standard may be changing. In a preliminary study, the Miami-Dade laboratory was successful in approximately 50% of the DNA submissions from burglary cases, and New York City is set to begin routine property crime casework.

DNA testing is not limited to blood and semen. In fact, DNA can be collected from a myriad of items and materials. Saliva may be deposited on beverage containers, stamps and envelope seals, gum, cigarettes, or food (Figure 44-2). Investigators have followed suspects to obtain "abandoned" specimens, such as facial tissues or stains from the sidewalk where a suspect spat. Cords used as a murder weapon for strangulation can yield both victim and perpetrator DNA (Figure 44-3). Shed hairs, which contain mitochondrial DNA, also can be used for identification purposes. Fingernail scrapings occasionally yield DNA from the perpetrator if a victim struggled and scratched

Figure 44-2. DNA testing identified a masked bandit when his peach strudel was left at the scene of an armed robbery and used for DNA testing.

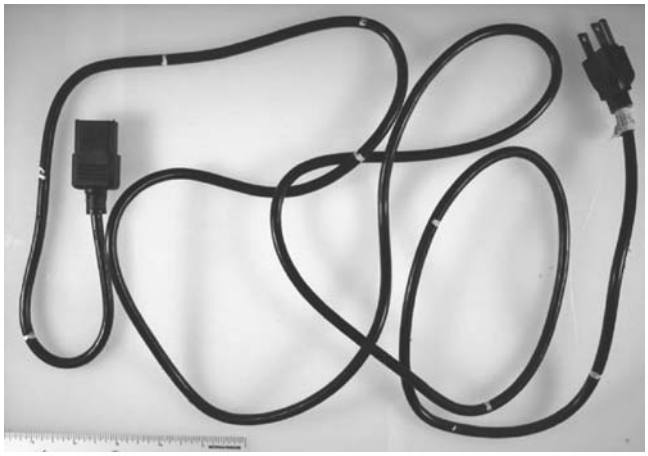


Figure 44-3. This vacuum cleaner cord was used as a ligature for a strangulation murder. Swabbings of the cord along its length revealed the victim's DNA in the center and a mixture of the victim and accused on outer areas of the cord.

the perpetrator. Bite marks can be swabbed for DNA. Reference samples may come from toothbrushes, razors, combs, clothing, and medical specimens.

"Trace DNA" or "low copy number" (LCN) DNA evidence became an issue when van Oorschot reported in 1997 that minute quantities of DNA can be recovered from fingerprints.[7] In 1999, the United Kingdom's Forensic Science Service (FSS) became the first laboratory in the world to use LCN DNA analysis to investigate crimes. Conventional laboratory testing will successfully type DNA from approximately 150 cells, but the FSS laboratory types DNA from 30 to 50 cells. The High Sensitivity Laboratory of the New York City Office of the Chief Medical Examiner (NYC OCME) became the first laboratory in the United States to routinely perform LCN analysis it and claims an ability to detect DNA from as few as six cells. Such testing involves minimizing buffer volumes and increasing thermal cycling from 28 to 32 cycles. The High Sensitivity Laboratory anticipates useful yields in only 10% to 20% of its specimens. The implications of being able to test DNA from "handled objects" are enormous. Analysts hope to obtain pure (single) profile results that are assumed to be from the last person who pulled the trigger of a gun or handled a knife. However, contamination from prior handling, secondary transfer, and spurious results is a concern that needs to be addressed. These issues are compounded by the destructive nature of the testing, which eliminates the possibility of retesting. For these reasons, there is controversy over whether LCN analysis should be used beyond investigative applications, that is, for prosecutorial purposes. No national standards are yet in place.

Forensic DNA identity testing can be used in other contexts. Urine samples from drug testing may be tested to confirm that the sample is truly from the person who allegedly gave the sample. DNA testing is used in a forensic context for disaster victim identification,[8] parentage analysis, and even in wildlife cases (where speciation or herd origin may be determined by genetic testing to prove illegal poaching).

Genetic Systems and Methods for DNA Typing

Forensic DNA identity testing is based on the detection and comparison of polymorphisms, or differences, in the DNA between individuals.[9,10] These differences can be due to single nucleotide polymorphisms (SNPs [pronounced "snips"]) or variations in length of a specific region or locus in the genome, that is, length polymorphisms. These polymorphisms result in different forms, or alleles, of genetic markers. Everyone has two alleles of each autosomal chromosomal genetic marker: one inherited maternally and the other paternally. DNA typing (profiling) most commonly involves length polymorphisms from regions of repetitive DNA, such as minisatellites, microsatellites, and interspersed elements. Regardless, forensic DNA testing employs only a small fraction of the differences in the human genome between different individuals.

Identity Testing by RFLP

The polymorphic nature of DNA can be visualized by conventional RFLP analysis, as used for clinical disease testing. DNA fragments of differing length are produced when DNA from different individuals is cleaved with a restriction endonuclease with a recognition sequence that is altered by a nucleotide polymorphism. Thus, the presence or absence of a given sequence within the restriction endonuclease cleavage site determines the size of DNA fragments produced. RFLP fragments in genomic DNA can be visualized by electrophoresis followed by capillary transfer to a membrane and hybridization with a radioactively labeled probe complementary to the specific locus being assessed using Southern blot analysis (SBA). For any one polymorphism, the population is divided into only three groups: variant homozygotes, heterozygotes, and wild-type homozygotes. Although the process is potentially useful, it simply requires too much work for too little information to be of any practical use for forensic identity testing. However, RFLP analysis can be used to visualize length polymorphisms. "Hypervariable DNA" regions, discovered by Ray White and then later rediscovered by Alec Jeffreys, produce many different bands within a population and even from a single family kindred using SBA. This discovery made the Southern blot useful for forensic purposes.

Hypervariable, or "minisatellite," regions are commonly located in subtelomeric chromosomal regions and are also known as variable number of tandem repeat (VNTR) loci. VNTR loci can generate RFLPs because each locus contains a core sequence that is repeated many times in tandem, but with a different number of total core repeats on each allele in different people. The core repeat unit length in VNTRs is 8 to 80 base pairs (bp), resulting in overall sizes of 0.5 to >20 kilobases (kb) at different loci. The variation is hypothesized to originate from slippage of DNA polymerase during DNA replication. When a VNTR locus is located between two restriction enzyme cleavage sites and the DNA fragments are analyzed by SBA, different-sized bands are produced depending on the number of core repeats at the locus probed.

The probes used originally by Jeffreys hybridized to multiple minisatellite loci in the genome. Thus, in the RFLP Southern blot, numerous bands were visualized per probe, leading to the term "multilocus probes." The resultant band pattern resembled a bar code, which Jeffreys claimed was unique to an individual (Figure 44-4). However the pattern was subject to slight changes depending on the hybridization stringency conditions; that is, bands would drop out or in. Another drawback was that the technique required a significant amount of high-molecular-weight DNA to perform, and, perhaps most distressing, because any given band could not be ascribed to any specific genetic marker, no allele frequency statistics could be determined.

In contrast to Jeffreys' multilocus probes, single-locus probes (SLPs), which hybridize to only a single region of the genome, were being pioneered in the United States by Lifecodes. Using SBA, SLPs generally yield one or two bands per probe (Figure 44-5). SLP genetic systems have many advantages over multilocus probe RFLPs, including cleaner bands, precise frequency statistics, less requirement for high-molecular-weight DNA, and less dependence on electrophoresis conditions. For these reasons, RFLP testing with SLPs became the predominant method of DNA typing in forensic laboratories in the United States throughout most of the 1990s. Initially, restriction enzymes with a 6 bp recognition sequence were used but were later abandoned in favor of the more robust and reliable 4 bp recognition restriction enzyme *Hae* III, pioneered by Bruce Budowle of the FBI. The forensic community settled on a group of six genetic loci for routine use (D1S7, D2S44, D4S139, D10S28, D14S13, and D17S79). The nomenclature used for hypervariable repeats is D for DNA, a number for the chromosome on which the locus is located, S for unique DNA segment, and the final number, a sequential number assigned for uniqueness.

VNTR loci are highly polymorphic and thus extremely discriminating. The power of discrimination, the ability to distinguish individuals from one another, for the set of 6 standard VNTR loci for Caucasian Americans is approximately 1 in 1.2×10.[11] However, testing of VNTR loci by SBA is expensive, labor-intensive, and difficult to automate. Typically, in the early days of DNA testing it took crime laboratories 6 weeks to perform testing for an identity testing case by RFLP analysis. Furthermore, RFLP does not produce discrete allelic designations but rather a fragment size with measurement imprecision. Lastly, RFLP analysis requires a large amount of intact high-molecular-weight DNA, making it impractical for many forensic applications. The advent of PCR to amplify regions of genomic DNA resolved many of the disadvantages of using SBA. PCR transformed the field of forensics and led to the abandon-

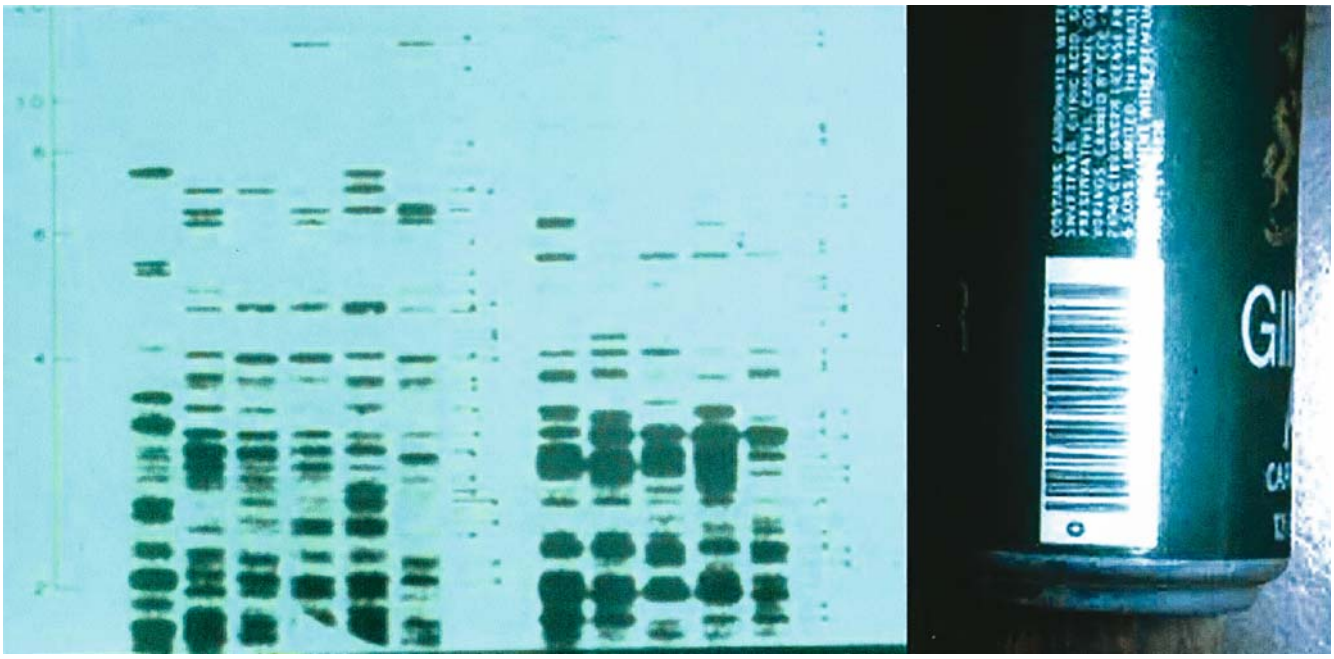


Figure 44-4. In 1985, Alec Jeffreys first described a DNA fingerprint. He used a multilocus minisatellite probe resulting in a band pattern similar to a bar code, such as the one shown on the can to the right. The various lanes of the autoradiograph are from different individuals, demonstrating that each shows a different pattern of bands. This multilocus probe method of DNA typing has since been abandoned for forensic use.

ment of RFLP-based systems in forensic analysis in favor of PCR-based systems. Surveys conducted by the BJS indicate that 43 of 110 laboratories in 1998 reported using RFLP for casework and 14 for database information, but by 2000, only one laboratory reported using RFLP for casework and none for database information.[5]

Identity Testing by PCR

Today, all forensic DNA testing begins with amplification of the DNA target by PCR (Table 44-1). PCR amplification is relatively easy to perform, inexpensive, quick, and amenable to automation. PCR also permits labeling of the amplified fragments, and multiplex or simultaneous amplification of several loci in a single reaction. PCR amplification also allows the routine testing of nanogram quantities of DNA, and can be optimized to permit testing of trace, picogram quantities of DNA, enabling the use of new classes of evidentiary specimens. However, such a level of sensitivity requires extreme care to prevent contamination, including laboratory facilities with separate pre- and postamplification areas. Lastly, PCR permits testing of evidentiary material in which the DNA has become degraded and only a few mere fragments with the intact target sequence remain. Although amplification methods other

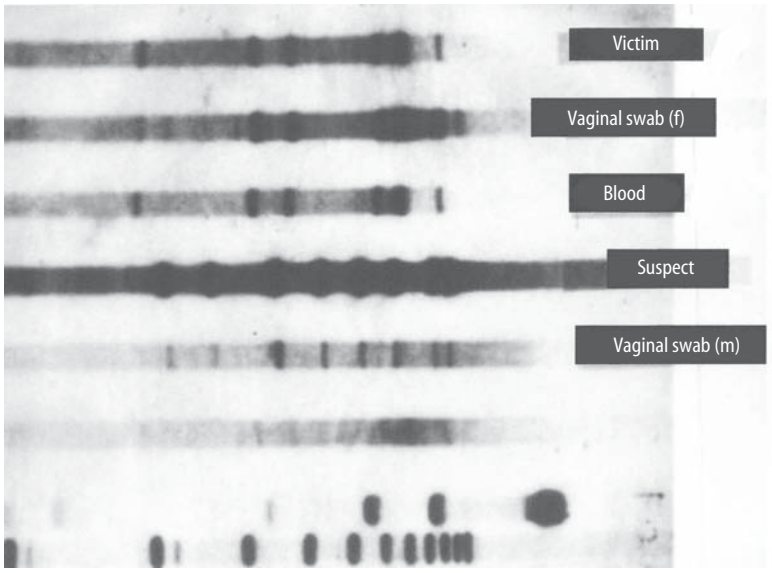


Figure 44-5. RFLP autoradiograph with five analytical lanes and three control lanes. The DNA profile of the reference sample from a female rape victim matches the DNA profile of blood found at the scene and that of the female fraction of a vaginal swab. The DNA profile of the suspect reference specimen matches the male fraction of a vaginal swab but does not match the DNA profile of the female victim.

Table 44-1. Summary of DNA Typing System Usage in Crime Laboratories

Typing Method	PCR-Based	Late 1980s	1990s	2000s	Utility
RFLP	No	**Dominates**	**Dominates**	Abandoned	Routine casework
Dot blots	Yes	Used	Used	Abandoned	Routine casework
STRs	Yes	In research	Used	**Dominates**	Routine casework
mtDNA	Yes	In research	Used	Used	Hairs, degraded samples
Y-STRs	Yes		In research	Used	Vaginal swabs in rape cases
SNPs	Yes			In research	Very degraded samples

mtDNA, mitochondrial DNA; RFLP, restriction fragment length polymorphism; SNPs, single nucleotide polymorphisms; STRs, short tandem repeats; Y-STRs, Y chromosome short tandem repeats.

than PCR exist, the conservative forensic community will not likely be quick to adopt an alternative to PCR unless there is a very good reason to do so (e.g., chip-based technologies).

The Human Leukocyte Antigen DQA1 Gene and the Polymarker System

The first application of PCR for identity testing was based on the detection of sequence polymorphisms at the DQ alpha (later called DQA1) locus of the human leukocyte antigen (HLA) histocompatibility region. The DQA1 system is a single but complex system with 21 genotypes that give an average discriminatory power of about 1 in 20. Detection of the DQA1 polymorphisms was commercialized for forensic identification purposes in a reverse dot-blot format in which amplified target DNA fragments, called amplicons, bind to complementary immobilized probes. Although the discriminatory power of the DQA1 locus is less than a VNTR locus, many smaller forensic laboratories initially preferred DNA testing by PCR of the HLA DQA1 locus because the dot-blot format was easier to perform, required less training, did not require sophisticated statistical analysis, and was less expensive to implement. In fact, the first case in the United States in which DNA testing was used, *Pestinikas v Pennsylvania* (1986), involved DQA1 testing.

To improve its discriminatory power, the DQA1 system later was incorporated with a "polymarker" system (Roche Molecular Systems Amplitype PM + DQA1) composed of five additional genetic loci (low-density lipoprotein receptor, glycophorin A, hemoglobin gamma-globin chain, D7S8, and group-specific component; Figure 44-6). The discriminatory power of the polymarker system combined with HLA DQA1 is approximately 1 in 2,000. Although analysis of the DQA1 and polymarker polymorphisms is technically simple, it has a few drawbacks. First, the power of discrimination is significantly lower than that of RFLP analysis because fewer alleles are used. This is a problem for forensic analysis, since a high discriminatory power is required. Second, the dot-blot method is not optimal for the interpretation of mixed samples containing DNA from more than one individual. For these reasons the HLA DQA1 and polymarker dot blot have been replaced by PCR analysis of short tandem repeats (STRs). However, the dot-blot method may reemerge for analysis of mitochondrial and SNP phenotypic marker systems as a practical hybridization assay for arrays of sequence polymorphisms.

Short Tandem Repeats

Short tandem repeats (STRs), also known as microsatellites, are a class of length polymorphisms throughout the genome that have become the mainstay of current forensic

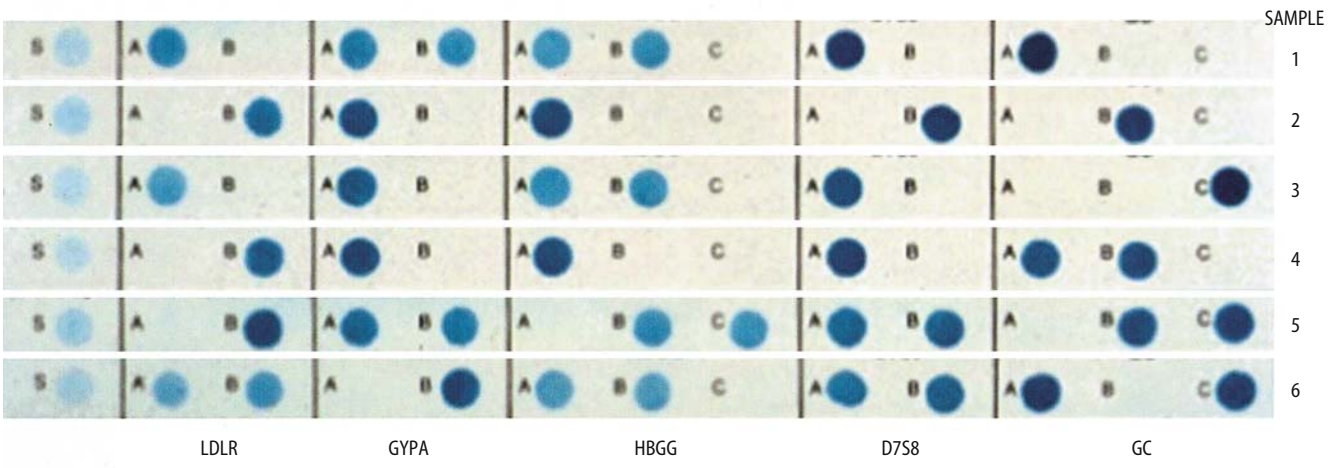


Figure 44-6. Polymarker strips from different individuals using five genetic systems detected by PCR amplification and reverse dot-blot hybridization probes. LDLR, low-density lipoprotein receptor; GYPA, glycophorin A; HBGG, hemoglobin gamma-globin chain; GC, group-specific component.

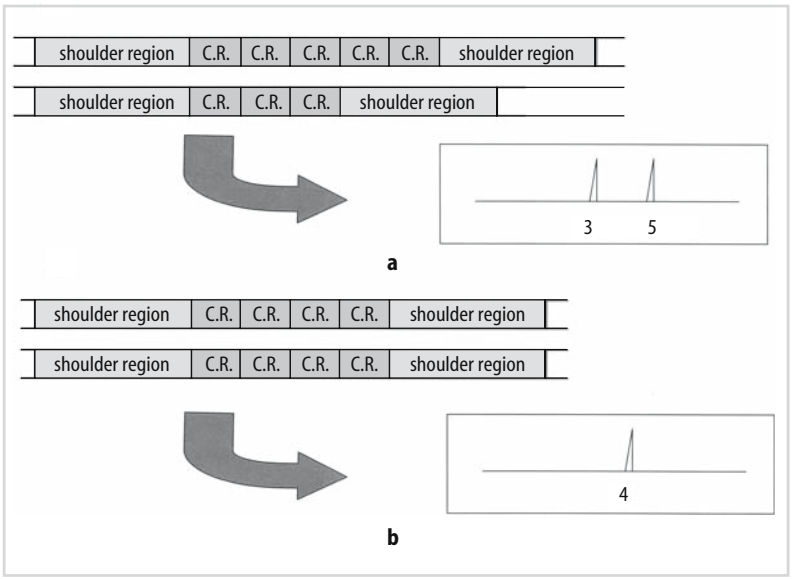


Figure 44-7. Diagram of short tandem repeat DNA segments composed of varying numbers of core repeats (CR) and accompanying electropherograms showing the corresponding allele peaks: (a) Heterozygous pattern with alleles of 3 and 5 repeats. (b) Homozygous pattern with allele of 4 repeats. The shoulder region is the flanking constant region to which PCR primers hybridize.

identity profiling around the world.[11,12] The core repeat unit in STRs is between 2 and 7 bp in length, in contrast to 8 to 80 bp in VNTRs. While the majority of VNTRs are too large to be successfully amplified by PCR, STR loci are amenable to PCR amplification (Figure 44-7). Thus, STR testing involves amplified fragment length polymorphisms with resulting amplicon sizes of approximately 100 to 350 bp.

STR analysis is faster and easier than RFLP, is amenable to automation, is more sensitive than RFLP, is successful despite DNA degradation, and yields discrete allele types. Although individual STR loci are less discriminating than individual RFLP loci, several STR systems can be multiplexed to achieve discriminatory powers similar to or greater than those of RFLP testing (typically $>10^{-12}$) (Table 44-2). As a result, PCR-based STR testing has become the dominant method of routine forensic DNA typing today and will likely continue to be so for the foreseeable future. However, other systems, described below, using mitochondrial DNA, SNPs, and Alu repeats, under certain circumstances may be performed as adjuncts or when STR tests are not successful.

In the late 1980s, Tom Caskey, then at Baylor College of Medicine in Houston, Texas, was funded by the NIJ to develop STR systems for forensic applications.[13] Subsequently, in 1991, STRs were first used by the Armed Forces DNA Identification Laboratory, through a subcontract with Cellmark Diagnostics, to identify service members who died in the first Persian Gulf War. However, it was Peter Gill at the United Kingdom's FSS who first applied analysis of STR loci (using in-house systems) to routine criminal casework in the mid-1990s.[14]

Recognizing the importance of cross-jurisdictional matches, the FBI convened a panel of forensic scientists in 1998 to choose a panel of STR loci for use in their National DNA Index System (NDIS). Thirteen STR loci, all containing tetranucleotide repeats (i.e., four nucleotides in each core repeat), were chosen: D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, CSF1PO, FGA, THO1, TPOX, and vWA (Table 44-2).[15] These 13 core loci, which have become the standard loci for forensic casework in much of the world, are referred to as "CODIS" loci, after the Combined DNA Index System (CODIS) software

Table 44-2. Comparison of Polymorphism and Discrimination of RFLP and STR Systems

Method	Locus	Bins/Alleles	Heterozygosity	Match Probability
RFLPs	D1S7	28	0.945	0.0058
	D2S44	26	0.926	0.0103
	D4S139	19	0.899	0.0184
	D10S28	24	0.943	0.0063
	D14S13	30	0.899	0.0172
	D17S79	19	0.799	0.07
Average			0.902	8.26×10^{-12}
STRs	CSF1PO	11	0.734	0.112
	TPOX	7	0.621	0.195
	THO1	7	0.783	0.081
	vWA	10	0.811	0.062
	D16S539	8	0.767	0.089
	D7S820	11	0.806	0.065
	D13S317	8	0.771	0.085
	D5S818	10	0.682	0.158
	FGA	19	0.86	0.036
	D3S1358	10	0.795	0.075
	D8S1179	10	0.78	0.067
	D18S51	15	0.876	0.028
	D2S11	20	0.853	0.039
Average			0.781	1.7×10^{-15}

Genetic systems are listed along with the number of bins used or alleles seen for the system, and the system's heterozygosity and average match probability. Bins are fragment-size ranges over which fragments are treated as a single allele. Heterozygosity is a measure of discrimination based on the distribution of alleles in a population. The average match probability is a statement of the likelihood of a random match in a population.

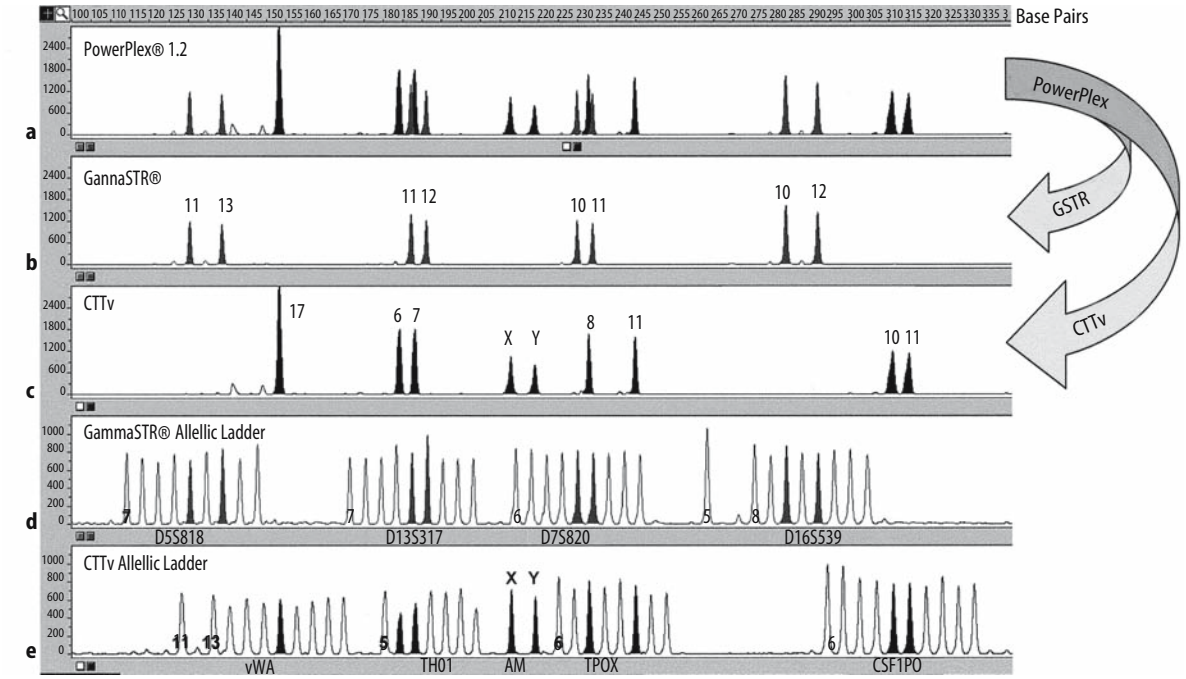


Figure 44-8. Electropherogram of multiplexed fluorescently-labelled PCR amplicons from STR loci. DNA from one individual was amplified with an early commercial STR (PowerPlex 1.2 [Promega]) multiplex primer set, which allows for the simultaneous amplification and two color detection of eight polymorphic STR loci plus the sex marker amelogenin (AM). Primers for four of the loci, CSF1PO, TPOX, TH01, and vWA ("CTTv"), are labeled with tetramethylrhodamine and can be visualized as the blue peaks (A and B). The remaining four primers sets that amplify the D16S539, D7S820, D13S317, and D5S818 loci, also known as "GammaSTR," are labeled with fluorescein, which can be visualized as the black peaks (A and C). The use of the two colors permits the discrimination of the alleles at each locus, even though the product sizes overlap. For allele assignment, one color is visualized at a time (B and C). The alleles can be identified by comparison to allelic ladders, shown in D and E, composed of the common population alleles for each locus. The repeat numbers of the initial alleles for each locus are shown. While this multiplex is no longer commonly used for forensic applications, it is still used for other identity-testing applications, as described in subsequent chapters in this section.

into which DNA profiles are entered. "SGM plus" is a ten locus system used in the United Kingdom, and in Germany an eight locus system is used. Databases are maintained of the STR alleles of convicted felons, casework profiles, and missing persons. The commonality of genetic systems (i.e., STR loci) used in forensic casework enables computer searches for matches across jurisdictions.

The selection of STRs with tetranucleotide repeats for the CODIS loci was based on the need to minimize stutter and optimize amplification. Stutter peaks are produced when the DNA polymerase slips during amplification, resulting in PCR products that have fewer and more repeat units than the starting template. Substantial stutter can confuse the interpretation of alleles from amplified STRs and certainly is problematic in presentations to a jury. The amount of stutter products is inversely related to the number of nucleotides in the core repeat, with di- and trinucleotide repeats having the most stutter and tetranucleotide repeats having less stutter, with the major stutter product having most commonly one repeat unit less than the template. Pentanucleotide repeat STRs have even less stutter than the tetranucleotide repeat STRs; however, few pentanucleotide loci had been characterized at the time the CODIS loci were selected. Although larger repeat units have less stutter, a disadvantage of these larger repeat regions is greater susceptibility to DNA degradation. Furthermore, preferential amplification of shorter fragments to the near exclusion of larger fragments may be problematic.

Commercial kits for amplification of the CODIS core STR loci are available in various combinations of multiplex primer sets from two companies: Promega Corporation (Madison, WI) and Applied Biosystems (ABI, Foster City, CA). Primer sets are available with either a fluorescein label to detect the presence of amplified STR products with fluorescence-based detection instrumentats (Figure 44-8) or unlabeled for detection by silver staining following separation by denaturing polyacrylamide gel electrophoresis (PAGE). While a variety of multiplex and monoplex primer set combinations are available from both manufacturers for identity testing applications, forensic identity testing requires amplification of the CODIS loci. To this end, ABI offers the AmpFLSTR Profiler (with loci D3S1358, vWA, FGA, amelogenin, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820), which can be combined with the AmpFLSTR Cofiler (with loci D3S1358, D16S539, amelogenin, THO1, TPOX, CSF1PO, D7S820) to amplify all 13 loci. Alternatively, Identifiler amplifies all 13 core loci as well as amelogenin, D2S1338, and D19S433 in a single reaction. Promega offers the PowerPlex 16, which includes the 13 core loci, plus amelogenin and two pentanucleotide repeat loci, Penta D and Penta E. The commercial availability of these reagents for amplification of STR loci simplifies the optimization and validation of amplification to allow the ready application of these markers for additional identity-testing purposes as described in chapters 45 to 47.

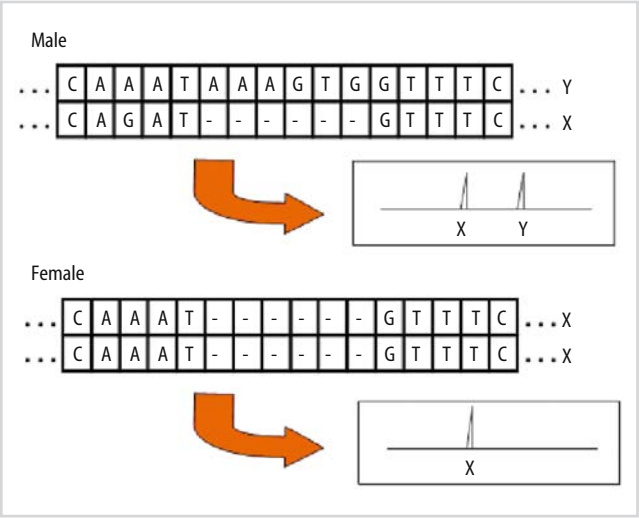


Figure 44-9. The amelogenin locus is 6 bp longer on the Y chromosome than on the X chromosome. Thus, a male will have two peaks and a female will have only one peak in the electropherogram.

A significant resource for information on STRs, including variant alleles, is available on the Web site of the National Institute of Standards and Technology (NIST): http://www.cstl.nist.gov/biotech/strbase.[16]

Amelogenin

The amelogenin locus is included in current commercial STR amplification kits as a gender marker. The amelogenin gene is present on both the X and Y chromosomes, but is 6 bp longer in the Y chromosome than in the X chromosome. Therefore, after amplification of the gene, males will manifest two peaks whereas females will manifest a single peak of twice the intensity (Figure 44-9). The amelogenin marker system is robust and reliable. However, in rare individuals, gender discrepancies have been noted.

Identity Testing by Mitochondrial DNA

Mitochondrial DNA (mtDNA) is a useful alternative to nuclear DNA systems for identification purposes because it is present at high copy number in each cell, has a different inheritance pattern, and is present in some tissues lacking a significant nuclear DNA presence.[17] Mitochondria may be derived from an ancient symbiotic parasite, which may explain the presence of DNA within the organelles of the cytoplasm. Each of the tens to thousands of mitochondria has one or more DNA particles. Thus, there are generally 500 to 2,000 copies of mtDNA per cell, compared to one set of diploid chromosomal, nuclear DNA. When a DNA specimen is substantially degraded, the high copy number mtDNA is more likely to yield a result than nuclear DNA. This is particularly important in the identification of older skeletal remains, such as human remains from past wars (i.e., the Vietnam War, Korean War, and World War II).

MtDNA has an inheritance pattern different from that of nuclear DNA, being maternally transmitted without recombination (Figure 44-10). The small numbers of mtDNA particles that pass into the fertilized egg from the sperm are destroyed by the enzyme ubiquitin, leaving the maternal mtDNA intact. Thus, all mitochondria are derived from the egg from the mother. MtDNA, unlike the paired chromosomal DNA, does not undergo meiosis and does not participate in genetic recombination events. MtDNA remains unchanged through generations, until the occurrence of a mutational event. The rate of mutations in mtDNA is 10 to 20 times greater than that in nuclear DNA, presumably due to the exposure of mtDNA to oxygen free radicals. Hence, analysis of mtDNA may be important when only a distant maternal relative is available for a reference specimen. In contrast, analysis of identity by nuclear DNA requires the availability of DNA from multiple close kindred.

A major advantage of mtDNA is that it may be present in tissues without nuclear DNA, such as hair and fingernails. Shed hairs (telogen hairs), which do not have roots, generally contain mtDNA but not nuclear DNA. On average, an individual loses 200 hairs per day, and thus it is not surprising that shed hairs constitute a significant trace evidential specimen.

Human mtDNA is a circular, double-stranded DNA containing 16,569 bp (Figure 44-11). For identity testing, only sequence-based polymorphisms are of practical utility in mtDNA, since little repetitive DNA is present. The mtDNA sequence obtained from a sample is compared to the Revised Cambridge Reference Sequence.[18] (Anderson first sequenced the entire mitochondrial genome in 1981 but a later updated consensus sequence is used.) The sequence polymorphisms are concentrated in two hypervariable regions, which are located in the noncoding displacement loop (D-loop), or control region (15971–579). This 1200 bp control region spans the transcription origin

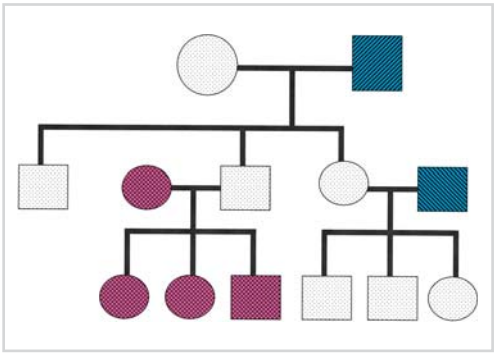


Figure 44-10. Mitochondrial DNA (mtDNA) is maternally inherited without recombination. The mtDNA sequence will be exactly the same in all children of the grandmother and all children of her daughters. However, all grandchildren will inherit 25% of the nuclear DNA of their grandmother. In this pedigree diagram, all females (circles) and males (squares) filled with the same pattern have maternally inherited the same mtDNA.

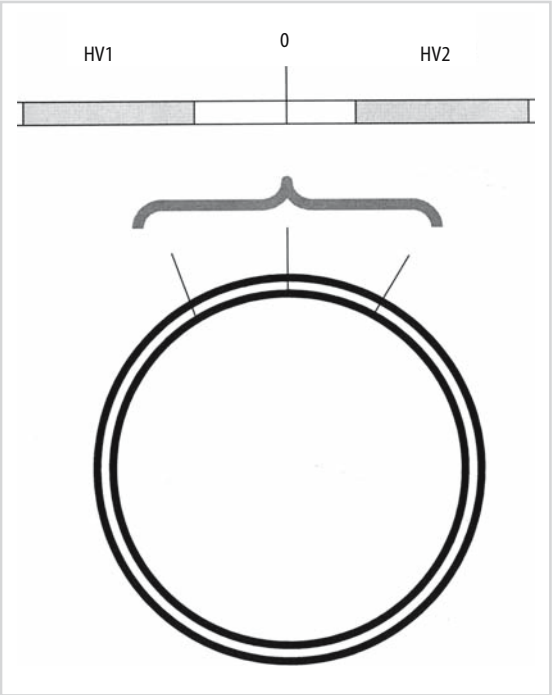


Figure 44-11. Mitochondrial DNA is a circular length of DNA of 16,569 bp. The "control region," or "displacement loop," is a segment of approximately 1,200 bp that contains the origin and two hypervariable segments (HV1 and HV2) that are used for forensic purposes.

(base pair = 0), so that hypervariable region I (HVI) is numbered 16024 to 16365 and hypervariable region II (HVII) is numbered 73 to 340 (Figure 44-11). Polymorphisms in the coding region exist and are useful but are too disparate to be practically interrogated by sequencing.

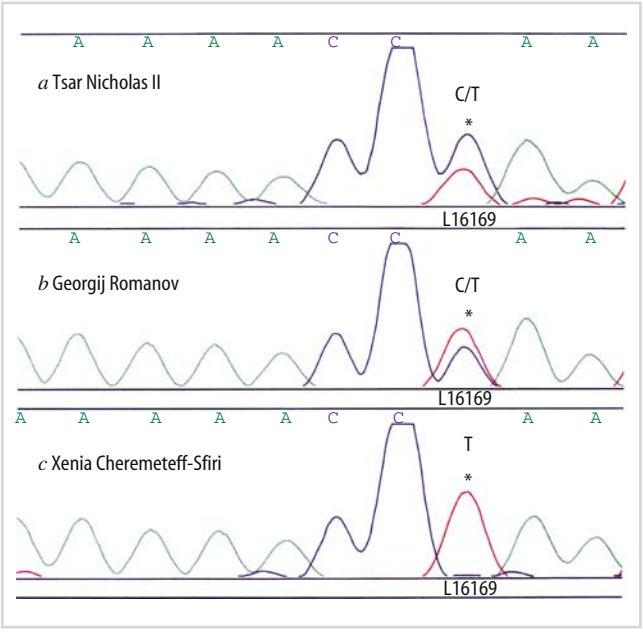


Figure 44-12. The first description of mtDNA heteroplasmy in humans was in the case of Czar Nicholas Romanov II, the last imperial Russian monarch. DNA sequence analysis shows that the czar (a) shares the heteroplasmy (C/T marked with an asterisk) at position 16,169 with his brother Georgij (b) but not with his distant relative Xenia Cheremeteff-Sfiri (c) (five generations removed) who is homozygous for the T allele.

Heteroplasmy refers to the presence of more than one mtDNA sequence in the same organism. Although heteroplasmy was well known in plants and nonhuman animals, it was first described during the identification of Czar Nicholas Romanov II (Figure 44-12).[19] A low level of heteroplasmy may, in fact, be present in all individuals. To be detected using standard DNA sequencing of mtDNA, the level of heteroplasmy must be above approximately 30% of the mtDNA sequence. Otherwise the presence of sequence variations as secondary peaks of lower peak height will not be distinguishable from the background noise. Heteroplasmy is not uniform through the body and appears to be somewhat tissue specific. In addition, heteroplasmy may be rapidly lost (reversion to homoplasmic state) in family lineages because of the bottleneck phenomenon that occurs during reproduction from a single fertilized egg. Heteroplasmy complicates forensic analysis because, for example, two hairs cannot be assumed to be from different individuals if they differ by a single mismatch.

Forensic mtDNA testing is performed by only a few specialized laboratories because the standard analytical method is DNA sequencing, which is expensive, labor-intensive, relatively slow, and highly sensitive to contamination. The exquisite sensitivity of the testing mandates special laboratory facilities and procedures. Also, interpretation is less straightforward than for routine STR results. The FBI has recently designated four state laboratories as regional mtDNA laboratories.

A dot-blot technology has been commercialized (Roche Molecular Systems, Alameda, CA) that enables mtDNA testing for screening purposes by most laboratories. This linear array assay uses 27 sequence-specific oligonucleotide probes that capture most, but not all, polymorphic information of the HVI/HVII control region.[20,21]

Additional Genetic Systems for Forensic Analysis

While STRs are not likely to be replaced for fundamental routine casework anytime soon, other genetic systems, described below, are being developed and slowly introduced into use as adjuncts to STR forensic analysis or in place of STRs in certain situations.

Male-Specific DNA Typing with Y Chromosome Markers

A new class of genetic system, the Y chromosome markers, is being introduced into forensic laboratories. Currently, these markers are STR systems located in the male-specific region (MSR) of the Y chromosome, but SNP systems from the same region also are under development. Y chromosome markers, other than amelogenin, are not gender markers but rather male-specific identity systems that permit typing of spermatozoal DNA in vaginal swabs from rape cases without concern for contaminating DNA from

female epithelial and white cells. Currently, to separate the male cells from the female cells, forensic laboratories perform differential extractions. The female cells are gently lysed, followed by treatment with dithiothreitol (DTT) to break the disulfide bonds in the sperm capsules and release the male DNA fraction. The technique works well for RFLP testing, but more-sensitive PCR tests often have contaminant female DNA in the male fraction. The amplification of Y chromosome–specific markers eliminates the need for separation of male and female cells resulting in unequivocal identification despite female admixture. As with any extraction procedure, some target DNA may be lost from the differential extraction procedure, and thus Y chromosome markers may result in successful DNA typing when conventional STRs after differential extraction fail.

Y chromosome markers are useful precisely because they are inherited paternally. The absence of recombination means that the exact same Y chromosome DNA allele types will be present in distant paternal relatives of an individual. For example, Y chromosome markers were used in determining the paternity of US President Thomas Jefferson among his distant descendants.

To develop Y chromosome–specific markers, attention was initially focused on STR loci located on the Y chromosome (Y-STRs). Since markers on the Y chromosome are inherited together, the alleles at several loci form a haplotype, that is, a group of genetic markers in linkage disequilibrium. European investigators established a 9-locus minimal haplotype (DYS19, DYS385a/b, DYS 389 I/II, DYS390, DYS391, DYS392, DYS393), which was endorsed by the International Society of Forensic Genetics, and an extended Y-locus haplotype with the addition of the YCAIIa/b loci. The Scientific Working Group on DNA Analysis Methods in the United States endorsed a similar 11-locus U.S. haplotype composed of the European minimal haplotype loci plus DYS348 and DYS349. Amplification kits for Y-STRs are available commercially through ABI and Promega. The amplified Y-STR amplicons can be analyzed using the same instruments used for other forensic STR systems. A 20 Y-STR locus multiplexed system is under development at the NIST. More information on Y-STR haplotypes is available from the Y-STR Haplotype Reference Database Web site (http://www.ystr.org).[22]

SNPs for Typing of Highly Degraded DNA

By far the most common polymorphisms in the human genome are simple isolated single base differences, SNPs. Very small DNA fragments can be interrogated for SNP alleles; thus, SNP genotyping can be applied for forensic identification despite extreme DNA degradation, since small DNA targets are more readily amplified than larger targets in fragmented DNA. Most SNPs are biallelic; that is, there are only two alleles, a dominant allele and a non-dominant allele, despite the fact that there are four possible nucleotide bases. Therefore, a large set of SNPs must be used to obtain significant discriminatory values. As a result of the technical demands of interrogating a large number of SNP sites, SNPs are not used commonly for forensic analysis at this time but are being developed by a number of laboratories. The identification of human remains recovered from the World Trade Center disaster is one scenario in which SNPs showed great advantage (OrchidGeneScreen, Dallas, TX), since the DNA was severely degraded in a large percentage of the specimens. Analysis of SNPs is particularly amenable to automation and analysis with chip technologies using hybridization, polymerase extension, or ligation reaction assays.

Ethnic and Phenotypic Markers: Alu Repeats, LINEs, SNPs

Large, high-copy-number genomic elements can be used as markers of ethnic origin.[23] A significant portion of the human genome consists of self-replicating sequences, known as "mobile elements," that replicate and insert randomly throughout the genome. The Alu family of mobile elements comprises 5% of the human genome. Once inserted, these elements are inherited in a Mendelian fashion. Mutations occur with sufficient frequency that they can be used as markers of descent. Since mutations collect over time, the age of a given insertion generally can be dated. Also, the insertion sequence tends to be lost with time. Similarly, the L1 family of Long Interspersed Elements (LINES) can be used to trace evolutionary ancestry. In combination with other genetic systems (microsatellites, SNPs, Y-STRs, and mtDNA), Alu and LINES markers will provide some statistical inference about human evolution as well as race and ethnicity that may be helpful to investigators. Certainly, population admixture is making this effort more risky. Others are pursuing SNP genetic determinants of eye, hair, and skin characteristics, a genetic version of "driver's license data," also in an effort to assist investigators. Although some progress has already been achieved, much work still is needed.

Instruments for DNA Typing

Heavy forensic caseloads demand that the forensics community use methods of analysis with high throughput. All current instrumentation for analysis of genetic markers of identity employ electrophoretic separation followed by a detection method with laser-induced fluorescence. This method of detection is sensitive and permits the incorporation of fluorescent labels with different emission wavelengths into the PCR products, thus permitting high-order multiplexing of different markers with overlapping product sizes in the same amplification reaction (Figure 44-8).

Capillary electrophoresis (CE) instruments size fractionate and detect fluorescently labeled DNA fragments

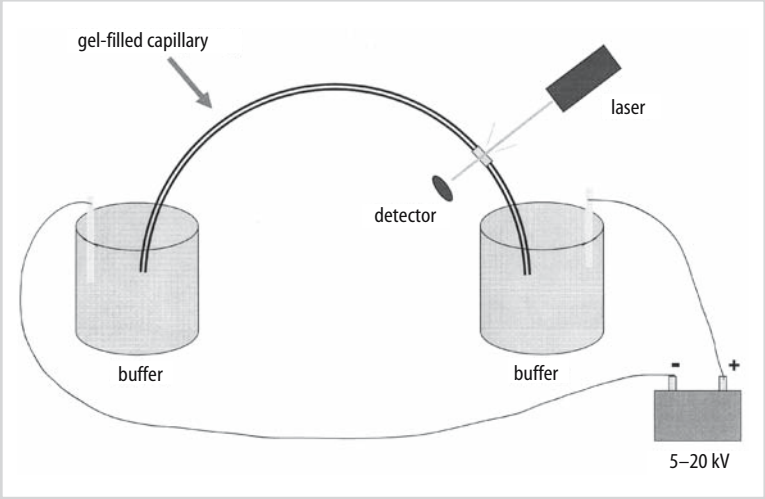


Figure 44-13. This simplified cartoon shows how capillary electrophoresis separates DNA fragments by size as an applied voltage results in the movement of electronegatively charged DNA fragments through a polymer gel-filled capillary tube toward a positive charge. As fluorescently labeled fragments pass a window through which a laser beam is directed, the fluorophores are excited and a detector captures the emitted signal. The emitted signal is then used, within a linear range, to detect the quantity of DNA fragments of each size.

using a polymer gel-filled capillary tube (Figure 44-13). A denatured DNA sample is "injected" into one end of the capillary by application of a voltage that results in movement of the negatively charged DNA toward the cathode. These CE instruments have replaced slab gel electrophoresis systems because there is no need to clean glass plates and pour the gels, and because they have inherently faster run times and use smaller (hence less expensive) sample volumes but achieve similar results. The CE capillary, with a typical internal diameter of 75 µm, is filled with a linear polyacrylamide or similar gel matrix.

Typical casework calls for on-demand instrumentation that can handle relatively few specimens but with fast run times. High-throughput CE instruments are used as batch instruments for DNA data-banking operations. The ABI Prism 3100 Genetic Analyzer is available as a 4- or 16-parallel-capillary instrument, and the ABI Prism 3700 DNA Analyzer is a 48- or 96-capillary instrument. Amersham Biosciences sells the Megabace series of instruments, which can be configured with 48, 96, or 384 capillaries.

New technologies for use in forensics are being developed, in large part due to funding efforts by the NIJ.[24] Network Biosystems is producing a capillary array fabricated by etching microchannels into glass chips, resulting in an ultrafast CE array system that has run times of seconds and produces sharper bands than conventional CE instruments. Nanogen (San Diego, CA) has produced chip technology that can detect STR and SNP systems. Integrated microsystems are being developed by Richard Mathies at the University of California at Berkeley, Jerome Ferrance at the University of Virginia, and Frederic Zenhausern at the Arizona State University Biodesign Institute. In each case, the potential for eventually developing a field unit could be contemplated. Testing at the crime scene may yield an immediate match in a DNA database or immediately exclude a set of available suspects.

Interpretation of Results

Defense attorneys faced with evidence that points to their client, with odds of one trillion to one that the perpetrator could be another person, began by directly attacking this new scientific evidence. The early days of the challenge to DNA evidence in court are sometimes referred to as the "DNA wars." The attacks were centered primarily on the issue of statistical interpretation.

Due to measurement imprecision, RFLP testing resulted in continuous allele distributions, in which a 6.72 kb band might or might not be the same as a 6.73 kb band. To address this issue, the community adopted conservative statistical measures, including the use of sizing bins (19–30 for the six commonly used RFLP systems). Nonetheless, the defense challenge was striking. Today, STR systems yield discrete alleles, in which a heterozygous 7-repeat and 8-repeat allele at the TH01 locus in an individual (TH01 7,8) is analogous to an O blood type. This is because the measurement imprecision using high-resolution acrylamide gels (generally about one tenth of a base) is well within the 4-base core repeat unit that defines the allele. Therefore, with the use of STRs, binning and match criteria are no longer relevant for statistical interpretation of DNA identity testing. Microvariants, due to incomplete or variant core repeats in some individuals, must be recognized. In addition, certain pitfalls must be considered in analyzing forensic DNA profiling results. For example, if samples are partially degraded, larger alleles may drop out, but smaller alleles may remain recoverable (Figure 44-14). Mixtures of DNA from two or more individuals also can be problematic to interpret.

Identity testing reports generally state the frequency with which a given genotype of all the alleles tested will be found in a white, black, or Hispanic population. The calcu-

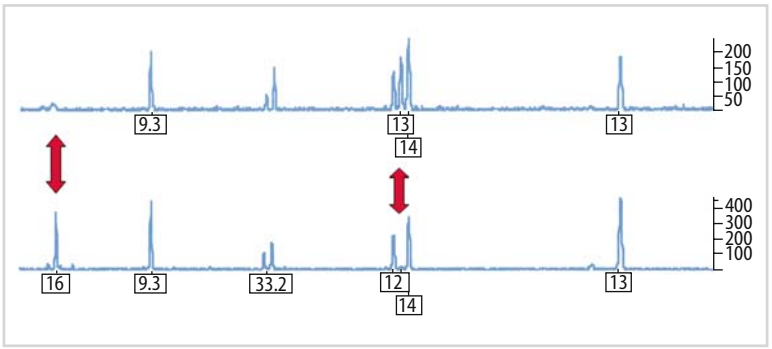


Figure 44-14. Allelic drop out from sample degradation or primer site polymorphism is one of the few interpretative pitfalls in the analysis of STRs. Allelic drop in can occur from contamination. This figure demonstrates a drop out of allele 16 and a drop in of allele 13 in the upper tracing compared to the lower tracing.

lation is achieved by multiplying the allele frequencies of individual alleles in the given ethnic population.[25] In this way, very high powers of discrimination are achieved. Discriminatory power should not be confused with accuracy, since ABO blood group typing is accurate but has low discriminatory power (perhaps one in three).

Certain features of genetic systems were cause for initial attack, including independence of genetic loci, whether or not a polymorphism was in Hardy-Weinberg equilibrium, and whether substructure was present, indicating subgroup predilection. Current forensic genetic systems show greater intergroup than intragroup diversity, allaying many of the concerns over appropriate reference population allelic frequency data.[26] A National Research Council (NRC) report, NRC I,[27] was issued in part to address these statistical issues. The NRC I report introduced the "ceiling principle," in which a minimum allele frequency was set to prevent an overly high estimate of the discriminatory value, but this itself generated considerable controversy. This led to NRC II report,[28] which withdrew the ceiling principle and has, in fact, largely settled most statistical issues.

Since mtDNA is considered a single genetic system and the polymorphisms are linked, the individual frequencies cannot be multiplied together to generate large discriminatory values. Instead, mtDNA statistics involve the mere counting of similar genotypes in a large database to derive a frequency statistic. Most mitotypes in the database are unique. Since the reference forensic database maintained by the FBI has more than 600 entries, it can be fairly stated that most mitotypes have a discriminatory value of greater than 1 in 600. On the other hand, 20% of the time, the DNA sequence will be one of 18 common haplotypes with a population frequency of greater than 0.5%, and the most common haplotype is present in 7% of the population.

Similar to mtDNA testing, the discriminatory power of Y chromosome markers is not as high as that of autosomal markers, because the frequency statistics cannot be multiplied together as independent variables, but rather are inherited as linked genetic systems. Thus, their frequencies are determined by the counting method. Consequently, a greater number of systems and a large population database are desirable.

Convicted Offender Databases

The BJS reports that within 3 years of release, 67.5% of prisoners are rearrested, 46.9% are reconvicted for a new crime, 25.4% are resentenced to prison for a new crime, and 51.8% are back in prison for either a new crime or violation of their parole (failing a drug test, missing an appointment, etc.). Rearrest rates include: homicide, 40.7%; rape, 46.0%; sexual assault, 41.4%; robbery, 70.2%; and burglary, 74.0%.[29] Due to this documented recidivist nature of criminals, it is logical to collect DNA from certain offenders to assist with their recapture should they perpetrate another crime. Indeed, this is not very different from collection of a fingerprint.

Some privacy advocates argue that, in fact, the information in DNA amounts to a window into more personal information about the individual than should be accessible to law enforcement. In fact, law enforcement has little interest in medical information available through DNA analysis nor the funding or expertise to collect or analyze it. Some are concerned that genetic predisposition to violence might be of interest to law enforcement, but given the multigenic nature of such a characteristic, it would appear that the best information on the future criminal behavior of an individual is his or her own criminal record, which is already in the hands of the government.

The United Kingdom has been the leader in forensic DNA data banking.[30] The United Kingdom more aggressively uses DNA testing in casework and collects DNA from a broader range of individuals, including arrestees. The UK database currently contains approximately 2.5 million profiles, representing 3.4% of the UK population. Accordingly, officials claim that they can find matches in approximately 50% of biologic specimens from crime scenes. They believe that it is more cost-effective to perform the DNA testing than to conduct traditional police canvassing of a neighborhood. Their experience has convincingly demonstrated that a relatively small group of professional criminals perpetrate most crimes.

Virginia passed the first DNA data bank law in 1989. The law initially applied to all convicted sexual offenders but was expanded in 1990 to include all felons, and expanded

again in 2003 to include persons arrested for a violent felony after a finding of probable cause by a magistrate. In June 2004, the Virginia laboratory had achieved 2,000 identifications, or "hits," after more than 216,000 samples had been collected. It took 8 years for the first 1,000 hits, but only 18 months before the second 1,000 hits were obtained. The hits assisted or solved criminal investigations of the following types: 11% homicides, 21% sex offenses, 59% property crimes, and 9% miscellaneous other (i.e., indecent exposure, aggravated assault, arson, etc.).

Policymakers from all states in the United States have concluded that maintaining a database of DNA types of convicted offenders is worthwhile and appropriate. All 50 states have required the collection of DNA samples from convicted offenders since May 2004. These databases have generally expanded from registries of sexual offenders to registries of all felons. They also have expanded to include those who commit lesser crimes, as perpetrators of lesser crimes often progress to the commission of more serious crimes. Today, nearly 2 million convicted offender profiles are in the national database, representing approximately 0.6% of the US population. Other than convicted offenders, databases also are maintained of active casework, missing persons, and relatives of missing persons. Louisiana, Texas, and Virginia have promulgated authorizing legislation for the obtaining of DNA specimens on arrest. Some commentators have argued for a database of all citizens, primarily to preclude system bias.

The DNA Identification Act of 1994 (Public Law 103-322) authorized the creation of the FBI's NDIS. DNA profiles are loaded onto CODIS software.[31] Searches can be performed locally through a Local DNA Index System (LDIS) or State DNA Index System (SDIS) and across state lines through the NDIS. Each state has a single designated state CODIS operation, SDIS, that enters DNA casework and databank information into the NDIS. A match from an NDIS search will result in the local crime laboratory of one state being put into contact with the local crime laboratory in another state to discuss details. Identifying information other than the DNA profile is not entered into the system. Furthermore, uploading of DNA profiles triggers federal regulatory requirements on the use of the DNA specimens and profiles.

Quality Assurance and Laboratory Issues

Since an individual may be executed on the basis of forensic testing, the forensics community feels a heavy burden for ensuring the accuracy of test results. Precisely because DNA testing is so decisive, it must be performed with great care. Unfortunately, some egregious lapses in quality have been exposed in forensic testing. Pressure to find someone guilty seems to be the basis for a few well-known examples of analysts that intentionally reported findings consistent with those of detectives and prosecutors. The more-common situation is due to poor laboratory resource support. The Houston Police Department (HPD) DNA unit was shut down in December 2002 and its entries into the CODIS database removed after an audit exposed deficiencies and outside testing verified that HPD testing had incorrectly resulted in a conviction. The caseload per analyst in the HPD laboratory was more than four times the national average. Such instances are not characteristic of the community, which emphasizes scientific integrity and justice. Moreover, testing generally is performed using commercial kits that have undergone quality control procedures, with additional extensive internal validation by the laboratory.

The experience of the College of American Pathologists from proficiency testing surveys is that the forensic community is able to produce remarkably high-quality data. The correlations of variation for fragment sizes from RFLP tests of the forensic science laboratories were far tighter than those of the clinical laboratory community, and today surveys often show complete agreement in the typing of STR alleles. However, comparisons between clinical and crime laboratories are unfair in that fragment sizing is of critical importance to crime laboratories but not to clinical laboratories. Furthermore, the forensics DNA typing community is small and close-knit. Early on, the FBI formed the Technical Working Group on DNA Analysis Methods (TWGDAM) to allow analysts from different laboratories to critique one another's protocols, autoradiographs, and other results. There are several other forums for technical discussions each year, including the Promega Annual Symposium on Human Identification, the American Academy of Forensic Sciences' criminalistics section, CODIS meetings, the American Society of Crime Laboratory Directors' annual meeting, and meetings of the regional forensic science organizations.

Crime laboratories are unregulated, except in the case of submission of DNA results into NDIS.[31] The DNA Identification Act of 1994 gave the FBI regulatory oversight of DNA profiles entered into the national database. The legislation called for a DNA Advisory Board (DAB) that produced recommended standards, based largely on guidelines of the TWGDAM, which were adopted with little change by the FBI director. TWGDAM has since been renamed the Scientific Working Group on DNA Analysis Methods and continues to recommend new standards to the FBI. The FBI conducts audits of laboratories to verify and enforce compliance at least with respect to profiles that are generated and uploaded into the NDIS.

The FBI/DAB standards require accreditation. The American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) and, more recently, the Forensic Quality Services (FQS) accredit laboratories. ASCLD/LAB requirements include minimal educational credits and experience, proficiency testing twice a year per analyst, and annual audits, among other requirements. All testing requires a technical and an administrative review.

The accreditation requirements and audits are rigorous. The FQS has based its accreditation process on the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) Guide 25 and Standard 17025, and ASCLD/LAB is moving toward use of these standards. The BJS, however, found that as of January 2001, only 63% of laboratories were accredited and 19% had applied for accreditation or had undergone a preaccreditation inspection.[5]

Standard reference materials from the NIST are available for RFLP profiling (SRM 2390), PCR-based profiling (SRM 2391b), Y chromosome testing (SRM 2395), and mtDNA testing. Standards require annual comparisons with NIST-traceable standard materials.

In addition to these forensic science laboratory measures, judicial scrutiny provides additional review of those cases that go to court. However, defense review and challenges vary greatly.

Legal Issues and Future Directions

The "DNA wars" are largely over. The scientific basis of forensic DNA typing was never seriously questioned, but rather vitriolic challenges were launched at laboratory procedures and statistical interpretation. The admissibility of DNA evidence was not challenged in the 1995 O.J. Simpson case despite the presence of a well-funded and experienced defense team; instead, merely the "weight" of the evidence was challenged, on the theory that police investigators had intentionally planted Mr. Simpson's blood. The most common challenges today are to sample collection, preservation of the evidence, chain of custody, documentation, and validation studies.

New genetic testing systems and technologies will undergo renewed judicial scrutiny, with particular scrutiny given to LCN DNA testing and ethnic determinations. But such is the nature of the justice system, and we are better off for the scrutiny despite the difficulties that are endured in the process. DNA evidence is too powerful not to ensure its accuracy.

Postconviction DNA testing is the use of DNA tests for the purpose of exoneration after an individual has been convicted. Between 100 and 200 individuals, many on death row, have been released from prison by DNA testing.[32,33] Barry Scheck's Innocence Project has been the leader in this movement, though many other subsequent programs have been established around the country.[34] Since our judicial appellate system has been based on showing a lack of fairness at trial rather than a showing of innocence, new legislation has been promulgated in many states to specifically authorize the funding and use of DNA tests to exculpate the falsely convicted. As could be expected, many thousands of inmates claim to be falsely convicted. Forensic DNA testing will continue to increase in use and scope, and legal challenges will continue to evolve.

References

1. Jeffreys AJ, Wilson V, Thein SL. Hypervariable "minisatellite" regions in human DNA. *Nature.* 1985;314:67–73.
2. Jeffreys AJ, Wilson V, Thein SL. Individual specific "fingerprints" of human DNA. *Nature.* 1985;316:76–70.
3. Wambaugh J. *The Blooding*. New York: Bantam Books; 1989.
4. Weedn VW, Roby RK. Forensic DNA testing. *Arch Pathol Lab Med.* 1993;117:486–491.
5. Steadman GW. Survey of DNA Crime Laboratories, 2001. NCJ 191191. Bureau of Justice Statistics. January 2002. Available at: http://www.ojp.usdoj.gov/bjs/abstract/sdnacl01.htm.
6. The President's DNA Initiative: Advancing Justice Through DNA Technology. Available at http://www.dna.gov.
7. van Oorschot RAH, Jones MK. DNA fingerprints from fingerprints. *Nature.* 1997;387:767.
8. Weedn VW. Postmortem identification of remains. *Clin Lab Med.* 1998;18:115–137.
9. Weedn VW. Forensic DNA tests. *Clin Lab Med.* 1996;16:187–196.
10. Rudin N, Inman K. *An Introduction to Forensic DNA Analysis*. 2nd ed. Boca Raton, FL: CRC Press; 2001.
11. Butler JM. *Forensic DNA Typing: Biology and Technology Behind STR Markers*. London: Academic Press; 2000.
12. Budowle B, Moretti T, Smith J. *DNA Typing Protocols: Molecular Biology and Forensic Analysis*. Natick, MA: Eaton Publishing; 2000.
13. Puers C, Hammond HA, Jin L, Caskey T, Schumm JW. Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTHO1 $[AATG]_n$ and reassignment of alleles in population analysis by using a locus-specific ladder. *Am J Hum Genet.* 1993;53:953–958.
14. Lygo JE, Johnson PE, Holdaway DJ, et al. The validation of short tandem repeat (STR) loci for use in forensic casework. *Int J Legal Med.* 1994;107:77–89.
15. Weedn VW, Rogers GS, Henry BE. DNA testing in the forensic laboratory. *Lab Med.* 1998;29:484–489.
16. Butler JM, Reeder DJ. STR DNA [database online] 2006. Available at: http://www.cstl.nist.gov/biotech/strbase.
17. Holland MM, Parsons TJ. Mitochondrial DNA sequence analysis: validation and use for forensic casework. *Forensic Sci Rev.* 1999;11:21–50.
18. Wallace DC, Lott MT. MITOMAP: A Human Mitochondrial Genome Database [database online]. 2003. Available at: http://www.mitomap.org (Revised Cambridge Reference Sequence available at: http://www.mitomap.org/mitoseq.html).
19. Ivanov PL, Wadhams MJ, Roby RK, Holland MM, Weedn VW, Parsons TJ. Mitochondrial DNA sequence heteroplasmy in the Grand Duke of Russia Georgij Romanov establishes the authenticity of Tsar Nicholas II. *Nat Genet.* 1996;12:417–420.
20. Roche Applied Science Linear Array Mitochondrial DNA HVI/HVII Region-Sequence Typing Kit. Available at: http://www.roche-applied-science.com/announcements/mito_dna.pdf.
21. Gabriel MN, Calloway CD, Reynolds RL, Primorac D. Identification of human remains by immobilized sequence-specific oligonucleotide probe analysis of mtDNA hypervariable regions I and II. *Croation Med J.* 2003;44:293–298. Available at: http://www.cmg.hr/2003/44/3/12808721.pdf.
22. Willuweit S, Roewer L, Krawczak M, Kayser M, de Knijff P. Y Chromosome Haplotype Reference Database, Institute of Legal Medicine, Charite'—Univ of Medicine Berlin. Available at: http://www.ystr.org/index.html.
23. Batzer MA, Alegria-Hartman M, Bazan H, et al. Alu repeats as markers for human population genetics. In: *Proceedings of the Fourth International Symposium on Human Identification.* Madison, WI: Promega Publishing; 1993:49–57.
24. Weedn VW. DNA(e): future analytical techniques. In: Siegel J, ed-in-chief. *Encyclopedia of Forensic Sciences*. London: Harcourt, Brace and Co; 2000.
25. Evett IW, Weir B. *Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists.* Sunderland, MA: Sinauer Associates; 1998.

26. Lander ES, Budowle B. DNA fingerprinting dispute laid to rest. *Nature.* 1994;371:735–738.
27. National Research Council. *DNA Technology in Forensic Science* (NRC I). Washington, DC: National Academy Press; 1992.
28. National Research Council. *The Evaluation of Forensic DNA Evidence* (NRC II). Washington, DC: National Academy Press; 1996.
29. Lanagan PA, Levin DJ. Recidivism of Prisoners Released in 1994, Washinton, DC: U.S. Department of Justice Statistics, NCJ 193427. Available at: http://www.ojp.usdoj.gov/bjs/abstract/rpr94.htm.
30. Weedn VW, Hicks JW. The unrealized potential of DNA testing. *National Institute of Justice Journal.* 1997;234 [republication of Weedn VW, Hicks JW. The unrealized potential of DNA testing. *Research in Action*. NCJ 170596. Washington, DC: National Institute of Justice; 1998. Available at: http://www.ncjrs.org/pdffiles/170596.pdf.]
31. Federal Bureau of Investigation. CODIS Web site. Available at: http://www.fbi.gov/hq/lab/codis/index1.htm.
32. Connors E, Lundregan T, Miller N, McEwen T. *Convicted by Juries, Exonerated by Science: Case Studies in the Use of DNA Evidence to Establish Innocence After Trial*. Research report NCJ 161258. Washington, DC: National Institute of Justice; June 1996. Available at: http://www.ncjrs.org/pdffiles/dnaevid.pdf.
33. Scheck B, Neufeld P, Dwyer J. *Actual Innocence: When Justice Goes Wrong and How to Make It Right*. Signet Books, New York; 2001.
34. Benjamin Cardozo Law School. Innocence Project Web site. Available at: http://www.innocenceproject.org.

Chapter 45

Parentage and Relationship Testing

Herbert F. Polesky

Historical Overview

Methods to determine parentage and other relationships between individuals are not new. An early reference is found in Kings (3: 16–27), in which Solomon uses the threat of dividing the child to determine maternity. The application of scientific methods to this problem followed the discovery of the ABO blood group. Laws recognizing the possible exclusion of parentage by blood group testing were enacted by some states as early as 1935. Between 1940 and 1960, as additional blood group systems were defined and shown to follow Mendelian inheritance, more tests to exclude paternity were introduced. When appropriate reagents and methods were used, these red blood cell (RBC) group marker systems were reliable; however, the distribution of these markers in most populations limited the chance that a falsely accused man could be excluded. Subsequently, in the 1960s and 1970s, the discovery of polymorphic protein and red cell enzyme systems resulted in new markers that increased the power of the testing. The introduction of human leukocyte antigen (HLA)-A,B typing further expanded the possibility of excluding most falsely accused men.[1]

Prior to the 1970s, the information from genetic marker testing was largely limited to providing evidence that a relationship did not exist. As more test systems were used it was recognized that when testing failed to exclude a relationship, the probability of a relationship between the tested individuals could be reliably calculated from the genetic information obtained. An amendment to the US Social Security Law (title IV D) in 1974 required that women requesting financial aid for dependent children identify the father. This change in the law increased the importance of providing estimates that a tested individual could be related.

Reports suggesting the use of DNA markers based on restriction fragment length polymorphisms (RFLP) for the determination of parentage first appeared in 1986.[2] Over the next decade most laboratories performing relatedness testing applied DNA-based systems. The development of polymerase chain reaction (PCR) methods that easily allowed testing of a minimal sample led to a shift from testing by the classic marker systems (RBC antigens, HLA, etc.) to testing by only DNA marker systems.[3] The evolution of testing methods offered by laboratories subscribing to the College of American Pathologists (CAP) parentage proficiency testing program between 1993 and 2003 is shown in Figure 45-1. In conjunction with a doubling of the number of laboratories performing parentage testing (from ~50 to 100 laboratories), there was a sharp increase in PCR testing with a concomitant drop in all other testing methods.

Indications for Relatedness Testing

Parentage

Attempting to determine whether a man is or is not the biologic father of a child is the most frequent reason for performing relatedness testing. As indicated above, child support programs are mandated to identify fathers of children born out of wedlock. In many situations the mother is unsure of who is the father. By establishing paternity, the mother is able to obtain financial support for the child. Even when a child is born to a married couple, paternity is not a certainty. Family studies performed as part of medical trials have frequently reported that a small percentage of participants' children have extramarital genetic origins. Testing may be requested when a divorce is pending or when a father has doubts about his relationship to a child. Relatedness testing also is useful when it is suspected that babies have been mixed up in the nursery or to confirm the identity of a kidnapped child. Genetic marker tests can establish parentage of a child born to a surrogate mother or in some cases when assisted reproductive technologies, such as in vitro fertilization, are used.

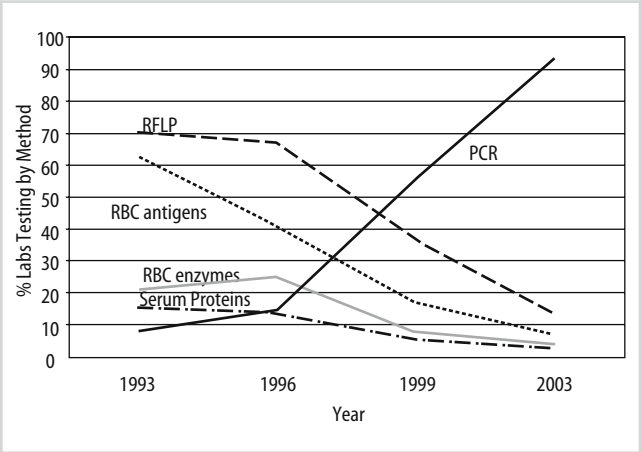


Figure 45-1. The evolution of testing methods offered by laboratories subscribing to the College of American Pathologists (CAP) parentage proficiency testing program between 1993 and 2003. There was a sharp increase in PCR testing with a concomitant drop in all other testing methods by 2003.

Adoption and Absent Parents

A recent change in laws concerning availability of adoption records has increased requests for testing to identify biologic parents. Genetic evidence can confirm whether a relationship exists. Testing may be requested when individuals who were raised separately desire to establish whether they are siblings.[4] Similarly, testing may be useful to identify a person suspected of being an absent parent. In some cases, the establishment of parentage may be important with regard to resolution of inheritance claims. Proving membership in a Native American tribe to share in gambling revenues is another reason for determining parentage.

Immigration

The current US immigration quota system gives preference to relatives of US residents applying for immigrant status. Genetic testing evidence is considered in establishing the relationship between a US resident and a person desiring to immigrate to the United States. Often these cases involve putative children or parents with incomplete documentation of their relationship.

Forensic Parentage

The usual use of genetic testing in forensic cases is to compare results from a suspect with material found at the crime scene or on a victim. This testing generally involves determining whether two samples match or do not match. In contrast, relatedness testing depends on the comparison of samples from two or more individuals. Determining the identity of victims in the World Trade Center disaster on September 11, 2001 and several aviation accidents has required comparison of the genetic information from identified remains with samples from one or more of a victim's family members. Reconstruction of possible family relationships has been useful in the identification of recovered material at grave sites.[5] Analysis of products of conception can be used to determine the paternal contribution, which may help identify the perpetrator of a sexual assault.

Sample Considerations

Samples

Relationship testing can be done using any sample from which DNA can be extracted or amplified, or both. Most laboratories routinely use carefully collected buccal swabs. The advantages of buccal swabs include the ease of collection from infants, the ability to be collected by individuals with minimal training, and ease of sample shipping with a low risk for breakage and and no need for special shipping containers.[3] Additional sample types include whole blood, cord blood, tissue samples, and dried blood spotted on filter paper. The use of specially treated filter paper permits samples to be stored for long periods of time without DNA degradation. Fetal DNA can be obtained from amniocytes, chorionic villus samples, and placental scrapings; however, it is important to rule out maternal contamination (see chapter 47) when using these samples. The isolation of fetal cells from maternal peripheral blood samples also has been reported. Mitochondrial DNA obtained from bones found at grave sites has been compared to samples from presumed maternal relatives to identify individuals missing in action. Proof that the bones found in a mass grave site were those of Czar Nicholas II and his family was based on comparison with the mitochondrial DNA of living maternal relatives[5] (see chapter 44).

Sample Collection

A testing facility must ensure that samples are identifiable and traceable from the time of collection through testing and reporting of results.[6] Samples need to be inspected and must meet acceptability criteria at all stages of testing. Collection of the samples must be done by individuals with no interest in the outcome of the testing, usually an employee or an agent of the laboratory; however, in some cases submitted samples are collected "at home." Under these circumstances it is very important to have a second person witness and verify the collection procedure.

The American Association of Blood Banks (AABB) has developed very specific requirements for the information that must accompany samples used for parentage testing,[7] including records of the name, alleged relationship, date of birth, and race or ethnic background of the individuals to

be tested. Records must include information about the place, date, and person(s) collecting the sample. At the time of sample collection, the individual to be tested should be queried about a history of hematopoietic progenitor cell transplantation at any time in the past, as well as about transfusion in the past three months. If only DNA testing is to be done, it is unlikely that transfused donor DNA will be detected. In multiply transfused patients, DNA-based blood group testing is used to determine the phenotype of the recipient when it is difficult to find compatible donor blood because of the presence of antibodies. For routine parentage testing, a photograph taken at the time of sample collection or a legible copy of a government photo ID, or both, should accompany the specimen.

Relationship Testing

General Issues

RBC antigens, polymorphic red cell enzymes, serum proteins including Gm and Km, and serologic tests for HLA-A and HLA-B markers have been very useful systems for relatedness testing, although their use is declining in favor of DNA-based identity testing[8] (Figure 45-1). The molecular methods used for testing parentage samples are identical to the methods for forensic testing described in chapter 44; therefore, this discussion is limited to method-specific requirements and quality control issues, some of which are particular to parentage testing. As is true for all molecular tests, the reliability, reproducibility, and accuracy of the method should be evaluated prior to implementation. Any new method or system used by a testing facility should be validated in several ways, for example, by performing a literature review, by performing the method in parallel with existing laboratory methods, by sample exchange with another laboratory, or by a combination of these.[7] In addition, for amplification-based methods, inclusion of negative controls to monitor for sample contamination is important. Regardless of the method used, appropriate documentation of all sample handling and testing steps is critical for parentage testing as mandated by the AABB standards.

RFLP

The interpretation of an RFLP using Southern blot analysis requires identification of DNA fragments for each locus analyzed for each individual tested. The procedure includes digestion (restriction) by a specific restriction endonuclease of the DNA isolated from the sample, electrophoretic separation of the restriction fragments on a membrane, hybridization with a labeled DNA probe, washing to remove excess nonspecifically bound probe, and measurement of the resulting labeled DNA fragments. The method should include evaluation of the completeness of the restriction enzyme digestion. Documentation in control populations of the sizes of the variable and constant fragments associated with the DNA locus used for testing is important. Inclusion of a human control of known genotype and a ladder composed of DNA fragments of known size that span the range of expected sizes for DNA fragments at the locus is required to accurately determine the allele sizes of tested individuals. In addition for parentage testing, coelectrophoresis of a mixture of DNA from the child and parent in question is recommended to determine if allelic DNA fragments are identical or a closely spaced mismatch.[7] For each RFLP system (locus/enzyme combination, i.e., D2S44/*Hae* III or D12S11/*Pst* I), the laboratory must determine the limits of resolution for two closely spaced bands in the same lane. This value, delta, is expressed as a percentage and establishes the size of the floating bin (x, the allele size ± delta) used to calculate the frequency of an allele.[7,8] Before reporting, AABB standards require that RFLP results be interpreted twice, independently.[6] The second interpreter must determine whether there is a visual match between the tested parties and whether the reported sizes agree with the sizing fragments in the DNA ladder and human control.

Short Tandem Repeats

PCR amplification of short tandem repeat (STR) markers is the method most frequently used for relationship testing. A variety of commercial kits are available, as described in chapter 44, that amplify many STR markers, including the CODIS loci used for felon database files.[9] For parentage testing, unlike for criminal forensics, there is no specific recommended panel of loci. Although the CODIS loci can be used, alternative or additional loci, or both, can be selected at the discretion of each laboratory as long as the final results achieve or exceed statistical significance (see "Reporting of Test Results," below) in almost all cases.

Determination of correct allele sizes in STR amplifications is of critical importance. A human control of known genotype, for example, a registered cell line such as K562, NIST standards, or a well-documented in-house sample, can be used to evaluate the accuracy and reproducibility of the test system. Allelic ladders consisting of DNA fragments of the size of commonly encountered alleles for each locus must be used. Allele assignments can be made using computer software; however, correct allele assignment should be validated by visual inspection of peaks. Procedures such as coelectrophoresis of the amplicons from the child and parent in question or testing of additional loci, or both, may be necessary to resolve closely spaced alleles (<1 tandem repeat apart), particularly when the result may alter the final interpretation. STR phenotypes may be assigned by a single automated reading or duplicate manual readings. Regardless of the reading method, AABB standards require that PCR results be interpreted twice, independently.[6]

Nucleotide Sequence Determination

Nucleotide sequence determination can be used for relationship testing either by a traditional sequencing method or by a sequence-specific method, such as analysis of HLA alleles by sequence-specific oligonucleotide probes (SSOP) or sequence-specific primers (SSP).[10,11] Single nucleotide polymorphism (SNP) analysis also can be applied using a variety of methods.[12] For all methods, inclusion of appropriate controls is important. Ideally, DNA sequencing should include analysis of both forward and reverse sequences. Comparison of results with a consensus sequence can be used for validation.[7] SSOP determination of HLA alleles should include a positive control probe that is complementary to a conserved sequence present in all alleles; likewise, the use of SSP requires that a positive internal amplification control of a nonpolymorphic sequence be included with each reaction. While SNP analysis can be highly automated, its use for relationship testing requires the analysis of a large number of loci to achieve statistical significance of the interpretations.

Y Chromosome and Mitochondrial DNA Markers

Sex-linked markers are not routinely used for relatedness testing; however, amplification of the amelogenin locus is helpful to verify that male and female samples have not been switched. Mitochondrial DNA, which is maternally inherited, and Y chromosome markers (Y-STR) are useful for family and population reconstructions.[13] More than ten Y-STR loci are available and can be helpful in determining the relatedness of two or more males.

Interpretation of Parentage Results

Overview

Parentage testing seeks to determine whether a particular relationship exists between a set of individuals. If the results of the testing provide sufficient evidence that the purported relationship does not exist, an exclusion is said to occur. When the results do not exclude a relationship, the situation is termed an inclusion, and a statistical analysis to calculate the probability or odds ratio that the relationship exists is performed using the testing results. Several mathematical expressions are generally used to express the inclusionary results: paternity index (PI), combined paternity index (CPI), and the probability of paternity.

Exclusion

Test results from each of multiple independent systems are compared to establish whether the tested parties are related or unrelated. Table 45-1 shows examples of test results for several situations in which the test results indicate that a relationship does not exist. When an allele is found in a child that is not in the child's mother, it must have come from the biological father. Therefore, in a standard trio paternity case with a mother, father, and child, a direct exclusion for the system tested occurs when an alleged father does not have an allele that is found in the child but is not present in the child's mother (Table 45-1). An indirect exclusion occurs when the exclusion is based on the observation that the child and the alleged father are homozygous for different alleles (Table 45-1). This type of exclusion is not conclusive, since it is possible that the child and alleged father actually share a null gene that is not detected. Therefore, before an opinion of nonpaternity is rendered based on an indirect exclusion, a search is made for additional direct or indirect exclusions. Exclusions in two independent tests are sufficient to exclude an individual as the biological father of a child, as long as the chance of mutation has been considered.

Table 45-1. Examples of Phenotypes That Exclude Parentage

Type of Exclusion	Mother	Child	Tested Man	Obligatory Allele(s)
Direct	A	AB	C	b
	AB	AC	AD	c
Two-haplotype	A	A	BC	a
	AB	AB	CD	a or b
	AB	A	BC	a
	Unknown	AB	CD	a or b
Indirect*	A	A	B	a
	AB	B	A	b
	Unknown	B	A	b

The obligatory allele is based on the required paternal genetic contribution given the mother's and the child's phenotypes. If two alleles or two groups of linked markers are defined in one individual, but neither is present in the second tested person, the exclusion is defined as a "two-haplotype" exclusion.
*Mutation must be considered if indirect exclusion is in a single system.

Since mutations can occur, albeit rarely, in many of the DNA test systems, a single inconsistency in expected inheritance is not adequate to reach a conclusion of nonparentage, particularly if the inconsistency is only one tandem repeat size different from the child's.[6] Mutation rates of commonly used loci are available, as well as statistical methods for incorporating the information into the report.[7] To avoid misinterpretation, results should be reported as phenotype rather than as assumed genotype when a test indicates a single marker at a locus. In the parentage testing field, the term "phenotype" is used to represent the measurable genetic markers in a given individual at a specific locus. Although confusing to those more familiar with the term "genotype" in this setting, the terminology for parentage testing has its origins in RBC blood typing in which a person with type A blood is said to have the A phenotype but could have the AA or AO genotype. Therefore, to avoid the assumption of homozygosity, test results are reported as phenotypes, which may or may not

reflect the genotype. While it is less common for a genetic test result not to match the actual genotype, null alleles and primer-site polymorphisms have been reported. In addition to mutation, technical errors, sample mix-ups, and reagent problems such as primer-binding site differences can lead to false exclusions. When an individual is excluded, his or her phenotype must be confirmed by an independent isolation of the DNA.[6]

Maternity is presumed when samples are submitted for testing. However, occasionally a different child is substituted for the biologic child. When testing is completed on a trio (mother, child, and alleged father), the results on the mother and child should first be evaluated to determine that at each locus there is one common allele. This comparison is one of the ways that the laboratory can validate that the correct samples were tested. On rare occasions, inconsistencies between a mother and a child are observed. Usually this is a finding of apparent reverse homozygosity due to a mutation and is not a maternal exclusion.

Inclusion

The most frequent use of relationship testing is to determine paternity. Each genetic system used in relationship testing should be well documented in the literature and must have been shown to follow Mendelian inheritance. Ideal genetic systems for relationship testing have multiple alleles, a low rate of mutation, and fairly uniform frequency distribution of alleles in the population.[8] The power of exclusion, which can be calculated for each locus based on the number of alleles in the population and their frequencies, represents the chance that an unrelated individual can be excluded by the locus. If a system has a few alleles that are common to most individuals, the chance that two unrelated individuals will have the same allele is high. If the system has multiple alleles, none of which occurs with a high frequency, then the power of exclusion is larger.

Based on Mendelian genetics, if a child has an allele that was not inherited from the mother, then it must have come from that child's biologic father. If the alleged father has that allele, then he is included among the group of men who could be the child's father. The examination of alleles at multiple genetic loci is used to statistically reduce the number of men who could be the biologic father. AABB standards provide guidelines for the statistical significance that must be achieved to report the results, as described below.

Parentage Index

The parentage index (PI) represents how many times more likely it is that the tested individual is the biologic father of the child than a randomly selected man. Thus, the PI is an odds ratio comparing the probability that the child is the product of the mother and the alleged father to the probability that the child is the product of the mother and a random man of the same racial group. At each of the multiple loci tested, the chance that the tested individual passed the obligatory allele is compared to the chance that the allele could have been contributed by a random individual.[14] A formula to calculate the PI can be derived using the laws of probability for each potential combination of alleles for a mother, child, and alleged father trio at each locus, based on allele frequencies (Table 45-2). A PI must be calculated separately for each locus tested.

Table 45-2. Phenotype Results and Formulas Used to Calculate the Parentage Index by Allele Combination

Mother	Child	Tested Man	PI
A or AB	A	A	1/a
A or AB	A	AB or AC	0.5/a
B	AB	A	1/a
B	AB	AC or AB	0.5/a
AB	AB	AB or A	1/(a + b)
AB	AB	AC	0.5/(a + b)
BC	AB	AB or AD or AC	0.5/a
BC	AB	A	1/a

A, B, C, and D represent different alleles. A is always used to represent an obligatory paternal allele; a and b are frequencies of alleles A and B, respectively, in the same racial group as the tested man (alleged father).

Since the PI calculations require allele frequencies by racial group, for each locus the frequency of all the possible alleles must be determined from the phenotypes of at least 100 unrelated individuals from the same racial group. The larger the available database, the smaller the error of the estimate. The ideal population frequency is one calculated from testing done by the specific laboratory. These data should be evaluated by comparison with published frequencies. When a laboratory has accumulated additional test data or introduces a new method, new values should be calculated and compared with existing frequencies. This is helpful in monitoring changes in the laboratory. For STR systems with discrete alleles, it is fairly easy to establish frequency tables that can be validated and compared with those of other facilities. RFLP markers occur over a continuum and are influenced by measurement variables. Thus, determining the frequency of each band requires the use of the delta value for the test system.[8] Computer programs and manual calculations must be periodically reviewed and validated.

Cases may involve individuals from racial or ethnic groups for which there is only a limited sample from the testing laboratory. In these cases, published tables may be used.[15] Table 45-3 shows the variation in PI values by racial group for several STR markers using published allele frequencies.

Combined Parentage Index

Assuming each locus tested is inherited independently, a combined parentage index (CPI) is calculated by multiplying the PI values of all the individual loci to obtain

Table 45-3. Calculation of the PI Using Allele Frequencies for Different Racial Groups

					Parentage Index		
	M	Child	Man	X/Y	Caucasian	African American	Native American
CSF1PO	9, 11	9, 11	11	1/(a + b)	2.98	4.10	3.29
D5S818	10	10, 13	13	1/a	6.46	4.64	10.62
TH01	7, 8	8	8, 9	0.5/a	4.36	2.41	9.23
TPOX	8	8	8	1/a	1.86	2.77	2.63
Combined Parentage Index					**156**	**127**	**848**
Probability of Paternity (%)					**99.36**	**99.22**	**99.88**

X/Y, odds ratio of tested man versus a random man being the father of this child (see Table 45-2); M, mother; a, frequency for the obligatory paternal allele(s) in each population; b, frequency of the child's maternal allele when the mother and child are heterozygous for the same two alleles.

a combined odds ratio, as specified by the third law of probability for completely independent events. AABB standards require that multiple independent genetic systems be tested that will provide a CPI of at least 100, which indicates that it is 100 times more likely that "this man," rather than "some random man," is the father of this child.[6] Table 45-4 is an example of a case meeting this requirement.

Residual Paternity Index

The residual paternity index is a calculation that is useful when there are inconsistencies in inheritance in some systems. This index is calculated by multiplying the PI for all systems that fail to exclude. A high residual PI suggests that the true parent is a relative of the tested individual. Since a father and son share at least one marker at each locus, and brothers have more markers in common with each other than either would with a random man, if the true father is related to the tested individual, the alleged father is less likely to be excluded in most systems. An avuncular index is calculated when it is suspected that the true father is the brother of the tested man.[16] A high residual PI may occur if the single inconsistency is due to a mutation or if there has been a laboratory error in testing or recording results of a single system. If the inconsistency is considered to be a mutation, the PI for the system can be calculated using a value for the mutation frequency, if known.[17] In these cases the CPI often exceeds 100 even though the PI for the system with the mutation is very small.

Probability of Parentage

If the individual in question is included as the parent by the genetic testing, this information is combined with the nongenetic evidence to calculate the probability of parentage (W) using Bayes' theorem:

Table 45-4. Detailed STR Parentage Test Results for a Caucasian Trio

	STR Test Results			Calculations		
Locus	Mother	Child	Tested Man	Frequency of OG(s)	Formula X/Y	Locus PI
CSF1PO	10, 11	10, 12	10, 12	12 = 0.33	0.5/0.33	1.52
D3S1358	15	15	15	15 = 0.26	1/0.26	3.85
D5S818	10, 11	10, 12	11, 12	12 = 0.35	0.5/0.35	1.43
D7S820	8, 11	8, 11	8, 11	8 = 0.16 11 = 0.18	1/(0.16 + 0.18)	2.94
D8S1179	12, 13	13, 14	14	14 = 0.21	1/0.21	4.76
D21S11	28	28, 30	29, 30	30 = 0.25	0.5/0.25	2
FGA	22, 24	22, 23	23	23 = 0.15	1/0.15	6.67
TH01	9	7, 9	7	7 = 0.22	1/0.22	4.55
vwa31/A	18, 19	18, 19	17, 19	18 = 0.23 19 = 0.08	0.5/(0.23 + 0.08)	1.61
				Combined Parentage Index: 11,444		
				Probability of Paternity: 99.99%		

OG(s), obligatory paternal allele(s). Frequency of OG(s) refers to the allele frequency of the given OG allele in the Caucasian population. X/Y, odds ratio of tested man versus a random man being the father of this child. Formula X/Y can be found in Table 45-2. Locus PI is calculated for each locus, and then the individual locus PIs are multiplied together to obtain the combined Parentage Index (PI). The probability of paternity is calculated using Bayes' theorem and a neutral prior probability of 0.5.

$$W = \frac{(p)(\text{CPI})}{(p)(\text{CPI}) + (1-p)}$$
$$W * 100 = \%$$

where p is the prior probability of nongenetic events and CPI is the combined parentage index.[18] The nongenetic information is referred to as the prior probability and includes information about access at the appropriate time, fertility, and so on. Since the nongenetic information is unknown to the testing facility, it is general practice to assign a neutral value of 0.5 (50%) to perform this calculation. Alternatively, the probability of parentage can be calculated with prior probabilities of 0.1, 0.5, and 0.9 to give a range of probabilities. The number calculated, referred to as the posterior probability, is expressed as a percentage. If the CPI is 100, then the probability of parentage is 99%. The higher the value of the CPI, the closer the probability of parentage is to 100%. If the genetic evidence indicates that no relationship exists (multiple exclusions), then the probability of parentage is reported as zero.

Random Man Not Excluded

Random man not excluded (RMNE) is a calculation that estimates the fraction of the population that would not be excluded by a test system based on the phenotypes of the mother and child.[19] The results of the tested man are not considered in determining this value. The RMNE value for several systems used as a battery of tests gives a power of exclusion for a random man that is compared to the mother-child pair. RMNE adds little or no additional information to the PI and is not routinely reported by most facilities.

Reconstruction Cases

In many circumstances it is possible to estimate the likelihood of a relationship when only some of the persons in the family are available for testing.

Motherless Cases

If only the man and child are tested, a PI can be calculated using formulas that compare the chance that the tested man will transmit the shared allele(s) with the chance that the allele(s) are from unknown parents[20] (Table 45-5).

Table 45-5. Phenotype Results and Formulas Used to Calculate the PI by Allele Combination in Motherless Cases

Child	Tested Man	PI
A	A	**1/a**
A	AB	**0.5/a**
AB	A	**0.5/a**
AB	AB	**(a + b)/4ab**
AB	AC	**0.25/a**

A, B, and C represent different alleles; a and b are frequencies of alleles A and B, respectively.

Fatherless Cases

Formulas are available to calculate the possibility of parentage when the mother, child in question, and assumed relatives of the unavailable father are tested.[21] A common situation, shown in Figure 45-2, is to test one or both parents of a deceased man (i.e., the presumed grandparents). If the obligatory paternal allele in the child is not present in either of the child's presumed grandparents, the man is excluded. If it is present, then the PI can be determined. It is also possible to use tests of other children of the deceased man with or without testing their mother(s), his siblings, or other related individuals. In doing these reconstructions, one assumes that the surrogate individuals are biologically related to the father.

Kinship Analysis

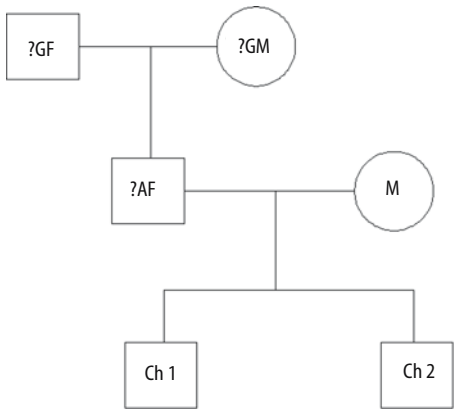
Estimation of whether two tested individuals are full siblings, are half-siblings, or have some other relationship is possible using appropriate kinship coefficients.[4,22] Although full siblings will share one or both alleles by descent in most systems, 25% of the time they will not share a common allele. Molecular genetic markers also are very useful in determination of the zygosity of twins. Table 45-6 shows the testing results for a mother and three children. From this data it is possible to determine whether the children have the same or different fathers and thus are full or half siblings. For the case shown in Table 45-6, the data indicate the existence of multiple fathers.

Reporting of Test Results

Many of the reports of relatedness testing may become part of legal proceedings. In some cases, one of the tested parties may challenge the results, and thus it is very important to have careful documentation of all steps from sample collection to reporting of results. AABB standards define the required information for a report of parentage testing.[6,7] The date of collection of each sample, the name of each person tested, and his or her relationship to the child should be stated in the report. The phenotype observed for each individual using standard nomenclature and the PI for each genetic system tested must be included. The report must explain whether a relationship can or cannot be excluded. The racial or ethnic background used for calculations, the combined parentage index, and the probability of parentage as a percentage must be stated. Any unusual findings such as mutations must be explained.

System	Mother	Child 1	OG	Child 2	OG	?GF	?GM	Possible phenotype of ?AF
D2S44 *Hae* III	1.79, 2.09	1.59, 2.09	**1.59**	2.09, 2.56	2.56	1.59, 1.79	1.82	**1.59,1.82** 1.79,1.82
D10S28 *Hae* III	1.46, 4.08	4.08	**4.08**	1.46, 2.05	**2.05**	2.05, 3.44	1.71, 4.08	**1.71,2.05** 1.71,3.44 **2.05,4.08** **3.44,4.08**
D3S1358	15, 17	16, 17	**16**	15, 17	15, 17	16, 18	18	**16, 18** 18
D21S11	28, 29	28, 29	**28, 29**	28, 30	30	28	27, 29	**27, 28** **28, 29**
FGA	22	21, 22	**21**	20, 22	**20**	21, 22	20, 21	**20,21** **20,22** **21**

a



b

Figure 45-2. Reconstruction of a family with a deceased alleged father (?AF). (a) Phenotype results from two RFLP systems and three STR systems on members of the family shown in (b). (b) Pedigree of family relationships. The alleged father (?AF) is deceased. His presumed parents (?GF and ?GM) were tested, and the results are given in (a). Shown in bold are the obligatory paternal allele(s) (OG) that could be found in the alleged father if ?GF and ?GM are his parents. These results show that Child 1 (Ch 1) could be the product of a mating between the mother and the deceased man (?AF). Child 2 (Ch2) is excluded as being the child of ?AF and the mother in several systems (shaded boxes), since the OGs are not in the presumed parents of ?AF.

Table 45-6. Relationship Testing to Determine Sibship

Sample/ System	Mother	Child 1	Child 2	Child 3	Inferred Paternal Alleles
D3S1358	16	16	16	16	16
D5S818	11, 12	10, 11	12, 13	11	**10, 11, 13**
D7S820	12	11, 12	8, 12	9, 12	**8, 9, 11**
D8S1179	13, 15	13, 14	14, 15	10, 13	10, 14

The paternal alleles can be inferred from each child. The existence of more than two obligatory paternal alleles in loci D5S818 and D7S820 indicates the existence of multiple fathers of these children.

Conclusion

Testing of multiple genetic systems by validated molecular methods can provide information to resolve questions of relatedness. Though these tests are powerful tools that can exclude almost all falsely accused parents, they alone do not prove absolutely that a relationship exists between two individuals.

References

1. Miale JB, Jennings ER, Rettberg WAH, et al. Joint AMA-ABA guidelines: present status of serologic testing in problems of disputed parentage. *Fam Law Q.* 1976;10:247–285.
2. Baird M, Blazas I, Guisti A, et al. Allele frequency distribution of two highly polymorphic DNA sequences in three ethnic groups and its application to the determination of paternity. *Am J Hum Genet.* 1986;39:489–501.
3. Polesky HF. Impact of molecular (DNA) testing on determination of parentage. *Arch Pathol Lab Med.* 1999;123:1060–1062.
4. Wenk RE, Traver M, Chiafari FA. Determination of sibship in any two persons. *Transfusion.* 1986;36:259–262.
5. Ivanov PL, Gill P, Sullivan KM, et al. DNA-based identification of the last Russia's royal family. In: *Proceedings from Fourth International Symposium on Human Identification.* Madison, WI: Promega Corporation; 1993:37–47.
6. AABB. *Standards for Relationship Testing Laboratories 7th* ed. Bethesda, MD: AABB; 2005.
7. American Association of Blood Banks. *Guidance for Standards for Parentage Testing Laboratories.* 5th ed. Bethesda, MD: American Association of Blood Banks; 2002.
8. Polesky HF. Parentage testing: use of DNA polymorphism and other genetic markers. In: Henry JB, ed. *Clinical Diagnosis and Management by Laboratory Methods.* Philadelphia: WB Sanders; 2001: 1390–1401.

9. Budowle B, Moretti TR, Niezgoda SJ, et al. CODIS and PCR based short tandem repeat loci: law enforcement tools. In: *Second European Symposium on Human Identification, 1998.* Madison, WI: Promega Corporation; 1998:73–88.
10. Allen M, Lui L, Gyllensten U. A comprehensive polymerase chain reaction oligonucleotide typing system for the HLA class I A locus. *Hum Immunol.* 1994;40:25–32.
11. Bunce M, Fanning GC, Welsh KI. Comprehensive, serologically equivalent DNA typing for HLA-B by PCR using sequence-specific primers (PCAR-SSP). *Tissue Antigens.* 1995;45:81–90.
12. Nikiforov TT, Rendle RB, Goelet P, et al. Genetic bit analysis: a solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Res.* 1994;22:4167–4175.
13. Roewer L. The use of the Y chromosome in forensic genetics—current practices and perspectives. In: Brinkmann B, Carracedo A, eds. *Progress in Forensic Genetics.* Vol 9. Amsterdam: Elsevier Science; 2001:279–280.
14. Walker RH. Probability in the analysis of paternity test results. In: Silver H, ed. *Paternity Testing.* Washington, DC: American Association of Blood Banks; 1978:69–135.
15. Budowle B, Shea B, Niezgoda S, et al. CODIS STR loci data from 41 sample populations. *J Forensic Sci.* 2001;46:453–489.
16. Morris JW, Garber RA, d'Autremont J, et al. The avuncular index and the incest index. In: Mayr WR, ed. *Advances in Forensic Haemogenetics.* Vol 2. Berlin: Springer-Verlag; 1988:607–611.
17. Gjertson DW. The effect of an isolated single-locus inconsistency in the statistical evaluation of paternity. In: American Association of Blood Banks. *Guidance for Standards for Parentage Testing Laboratories.* 5th ed. Bethesda, MD: American Association of Blood Banks; 2002:143–144.
18. Essen-Möller E. The conclusive force of similarity in paternity proof: theoretical basis. *Mitt Anthrop Ges (Wien)* 1938;68:9–53.
19. Salmon D. The random man not excluded expression in paternity testing. In: Walker RH, ed. *Inclusion Probabilities in Parentage Testing.* Arlington, VA: American Association of Blood Banks; 1983:281–292.
20. Brenner CH: A note on paternity computation in cases lacking a mother. *Transfusion.* 1993;33:51–54.
21. Traver M. Numeric statements of the strength of the genetic evidence. In: American Association of Blood Banks. *Guidance for Standards for Parentage Testing Laboratories.* 5th ed. Bethesda, MD: American Association of Blood Banks; 2002:124–129.
22. Brenner CH. Symbolic kinship program. *Genetics.* 1997;145:535–542.

Chapter 46

Assessment of Chimerism in the Setting of Allogeneic Hematopoietic Cell Transplantation

Vivianna M. Van Deerlin and Erica Williams

Introduction

Hematopoietic cell transplantation (HCT) has become a well-established treatment option for a variety of malignant and nonmalignant diseases. Molecular analysis of chimerism is used to monitor the levels of donor and recipient cells in patients after HCT. The clinical utility and interpretation of chimerism analysis depend on the type of bone marrow transplant used and the underlying disease. After transplantation, chimerism analysis is used to confirm engraftment of donor hematopoiesis and identify and quantify the percentage of recipient cells to guide patient management aimed at helping to prevent graft failure or relapse.

Terminology

A chimera is a mythological creature consisting of the head of a lion, the body of a goat, and the tail of a serpent. In medicine, a chimera is an individual whose cells derive from two or more genetically distinct individuals. While this can occur naturally by fetal-maternal transfer, in the context of hematopoietic transplantation, chimerism is the desired outcome, with the blood genetically of donor origin. A variety of terms are used to describe posttransplant chimerism (Table 46-1).[1–4] "Complete chimerism" (also referred to as "full chimerism"), or "complete (full) donor engraftment" indicates that only donor hematopoietic cells are identified. In contrast, when both recipient and donor cells are present in a cellular compartment, this is referred to as "mixed chimerism." Typically peripheral blood or bone marrow is analyzed for chimerism; the chimeric status of cellular fractions such as T cells or B cells is called "subset chimerism."[5] The presence of complete chimerism (100% donor) in one or more cellular fractions, and mixed chimerism or 100% recipient origin in another cellular fraction is often called "split chimerism." While detection of mixed chimerism is important, identification of dynamic change (increasing or decreasing mixed chimerism over time) is most useful clinically. The sensitivity of chimerism analysis is important. The most commonly used methods have a lower limit of sensitivity of ~1%; however, alternate methods with much higher sensitivity are available. Since the significance of chimerism analysis results may be different depending on the clinical situation, detection of less than 1% recipient is frequently referred to as "microchimerism," a term that arose from the solid organ transplant field.

Clinical Utility

Hematopoietic Cell Transplantation

Allogeneic HCT was developed as a rescue for bone marrow function after treatment with high-dose chemotherapy or total body irradiation aimed at eradication of malignancy and immunosuppression to permit donor cell engraftment, or both.[6] Hematopoietic rescue after myeloablation is accomplished using donor stem cells from bone marrow or granulocyte-colony stimulating factor (G-CSF) stimulated peripheral blood. The donor can be either a related individual or an HLA-matched unrelated donor. A better understanding of the therapeutic mechanism and improved patient outcomes using myeloablative HCT over the last several decades has led to HCT becoming the treatment modality of choice for a variety of disorders.[6] Allogeneic HCT is used traditionally to treat hematopoietic malignancies such as acute and chronic leukemias, as well as myelodysplastic syndromes (MDS). Inherited or acquired nonmalignant diseases, such as thalassemia, severe aplastic anemia, immunodeficiency disorders, and enzyme deficiencies, also are successfully treated with allogeneic HCT.

The efficacy of HCT for treatment of malignant disorders includes mediation of a therapeutic antitumor effect known as graft versus tumor (GVT) effect by the donor allogeneic immunocompetent cells, which is

Table 46-1. Terminology of Posttransplant Chimerism Analysis

Term	Definition	Significance
Complete or full chimerism	100% donor detected	Suggests complete hematopoietic engraftment, but very small populations of recipient cells may not be detected
Mixed chimerism	Both donor and recipient detected	May indicate increased risk of relapse; dynamic change most important (increasing, decreasing, or stable); for nonmyeloablative transplants, mixed chimerism is an expected outcome
Subset chimerism	Mixed chimerism in a cellular subpopulation, e.g., T cells	Patterns of T-cell and natural killer (NK)-cell chimerism may be predictive of GVHD or graft loss in nonmyeloablative transplant patients
Split chimerism	Recipient not detectable in all cell lineages	Patterns of T-cell and NK-cell chimerism may be predictive of GVHD or graft loss in nonmyeloablative transplants
Microchimerism	<1% recipient detected	Associated with solid organ transplantation; unclear significance in setting of HCT

analogous to graft-versus-host disease (GVHD).[6] In fact, evidence for the GVT or graft-versus-leukemia (GVL) effect arose from noting that T-cell depletion strategies used to decrease GVHD, as well as syngeneic (identical twin) transplants, resulted in a higher relapse rate. Further evidence for the GVT effect came from observations that donor lymphocyte infusions (DLIs) can induce a complete remission after relapse following HCT in patients with hematologic malignancies, used most successfully in chronic myelogenous leukemia (CML). The theory is that infused donor lymphocytes become sensitized to surface antigens expressed on the leukemic cells, thereby transforming them into cytotoxic lymphocytes targeted against the host leukemic cells.

Thus, allogeneic HCT destroys malignant cells not only by the chemoradiotherapy but also by the GVT effect. This observation led to the development of nonmyeloablative approaches to allogeneic transplantation, which use less-intense conditioning regimens, resulting in reduced toxicity.[7,8] The lower toxicity of nonmyeloablative transplants permits use of HCT as a treatment modality in patients previously deemed ineligible for myeloablative HCT for reasons of advanced age (>50 years old) or comorbid conditions. The immunosuppression induced by the reduced-intensity conditioning regimen permits development of a mixed chimeric state, which may convert to complete donor chimerism spontaneously or after repeated DLIs.[9]

Chimerism in Myeloablative HCT

The significance of engraftment and mixed chimerism depends on the underlying disease and the type of transplant (Table 46-2). Most patients undergoing conventional myeloablative HCT rapidly achieve complete chimerism, which is considered to be the goal essential for the success of the therapy. Subsequently, the detection of small amounts of mixed chimerism is evidence of either remaining or reappearing recipient cells, which may represent normal or malignant cells. Relapse is the most common cause of treatment failure; therefore, it has been of central interest to have a test that predicts relapse to allow for institution of preventive treatments, such as DLI or a rapid taper of immunosuppression. While some studies have shown that identification of mixed chimerism is associated with a higher risk of relapse, others have not confirmed this relationship. Thus, the significance of mixed chimerism in the setting of myeloablative HCT is controversial.[1–4]

There are many possible reasons for this controversy, including the variability and sensitivity of the methods

Table 46-2. Comparison of Chimerism Analysis in Myeloablative and Nonmyeloablative HCT Based on Recommendations of the National Marrow Donor Program and the International Bone Marrow Transplant Registry 2001 Workshop[10]

	Myeloablative	Nonmyeloablative
Recommended Source of DNA	Peripheral blood	Peripheral blood and cell lineage subpopulations
Expected Chimerism	Complete engraftment	Mixed chimerism for approximately 6 months; may convert to complete chimerism with or without DLIs
Frequency of Analysis	1, 3, 6, and 12 months after HCT; after complete chimerism, based on change in clinical condition	Every 2–4 weeks until DLIs
Significance	Mixed chimerism, particularly increasing mixed chimerism, may help predict relapse; significance depends on disease and cell lineage of chimerism	Early patterns of chimerism may predict either GVHD or graft loss
Method	Tandem repeat loci sensitive for detection of recipient allele(s)	Tandem repeat loci sensitive for detection of both recipient and donor allele(s)

used to identify chimerism, the source of sample used for the analysis, the underlying disease, the age of the patient, and the timing of the measurements. When very sensitive molecular methods are employed, detection of mixed chimerism is common. In fact, transient mixed chimerism has been well documented, particularly in the first 9 months after transplant.[3] Underlying disease is also a factor. For example, in patients with CML, the reappearance of recipient cells has been associated with increased risk of relapse; however, for acute leukemias and myelodysplastic syndrome, the same association has not been found. However, it is generally true that identification of increasing levels of mixed chimerism does predict a high risk of relapse. In these patients, preemptive immunotherapy with DLI could be instituted.[11] The promise of early detection of relapse by chimerism analysis is tempered by the fact that the kinetics of relapse is generally very rapid, particularly with acute leukemias, so the likelihood of detecting clinically unsuspected relapse within the testing interval is unlikely.

Minimal residual disease (MRD) detection can be performed by amplification of disease-specific markers such as *BCR/ABL* transcripts.[2] The sensitivity of MRD analysis (generally between one cell in 1,000 and one cell in 1,000,000) is greater than that of standard methods for chimerism analysis (~1% to 5%). In addition, MRD detection with a disease-specific marker detects only the malignant cells, while chimerism analysis detects all recipient cells without distinguishing between malignant and normal host cells. However, when a disease-specific marker is not available, chimerism analysis can serve as a surrogate test for disease. This is especially true in light of newer, more sensitive methods of detection of microchimerism, which use more sensitive markers (such as targets on the Y chromosome) or target specific cell populations, which increases the sensitivity significantly over unsorted peripheral blood. For example, mixed chimerism in CD34-positive cells from peripheral blood appears to be predictive of relapse in patients with various leukemias and appears to correlate with the results of reverse transcription–polymerase chain reaction (RT-PCR) for the *BCR/ABL* fusion messenger RNA (mRNA) in CML patients.[4,11] The drawback of this approach is that it is limited to the adult population because of the amount of blood required for the CD34 enrichment, and is costly to perform on a regular basis. When available, testing using a disease-specific marker is the method of choice for MRD detection; in all other cases, chimerism analysis is an appropriate alternative.

As research continues in the area of engraftment analysis, the best choice of strategy for its clinical application is often under debate and varies among clinicians. To develop consistent criteria for the use and interpretation of chimerism analysis, the National Marrow Donor Program and International Bone Marrow Transplant Registry (NMDP/IBMTR) sponsored a workshop in 2001 to establish a rational and consistent approach to the measurement of chimerism (Table 46-2).[10] For myeloablative HCT, analysis of chimerism can be used to document engraftment on a schedule of 1, 3, 6, and 12 months after transplant. Once complete chimerism is established, repeat testing should be used to investigate changes in clinical condition, for example, to distinguish relapse from other causes of clinical decline such as viral infection.

Chimerism in Nonmyeloablative HCT

In contrast to conventional myeloablative transplants, chimerism is the initial expected outcome of nonmyeloablative or reduced-intensity conditioning HCT. A mixed chimeric state is frequently present for the first 6 months after nonmyeloablative HCT.[12] The use of DLI for conversion from mixed chimerism to complete donor chimerism is guided by the determination of the chimeric status of the patient.[6] Therefore, chimerism analysis plays a critical role in the management of the nonmyeloablative HCT patient. Chimerism analysis can be performed on unfractionated peripheral blood or bone marrow; however, subset chimerism analysis of specific cellular lineages, such as T cells, appear to be the most useful.[5,13]

Many studies have examined the prognostic value of the level of mixed chimerism in different cellular subsets for early prediction of graft failure.[14] Different cell subpopulations engraft at different rates, with myeloid engraftment usually occurring prior to T-cell engraftment.[12] The kinetics of T-cell engraftment vary depending on the underlying disease, conditioning regimen, and graft source; however, high donor T-cell and natural killer (NK)-cell chimerism appears to be critical for successful engraftment, which in turn is strongly associated with sustained antitumor response.[13] In contrast, low donor T-cell and NK-cell chimerism (below 20% to 50%) has been associated with an increased risk of graft rejection. Rapidly increasing T-cell donor chimerism by 28 to 30 days after transplant appears to be associated with an increased risk and severity of acute GVHD. Determination of myeloid-specific chimerism is important in patients with malignancies of myeloid origin. Since bone marrow is composed predominantly of myeloid precursors, chimerism analysis of bone marrow is most representative of myeloid lineage chimerism but correlates poorly with T-cell chimerism.

Frequent monitoring of mixed chimerism early after nonmyeloablative HCT is clinically useful since early patterns of chimerism may be predictive for GVHD (rapidly increasing T-cell chimerism) or graft loss (decreasing T-cell chimerism to ≤20% donor). The recommendations of the NMDP/IBMTR workshop for nonmyeloablative HCT are analysis of chimerism every 2 to 4 weeks using peripheral blood, preferably including lineage-specific testing, until therapeutic DLI treatment is instituted (Table 46-2).[10] DLI is indicated for declining chimerism, stable mixed chimerism for more than 2 weeks, or persistent or progressive disease. In patients with GVHD or complete

Table 46-3. Comparison of Methods of Chimerism Analysis

Method	Sensitivity	Advantages	Disadvantages
Karyotyping	5–10%	Monitors chromosomal abnormalities as well	Requires dividing cells; low sensitivity; high false-positive rate
RFLP	5–10%	Highly informative	Requires Southern blot analysis; alleles not distinct
Red blood cell phenotyping	0.04–3%	Simple	Not very informative; can be affected by transfusion; not useful for rapidly proliferating diseases
Short tandem repeat PCR	1–5%	Highly informative; quantitative	Only moderate sensitivity; multiplex amplification may decrease sensitivity
FISH for X and Y chromosomes	0.1–0.5%	Sensitive; feasible to screen large numbers of cells for quantitative result	Only for sex-mismatched transplants
Single nucleotide polymorphism (SNP)	0.01–1%	Highly sensitive; quantitative	Less quantitative with higher chimerism; need to evaluate many SNPs to find informative SNP
Y-chromosome PCR	0.001–0.1%	Sensitive; highly informative in sex-mismatched transplants	Only for sex-mismatched transplants with male recipient and female donor
Cell sorting + STR-PCR	0.0001–0.1%	Informative; sensitive; cell lineage information important	Requires large volume of blood; technically demanding

RFLP, restriction fragment length polymorphisms; STR-PCR, short tandem repeat PCR; XY-FISH, X and Y chromosome fluorescence in situ hybridization.

chimerism, analysis is recommended every 3 to 6 months.

HCT for Nonmalignant Disease

Allogeneic HCT is a potentially curative treatment for some nonmalignant diseases, such as enzyme deficiencies, hemoglobinopathies, and aplastic anemia. In these cases, the aim of HCT is to correct the cellular defect, for example, to improve hematopoietic function or normalize the enzyme deficiency. Since replacement of the bone marrow is not necessary to correct the defect, the less-toxic nonmyeloablative approaches are being applied. In addition, a stable mixed chimeric outcome may be sufficient to achieve significant clinical improvement of the defect.[1] The percentage of donor chimerism required to correct the defect is disease dependent.[10]

Graft rejection is the most common cause of treatment failure and may be heralded by detection of decreasing donor cells. Therefore, monitoring of mixed chimerism after HCT for some nonmalignant disorders is recommended at 1, 2, and 3 months after transplant.[10] Ongoing monthly follow-up is warranted if the percentage of donor chimerism is noted to be declining.

Methods

The assessment of chimerism is based on the ability to distinguish donor cells from recipient cells. The methods used for chimerism analysis have historically paralleled the technologies and advancements for identity testing in forensics, which are described in chapter 44. Many methods have been applied for chimerism analysis, including red blood cell phenotyping, HLA typing, genomic DNA restriction fragment length polymorphism (RFLP) analysis, and classical cytogenetics. These methods have been largely superseded by fluorescence in situ hybridization (FISH) and PCR-based methods, which are more informative, more sensitive, more robust, and faster to perform (Table 46-3).

Fluorescence In Situ Hybridization

FISH analysis for the X and Y chromosomes (XY-FISH) can be used for chimerism analysis by hybridization of interphase preparations with two differentially labeled probes for sequences on the X and Y chromosomes, respectively.[4] Classic karyotyping methods could be used for detection of X and Y chromosomes but would be labor-intensive, slow, and limited by the number of cells analyzed. For example, to attain a sensitivity of 1%, more than 300 cells would need to be evaluated, which is not feasible. FISH allows the rapid screening of a large number of cells. The method can yield quantitative results by rapidly counting between 500 and 1000 cells. A sensitivity of approximately 0.1% to 0.5% can be achieved. Standardization of XY-FISH has made it suitable for clinical use.[14] However, the major limitation of XY-FISH is the applicability only for sex-mismatched transplants.

Tandem Repeats

Tandem repeats are a class of polymorphism consisting of a variable number of tandemly repeated core sequences.[15] Tandem repeats are classified according to the number of nucleotides in each core sequence. Minisatellites have core sequences of approximately 8 to 80 base pairs (bp) and are also known as variable numbers of tandem repeats (VNTR). In contrast, microsatellites, also known as short tandem repeat (STR) loci, have core sequences with 2 to 7 bp.[16] Allelic variation at each VNTR and STR locus arises from the number of core repeats present. STR alleles have

between 3 and 40 tandem core repeats; each STR locus typically has 10 to 20 different allelic variants. Variant alleles with imperfect repeats also occur.[16]

STR loci and some VNTR loci are amenable to PCR amplification using primers that flank the repeating sequence. After amplification, the number of repeats is calculated from the size of the PCR product, or, more simply, the products are compared to an "allelic ladder" constructed to contain the majority of known alleles in the human population of a particular locus.[16] PCR amplification of tetranucleotide STR loci (4 bp in each core repeat) has become the predominant method for DNA-based human identification. Comprehensive information about STR loci and their application in human identification is available from a Web site sponsored by the National Institute of Standards and Technology (http://www.cstl.nist.gov/div831/strbase).

Amplification of STR loci has become the most frequently applied technology for the study of engraftment and mixed chimerism after allogeneic HCT for several reasons.[17] By examining multiple STR loci, which is made easier by multiplex PCR, it is possible to find at least one STR locus for each donor-recipient pair that can be used to distinguish the donor from the recipient; thus, STR-PCR is essentially universally informative. In addition, STR-PCR can be used to quantitatively discriminate mixtures of DNA based on relative amplification of alleles. The sensitivity of STR-PCR for chimerism analysis is approximately 1% to 5%. STR alleles are relatively similar in size due to the small repeat unit; thus, compared with VNTR alleles, which have larger differences, STR-PCR is theoretically better for quantification because preferential amplification of smaller alleles is less likely to occur. The advent of capillary electrophoresis and automated fluorescent detection has made the analysis of PCR products technically straightforward, and the availability of commercial reagents for STR-PCR simplifies assay development.

Y Chromosome Analysis

Amplification of Y chromosome–specific sequences can be used for assessment of chimerism in sex-mismatched HCT involving a male patient and a female donor, which represents approximately 25% of all HCT cases. Because the targeted sequences are present only on the Y chromosome, the method is very sensitive. Sensitivities as high as 1 male in 100,000 female cells (<0.001%) have been reported, although the clinical relevance of chimerism detection at such high levels of sensitivity is not entirely clear.[18,19] Use of quantitative real-time PCR has the advantage of being not only sensitive but also quantitative for very low levels of mixed chimerism, useful for detecting an increasing or decreasing signal over time. Amplification of the amelogenin gene, which is present on both the X and Y chromosomes, but is 6 bp longer on the Y chromosome, is used for sex typing in forensic analysis and also has utility for chimerism analysis.[16] The major disadvantage of all Y chromosome PCR methods is the limited application only to a subset of sex-mismatched transplants.

Single Nucleotide Polymorphisms

Single nucleotide polymorphisms (SNPs) are biallelic variants at a single nucleotide position. SNPs occur on average every 1.3 kilobases (kb) in the human genome. Since SNPs are sequence polymorphisms rather than length polymorphisms, their utility for chimerism analysis is based on the frequency of heterozygosity in the particular population under study. With only three potential genotypes for each SNP locus (AA, AB, BB), the chance of distinguishing two individuals is much lower than for STR loci, which have 10 to 20 alleles per locus. Consequently multiple SNPs must be analyzed to successfully distinguish all donors and recipients. One investigator statistically calculated the number of SNPs required to identify an informative locus in 96.5% of cases to be 25 SNPs.[20] In one study using a panel of 51 SNPs, an informative locus was found in 100% of cases; however, this study used a microarray-based minisequencing system to enable high throughput testing.[21] While the number of loci necessary is a disadvantage, an advantage of SNP detection over tandem repeat amplification is that SNPs are less susceptible to preferential amplification of alleles based on length, and stutter peaks (described later) are not a concern.

While there are many methods for analysis of SNPs, real-time PCR technology is commonly used for SNP-based chimerism analysis.[22] The real-time PCR analysis permits sensitive detection of mixed chimerism, reported to be approximately 0.1%. Although real-time PCR is quantitative, the coefficient of variation (CV) at chimerism levels above 5% is as high as 30% to 50%, compared with a CV of about 5% for STR-PCR methods.[4] Thus, while SNP real-time PCR is more sensitive and quantitative, particularly at low levels of mixed chimerism, the variability of the quantitation when the level of recipient is above 5% does not permit accurate determination of dynamic change in mixed chimerism levels. Real-time SNP analysis of mixed chimerism is a promising method; however, more studies are needed to determine its utility for HCT chimerism analysis.

Laboratory and Testing Issues

PCR amplification of STR loci is currently the most commonly accepted and most widely used method for assessment of mixed chimerism after HCT.[10,23] Therefore, the remainder of this chapter details the use and interpretation of STR-PCR for analysis of mixed chimerism in a clinical laboratory.

STR Amplification Reagents

Chimerism analysis requires the identification of at least one informative STR locus that distinguishes recipient

from donor cells prior to transplant, which is used to monitor recipient cells, donor cells, or both in the recipient after transplant. A variety of commercial reagents are available for amplification of STR loci. The majority of commercially available STR amplification reagents are designed for forensic analysis in which a maximum amount of information is desired from frequently limited samples.[16] Multiplex PCR, in which multiple loci are amplified simultaneously by including more than one set of PCR primers in the reaction, is ideal for this purpose. The PCR product sets from the multiplexed reactions are distinguished by the expected size range of the products (designed to be nonoverlapping) and by the use of different fluorophores, allowing amplification of loci with overlapping product sizes. High-order multiplex primer sets have been designed that amplify 8 to 16 STR loci. The two major suppliers of STR primer sets are Promega Corporation (Madison, WI) and Applied Biosystems (ABI, Foster City, CA), each with a variety of multiplex combinations, some designed for detection with specific methods or instruments. Beckman Coulter Inc (Fullerton, CA) has developed the first 12-locus STR multiplex primer analyte-specific reagent (ASR) specifically designed for chimerism analysis in the clinical laboratory on its capillary electrophoresis instrument. Issues relating to selection and use of the appropriate reagents for chimerism analysis are discussed in the context of pretransplant analysis.

Detection Methods

Amplified STR products can be detected in numerous ways, as described for forensic analysis in chapter 44. Separation by sequencing gel electrophoresis with fluorescent or silver stain detection, either after standard electrophoresis or with an automated instrument, is acceptable but slow and labor-intensive. In the absence of an automated fluorescence detection gel electrophoresis instrument, postelectrophoresis detection of fluorescently labeled DNA fragments can be performed with a flatbed scanner. By far the most popular method for analysis of STR amplification products is fluorescent detection by capillary electrophoresis (CE), which has the advantage that a gel does not need to be poured and analysis of samples is rapid.[16] CE can be performed with a single capillary instrument or with a multicapillary instrument such as the ABI 3130, which has 16 parallel capillaries, or the Beckman Coulter Vidiera NsD, which has 8 parallel capillaries. The advantage of multiple parallel capillary systems is that analysis can be performed more rapidly than with a single capillary instrument. Capillary arrays with 96 or more capillaries are available; however, their instrument and reagent costs are greater, and thus they are cost-effective only for very high volume testing.

In the CE instruments, DNA fragments migrate through the liquid polymer or linear polyacrylamide within a thin capillary with applied voltage. The fluorescently labeled products are excited by laser-induced fluorescence when the fragments pass the optical reading window and are detected by the wavelength of the emission. Instruments that can detect multiple fluorophores permit not only multiplex amplification of loci resulting in PCR products of overlapping in size, but also inclusion of a size marker for accurate sizing of the fragments. Of note, it is important to use fluorophores that are detectable by the instrument being used. For example, the fluorophores used in the Promega Corporation and Applied Biosystems STR amplification reagents are not detected by the Beckman Coulter instrument.

Pretransplant Testing

Sample

The goal of pretransplant analysis is to identify one or more informative loci that will be used to monitor chimerism after HCT. This involves determining the STR alleles of the recipient and the donor at each STR locus used by the laboratory. Usually peripheral blood samples from the donor and recipient are used to extract DNA for pretransplant analysis. Analysis of a second sample from the recipient, such as buccal cells, controls for switching of the donor and recipient samples in the preanalytic stage. Confirmation of the same allele pattern from the recipient blood and buccal cells ensures that the identity of the samples is correct.

On occasion, a recipient is transplanted prior to the STR evaluation of a pretransplant sample. In this case, a posttransplant blood sample will not be useful for identification of the recipient's genotype. Alternatively, buccal cells, skin biopsy, paraffin-embedded tissue samples, hair follicles, urinary sediment, or a stored blood smear slide from before the HCT can be used. While these tissue samples will be of mostly recipient genetic origin, any sample obtained after transplant may contain some hematopoietic cells within the sample, thereby generating a mixed chimeric pattern. If the recipient DNA predominates, the recipient alleles can be deduced if the donor alleles are known.

Identification of an Informative Locus

To identify an informative locus for each recipient-donor pair usually many loci must be examined. This is particularly true when the individuals are related and thus may share many alleles. Multiplex amplification increases the speed and decreases the labor associated with the analysis of multiple loci. Of the loci found to be informative, not all will be appropriate for chimerism analysis. This is because the goal is not to differentiate the individuals as in forensic identity testing but to detect small amounts of recipient or donor alleles in a mixed specimen.[24]

The main factor limiting the utility of some informative alleles is the presence of a stutter peak one repeat smaller ($n - 1$) and, to a much lesser extent, one repeat larger

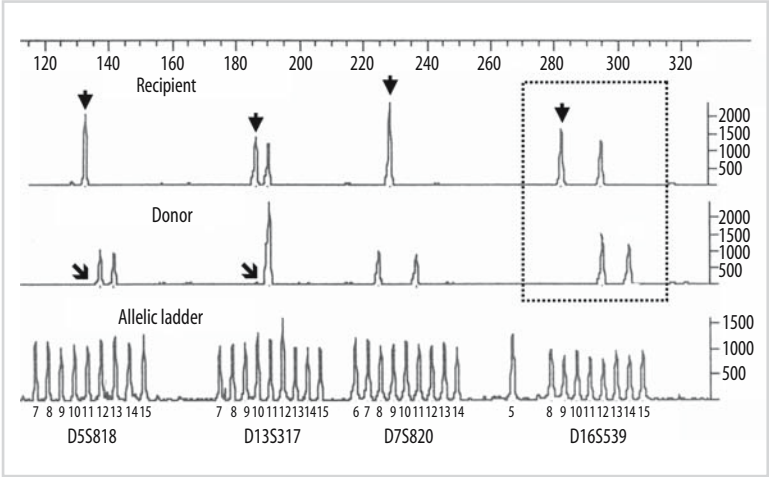


Figure 46-1. Identification of informative loci using STR-PCR and CE analysis. The STR alleles at four loci (GammaSTR, Promega Corporation) are compared for a recipient-donor pair to identify the best informative locus. The number of repeats in each allele is determined by comparison to the allelic ladder for which the sizes are indicated. While all 4 loci are informative for distinguishing the recipient from the donor, the D16S539 locus is best because the recipient-specific allele (marked with vertical arrows) is separated from the nearest donor allele by at least two repeats. In contrast, for the D5S818 and D13S317 loci, the recipient-specific allele would colocalize with a donor stutter peak, the location of which is indicated by the diagonal arrow. The D7S820 recipient-specific allele is acceptable, but, if present, an upper ($n + 1$) stutter from one of the donor alleles may interfere with its detection. The allele combination at the D16S539 locus also has the advantage that both donor- and recipient-specific peaks are present.

(n + 1) than each allele peak. When tandem repeats are amplified, stutter peaks are thought to result from slippage of the polymerase, causing mispairing of the template, which is shifted by one repeat unit, resulting in an additional PCR product predominantly one repeat smaller than the expected PCR product. The intensity of this $n - 1$ peak is usually approximately 2% to 10% of the area of the corresponding STR allele. The amount of stutter is influenced by the locus (some loci tend to have more stutter than others), the core repeat unit length (loci with shorter core repeat size having greater stutter), and the allele size (larger repeat alleles have slightly more stutter than smaller ones).[16]

For chimerism analysis, optimal informative loci have a recipient-specific allele that does not colocalize with a donor allele or stutter peak (Figure 46-1).[25] Ideally, the recipient-specific allele should be at least two repeats larger or smaller than the donor allele(s).[23] Having two unique recipient alleles that meet these criteria is preferable to having only one recipient-specific allele for the purpose of confirmation that a recipient allele peak is present, which is usually only of concern when the recipient signal is low. In addition, due to the theoretical possibility for preferential amplification of shorter fragments, the sensitivity for detection of a minor recipient-cell population may be greater if the informative recipient allele(s) have fewer repeats than any donor allele.[23] On the other hand, a disadvantage of selecting a recipient-specific allele smaller than the donor allele(s) is that this downstream region tends to have more background peaks. Because of the endless combinations of alleles between recipients and donors, each HCT pair presents a unique situation, and each must be evaluated on a case-by-case basis. These are only guidelines for the selection of informative loci; how useful a particular locus will be is ultimately determined empirically by use in the posttransplant setting. In some instances, no optimal locus can be identified. In this case, posttransplant analysis still can be performed with the caveat that the sensitivity of detection will be decreased because of colocalization of the recipient-specific allele with a stutter or other background peak.

Another factor considered when selecting an informative locus is whether there are both specific recipient and specific donor alleles (Figure 46-1). While this is not essential for myeloablative HCT, for which the combined donor and recipient allele pattern must be optimized to enhance detection of a small percentage of recipient in a donor background, it is very important for nonmyeloablative HCT, since any level of mixed chimerism (low donor or low recipient) may be seen. If no such locus can be found, then at least two loci should be selected, one optimized for detection of recipient allele(s) and the other for donor allele(s). As the level of mixed chimerism in the patient changes from low to high donor chimerism, the informative locus may need to be changed to optimize sensitivity (Figure 46-2).

Posttransplant Testing

Approaches to Posttransplant Amplification

Peripheral blood, bone marrow, or lineage-specific cell populations from the recipient after transplant can be analyzed for mixed chimerism or to confirm engraftment. Whereas it is ideal to use a high-order multiplex PCR assay for pretransplant identification of an informative locus, the multiplex assays have reduced sensitivity for detection of low levels of chimerism in the posttransplant setting.[17] The reduced sensitivity of multiplex PCR is due to competition for reagents, leading to overall decreased synthesis of each PCR product type compared to an individual PCR. Thus, a minor template, for example, DNA from recipient present at 1%, may not be sufficiently amplified to be detectable above the background signal at some loci. The use of more DNA in the reaction may overcome this, but high amplification background with nonspecific peaks can occur if too much DNA is used. From the laboratory perspective, using a highly multiplexed reaction simplifies the set-up and performance of the assay since the same reagents and controls are used for all recipient-donor pairs. The cost of using high-order multiplex primer sets for

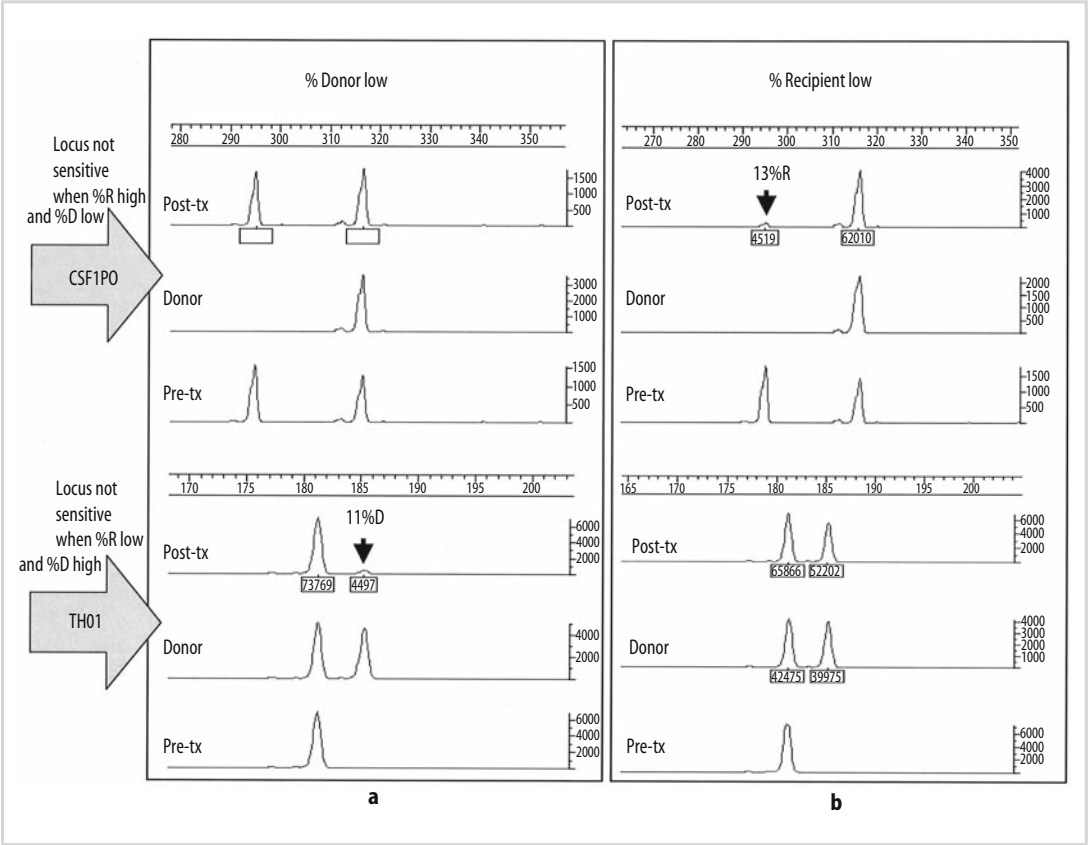


Figure 46-2. Different STR loci may be optimal depending on the level of recipient. Early after a nonmyeloablative transplant, the percentage of recipient DNA may be high and the percentage of donor DNA low. Accordingly, the locus used for assessment of chimerism must be optimized to detect low levels of donor (a). Conversely, as the percentage of donor DNA increases, the locus must be optimized for detection of low levels of recipient (b). The analysis results for two loci (CSF1PO and TH01) at two posttransplant (post-tx) time points (a and b) for a recipient-donor pair are shown. For each analysis, the donor and pretransplant (pre-tx) alleles are shown, as well as the post-tx sample showing mixed chimerism. For this recipient-donor pair, when the percentage of donor DNA is low (a), the CSF1PO locus is not sensitive for detection of donor. The TH01 locus, which has a donor-specific allele can be used. Similarly, when the percentage of donor DNA is increased (b), the situation is reversed. The CSF1PO locus is optimal for detection of lower amounts of recipient, and the TH01 locus is not useful because the post-tx peaks cannot be distinguished from the donor pattern. The normal variation in amplification of the two donor alleles is greater than the percentage of recipient DNA present.

posttransplant analysis is a drawback, particularly if only one or a few loci are informative, resulting in a high ratio of cost to information.

One approach is to use one or more multiplex STR reagents for the pretransplant analysis to genotype many loci and then use a multiplex primer set with a maximum of 4 multiplexed loci that contains the best informative locus or loci for the posttransplant analysis (Figure 46-1).[17] All things being equal, a primer set containing more than one optimal informative locus is preferable. The advantage of this approach is maintaining sensitivity because fewer loci are being simultaneously amplified, while simplifying the assay technically for the laboratory staff by minimizing the number of different primer sets stocked and utilized when the assay is performed. For example, if 12 loci are analyzed for selection of informative loci before transplant, all posttransplant samples could be accommodated with three quadriplex primer sets. This method has the added benefit of limiting the reagent cost of the posttransplant analysis.

Optimizing Sensitivity

Irrespective of the primer set used, the sensitivity of chimerism analysis can be optimized. As mentioned earlier, the amount of DNA is important for sensitivity because it directly reflects the number of cells being analyzed. Since most commercial STR reagents are designed for forensic analysis of limited samples, the amount of input DNA indicated in the forensic protocols may be less than optimal for chimerism analysis. Each laboratory should determine the maximum amount of DNA that can be used with its chosen reagents without increasing the nonspecific background. The injection time for CE instruments, which is the time of applied voltage during which sample migrates into the capillary and is essentially equivalent to the amount of product loaded on a gel, also must be optimized. Appropriate injection times are specific to each instrument. Since there is variability of amplification from sample to sample, each sample can be analyzed in duplicate, using two different injection times to ensure

that at least one of the injections is within the linear detection range of the instrument, allowing for accurate quantification (Figure 46-3a and b). In some cases, longer injection times may allow detection of very small percentages of recipient, even though the predominant donor alleles are not within the linear range for quantification. The "overloaded" donor allele peak areas would be underrepresentative of the actual amount of donor PCR product in the sample. Such "overloaded" peaks are distinguishable because they generally have flat, blunted tops rather than a sharp peak (Figure 46-3c and d).

A laboratory can estimate the sensitivity of its assay using artificial mixtures of two DNA samples. Reported sensitivities of chimerism analysis from different laboratories

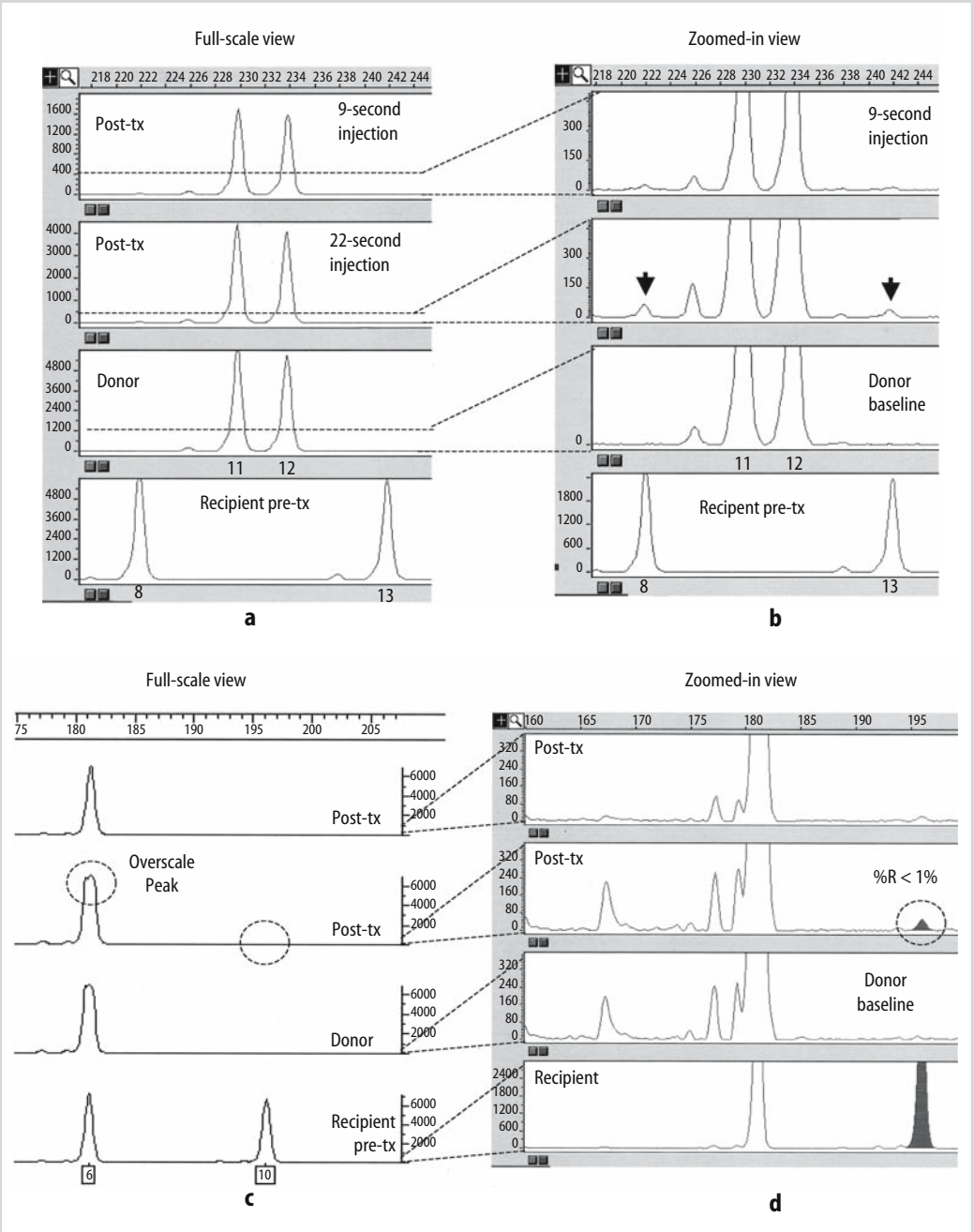


Figure 46-3. Optimizing sensitivity of STR-PCR with CE analysis. The posttransplant (post-tx) analysis of a recipient-donor pair at the D7S820 locus is shown at two injection times (9 and 22 seconds) on an ABI3100 CE instrument (a). The donor and recipient pretransplant alleles are shown. Because the recipient-specific peaks are small (arrows), the y-axis is expanded (zoomed-in view) to examine the baseline in greater detail (b). The zoomed-in view shows the presence of recipient-specific alleles in the 22-second injection, but they are not clearly visible in the 9-second injection. To evaluate the post-tx sample, a comparison must be made to the donor baseline at the same y-axis scale. Similarly, increasing the injection time may lead to overloading of the sample, which compromises quantification (large overloaded peak areas are underestimated) but may increase the sensitivity for detection of recipient alleles. The full-scale view (c) shows an overloaded peak with a blunted tip (circled) at the TH01 locus. The zoomed-in view shows detection of a recipient-specific peak (shaded) only for the post-tx sample with a blunted peak. A percent recipient can be calculated; however, it will be an overestimate of the percent recipient because the donor allele area is underestimated. Therefore, this result can be reported as less than 1%.

range from <1% to 10%. This variability may be explained not only by differences in methodology (number of loci coamplified, input DNA, detection method) but also by differences in sensitivity for each recipient-donor pair based on each locus and its constellation of alleles.[25]

Interpretation of Results

Capillary electrophoresis results from posttransplant STR analysis are analyzed to determine whether recipient-specific allele(s) are present, and if they are present, the percentage of recipient is calculated. Absence of recipient-specific alleles is consistent with complete engraftment at the sensitivity of the assay. Dynamic comparisons should be made only across the same type of sample, for example, peripheral blood, bone marrow, or cell lineage.

Analysis of Loci

If many loci are amplified, all optimally informative loci can be analyzed and reported individually or as an average value. Analysis of more than one locus provides independent confirmation of the presence and amount of recipient. In addition, if a chromosomal aberration occurs, resulting in loss of one allele, the analysis of several loci could easily detect this situation and prevent an interpretive error. Alternatively, a single optimal locus can be analyzed and reported. Analyzing and reporting results from only a single locus for each HCT patient provides a standardized, consistent, and simple approach for analysis and reporting of results.

Standard recommendations for analysis of chimerism do not exist. However, a survey of clinical laboratories performing chimerism analysis by the Association for Molecular Pathology (AMP) found that posttransplant specimens were analyzed at an average number of ten loci, with a range from 4 to 16 loci (S. Schichman, personal communication, 2005). The number of loci used to report results was variable, with some laboratories reporting a single or a few loci, but many reporting the results of all the informative loci.

Quantitative Analysis

If both donor- and recipient-specific allele(s) are clearly identified in the posttransplant sample, then the percentage of recipient and donor DNA in the sample can be calculated.[17,25] For each allele peak, the intensity of the fluorescent signal, indicated by the peak area, is proportional to the amount of PCR product. The percentage of recipient-specific DNA can be calculated as the quotient of the recipient-specific peak area(s) to the sum of the recipient- and donor-specific peak areas, as shown in Figure 46-4a and b. If the donor and recipient are both heterozygous and do not share

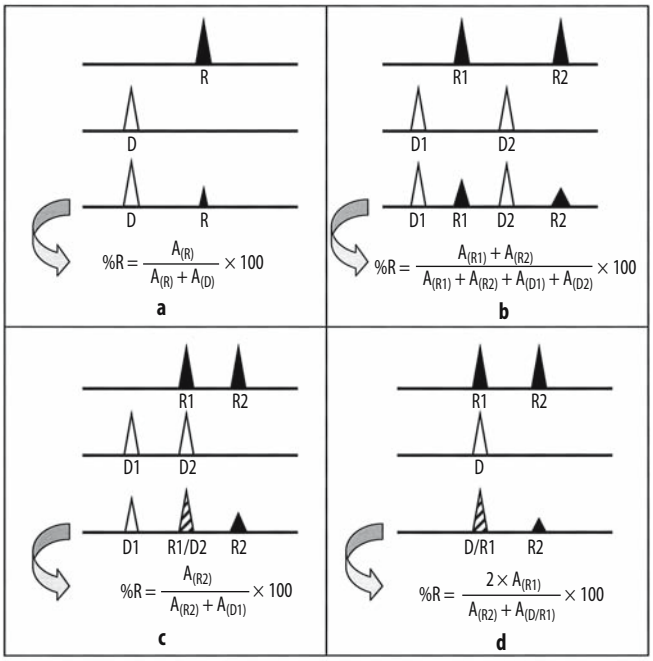


Figure 46-4. Calculation of percent recipient. Four representative scenarios (a-d) for combinations of informative recipient (R) and donor (D) alleles are shown. When heterozygous alleles are present, the subscripts 1 and 2 indicate the two alleles. For each scenario, the third line is a hypothetical posttransplant sample showing a mixed chimeric pattern. The formula that would be used to calculate the percent recipient is given, in which "A" represents the peak area of the appropriate peak designated by the subscript.

any alleles, then the peak areas of both recipient alleles are added together for the numerator and the areas of all four peaks (donor and recipient) are added for the denominator (Figure 46-4b). When the recipient and donor share an allele, the percent recipient is calculated based only on the unique alleles; the peak area of the shared allele does not enter into the calculation because it stays constant; that is, 100% of the cells in the sample (donor and recipient) will have the shared allele (Figure 46-4c). When the donor is homozygous for an allele and the recipient is heterozygous for the same donor allele and a second unique allele, as shown in Figure 46-4d, the recipient area of the shared allele can be estimated by the area of the unique recipient allele. Thus, the percent recipient is calculated as two times the unique recipient allele peak area divided by the sum of all peak areas (Figure 46-4d). In all cases, the fraction is converted to a percentage by multiplying by 100, and the percent donor is calculated as 100 minus the percent recipient.

If recipient-specific allele peak(s) are not visible initially, then it is critical to closely examine the baseline in the region of the expected recipient-specific alleles to identify small recipient peak(s) for each informative locus being analyzed. Depending on the analysis software being used, this can be done by changing the scale of the y-axis to "zoom in" on the baseline (Figures 46-3 and 46-5). It is useful to compare the posttransplant sample baseline at the location of the recipient alleles to the corresponding baseline of a pure donor sample (pretransplant donor DNA). By this comparison, any peaks that are present in the donor

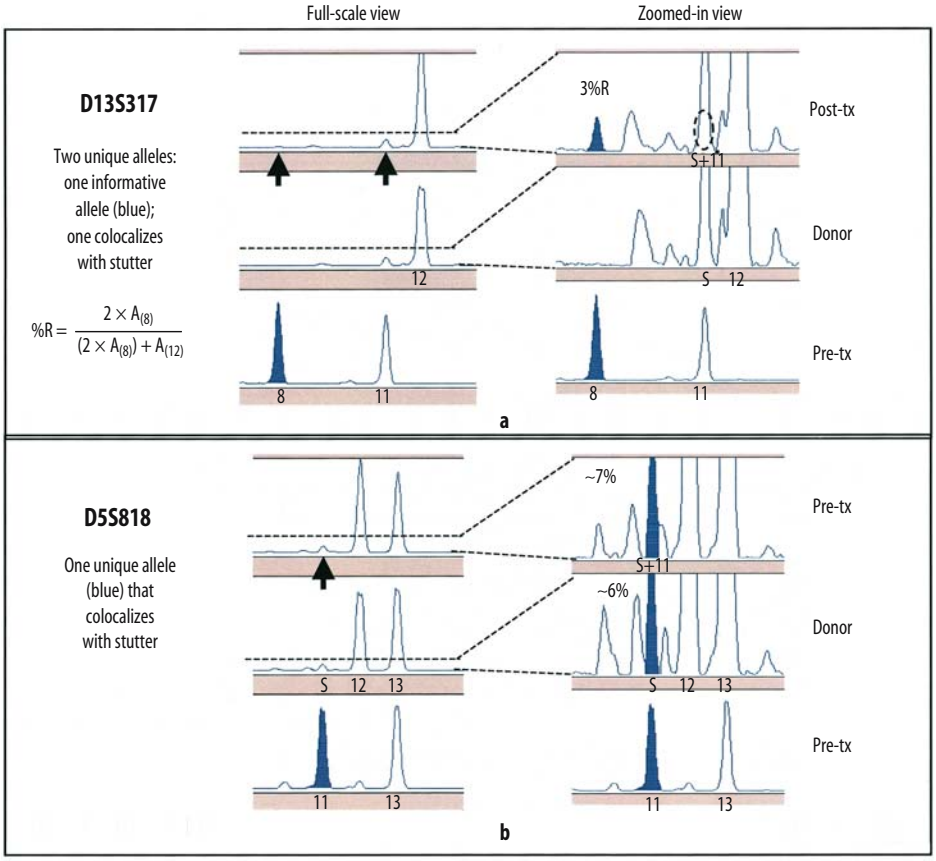


Figure 46-5. Stutter peaks from STR-PCR with CE analysis. This figure depicts the analysis of a posttransplant (post-tx) sample with a low level of mixed chimerism at two loci, as indicated using the GammaSTR multiplex primer set (Promega Corporation). Full-scale and zoomed-in views of the baseline are shown for each locus. (a) The D13S317 locus has two recipient-specific alleles, one of which is unique (8 repeats, blue) and one of which colocalizes with a donor stutter peak (11 repeats). In the posttransplant sample the recipient-specific allele (blue) is clearly seen in comparison to the donor baseline. The second allele is hidden within the area of the stutter (S + 11, dashed oval), and using the area of the stutter peak would overestimate the percent recipient. Instead, the peak area of the unique allele can be used as an estimate of the second recipient allele, as shown in the calculation formula. (b) The D5S818 locus was analyzed simultaneously with the D13S317 locus and has a single recipient-specific allele that colocalizes with a stutter peak (S + 11). The percentage of allele 11 can be calculated for both the posttransplant sample and the donor sample, and then the percentage of donor stutter can be subtracted. Whether the difference represents the presence of recipient is difficult to assess. To be definitive, the difference must be greater than the expected standard deviation of the stutter peak. In this example, the difference is likely to be real because of the clear results on the same sample at the D13S317 locus.

background in the location of a recipient-specific allele will not be mistaken for a recipient signal. In cases of graft loss, it is useful to compare the posttransplant sample baseline at the location of the donor alleles to the corresponding baseline of a pure recipient sample (pretransplant recipient DNA) to effectively distinguish low levels of donor- from nonspecific background signals in the recipient.

Some investigators advocate the use of a patient-specific standard curve for each locus created by making serial dilutions of pretransplant recipient DNA in donor DNA to aid in the quantification.[23] The associated cost and time required are prohibitive for most clinical laboratories. In addition, it is more important to obtain a reproducible result that can be used to assess dynamic change over time than to determine an absolute quantitative value. Artificial chimerism mixing studies looking at the accuracy of quantification of the same samples using different loci indicate that each locus yields slightly different quantitative results but that the results within each locus are reproducible.[17] Since dynamic change of mixed chimerism may be more important than any single quantitative value, it is more important to be consistent in the analysis, that is, always using the same single or set of loci, such that serial results can be compared, than it is to produce a standard curve. In general, the reproducibility of chimerism results is good, with low standard deviations and interassay CV generally less than 5%, although a higher CV of 10% to 15% has been noted for low levels of chimerism (less than 10% donor or recipient).[23]

Potential Pitfalls

When a Stutter Peak Cannot Be Avoided

If the only informative locus is one in which the recipient allele colocalizes with a donor stutter peak, then the sensitivity for detection of recipient will be decreased (Figure 46-5). This does not occur often but is more common for HCT between related individuals. Comparison of the posttransplant recipient peak to the stutter peak in the pretransplant donor sample is useful to determine whether the recipient signal is real. A calculation can be made for the donor stutter peak from the pretransplant donor sample analysis (using the same formula that would be used for the recipient if the stutter were not present), and this value can be subtracted from the value of the posttransplant recipient or stutter peak. Differences of less than 1% are consistent with absence of mixed chimerism, but it is important to note in the report that the sensitivity for detection of recipient DNA is decreased. Differences greater than 10% clearly indicate presence of recipient DNA, because the contribution of recipient PCR product will be incremental and thereby additive to the stutter peak area. Differences between 1% and 10% must be assessed on a case-by-case basis because there is normal variation of <1% to 5% in the stutter peak as a percentage of donor, depending on the allele size and locus.[16] To be clearly present, the recipient contribution must be greater than the standard deviation of the stutter

percentage. Recognizing the limitations of using a locus in which the recipient-specific allele colocalizes with a stutter peak, reporting such results as indeterminate or suspicious for mixed chimerism is an acceptable approach. In these cases, some information is better than none at all.

When a heterozygous recipient has a unique informative allele and a second one that colocalizes with a donor stutter peak, the accuracy of quantification is affected rather than the sensitivity. The area of the peak colocalizing with the stutter is greater than the recipient contribution due to the additional stutter signal. To avoid an overestimate of the percent recipient, the peak area of the unique allele can be used as an estimate of the second recipient allele (Figure 46-5). Therefore, the area of the unique allele is multiplied by two (instead of including the area of the peak that colocalizes with the stutter peak) and used in the appropriate formula (Figure 46-4b).

Chromosomal Abnormalities

Chromosomal abnormalities are common in hematologic malignancies and may affect identity testing results if the locus is present on the aberrant chromosome. The presence of such abnormalities may be identified by diagnostic karyotyping. If loci on the aberrant chromosome are used for chimerism analysis, the loci should not be used for the specific donor-recipient pair. Alternatively, unusual amplification patterns may provide evidence of abnormalities such as unequal amplification due to extra chromosomes or chromosomal deletions (Figure 46-6a). Chromosomal abnormalities also can arise after treatment. Such abnormalities have been reported and produce unexpected or unusual posttransplant allele patterns that are not consistent with the clinical status of the patient.[26]

Doublet Peaks

Nontemplate addition of an extra nucleotide, usually an adenosine (A), at the 3′ end of PCR products is a known occurrence with DNA polymerases, particularly Taq.[16] This adenylation results in a PCR product that is 1 bp longer than the DNA template sequence. The degree of adenylation is dependent on the sequence of the template strand. Incomplete addition of A nucleotides can occur if too much DNA template is used. Addition of the A can be favored with a final incubation step at 60°C at the end of the PCR reaction. If not all the PCR products have an extra A, this will lead to the appearance of split, or doublet, peaks. Usually this affects all the allele peaks at a particular locus. It is best to optimize reaction conditions to prevent incomplete adenylation, but if this is not possible and it is necessary to use the peak areas of such doublet peaks to calculate the percent recipient, the areas of both parts of each peak should be summed and used for calculations (Figure 46-6b).

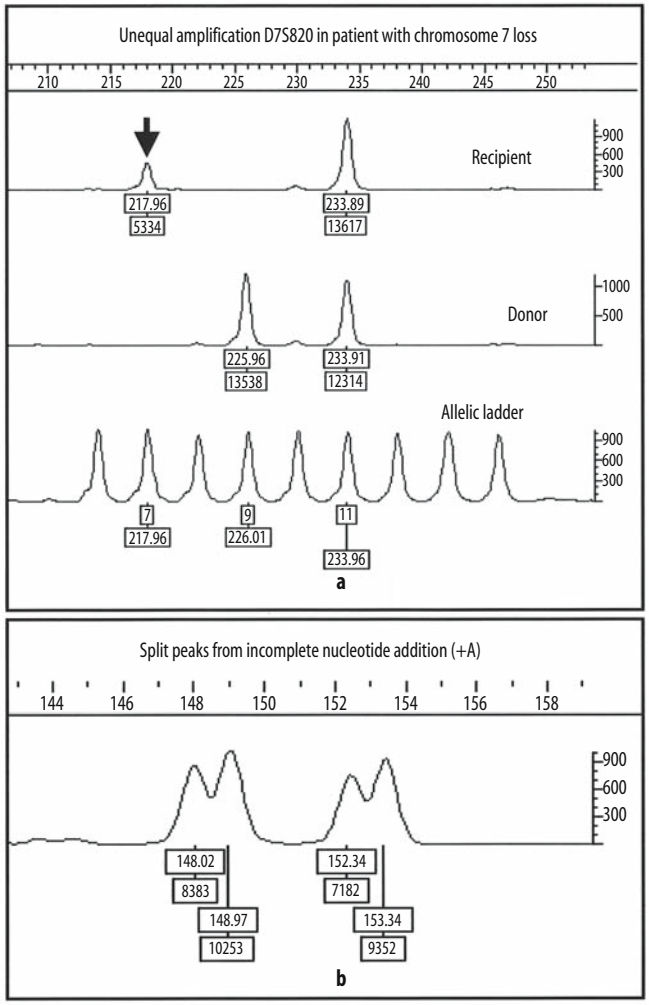


Figure 46-6. Chromosomal abnormalities and split peaks affecting SRT-PCR with CE analysis. (a) Unequal amplification of the recipient alleles at the D7S820 locus was noted (arrow) during pretransplant informative locus evaluation. The patient was known to have myelodysplastic syndrome with loss of one chromosome 7. Accordingly, the peak height of one allele is about one half the height of the other allele, indicating that some, but not all, of the patient's cells lacked one chromosome 7. (b) Nontemplate addition of an extra nucleotide, usually an adenosine (A), at the 3′ end of PCR products results in a PCR product that is 1 base pair longer than the template sequence. If not all the PCR products have an additional A, then the PCR products will appear as split or doublet peaks separated by one nucleotide in size, as shown. The size of the PCR products is indicated by the upper boxed number. The lower boxed number is the peak area.

Reports

There are no specific guidelines for reporting of chimerism results. Furthermore, the AMP survey results indicate that there is considerable variability in how laboratories report chimerism analysis results. Some, but not all, laboratories issue a separate pretransplant report to indicate the identification of informative loci. The following relevant components were included by most laboratories in their reports: specimen information, sample type, description of method, loci analyzed, sensitivity for detection of recipient, quantitative result, and interpretation of result. A minority included clinical and transplant information. Since dynamic change between time points is important, consideration should be

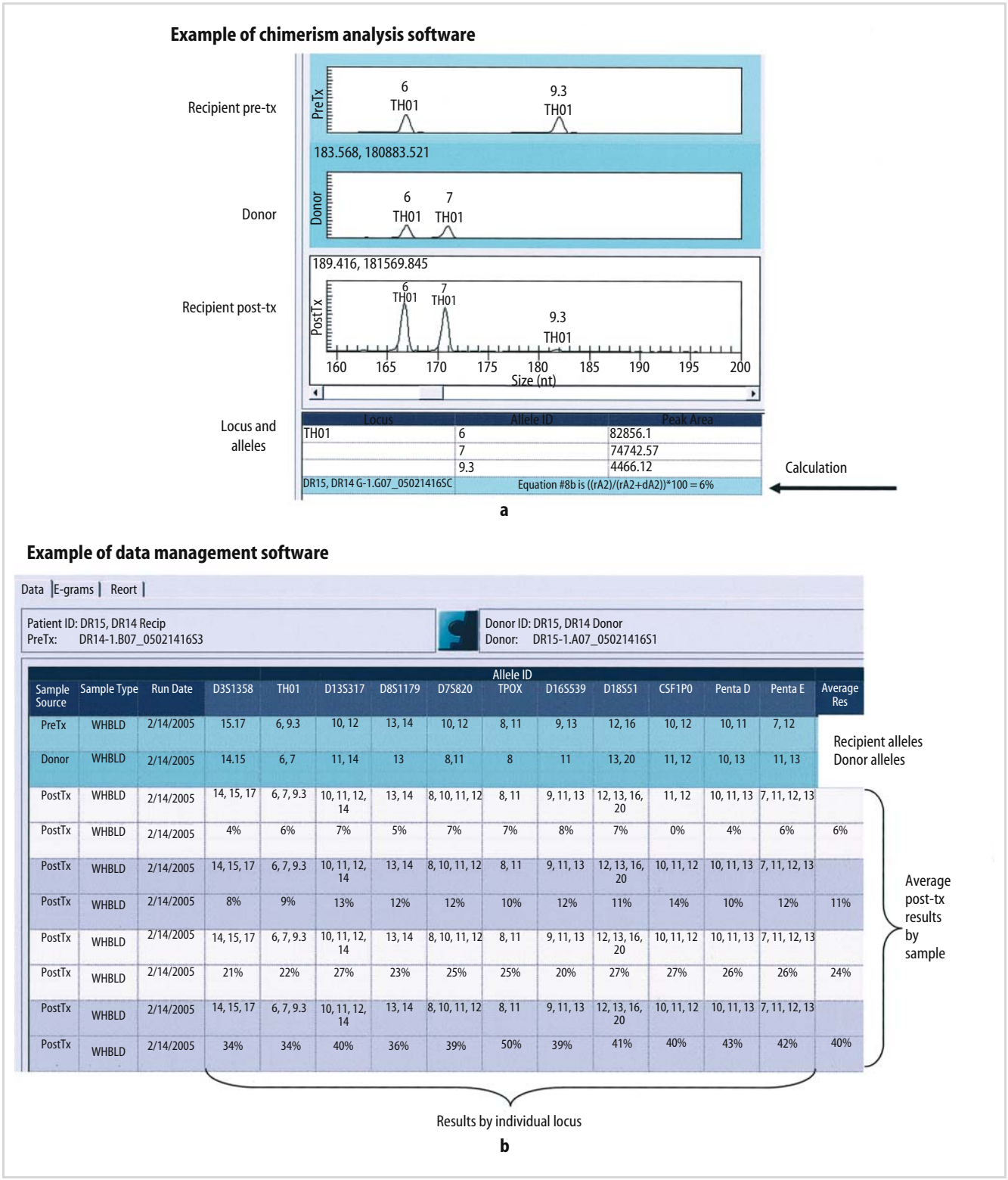


Figure 46-7. Analysis and data management software for chimerism analysis. (a) Computer screen image of the analysis of one locus. Software designed for chimerism analysis (Vidiera, Beckman Coulter) permits inspection of the electropherogram, the alleles identified at that locus, the peak areas for each allele, and the calculation used to generate the percent recipient DNA. (b) Result analysis and tracking of patient samples over time are important components of chimerism analysis. Shown here is a computer screen image for several posttransplant samples tested at multiple loci. The chimerism results for the loci are indicated individually and as an average. The results of any locus or loci may be included or excluded from the final average.

given to including in the report at least the prior result for the same sample type (i.e., comparing bone marrow to bone marrow) and giving an assessment of the degree of change in comparison with the CV of the assay. For example, stating in the interpretation whether there has been a significant increase or decrease, a slight increase or decrease, or no significant change is potentially useful. If feasible, a graphic representation of a patient's chimerism results over time would be ideal, but most clinical laboratories are constrained in report format by the laboratory information system. Ultimately the report is a method of communication between the laboratory and the clinicians; therefore, it behooves each laboratory to determine the information and format that are most valuable to their clinicians.

Data Management and Quality Control

Chimerism analysis poses several unique clerical and data management issues for the clinical laboratory. The laboratory must have a way to match donors and recipients and tabulate the results of pretransplant testing. Subsequently, the results of repeated posttransplant testing for each patient must be easily accessible for review with each new sample. This can be accomplished with a well-organized paper folder system or with a paperless computer system. In a movement toward the latter, the Beckman Coulter Vidiera platform includes software specifically designed for chimerism analysis. The software performs the calculations of percent recipient at user-selected loci, prepares a customizable report, and allows easy tracking of patient results over time (Figure 46-7).

Proficiency testing for chimerism analysis is available from the College of American Pathologists. Quality control for specimen switching is inherent to identity testing for chimerism since patients are repeatedly tested over time, and any unexpected result, such as different alleles, is immediately noted and investigated.

Conclusion

The monitoring of chimerism after HCT has become routine to confirm engraftment and detect mixed chimerism. The use of chimerism analysis and its clinical utility are different for myeloablative and nonmyeloablative HCT. Chimerism analysis has clinical utility in guiding the use of therapeutic interventions such as DLI to prevent relapse or graft loss. Although many methods have been applied, PCR amplification of STR loci is the method of choice because it is informative, quantitative, relatively rapid, and sensitive. While the methodology is technically straightforward, the implementation, analysis, and reporting of results are more complicated. The clinical importance of chimerism results warrants the effort required by the laboratory for development of the test. In the future, establishment of guidelines for testing and reporting of chimerism will be highly valuable.

References

1. Bader P, Niethammer D, Willasch A, Kreyenberg H, Klingebiel T. How and when should we monitor chimerism after allogeneic stem cell transplantation? *Bone Marrow Transplant.* 2005;35:107–119.
2. Khan F, Agarwal A, Agrawal S. Significance of chimerism in hematopoietic stem cell transplantation: new variations on an old theme. *Bone Marrow Transplant.* 2004;34:1–12.
3. McCann SR, Crampe M, Molloy K, Lawler M. Hemopoietic chimerism following stem cell transplantation. *Transfus Apher Sci.* 2005;32:55–61.
4. Thiede C. Diagnostic chimerism analysis after allogeneic stem cell transplantation: new methods and markers. *Am J Pharmacogenomics.* 2004;4:177–187.
5. Winiarski J, Mattsson J, Gustafsson A, et al. Engraftment and chimerism, particularly of T- and B-cells, in children undergoing allogeneic bone marrow transplantation. *Pediatr Transplant.* 1998;2:150–156.
6. Baron F, Storb R. Allogeneic hematopoietic cell transplantation as treatment for hematological malignancies: a review. *Springer Semin Immunopathol.* 2004;26:71–94.
7. Slavin S. Graft-versus-host disease, the graft-versus-leukemia effect, and mixed chimerism following nonmyeloablative stem cell transplantation. *Int J Hematol.* 2003;78:195–207.
8. Slavin S. New strategies for bone marrow transplantation. *Curr Opin Immunol.* 2000;12:542–551.
9. Massenkeil G, Nagy M, Lawang M, et al. Reduced intensity conditioning and prophylactic DLI can cure patients with high-risk acute leukaemias if complete donor chimerism can be achieved. *Bone Marrow Transplant.* 2003;31:339–345.
10. 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 of the International Bone Marrow Transplant Registry and the American Society of Blood and Marrow Transplantation. *Biol Blood Marrow Transplant.* 2001;7:473–485.
11. Gardiner N, Lawler M, O'Riordan JM, Duggan C, De Arce M, McCann SR. Monitoring of lineage-specific chimaerism allows early prediction of response following donor lymphocyte infusions for relapsed chronic myeloid leukaemia. *Bone Marrow Transplant.* 1998;21:711–719.
12. Valcarcel D, Martino R, Caballero D, et al. Chimerism analysis following allogeneic peripheral blood stem cell transplantation with reduced-intensity conditioning. *Bone Marrow Transplant.* 2003;31:387–392.
13. Baron F, Little MT, Storb R. Kinetics of engraftment following allogeneic hematopoietic cell transplantation with reduced-intensity or nonmyeloablative conditioning. *Blood Rev.* 2005;19:153–164.
14. Dewald G, Stallard R, Al Saadi A, et al. A multicenter investigation with interphase fluorescence in situ hybridization using X- and Y-chromosome probes. *Am J Med Genet.* 1998;76:318–326.
15. Bennett P. Demystified . . . microsatellites. *Mol Pathol.* 2000;53:177–183.
16. Butler JM. *Forensic DNA Typing: Biology and Technology Behind STR Markers*. San Diego: Academic Press; 2001.
17. Van Deerlin VM, Leonard DG. Bone marrow engraftment analysis after allogeneic bone marrow transplantation. *Clin Lab Med.* 2000;20:197–225.
18. Fehse B, Chukhlovin A, Kuhlcke K, et al. Real-time quantitative Y chromosome-specific PCR (QYCS-PCR) for monitoring hematopoietic chimerism after sex-mismatched allogeneic stem cell transplantation. *J Hematother Stem Cell Res.* 2001;10:419–425.
19. Bader P, Beck J, Frey A, et al. Serial and quantitative analysis of mixed hematopoietic chimerism by PCR in patients with acute leukemias allows the prediction of relapse after allogeneic BMT. *Bone Marrow Transplant.* 1998;21:487–495.
20. Oliver DH, Thompson RE, Griffin CA, Eshleman JR. Use of single nucleotide polymorphisms (SNP) and real-time polymerase chain

reaction for bone marrow engraftment analysis. *J Mol Diagn.* 2000;2:202–208.
21. Fredriksson M, Barbany G, Liljedahl U, Hermanson M, Kataja M, Syvanen AC. Assessing hematopoietic chimerism after allogeneic stem cell transplantation by multiplexed SNP genotyping using microarrays and quantitative analysis of SNP alleles. *Leukemia.* 2004;18:255–266.
22. Alizadeh M, Bernard M, Danic B, et al. Quantitative assessment of hematopoietic chimerism after bone marrow transplantation by real-time quantitative polymerase chain reaction. *Blood.* 2002;99: 4618–4625.
23. Lion T. Summary: reports on quantitative analysis of chimerism after allogeneic stem cell transplantation by PCR amplification of microsatellite markers and capillary electrophoresis with fluorescence detection. *Leukemia.* 2003;17:252–254.
24. Thiede C, Bornhauser M, Ehninger G. Evaluation of STR informativity for chimerism testing—comparative analysis of 27 STR systems in 203 matched related donor recipient pairs. *Leukemia.* 2004;18:248–254.
25. Nollet F, Billiet J, Selleslag D, Criel A. Standardisation of multiplex fluorescent short tandem repeat analysis for chimerism testing. *Bone Marrow Transplant.* 2001;28:511–518.
26. Zhou M, Sheldon S, Akel N, Killeen AA. Chromosomal aneuploidy in leukemic blast crisis: a potential source of error in interpretation of bone marrow engraftment analysis by VNTR amplification. *Mol Diagn.* 1999;4:153–157.

Chapter 47

Specimen Identification Through DNA Analysis

Gregory J. Tsongalis and Andrew Ricci, Jr.

Introduction

Analysis of identity through DNA typing has its origins in forensic identity testing and parentage testing. These same analyses also have found tremendous clinical utility in pathology laboratories for a variety of sample identification applications.[1–9] These applications include identification of mislabeled specimens of all types, identification of histologic "floaters" in surgical pathology specimens, and detection of maternal cell contamination in prenatal specimens. The scientific principles and technologies of identity testing are described in the preceding chapters. This chapter focuses on sample issues and result interpretation, which are specific to sample identification and related applications of identity testing in the clinical laboratory.

Clinical Utility

DNA analysis of polymorphic markers used in forensic identity and parentage testing can also be applied to many situations in the clinical laboratory.[3,6] Chimerism analysis following bone marrow transplant (described in chapter 46) is the most common application of identity testing to a clinical setting and is used to accurately assess the presence or absence of recipient and donor hematopoietic cells. The accurate identification of clinical specimens is clearly essential to proper management of a patient, and misidentification leading to misdiagnosis could result in catastrophic overtreatment or withholding of necessary treatment.

Identification of Mislabeled Specimens

For specimen identification and quality assurance purposes, clinical specimens are labeled with and/or accompanied by patient and specimen information that includes patient demographic data, a unique identifier such as a medical record number, and date and time of collection. Despite written policies, training of personnel, and careful attention by clinical and laboratory personnel to maintenance of proper specimen identification, specimen mislabeling or mix-ups remain a potential problem for the clinical laboratory. Resolution of real or suspected specimen mislabeling is essential to prevent serious clinical consequences. With broader use of identity testing methods in the clinical laboratory setting, these methods are more readily available and can be used to clarify specimen identification questions.

Mislabeled or unlabeled clinical specimens are the most common identification issue for the clinical laboratory. According to federal regulatory standards, improperly labeled or unlabeled specimens cannot be tested by a clinical laboratory unless a firm chain of possession can be documented by the clinician. Consequences of this standard vary by specimen type. Usually, blood samples can be redrawn. In the anatomic pathology laboratory, however, specimen re-collection via re-biopsy may be difficult or impossible. Tissue biopsy specimens may be mislabeled at the time of collection (i.e., multiple biopsies from the same or different patients) or during tissue processing (labeling of tissue holders or glass slides in the laboratory).[1,8,9] In addition to the potential mislabeling of specimens, physicians may be challenged by patients who are in denial about their condition and do not believe that the tissue diagnosis pertains to them. DNA identity testing can be used to certify the origin of biologic specimens in all these scenarios.

Histologic "Floaters"

An uncommon but critical challenge for the anatomic pathologist is the interpretation of fragments of tissue, known as "floaters," that histologically do not appear to be associated with the case at hand. This is not a labeling error or specimen mix-up but rather results from small tissue fragments being "carried over" from one case to the next

during processing in the laboratory. These floaters can be carried over during gross dissection or later during processing of tissue biopsies, which are forms of specimen contamination. This contamination may occur despite rigorous quality control procedures.

In the majority of cases, the anatomic pathologist is able to dismiss the questionable fragment of tissue as not pertinent to the case during routine histologic examination. This is easily accomplished when the floater does not resemble the histology of the remaining tissue, is positively recognized as coming from another case, or presents incongruous histology.[4] However, when the floater is the same type of tissue as the case at hand, it becomes more difficult to simply disregard it as irrelevant. In addition, floaters can represent just a few cells or a cluster of cells that are difficult to classify and evaluate definitively. A common scenario is a floater with malignant potential that is only focally represented in the tissue section. Such tissue cannot be ignored; nor can it be positively interpreted and reported. In the past, when diagnostic certainty could not be applied because of potential floater contamination, the pathologist could only "advise close follow-up and short term re-biopsy if possible." With the advent of identity testing in combination with microdissection, floaters can be definitively resolved.

Maternal Cell Contamination of Prenatal Specimens

In this era of molecular genetic testing, many laboratories are performing prenatal diagnostic testing. Typically this involves analysis of DNA isolated from fetal cells from either chorionic villi or amniotic fluid. Contamination of the fetal cells by maternal cells is a major concern in prenatal diagnostic testing. Such contamination could lead to a misdiagnosis. Identity-testing methods can be used to determine whether maternal cell DNA is present in the fetal DNA.[2]

Evaluation of Hydatidiform Moles

Complete hydatidiform moles arise when one or two sperm fertilize an anucleate ovum, so that the proliferating gestational tissues are completely paternally derived (i.e., androgenetic). A complete mole can be either homozygous, if a single sperm fertilizes an empty ovum and that sperm's haploid genetic material is duplicated, or heterozygous, if two sperm fertilize the empty ovum. Therefore, polymorphic marker analysis of microdissected maternal decidua compared to chorionic villi will show only nonmaternally derived homozygous or heterozygous alleles at multiple loci in the hydropic villi.[5]

Partial hydatidiform moles arise from fertilization of an egg by two sperm and are triploid, with both maternal and paternal genetic material (usually 69XXX or 68XXY). In these cases, polymorphic marker analysis will show three different alleles at multiple loci, one of maternal origin and two of paternal origin.

Specimen Issues

Specimen Selection and Documentation

Selection and verification of the appropriate specimens, tissues, or cells from which to extract DNA are the first and most critical steps in identity testing. The laboratory performing the analysis must know the purpose for the testing, since this will, in part, dictate the samples that need to be tested. For example, to identify a particular individual as the source of a mislabeled tissue specimen, DNA from the specimen must be compared to DNA from a specimen known to be from the individual, such as a blood or buccal swab specimen. In addition, for mislabeled specimens, documentation of specimen handling and processing is useful. In contrast, for the analysis of tissue fragments that are not consistent with the remainder of the specimen section, DNA from the two microdissected tissue areas are tested and compared to each other as well as to an independent specimen known to be from the patient, if available. Analysis of maternal cell contamination of a fetal specimen requires a maternal blood sample in addition to the fetal sample.

Tissue Fixation

While specimen handling and processing are major issues for certain types of molecular genetic testing, the use of smaller polymorphic markers allows identity testing to be performed on specimens that would be less than ideal for other clinical molecular tests. Frequently, identity testing involves the use of formalin-fixed, paraffin-embedded (FFPE) tissue or tissue sections. In some cases, the tissue sections will be stained and mounted on glass slides. It is not difficult to obtain adequate DNA from FFPE tissues for identity testing. It is also feasible to obtain adequate DNA from various microdissection techniques. Fixatives containing heavy metals are not recommended, as these tend to inhibit subsequent enzymatic reactions.

Microdissection of Tissue Sections

Analysis of tissue fragments that are not consistent with the remainder of the specimen section, so-called floaters, should begin with review of the tissue sections by an anatomic pathologist, who can (1) verify the nature of the submitted tissue and (2) demarcate relevant areas of tissue for DNA extraction. Simple procedures are currently available to select and demarcate tissues in paraffin-embedded tissue blocks or from stained or unstained tissue sections

on glass slides. Floaters, other relevant tissues, or both, on a glass slide can be marked with a diamond etching or permanent marker pen. Similarly, faced paraffin blocks can be highlighted with a permanent marking pen. The most suitable tissues for microdissection and DNA extraction should be cellular, nucleated, and nonnecrotic.

Diamond etching pens and permanent marking pens are relatively imprecise instruments, but in many cases of identity testing they are all that is needed for highlighting relevant tissues to be assayed. When greater precision is required, more refined and specific microdissection techniques may be employed that are capable of dissecting one cell or a small group of cells from a complex tissue section. From crude dissecting microscope procedures to use of chemical extractions and laser capture microdissection (LCM), the goal has always been the same: to obtain the correct specimen for analysis.

Chemical Microdissection

The PinPoint Slide DNA Isolation System (Zymo Research, Orange, CA) utilizes a proprietary solution that can be applied to paraffin-embedded tissue sections on a glass slide. The solution can be applied macro- or microscopically. Once the solution has dried, a scalpel blade is used to remove the isolation solution and adherent target tissue for DNA extraction.

Laser Capture Microdissection

LCM permits reliable procurement of pure cell populations from tissue sections. Briefly, a histologic section is placed on the stage of a specialized microscope and an area of interest is brought into view. A cap coated with a thermoplastic film is placed over the cells to be dissected and a low-power infrared laser beam is activated. The laser melts the thermoplastic film, which then binds to the cells beneath it. When the cap is removed, the cells adhere to the cap surface. LCM can dissect large clusters of cells or single cells as a pure population free from contaminating stroma or other cells.

Available Assays

DNA typing harnesses the polymorphic differences between individuals to resolve questions of identity. For both forensic and sample identity testing, the ability to work with minute samples is important, as illustrated by the application of microdissection in resolution of tissue section floater cases. The use of the polymerase chain reaction (PCR) permits the interrogation of nucleic acids extracted from very small amounts of fluid or tissue, as well as from partially degraded samples such as from FFPE tissue. Short tandem repeat (STR) loci can be amplified and the length polymorphisms interrogated by PCR product size analysis. While there are many STR polymorphisms in the human genome, a subset has been commercialized for identity testing purposes.

STR Analysis

Several commercially available primer sets are available for amplification of STR loci, as described in chapters 44 and 46 (Applied Biosystems, ABI, Foster City, CA, and Promega Corporation, Madison, WI). The primer sets coamplify up to 16 STR loci in a multiplex reaction. Fragment-size analysis of the PCR products is performed by gel or capillary electrophoresis on high-resolution DNA sequence detection systems. Commonly, the primers are fluorescently labeled with different fluoro-phores, thereby permitting simultaneous detection of overlapping-size PCR products.

Reverse Dot-Blot

Another commonly used procedure for typing multiple loci simultaneously is the reverse dot-blot procedure, which uses an array of immobilized probes to detect sequence polymorphisms in amplicons. The AmpliType PM+DQA1 amplification and typing kit (Roche Molecular Systems, Indianapolis, IN) has been used routinely by many forensic laboratories (described in chapter 44). This procedure has been applied to numerous cases of clinical specimen identification and can be used with many different specimen types. This kit interrogates six polymorphic loci in a single multiplex PCR assay and provides extremely reproducible results.

Interpretation of Results

Sample Identification

Sample identification determines whether two samples originated from the same or different individuals. Comparison of a tissue or other specimen with an identified specimen, such as peripheral blood, from the potential source patient can be used to confirm the source of the specimen under question. The interpretation of the results involves a comparison of the alleles (STR or sequence polymorphisms) between the samples. The identification of nonmatching alleles, preferably in two or more loci, provides evidence to exclude the individual as the source of the sample or to conclude that the two samples did not come from the same individual. The possibility of microsatellite instability or loss of heterozygosity should be considered if malignant tissues are being compared to normal samples using STR analysis, as described further below. If an exact allele match is observed between two samples, then the likelihood that

Table 47–1. DNA Typing Results for Case 1
STR analysis (PowerPlex2.1 and FFFL, Promega Corporation) of a bronchoalveolar lavage (BAL) shows additional alleles (bold) compared to the patient's peripheral blood at almost all loci, indicating the presence of a second source of DNA. Since the patient was a female, the presence of a Y chromosome product in the amelogenin result for the BAL indicates that the contaminating DNA was from a male.

Locus	Peripheral Blood	Bronchoalveolar Lavage
vWA	16, 18	16, **17,** 18
TH01	6, 10	6, **7,** 10
TPOX	11	**8**, 11
CSF1PO	11, 12	11, 12, **13**
LPL	10	**9**, 10, **11**
F13B	8, 10	**6**, 8, **9**, 10
FESFPS	11, 12	**10**, 11, 12
F13A01	3.2, 7	3.2, 7
D5S818	11	11
D13S317	11, 12	**8**, 11, 12
D7S820	9, 11	9, 11
D16S539	12, 13	**11**, 12, 13
Amelogenin	X	X, **Y**

they are from the same person is high, and an exact probability can be calculated by multiplying the allele frequencies for each allele tested. Allele frequencies for different ethnic groups for commonly used STR loci are published.

Case 1

There are many scenarios in which sample identification can be important. One example is illustrated in the following case. A woman underwent a bronchoalveolar lavage (BAL), which showed carcinoma on cytologic evaluation. Follow-up clinical evaluation of the patient could not identify the source of the tumor cells, so the BAL specimen and a peripheral blood sample from the patient were used for identity testing. The results of DNA analysis indicated that the DNA from the BAL specimen contained a mixture of the patient's DNA and another person's DNA, compared with the patient's peripheral blood (Table 47-1). These results provided an explanation for the diagnosis of carcinoma with negative results on follow-up and prompted an investigation into the source of the additional DNA.

This analysis shows the utility of identity testing for resolution of unexplained findings. Many times a specimen clearly is mislabeled, for example, if there is no label, but other times there may be a suspicion that a specimen is mislabeled when the clinical findings and the pathological results are discordant.

Analysis of Floaters

The analysis of floaters compares the genotypes of the predominant tissue on the slide with the small extraneous tissue fragment to determine whether they match or not. As with other forms of identity testing, differences in at least two markers must be identified to exclude a single individual as the source of the two specimens. Histopathologic review of the case is important before identity testing is performed to verify the nature of the submitted material and to select the appropriate tissues for analysis. The following case illustrates the analysis of a floater.

Case 2

A histologic section of a transverse colon biopsy contained multiple fragments of colonic mucosa showing hyperplastic features and an isolated fragment of glandular-type mucosa showing invasive signet ring adenocarcinoma. The reviewing anatomic pathologist identified the isolated glandular mucosa as probably gastric in origin and suspected that this tissue fragment was a floater and not from the same patient as the remainder of the colonic biopsy tissue. PCR-mediated reverse dot-blot (PM+DQA1) analysis was performed on the fragment of glandular mucosa with signet ring carcinoma and a sampling of the hyperplastic (benign) colonic mucosa. The testing showed different genotypes in four of the six markers, including DQA1, LDLR, HBGG, and D7S8 (bold alleles in Table 47-2), leading to the conclusion that the two tissue fragments originated from different individuals.

Consideration of Mutation

The presence of an allele mismatch at two or more loci is sufficient to conclude that two samples are from two different individuals. When one of the tissues is malignant, it is important to consider that the differences may be due to mutations in the tumor tissue. Nucleotide polymorphisms are highly conserved between normal and tumor tissue from the same patient; therefore, mutations are unlikely to be the cause of a difference between two samples using nucleotide differences, such as the PM+DQA1 kit. On the other hand, microsatellite instability (MSI) and loss of heterozygosity (LOH) occur in tumors and should be considered as causes for allelic differences. MSI is a common feature of some cancers (see chapter 19). For example, approximately 20% of sporadic colon cancers are charac-

Table 47-2. DNA Typing Results for Case 2
DNA analysis of two tissue specimens shows different genotypes at four of six markers tested.

	Tissue	
Marker	Glandular mucosa	Colon
DQA1	**3**, 4.1	**1.3**, 4.1
LDLR	**A**, B	B, B
GYPA	A, A	A, A
HBGG	A, **B**	A, A
D7S8	**B**, **B**	**A**, **A**
GC	A, A	A, A

terized as having a mutator phenotype that results in MSI. Since identity testing relies on polymorphic STR sequences, MSI could be a confounding factor for the interpretation of identity-testing assays. Laboratories should document the genotypic conservation of the STR loci used in identity testing of normal and malignant tissues. Loci known to show LOH or MSI in cancers are not recommended for identity testing.

Analysis of Hydatidiform Moles

Identity testing methods can be used to characterize hydatidiform moles as complete or partial moles, as well as homozygous or heterozygous complete moles.

Case 3

Histologic sections of tissue from a uterine evacuation specimen showed hydropic villi and trophoblastic proliferation consistent with a hydatidiform mole. The patient denied that the uterine evacuation belonged to her. To confirm the identity of the molar tissue, identity testing using STR loci was performed on the molar tissue and a blood specimen from the patient.

The identity testing results showed that at least one maternal allele was present in the molar tissue at every locus tested (Table 47-3) and confirmed that the tissue was derived from the patient and supported a diagnosis of partial hydatidiform mole. Cytogenetic analysis confirmed a triploid 69XXX karyotype. In a partial hydatidiform mole, the female contributes one and only one allele to the partial molar pregnancy, while the remaining alleles are paternally derived. For this case, the D19S433 marker showed four different alleles, of which two match the maternal alleles. This could result from mosaicism in the mole that was missed cytogenetically. This case demonstrates the need for precise clinical information and histologic correlation for correct interpretation of the DNA typing results.

Table 47-3. DNA Typing Results of Case 3
DNA analysis of a hydatidiform mole compares maternal blood with molar tissue. The molar tissue has three alleles at multiple loci, with four alleles at D19S433. These results are consistent with a partial mole with possible mosaicism.

Marker System	Patient	Tissue
D3S1358	15, 16	15, 16, 17
VWA	17, 19	17, 19
D16S539	12, 13	9, 12
D2S1338	20, 25	17, 20, 22
D8S1179	12, 16	12, 14, 16
D21S11	30, 32	30.2, 32
D18S51	13, 14	14, 17
D19S433	14, 15.2	12, 14, 15, 15.2
THO1	9.3	7, 9.3
FGA	21, 26	26, 27

Conclusion

Identity-testing methods are now being used in the clinical laboratory and can be critically applied in cases of questioned clinical specimen identification. A clinical specimen identity issue can arise because of a potential specimen mix-up, a floater, or a patient in denial of his or her diagnosis. Identity testing also has important utility for hydatidiform mole analysis and for the identification of mixed specimens, such as the identification of maternal cell contamination in prenatal specimens. In each scenario, polymorphic DNA marker analysis combined with clinical information and histopathologic review can resolve the majority of cases.

References

1. Abeln ECA, Van Kemenade FD, Van Krieken JHJM, et al. Rapid identification of mixed up bladder biopsy specimens using polymorphic microsatellite markers. *Diagn Mol Pathol.* 1995;4:286–291.
2. Antoniadi T, Yapijakis C, Kaminopetros P, et al. A simple and effective approach for detecting maternal cell contamination in molecular prenatal diagnosis. *Prenat Diagn.* 2002;22:425–429.
3. Hammond HA, Jin L, Zhong Y, et al. Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am J Hum Genet.* 1994;55:175–189.
4. Hunt JL, Swalsky P, Sasatomi E, Niehouse L, Bakker A, Finkelstein SD. A microdissection and molecular genotyping assay to confirm the identity of tissue floaters in paraffin-embedded tissue blocks. *Arch Pathol Lab Med.* 2003;127:213–217.
5. Keep D, Zaragoza MV, Hassold T, Redline RW. Very early complete hydatidiform mole. *Hum Pathol.* 1996;27:708–713.
6. Micka KA, Sprecher CJ, Lins AM, et al. Validation of multiplex polymorphic STR amplification sets developed for personal identification applications. *J Forensic Sci.* 1996;41:582–590.
7. Shibata D, Kurosu M, Noguchi TT. Fixed human tissues: a resource for the identification of individuals. *J Forensic Sci.* 1991;36:1204–1212.
8. Tsongalis GJ, Berman MM. Application of forensic identity testing in a clinical setting: specimen identification. *Diagn Mol Pathol.* 1997;6:111–114.
9. Tsongalis GJ, Wu AHB, Silver H, Ricci A Jr. Applications of forensic identity testing in the clinical laboratory. *Am J Clin Pathol.* 1999;112:S93–S103.

Section VII

HLA Typing

Section Editor: Debra G.B. Leonard

Chapter 48

Molecular HLA Typing

Malek Kamoun and Thomas M. Williams

Molecular Genetics of the Human Major Histocompatibility Complex

Histocompatibility antigens were first described by Snell, using inbred strains of mice. Later, Dausset, Payne, Van Rood, and others discovered that in humans antibodies to white blood cell antigens can be induced after pregnancy or transfusion. These antibodies provided the tools to define the human leukocyte antigens (HLA) (References 1–3 and references therein). Recently, using molecular methods, researchers have defined hundreds of different HLA alleles. The HLA genes reside within a genetic complex referred to as the major histocompatibility complex (MHC). The MHC comprises about 4 megabases (Mb) on the short arm of chromosome 6 (6p21.3). The MHC contains the most polymorphic coding sequences in the human genome. HLA genetic diversity accounts for more than 10% of all genetic diversity observed in the human genome. HLA molecules provide the molecular basis for immunologic self-recognition, and many HLA genes encode proteins that function within immune regulatory networks.

A number of genes in the MHC, arranged in order from the telomeric to the centromeric end, are displayed in Table 48-1. The MHC can be divided into three major regions, each of which controls the production of molecules that have distinct biologic functions.[4] The class I region includes the HLA-A, HLA-B, and HLA-C and related loci, class II includes the HLA-DR, HLA-DQ, and HLA-DP loci, and class III contains genes that encode a variety of proteins including molecules of the complement system (C2, C4, and Bf) and tumor necrosis factors (TNFs). Class I and class II loci encode two distinct classes of highly polymorphic cell surface molecules whose function is to capture and display various antigenic peptides to CD8+ and CD4+ T lymphocytes, respectively. The class II region also includes the genes for the TAP1 and TAP2 peptide transporter, the LMP genes that encode proteosome subunits, and the genes encoding the DMα and DMβ chains (Reference 4 and references therein).

HLA Polymorphism and Nomenclature

HLA genes are codominantly expressed, and most individuals are heterozygous at the classical HLA class I and class II loci. The nomenclature for the alleles of the HLA system is decided by an international nomenclature committee. Antigen names are composed of a letter designating the HLA protein encoded by a locus or loci and a number, for example, A2, B7, Cw4, DR17, and DQ3. Allele names are based on nucleotide sequence variation among alleles identified using DNA-based techniques (discussed below). The number of alleles at the HLA class I and class II loci as well as a listing of serologically defined specificities and their allele equivalents are available online at several Web sites (for instance, the Anthony Nolan Trust Web site at http://www.anthonynolan.com/hig/index.html or the Immunogenetics (IMGT) HLA Database Web site at http://www.ebi.ac.uk/imgt/hla/).

Pattern of HLA Polymorphism

Most of the sequence diversity for the HLA class I loci is localized to the second and third exons and, for the class II loci, to the second exon. HLA diversity is generated through the selection of alleles with nonsynonymous substitutions encoding amino acid differences within the peptide-binding site. This contrasts with the bias against substitutions within the introns and other binding domains. Linear sequences identifying residues critical to the expression of allospecificities are found in the helices of the α1 and α2 domains as well as external portions of the β pleated sheet. Other polymorphic amino acid residues that are important for peptide binding but not defined by alloantibodies are identified by DNA-based techniques.[5–7]

The pattern of allelic sequence diversity for both the class I and class II loci is unusual; most alleles differ from their closest neighbor by multiple substitutions, with some alleles differing in the second and third exons by as much as 15%. This pattern is suggestive of segmental exchange of

Table 48-1. Names for Genes in the HLA Class I and Class II Regions, Ordered from Telomeric to Centromeric Location on Chromosome 6

Name	Molecular Characteristics
HLA-F	Class I–like molecule
HLA-G	Class I–like molecule
HLA-H	Class I pseudogene
HLA-A	Class I α chain
HLA-J	Class I pseudogene
HLA-E	Class I–like molecule
HLA-C	Class I α chain
HLA-B	Class I α chain
MICA	MHC Class I–like molecule
MICB	MHC Class I–like molecule
HLA-DRA	DR α chain
HLA-DRB9	DRB pseudogene, isolated fragment
HLA-DRB3	DR β3 chain determining DR52 and Dw24, Dw25, Dw26 specificities
HLA-DRB2	Pseudogene with DR β-like sequences
HLA-DRB1	DR β 1 chain determining specificities DR1, DR2, DR3, DR4, DR5 etc.
HLA-DQA1	DQ α chain as expressed
HLA-DQB1	DQ β chain as expressed
HLA-DQB3	DQ β-chain-related sequence, not known to be expressed
HLA-DQA2	DQ α-chain-related sequence, not known to be expressed
HLA-DQB2	DQ β-chain-related sequence, not known to be expressed
HLA-DOB	DO β chain
TAP2	ABC (ATP binding cassette) transporter (associated with antigen presentation)
LMP7	Proteosome-related sequence (role in loading class I molecules with peptides)
TAP1	ABC transporter (associated with antigen presentation)
LMP2	Proteosome-related sequence (role in loading class I molecules with peptides)
HLA-DMB	DM β chain (control peptide loading by class II molecules)
HLA-DMA	DM α chain (control peptide loading by class II molecules)
HLA-DPA1	DP α chain as expressed
HLA-DPB1	DP β chain as expressed
HLA-DPA2	DP α-chain-related pseudogene
HLA-DPB2	DP β-chain-related pseudogene

nucleotide motifs between alleles of the same locus. There are also a few examples of interlocus gene conversion, and most are the result of HLA-B/-C recombination. Thus, different HLA alleles of a locus are patchwork (i.e., mosaic) combinations of polymorphisms. The extensive allelic diversity at HLA loci is thought to have been generated by polymorphic sequence motifs generated by gene duplication and recombinational mechanisms such as gene conversionlike events, which have shuffled these polymorphic sequence motifs. Point mutations also contribute to this allelic diversity.[6,7] One consequence of this pattern of patchwork polymorphism is that the sharing of specific epitopes on the HLA molecules leads to antigenic cross-reactivity. Due to epitope sharing, HLA antigens may be arranged in cross-reactive groups, or CREGs.[2,3] CREGs are important in that they allow for greater latitude in donor-recipient HLA matching and at the same time may be used to predict potential problems with graft outcome. A problematic consequence of the combinatorial nature of HLA alleles is that a given pattern of sequence motifs may be consistent with more than a single genotype. This problem of ambiguity in DNA-based typing is discussed below.

HLA Haplotypes

The HLA genes are in linkage disequilibrium and are transmitted in families as haplotypes. An individual inherits one haplotype from each parent. In any family having more than one sibling, there is a 1 in 4 chance that the two siblings will be HLA identical and a 1 in 2 chance that they will share one haplotype. Given the number of existing alleles, the number of possible haplotypes is astronomical. However, due to linkage disequilibrium, the number of haplotypes found in a population is more restricted.[7,8]

The organization of HLA-DQ and HLA-DP loci is conserved within human and other primates, whereas the number of HLA-DR genes varies in different haplotypes. Within the HLA-DR region, the DRB1 gene encodes the DRβ1 chain, which determines the DR private specificities DR1, DR2, DR3, DR4, DR5, and so on. The DRB3, DRB4, and DRB5 genes encode the DRβ3, DRβ4, and DRβ5 chains, which determine the DR52, DR53, and DR51 specificities, respectively (Figure 48-1).

Genetic recombination or crossing over in the HLA region is a relatively rare event, occurring for the most part no more than 1% per meiosis between HLA-A and HLA-B, and between HLA-B and HLA-DR. Recombination also can occur between HLA-A and HLA-C, and between HLA-B and HLA-C (0.6% and 0.2%, respectively). Such recombination can have important clinical implications for transplantation.

Linkage Disequilibrium or Gametic Association

Linkage disequilibrium is an important feature of the HLA system. Gametic associations are regularly found between certain alleles of HLA-A and HLA-B, HLA-C and HLA-B, HLA-B and HLA-DR, and HLA-DR and HLA-DQ. In some so-called extended haplotypes, significant linkage disequilibrium extends over 3 Mb from HLA-A at the telomeric end of the HLA region, to HLA-DP at the centromeric end, including HLA class III genes (Reference 8 and references therein). It is this linkage disequilibrium that allows the identification of bone marrow or stem cell donors in population registries who are matched at multiple loci with a given patient (see Histocompatibility Assessment in Unrelated Donors below).

HLA Gene Frequencies in Human Populations

Although a very large number of alleles (e.g., >400 for HLA-B and HLA-DRB1) can be found in the global population,

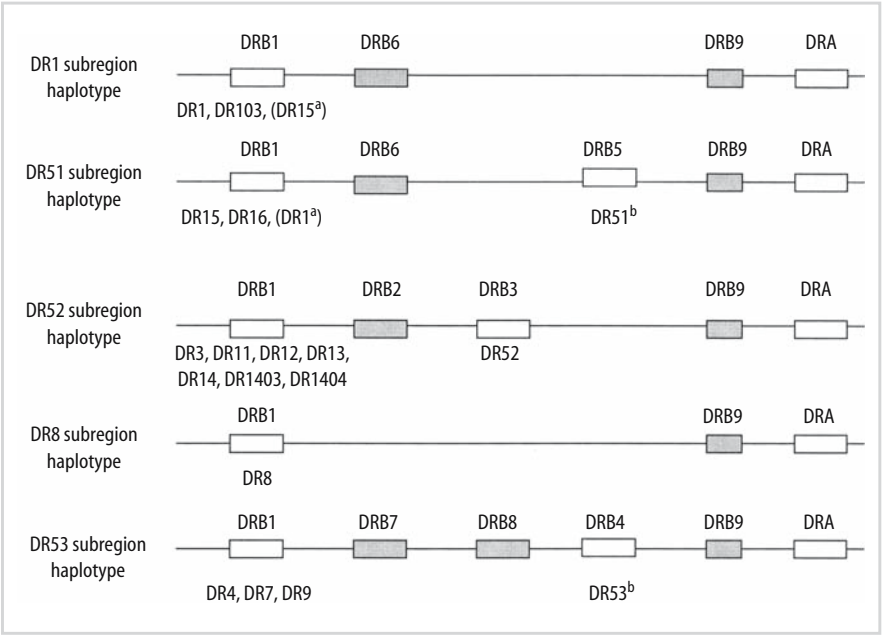


Figure 48-1. HLA-DR region and encoded allotypes. Pseudogenes are indicated by shaded boxes. Expressed genes are indicated by open boxes. The serologic specificity encoded by a gene is given underneath each expressed gene box. [a]Rarely observed haplotypes. [b]DR51 and DR53 may not be expressed on certain haplotypes. (Adapted from Little AM, Parham P. Polymorphism and evolution of HLA class I and II genes and molecules. *Rev Immunogenet.* 1999;1:105–123.)

a much smaller number (e.g., 30–50 for HLA-DRB1) is present in most individual populations. Importantly, different populations tend to have different frequency distributions of alleles and exhibit different patterns of linkage disequilibrium. This variability exists among both racial and ethnic groups.

Clinical Utility of HLA Typing

Solid Organ Transplantation

Kidney and Pancreas

Renal transplantation is carried out successfully using living and cadaveric donors.

HLA Matching of Living Related Donors

Transplants are performed using donor-recipient pairs sharing HLA-A, HLA-B, and HLA-DR antigens from one or two haplotypes. The analysis of HLA matching for kidney graft survival over the period of 1985–1999 is based primarily on serologic typing data. The half-life survival for a graft from an HLA-identical sibling is 23 years, compared to a one-haplotype-related donor, which has a half-life of 12.8 years.[6,9]

HLA Matching of Cadaveric Donors

Data excerpted from the United Network for Organ Sharing (UNOS) transplant registry[9] indicate that HLA serologic matching appears to have an effect on long-term graft survival of kidneys from cadaveric donors. Although the 1-year graft survival is not very different in cases involving a complete match from those who are completely mismatched, after 5 years the gap in percent survival widens significantly, suggesting that the immunosuppressive drugs are potent in avoiding early graft loss due to acute rejections. However, as the drug dosages are tapered over time, the HLA matching effect becomes significant. Patients receiving organs with zero mismatches for HLA-A, HLA-B, or HLA-DR antigens achieve a graft survival half-life of 11.3 years compared to a half-life of 6.3 years for grafts that were completely mismatched for these antigens. HLA matching is especially beneficial in second transplants and in patients with preformed antibodies. The effect of HLA-A, HLA-B, and HLA-DR matching remains significant even with the most recent forms of immunosuppression.[9]

UNOS HLA Matching Algorithm for Allocation of Cadaveric Kidneys and Pancreases

The selection of a recipient for any given random cadaveric donor is based on an HLA-matching algorithm defined by UNOS (http://www.unos.org). The current allocation policy considers the degree of HLA matching at HLA-A, HLA-B, and HLA-DR. The organs are distributed locally first, and if no match is found, they are offered regionally, and then nationally, until a recipient is found. A nationwide organ sharing is mandatory for identified donor-recipient pairs with zero HLA-A, HLA-B and HLA-DR mismatches; a local and regional organ sharing is based on one or two HLA-DR mismatches. All other match grades are allocated without regard to mismatching. The main purpose of this algorithm is to try to transplant more minority patients while at the same time not reducing overall graft outcome.

Transplantation of Other Solid Organs

The effect of HLA-A, HLA-B, and HLA-DR matching on the survival of heart and lung transplants is statistically significant. However, the effect of HLA matching for liver transplants is unclear.

Hematopoietic Cell Transplantation

HLA Typing Requirements

Typically, an HLA typing for hematopoietic cell transplantation (HCT) would initially include low-resolution typing for HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, and HLA-DQB1 loci. Allele-level typing is subsequently performed in the final identification of suitable donors for HCT.

Three different categories of donors are usually considered in the following order of preference: HLA genotypically identical siblings, HLA mismatched relatives, and matched or mismatched unrelated donors.

Histocompatibility Assessment in HLA-Matched Relatives

The goal when screening for an HLA-matched sibling donor is to identify which of any siblings have inherited the same HLA haplotype from their parents. This requires typing all siblings and, if possible, both parents for HLA-A, HLA-B, HLA-DRB1, and HLA-DQB1. Given the current average family size in the United States of 2.7 siblings, the approximate probability that a patient will have an HLA match within the family is 30% to 35%. For those patients who do not have an HLA-matched sibling within the family, an HLA-mismatched relative who shares one haplotype can be considered. However, the clinical data clearly indicate that with increasing disparity for HLA-A, HLA-B, HLA-DRB1, or HLA-DQB1 loci, there is increased risk of graft-versus-host disease (GVHD) following HCT.

Histocompatibility Assessment in Unrelated Donors

When a matched sibling donor does not exist for a patient requiring allogeneic HCT (70% of cases), searching for extended family members or donors from unrelated bone marrow registries would be the next option. Registries of volunteer bone marrow donors as well as cord blood exist in most developed countries. The largest of these is the National Marrow Donor Program (NMDP) in the United States, which has more than 4 million registered donors (http://www.marrow.org).

The probability of finding a matched donor at HLA-A, HLA-B, HLA-DRB1, and HLA-DQB1 depends on the ethnic origin of the patient and the composition and size of the donor registry being searched. Currently, an acceptable donor can be identified for approximately 65% of patients. However, patients belonging to racial groups that are not well represented in the registries have a considerably decreased probability of matched donor identification.[10]

Relative Impact of HLA Disparities

HLA compatibility affects not only the ability to achieve sustained engraftment of donor HCT but also the risk of developing acute and chronic GVHD.[6,11] Posttransplant risk of graft failure, GVHD, and mortality can be affected by quantitative and qualitative characteristics of donor-recipient HLA mismatching. Moreover, studies have reached different conclusions regarding the relative contributions of HLA class I and class II mismatching because of population-based differences in the specific HLA-mismatch combinations between patients and donors (Reference 11 and references therein). In an analysis of HCT for chronic myelogenous leukemia (CML), the risk of graft failure is affected primarily by donor disparity for HLA class I including HLA-C. The incidence of rejection correlates with the number of donor-incompatible alleles. The incidence of graft rejection is 0.7% for zero, 8% for a single, and 19% for multiple class I allele donor incompatibilities. Donor disparity for class II does not increase the risk of rejection.[6,11]

The level of matching is more critical for HCT than for solid organ transplantation (SOT) because of the risk of GVHD. GVHD is a major cause of mortality in HCT and is more frequent than allograft rejection. Analysis of HLA and GVHD focuses on mismatches in the host recognized by donor T cells. Generally, host disparity for class II is thought to convey greater risk for GVHD than class I disparity. In an analysis of HCT for CML, patients with a single class II mismatch at HLA-DR or HLA-DQ have a hazard ratio of 1.8 compared to HLA matches. Single class I mismatches are well tolerated with respect to GVHD; however, combined mismatching at class I and class II confers a hazard ratio of 2.0. Allele mismatches for HLA-A, HLA-B, HLA-C, HLA-DRB1, and HLA-DQB1 are also a significant factor for survival. Patients mismatched for more than one class I allele and those mismatched at both class I and class II alleles have a significantly lower survival than patients and donors fully matched for HLA-A, HLA-B, HLA-C, HLA-DRB1, and HLA-DQB1. A single class I mismatch or single class II mismatch does not appear to affect survival.[6,11] It is likely that different modes of patients' immunosuppression after transplant would have an impact on the relative risk associated with various HLA loci. Criteria for "permissible" mismatching in HCT are center dependent. However, very few centers would perform HCT across more than one HLA mismatch at HLA-A, HLA-B, HLA-DR, or HLA-DQ. The increased use of DNA-based HLA typing should help identify permissible mismatches.

Table 48-2. Examples of HLA Testing for Autoimmune Diseases

Disease	HLA Test	Comment
Ankylosing spondylitis (AS)	HLA B27	B27 is found in >95% of patients with AS but also is found in 7% of the white population. The presence of B27 is not diagnostic. The absence of B27 tends to exclude AS.
Narcolepsy	HLA-DR2, DQB1*0602	DR2 (DRB1*1501-DQB1*0602) is present in nearly all patients with narcolepsy but also is found in ~25% of the white population. The presence of DQB1*0602 is not diagnostic of narcolepsy, but the absence of DQB1*0602 tends to exclude narcolepsy.
Insulin-dependent diabetes mellitus (IDDM)	HLA-DQ2, DQ3	This association may be used to assess the clinical risk of IDDM in first-degree relatives of patient but is of no value in routine testing of patients.
Familial hemochromatosis	Family study HLA-A, HLA-B, HFE genotyping	Testing for disease risk can now be carried out by direct *HFE* genotyping.
Rheumatoid arthritis (RA)	HLA-DRB1 typing	HLA-DR4 is found in approximately 70% of patients with RA and 25% of the unaffected population. DRB1*0401, 0404, 0405 are the main susceptibility subtypes, whereas DRB1*0402, 0403, and 0407 are not associated with an increased risk of RA.

Interestingly, recent studies indicate a low incidence of GVHD associated with the use of umbilical cord blood (UCB) derived stem cells and show that two or less HLA mismatches were correlated with a high probability of survival, as long as the recipient received an optimal CD34 cell dose adjusted for the recipient body weight.

HLA Alleles and Disease Associations

Specific alleles of HLA molecules are associated with certain diseases. Associations have been shown to be of two types: those in which the disease is associated with class I antigens and those having strong associations with class II antigens (Table 48-2). Several autoimmune diseases are strongly associated with alleles of HLA class II loci, suggesting the possibility that some alleles of HLA class II molecules can selectively present autoantigens to T cells. Importantly, as shown in Table 48-2, some HLA disease associations are of diagnostic importance. The best known is the association of HLA-B27 with ankylosing spondylitis (reviewed in Reference 12).

Mechanisms for the Association of HLA Alleles with Autoimmune Diseases

The biological mechanisms underlying the associations between HLA antigens and various diseases remain mostly unknown. HLA-associated diseases are caused by interplay of multiple genes and environmental factors, in which HLA genes most often confer the strongest genetic predisposition. Based on results of a genomewide search for susceptibility genes, HLA accounts for more than 50% of the genetic risk in type 1 diabetes. Similarly, HLA accounts for more than 90% of the genetic risk in narcolepsy.[12]

Association of HLA Class II Alleles with Autoimmune Diseases

Susceptibility to rheumatoid arthritis (RA) is strongly associated with the DRB1*0401, *0404, *0405, and *0408 alleles. On the other hand, DRB1*0402, *0403, and *0407 are not associated with RA and may instead confer protection. A possible role of pocket 4 (residues 67, 70, 71, and 74 of the β chain) of DR4 molecules has been suggested.

Similarly, in type 1 diabetes it appears that both susceptibility and protection are associated primarily with some class II HLA molecules. Association is strongest with heterozygosity for HLA-DR3- and HLA-DR4-bearing haplotypes. It is currently believed that HLA-associated predisposition to type 1 diabetes is mainly determined by the combination of HLA-DQ or HLA-DR molecules or both.[12] This may explain differences among populations with respect to susceptibility associated with particular HLA-DQ or HLA-DR molecules. The HLA-DQ molecule associated with susceptibility often shares an amino acid other than aspartic acid at position 57 on the HLA-DQβ chain that is located in pocket 9 of the peptide-binding cleft. On the other hand, those HLA-DQ molecules associated with protection carry aspartic acid at this position (reviewed in Reference 12).

In celiac disease, a strong association is observed with the trans-pairing of HLA-DQα and HLA-DQβ observed in HLA-DR3-bearing haplotype or heterozygotes for HLA-DR5 and HLA-DR7.[12]

Other Applications

HLA-matching algorithms have been used for patients receiving repeated infusions of platelet concentrate who become refractory due to the generation of antibodies to

HLA class I antigens expressed on the surface of platelets. Other important applications include vaccine clinical trials for patients with melanoma and HIV infection, in which vaccines rely on defined HLA-restricted epitopes. Identification of the appropriate HLA type is an important step in identifying patients who could benefit from these therapies.

Available Assays

Virtually all methods used widely in clinical molecular laboratories have been applied to the problem of identifying HLA alleles (Reference 6 and references therein). Strategies based on allele-specific DNA amplification, oligonucleotide probe hybridization, and DNA sequencing have become the most common methods used in the field. Each of these methods relies on isolation of genomic DNA from the tested individual with polymerase chain reaction (PCR) amplification of relevant regions of an HLA gene. Samples for testing are usually peripheral blood but may be any nucleated cell or tissue. Laboratories typically amplify at least exons 2 and 3 of class I genes and exon 2 of class II genes, but may prepare larger amplification products that include exon 4 and 5 of class I genes for DNA sequencing. Amplification primers may be located within exons or introns. Intronic primers allow complete analysis of exons and inspection of exon-intron junctions for splice-site mutations. Primers must be chosen with care to achieve locus-specific amplification since the HLA loci are the products of gene duplication and divergence and retain substantial homology. Further, many HLA loci have closely related pseudogenes that do not encode functional polypeptides but may result in nonspecific PCR products. Each of the methods discussed below generates a large amount of data for each patient that must be compared to extensive lists of potential alleles, hybridization probe "hit" patterns, or allele sequence libraries to assign an allele type. While this may be done manually, most laboratories use software packages that must be updated frequently to include newly recognized alleles.

Resolving Power of Methods

Laboratories tailor HLA-typing assays to the specific clinical applications discussed above. For example, assays helpful in evaluation of narcolepsy may test only for presence or absence of DQB1*0602. Matching unrelated donors and recipients for HCT requires allele-level (high-resolution) typing at several class I and II loci. Conversely, low-resolution typing at the allele-group or serologic-equivalent level is typically performed for renal transplantation. Choice of the methods discussed below depends on typing volume, turnaround requirements, and the resolution needed.

Sequence-Specific Primer Polymerase Chain Reaction

The sequence-specific primer polymerase chain reaction (SSP-PCR) method employs pairs of PCR primers chosen so that their 3′-most nucleotide or nucleotides are complementary to a polymorphic position that distinguishes an allele or an allele group from other alleles (Reference 6 and references therein). If the individual possesses the allele(s) of interest, the PCR will lead to a product whose size and presence is typically identified by agarose gel electrophoresis or can be detected in real time.[6,13] If one or both primers are not complementary to their targets at their 3′ ends, the reaction will fail. By choosing many pairs of primers in independent reactions to cover all allele groups, laboratories can perform a typing. Laboratories include a second set of generic primers in each reaction for a control gene, as a DNA amplification control. Because of the large number of allele groups, about 100 PCR reactions are necessary to perform a low-resolution HLA-A, HLA-B, and HLA-DRB1 typing for a single patient. Additional reactions are required if allele-level typing is desired. Thus SSP-PCR typing requires considerable thermal cycler resources. SSP-PCR is suitable for low to moderate typing volumes in laboratories. Since it can be performed in a few hours, SSP-PCR can be used to identify HLA alleles in time-sensitive situations such as in deceased donors prior to renal transplantation.

Sequence-Specific Oligonucleotide Probe Hybridization

For sequence-specific oligonucleotide probe hybridization (SSOPH), PCR products including the relevant regions of a particular HLA gene are hybridized to a large panel of oligonucleotide probes that cover the known polymorphisms (Reference 6 and references therein). Hybridization and washing conditions of high stringency allow detection of single nucleotide differences. Most laboratories detect bound probes colorimetrically or via chemiluminescence. The reactivity of the entire panel is analyzed against expected results for particular alleles or groups of alleles to assign an allele type. Forward SSOPH is useful for large-scale, batched testing, with the PCR products from many individuals immobilized on a membrane, with subsequent hybridization with a single probe. By preparing many membranes, the laboratory can test the entire panel of probes. Reverse SSOPH assays employ multiple probes immobilized on a membrane strip for hybridization to the PCR products from a single individual.[6] Whereas reverse dot-blot and SSOPH methods typically employ enzyme labels and colorimetric substrates that require subsequent development, other SSOPH methods employ microspheres; up to 100 different populations of microspheres can be mixed together and distinguished by their unique fluores-

cence signatures when analyzed by a fluoroanalyzer. Thus all variants in the PCR products are analyzed simultaneously and the entire assay is carried out in a single vessel with the addition of a single reagent. Reverse SSOPH assays can be challenging to optimize since all probes must attain specificity under identical hybridization conditions. Reverse SSOPH reagents are generally available commercially. This type of method is very attractive for rapid low- to mid-volume testing. SSOPH testing can yield low-resolution results or allele-level results with larger panels of probes.

DNA Sequencing

DNA sequencing of PCR products is a method increasingly attractive for allele-level typing because of the large numbers of alleles known to exist at several of the HLA loci. In theory, any known or new allele should be detected by this method if defined by polymorphisms within the nucleotide sequence amplified by the two PCR primers. Recombinant DNA polymerases for sequencing reactions, improved fluorescent dyes for dideoxynucleotides, powerful analysis software, and capillary electrophoresis instruments have made DNA sequencing technically easier and less labor-intensive for complex clinical molecular tests such as HLA typing.[14] Careful selection of PCR primers to ensure roughly equal amplification of the two alleles present in an individual is a crucial requirement for this method. Otherwise heterozygous peaks in the electropherograms will not be reliably identified. DNA sequencing requires less modification in response to new alleles than other methods and is a powerful and general way of performing allele-level typing.

Other Methods

Early in the development of DNA-based HLA typing, restriction fragment length polymorphism (RFLP) analysis was employed. This approach declined in importance as laboratories realized that the complexity and number of HLA alleles outstripped the power of RFLP analysis. Reference strand conformation analysis is a method employing the analysis of heteroduplexes prepared by mixing PCR products from the tested individual with those prepared from a reference allele. The migration of the heteroduplexes is assessed on an acrylamide gel and compared to a library of migration profiles.[6] HLA alleles are likely to have unique migration patterns, especially when tested with several reference alleles. This is a fairly rapid method and is a simple and elegant way to assess whether a donor and a recipient are identical at a locus. Rare alleles not represented in the library with known migration patterns are difficult to identify. Other methods of HLA typing based on mass spectroscopy and DNA microarrays are in developmental forms.

Table 48-3. HLA Nomenclature

MHC	Major histocompatibility complex
HLA	Human leukocyte antigen
HLA-A	An HLA genetic locus within the MHC
HLA-A24	A serologically defined antigen encoded by HLA-A
HLA-A*24	A group of alleles that encode molecules with A24 specificity
HLA-A*2405	A specific HLA-A*24 allele
HLA-A*240301	An HLA allele that differs from A*240302 by a silent polymorphism
HLA-A*2409N	A null allele
HLA-A*24020102L	An allele differing from A*24020101 by a polymorphism outside the coding region and encoding an antigen with reduced expression

Source: Immunogenetics (IMGT) HLA Database Web site available at: http://www.ebi.ac.uk/imgt.

Interpretation of Test Results

Antigen-Level and Allele-Level HLA Typing

Some of the codes currently adopted in designating HLA alleles are shown in Table 48-3. The alleles of each of the HLA antigens are numbered based on the original serologic nomenclature. HLA alleles are designated by a superscripted asterisk after the locus of origin and a number corresponding to the particular allele (e.g., HLA-A*0201), as shown in Table 48-3. The parallel testing using serologic and DNA genotyping of HLA alleles has led to the use of one nomenclature for the description of low-resolution typing where the HLA assignment might include more than one possible related allele (e.g., HLA-A A2) and a nomenclature reflecting the high-resolution allelic typing. Thus, different HLA allele subtypes (e.g., for A2) can appear indistinguishable when tested by serology or with a limited panel of nucleic acid probes, so a generic or low-resolution typing is obtained. The naming of new HLA alleles is decided by an international nomenclature committee; however, the workshop prefix "w" and other designations often precede formal naming by the international nomenclature committee. As a result, the literature can be confusing for the uninitiated. Another feature of the nomenclature is the Bw4 and Bw6 specificities. All HLA-B alleles encode one of two possible epitopes located on the alpha 1 helix that were originally defined by serology as Bw4 and Bw6 specificities. The "w" prefix is retained in this case to distinguish them from true HLA alleles. Annual HLA nomenclature reports with frequent updates are available at several Web sites (HLA Informatics Group, http://www.anthonynolan.com/hig/index.html, or IMGT/HLA Database, http://www.ebi.ac.uk/imgt/hla/).

Correlation of Serologic and Nucleic Acid–Based HLA Typing

The HLA Dictionary provides a list of the World Health Organization (WHO)-assigned serologic designations and the actual serology test results for many alleles. For most of these alleles, the serologic types are known to correspond to the first two digits of the molecular type. For those molecular types for which no WHO assignment or serologic typing results are available, it is appropriate to use the two-digit low-resolution type for conversion. Most of the time molecular typing gives clear results that can readily be interpreted, even when serologic typing is ambiguous.

Conversion of HLA Molecular Types for UNOS Matching and Data Entry into UNet

For cadaveric kidney transplantation, matching at the antigen level is still considered the standard method by which donor kidneys are allocated through UNOS. UNOS provides a list of antigens that may be entered through UNet; this list forms the basis of the HLA-matching algorithm for cadaveric kidney and pancreas allocation (http://www.unos.org/). The list of antigens and criteria are reviewed annually. While molecular class I typing is more robust and more accurate than serologic typing, a definition of equivalent "antigens" is needed for the purpose of proper organ allocation. When allelic nomenclature was initiated, it was intended to reflect the existing serological nomenclature. Thus, for most types, low-resolution (two-digit) class I molecular typing provides the appropriate serologic equivalent. However, as new alleles were discovered, names of alleles were derived based on nucleotide sequence similarities that did not always result in similar serology. Thus, in many cases, it is difficult to derive a serologic designation based on an interpretation of a molecular HLA type. The low-resolution molecular types B*15 and B*40 represent complex examples. The use of the B15, B40, or other broad antigen designation can cause incorrect donor-recipient "matching" and can render molecular typing less specific and less accurate than serologic typing for the purpose of organ allocation.

Because of these complexities, UNOS has developed guidelines for assigning antigen equivalents (http://www.unos.org/). UNOS standards require laboratories to report the correct serologic equivalents for the UNOS application to ensure the most appropriate allocation and to allow screening for unacceptable antigens.

Another problematic issue involves the molecular detection of serologically null alleles (see next section). For instance, a patient with the null variant of A24 could make an antibody to A24. A donor with A24 should not be considered to be a zero mismatch for a recipient with an A24 null allele. Ideally, null alleles should be identified either by use of parallel serologic typing or by use of molecular kits that identify null alleles. All confirmed null alleles should be entered as blanks.

Null Alleles

Most HLA alleles are defined by polymorphisms that are essentially missense mutations leading to amino acid substitutions. Null alleles are variants that result in reduced or absent expression of mature class I and II proteins on the plasma membrane. Null alleles generally result from either alterations at exon-intron boundaries that cause aberrant RNA splicing or nonsense mutations that lead to truncated polypeptides. Most of the null alleles occur within the HLA-A and HLA-B loci; however, they are found at other class I and II loci as well. The nucleic acid sequences that characterize some of these null alleles may not be in the genetic region normally targeted by DNA typing; thus, without serologic information, these null alleles may go undetected. However, once a null allele is identified by serology and characterized by nucleic acid sequencing to identify the source of the lack of expression, the null allele can be readily detected in subsequent individuals by DNA testing procedures alone. Testing strategies that do not distinguish between a null allele and its closely related expressed counterpart can lead to clinically significant errors. Fortunately, the allele frequencies of most null alleles appear to be less than 0.001. However, exceptions occur: DRB4*0103101N is relatively common and C*0409N is associated with HLA-B*4403 carrying haplotype in some ethnic groups.

New Alleles

Laboratories performing clinical testing will encounter patients and donors who appear to have heretofore unknown alleles. New alleles, which appear to have arisen relatively recently, can be observed at relatively high rates, particularly in various indigenous populations.[7] New alleles may be suspected when SSP-PCR assays result in unexpected negative or positive reactions or reactions with products of unexpected size. For example, a typing with 7 positive reactions consistent with the presence of B*51 and B*40 alleles and a single reaction indicating a B*08 allele may suggest the presence of a new allele. Similarly, unusual SSOPH patterns for the panel of oligonucleotide probes used may indicate that a new allele is present. DNA sequencing assay results will indicate the presence of a nucleotide or nucleotides not known to be possible at a particular position in a library of known alleles. Finally, typings that indicate that an individual is heterozygous for two alleles that are each uncommon are sometimes alternatively explained by the presence of a known common allele and a new allele whose sequence is not present in a laboratory's library of "possible" sequences. Since typing strategies often rely on heterozygous PCR products, it is often not possible to determine

which of the two identified alleles is new. Thus, a method of isolating suspected new alleles in a hemizygous form via traditional cloning techniques, allele-specific amplification, or single-allele amplification is necessary for its unambiguous characterization. The sequences of potential new alleles should be communicated to GenBank to obtain accession numbers and then to the IMGT HLA Database (http://www.ebi.ac.uk/imgt/hla/). The WHO HLA Nomenclature Committee will review the data submitted and assign the allele a unique number. The IMGT will display an authoritative sequence for the allele and information about the ethnicity and HLA alleles identified at other loci in the individual with the new allele.

HLA Typing Problems

Ambiguities

Two major sources of ambiguity in HLA typing are polymorphisms outside the gene regions targeted by testing and the lack of information about phase in heterozygous individuals. Many laboratories choose tests designed to analyze exon 2 of class II genes and exons 2 and 3 of class I genes, since these regions include most of the known polymorphisms and encode domains that interact with bound peptides. However, polymorphisms occur in other exons as well. When a method does not test for a possible polymorphism, the typing is necessarily ambiguous. The laboratory should discuss with clients whether an ambiguity is likely to be clinically relevant, although the impact on transplant outcome of many known polymorphisms in other exons is unknown.

The second major reason for ambiguous typings derives from the fact that heterozygous typing data often is entirely consistent with two, three or more possible pairs of alleles rather than just one allele pair. If the PCR products prepared include the sequences of both alleles in an individual and contain several heterozygous possibilities, it is often not possible to know the cis/trans relationships of the heterozygosities. These ambiguous heterozygous combinations can be solved in several ways. Allele-specific amplification leading to PCR products from one of the two alleles will resolve ambiguities. Family studies will resolve heterozygosity ambiguities if appropriate relatives are available for testing. Traditional cloning techniques are a general solution but are time and labor consuming. Single-allele amplification is a general and rapid, but expensive solution.

Recombination

The fact that recombination within the MHC sometimes occurs during meiosis is discussed above. When recombination appears to have occurred within a family, interpretation of results can be difficult. Before accepting that recombination explains the results seen, laboratories should consider the possibilities of typing error and false paternity as alternatives.

False Paternity

HLA typing is a form of identity testing, and it may reveal false paternity when members of a family are tested. Our experience is that many transplantation clinicians do not believe it is helpful to disclose to families before testing that discovery of false paternity is possible. The ethical considerations associated with this issue are complex. When results are consistent with false paternity, the findings can be discussed with the ordering healthcare provider verbally.

Laboratory Issues

Quality Control

Controls

Laboratories find assembling a comprehensive array of positive controls for the 1400 known class I and II alleles difficult or impossible. Most laboratories acquire examples of the more-common alleles for use in quality control over time in the course of clinical and research typing. The NMDP and the American Society for Histocompatibility and Immunogenetics (ASHI) maintain a cell repository of a subset of known alleles (http://www.ashi-hla.org/). If controls for an allele cannot be obtained, alternative approaches include testing reagents such as SSOPH probes with related targets that have an identical sequence at a site of interest. Since HLA alleles in aggregate form a combinatorial array of possible sequences in the polymorphic regions, there are generally many opportunities to use surrogate alleles as positive controls in this manner.

Contamination

Potential contamination of pre-PCR work areas and reagents with extraneous genomic DNA or previously amplified PCR products is a problem HLA laboratories face in common with other clinical molecular laboratories. Appropriate measures to prevent contamination have been widely described (see chapter 49); such information is available at the ASHI Web site (http://www.ashi-hla.org/). Testing for contamination arising from PCR products prepared in SSP-PCR may be difficult. The products prepared in the many reactions necessary to type an individual vary in size and composition so that a general way to test for their presence may be difficult to devise.

Certifications of Laboratories and Professionals

Clinical histocompatibility laboratories are high-complexity laboratories that must be licensed under CLIA.

Laboratories may be inspected and accredited by the College of American Pathologists (CAP) or by the ASHI. UNOS, NMDP, and CLIA have all designated ASHI with deemed status for purposes of accreditation of HLA laboratories. Laboratories are generally directed by individuals with PhD or MD degrees or both. ASHI administers a program to assess the qualifications of doctoral-level individuals to direct ASHI-accredited laboratories. Directors can be certified by examination administered by the American Board of Histocompatibility and Immunogenetics (ABHI). The ABHI also certifies laboratory staff as histocompatibility technologists and specialists (http://www.ashi-hla.org/). The National Credentialing Agency and the American Society for Clinical Pathology have certification programs in molecular pathology for technologists.

Proficiency Testing

Three proficiency testing programs are available in the United States. The CAP, the ASHI, and the Southeastern Organ Procurement Foundation offer comprehensive programs to assess laboratories' ability to correctly identify HLA alleles. Proficiency testing samples are typically peripheral blood samples for which laboratories can perform low-resolution or allele-level typing at one or several of the HLA class I and class II loci. These organizations also offer challenges for other histocompatibility laboratory tests such as cross matching. Results from a number of years of DNA-based HLA allele identification proficiency testing challenges have demonstrated that participating laboratories generally achieve 90% to 100% consensus. The University of California, Los Angeles (UCLA) has offered an international cell exchange program for many years. Laboratories are challenged to correctly type samples that often include unusual or recently described alleles. An important attribute of the UCLA cell exchange program is its ability to allow correlation of serologic and DNA-based results for a tested sample.

References

1. Snell GD. The Nobel Lectures in Immunology. Lecture for the Nobel Prize for Physiology or Medicine, 1980: studies in histocompatibility. *Scand J Immunol.* 1992;36:513–526.
2. Amos DB. Fundamental antigens of HLA. *Hum Immunol.* 1991;30:236–246.
3. Van Rood JJ. HLA and I. *Annu Rev Immunol.* 1993;11:1–28.
4. Rhodes DA, Trowsdale J. Genetics and molecular genetics of the MHC. *Rev Immunogenet*. 1999;1:21–31.
5. Madden DR. The three-dimensional structure of peptide-MHC complexes. *Annu Rev Immunol.* 1995;13:587–622.
6. Erlich HA, Opelz G, Hansen J. HLA DNA typing and transplantation. *Immunology.* 2001;14:347–356.
7. Little AM, Parham P. Polymorphism and evolution of HLA class I and II genes and molecules. *Rev Immunogenet.* 1999;1:105–123.
8. Dawkins R, Leelayuwat C, Silvana G, et al. Genomics of the major histocompatibility complex: haplotypes, duplication, retroviruses and disease. *Immunol Rev.* 1999;167:275–304.
9. United Network for Organ Sharing (UNOS). *2000 Annual Report, the U.S. Scientific Registry of Transplant Recipients and the Organ Procurement and Transplantation Network. Transplant data 1992–2001.* U.S. Department of Health and Human Services, Health Resources and Services Administration. 2005. Available at: http://www.unos.org/
10. Beatty PG, Mori M, Milford E. Impact of racial genetic polymorphism on the probability of finding an HLA-matched donor. *Transplantation.* 1995;60:778–783.
11. Petersdorf EW, Anasetti C, Martin PJ, et al. Tissue typing in support of unrelated hematopoietic cell transplantation. *Tissue Antigens.* 2003;61:1–11.
12. Thorsby E. Invited anniversary review: HLA associated diseases. *Hum Immunol.* 1997;53:1–11.
13. Faas SJ, Menon R, Braun ER, et al. Sequence-specific priming and exonuclease-released fluorescence detection of HLA-DQB1 alleles. *Tissue Antigens.* 1996;48:97–112.
14. Wu J, Bassinger S, Griffith BB, et al. Analysis of HLA class I alleles via direct sequencing of PCR products. In: Hahn AB, Land GA, Strohman RM, eds. American Society for Histocompatibility and Immunogenetics. *ASHI Laboratory Manual*. 4th ed. 2000.

Section VIII

Laboratory Management Issues

Section Editor: Debra G.B. Leonard

Chapter 49

Molecular Pathology Laboratory Management

Hanna Rennert and Debra G.B. Leonard

Introduction

Currently more than 800 laboratories perform nucleic acid–based tests of human samples for the identification of genetic diseases, malignancies, infectious organisms, patient or sample identification, and human leukocyte antigen (HLA) typing.[1] The operation of a clinical molecular pathology laboratory requires integration of expertise in medical, scientific, and clinical molecular pathology, resources including facilities, equipment, and personnel, and skills in organization, administration, management, and communication. Quality service is achieved by adherence to clinical laboratory regulations, from specimen collection and processing to reporting of patient results. This chapter reviews fundamental knowledge important for the management and operation of a clinical molecular pathology laboratory.

The Role of the Molecular Pathology Laboratory

Molecular pathology laboratories perform tests for different clinical purposes, including genetics, cancer, and infectious diseases. Although all applications share the use of nucleic acids as the main analyte, the different types of testing require different management considerations (Table 49-1). Considerations relevant for each type of test performed in molecular pathology are reviewed.

Genetic Disorders

All diseases have a genetic contribution, whether it is a specific genetic disease or an increased likelihood for developing a medical condition. Genetic disorders are primarily caused by mutations present in every cell of an individual. Molecular testing for neoplasia or an infectious disease requires a sample of the tumor or the infected tissue, respectively, while molecular testing for a genetic disease can be performed on most accessible tissue types. In addition, because genetic testing examines germline mutations, it has important implications not only for the individual but also for family members that may have inherited the same mutation. Genetic testing also raises ethical concerns, including informed consent, potential for employment or insurance discrimination based on test results, and beginning-of-life issues with prenatal testing. Genetic test results often are not definitive, requiring complex risk-assessment calculations for interpretation. The American College of Medical Genetics (ACMG) and the American Society of Human Genetics (ASHG) have published guidelines and recommendations for many of the more common single-gene disorders (Table 49-2).

Genetic Testing

Molecular genetic testing is currently used for the following major clinical purposes: diagnostic, carrier, prenatal, and presymptomatic DNA testing.[2]

Diagnostic testing is performed on affected individuals for establishing or confirming a clinical diagnosis. Because these DNA tests are gene specific, the patient should have symptoms consistent with the disorder to justify performing the test. Genetic tests may be useful for diagnosis with an early atypical clinical presentation, or when other diagnostic procedures are more expensive or complex. For example, molecular testing for the absence of the survival motor neuron 1 (*SMN1)* gene for spinal muscular atrophy (SMA) in hypotonic newborns is a simple procedure performed on peripheral blood lymphocytes (PBL), replacing the need for a more painful and complex muscle biopsy.

Carrier testing is used to detect recessive mutations in healthy individuals to identify the risk of having an affected child. This application can be used for individuals with a family history of a genetic disorder, or for population screening. Testing of an affected family member can identify the specific mutation present in a family, thus

Table 49-1. Special Considerations in Molecular Pathology Testing Stratified by Clinical Application

Application	Considerations
Genetics	Ethical issues (presymptomatic and prenatal testing) Consequences for family members Informed consent requirements Requirements for family-related information (ethnicity, pedigree, specimens) Time sensitivity (prenatal diagnosis) Complex risk-assessment calculations
Oncology	Many types of samples Extensive use of paraffin-embedded tissue samples (familiarity with limitations) Need for diagnostic samples for optimal interpretation of minimal residual disease test results
Infectious diseases	High-volume testing Use of automated platforms High cost of commercial in vitro diagnostic test kits Increased need for quantitative testing with a wide dynamic range and low detection limit
Identity testing	Chain-of-custody documentation for specimens Special patient-identification requirements Complex calculations Special accreditation for paternity and forensic testing Special qualifications required for the laboratory director
HLA testing	Time sensitivity (≤24 hour-turnaround-time) Complex analyses Special accreditation by the American Society for Histocompatibility and Immunogenetics (ASHI)

allowing directed testing for other family members and improving the accuracy of the risk assessment for individuals with a negative test result. In contrast, population screening focuses on the most prevalent mutations, often with different sensitivity of mutation detection for different ethnic populations.

Prenatal testing refers to the detection of disease mutations in a fetus, using fetal cells obtained by amniocentesis or chorionic villus sampling (CVS). To overcome some of the problems associated with pregnancy termination of an affected fetus, some laboratories offer preimplantation genetic testing in the setting of in vitro fertilization for couples with a family history of a specific genetic disease. Preimplantation genetic testing is performed on a single blastomere from an early embryo, allowing selection of only unaffected embryos for implantation.

Presymptomatic testing is used primarily for the identification of adult-onset dominant disorders prior to the onset of clinical symptoms, in which the offspring of an affected parent has a 50% chance of inheriting the disease. This category includes neurological diseases such as Huntington disease and some types of cancer. Presymptomatic testing is the most problematic and challenging in terms of its psychological effect on the individual and hence requires extensive protocols for pre– and post–genetic test counseling.

Genetic testing requires special attention to informed consent issues, the appropriateness of the testing, and the urgency of testing. Many laboratories choose to require documentation of informed consent for the specific genetic test being requested prior to performing the genetic test. Informed consent can be documented by obtaining the completed consent form or a copy of the completed consent form, or by confirmation by the physician on the requisition form that informed consent is on record in his or her medical office. Until informed consent is documented, the laboratory can extract and store the appropriate nucleic acid, but only after confirmation of informed consent can the laboratory perform testing. The diagnostic laboratory also should review the requisition form to determine whether the test is appropriate for the specific patient. For example, carrier testing of a minor should be deferred until the minor is an adult, so should prompt a call to the referring clinician. Similarly, a request for testing of an asymptomatic individual for a dominant disease requires confirmation of adequate presymptomatic genetic counseling. For prenatal test requests, gestational age should be assessed to assure that the test result is available to allow for termination of the pregnancy if this is the family's choice. Finally, the laboratory should understand and communicate to healthcare providers the sensitivities and limitations of the tests performed.

Ethical Considerations, Genetic Counseling, and Informed Consent

Despite the tremendous advances in understanding of the human genome, the benefit of genetic testing for the patient is not always clear. Testing may be warranted, even if the results are inconclusive or preventive strategies or treatments are not available.[3] The availability of individual genetic information raises critical ethical, legal, and social issues because genetic testing examines the patient's own genetic makeup, rather than acquired (somatic) genetic abnormalities or infectious agents, with implications for other family members. Recently, after intensive lobbying by various advocacy organizations including health professionals and industry leaders, the Senate passed the Genetic Information Nondiscrimination Act of 2005 (S 306),[4] although the House of Representatives version of the nondiscrimination bill (HR 1227) has yet to be passed. This federal legislation would prevent health insurers and employers from using information to determine eligibility, set premiums, or hire and fire employees, and would ensure that genetic information is used for the benefit of the patient. The bills also encourage individuals to take advantage of genetic screening, counseling, testing, and new therapies that will result from the scientific advances in the field of genetics.

Table 49-2. Professional Organizations and Standards for Molecular Pathology Laboratories

Organization	Standards and Guidelines
Centers for Medicare and Medicaid Services (CMS; http://www.hcfa.gov) (formerly Health Care Finance Association, HCFA)	Mandatory federal guidelines for regulating laboratory testing via the Clinical Laboratory Improvement Amendments of 1988 (CLIA'88)
American College of Medical Genetics (ACMG; http://www.acmg.net)	Standards and guidelines for clinical genetic laboratories Gene patents and accessibility of genetic testing Policy regarding genetic testing of minors Storage and use of genetic materials Recommended standards for interpretation of sequence variations Recommendations and guidelines regarding genetic testing for fragile X syndrome, apolipoprotein testing for Alzheimer disease, Prader Willi-Angelman syndrome, Canavan disease, cystic fibrosis, Huntington disease, venous thromboembolism (factor V Leiden and prothrombin), congenital hearing loss, colon cancer, and breast and ovarian cancer
American Society of Human Genetics (ASHG; http://www.ashg.org)	Educational resources in human genetics Legislation for genetic information nondiscrimination
American Society for Histocompatibility and Immunogenetics (ASHI; http://www.ashi-hla.org)	Standards for molecular histocompatibility and immunogenetic testing
Clinical and Laboratory Standards Institute (http://www.clsi.org) (formerly National Committee for Clinical laboratory Standards, NCCLS)	Molecular diagnostic methods for genetic diseases (MM1-A2) Immunoglobulin and T-cell receptor gene rearrangement assays (MM2-A) Molecular diagnostic methods for infectious diseases (MM3-A2) Nucleic acid–amplification assays for hematopathology (MM5-A) Quantitative molecular diagnostics for infectious diseases (MM6-A)
The College of American Pathologists (CAP; http://www.cap.org)	Recommendations for in-house development and performance of molecular tests Molecular pathology checklist for laboratory accreditation
Food and Drug Administration (FDA; http://www.fda.gov)	Guidelines for industry registration and listing of analyte-specific reagents (ASRs) Guidelines for laboratory development and reporting of tests using ASR Premarket review templates for in-house-developed genetic tests Guidelines for industry for manufacturing and validating molecular tests for the detection of HIV-1 and hepatitis C virus (HCV)
National Institutes of Health (NIH; http://www.nih.gov)	Promoting safe and effective genetic testing in the Unites States: final report of the task force in genetic testing
Secretary's Advisory Committee on Genetics, Health and Society (SACGHS; http://www.4.od.nih.gov/oba/sacghs.htm) (formerly Secretary's Advisory Committee on Genetic Testing, SACGT)	Recommendations to the secretary through the assistant secretary for health on all aspects of the development and use of genetic tests
Clinical Laboratory Improvement Advisory Committee (CLIAC; http://www.phppo.cdc.gov/cliac/default.asp)	Setting new CLIA regulations for genetic testing
Centers for Disease Control and Prevention (CDC; http://www.cdc.gov)	Guidelines and recommendations for laboratory testing of HCV and HIV Data collection of genetic testing and results

Molecular genetic testing often requires interpretation using complex risk-assessment calculations.[5] Healthcare professionals must correctly interpret laboratory test results and be able to accurately convey the test results and interpretation to the patient, and the patient's family, as appropriate. Accurate communication of results to patients and families, however, can be complex and time-consuming and therefore may be performed by genetic counselors or clinicians trained in medical genetics, rather than physicians without genetics expertise. The benefits of this approach are illustrated by a study in which one third of physicians misinterpreted a negative genetic test result for familial adenomatous polyposis (FAP) and inappropriately suggested that the patients discontinue recommended aggressive surveillance.[6] For these reasons, most genetic tests should be ordered through healthcare professionals trained in genetics to ensure that the benefits and risks of testing have been explained to the patient during the informed consent process.[7] This also ensures that genetic testing is voluntary. Although obtaining consent is primarily the responsibility of the referring clinician, the diagnostic laboratory should consider requiring documentation of informed consent.[2] There are now proposed changes to the Clinical Laboratory Improvement Amendments of 1988 (CLIA) regulations specific for genetic testing that include a requirement for laboratories to document informed consent prior to performing a genetic test.[8]

Confidentiality

In view of the risks of discrimination in health insurance and employment associated with genetic testing, confidential handling of genetic test results is important. Discussion and communication of results must be limited to authorized healthcare providers. Genetic test results should be communicated verbally or faxed only to the referring physician or genetic counselor. Faxing of results is, in general, considered to be less confidential.[2] Release of any patient information, especially to nonhealthcare entities, must be authorized by the patient and documented. To ensure the confidentiality of genetic test results, some patients choose to personally pay for genetic testing rather than use health insurance, which should be accepted by the laboratory. The US Department of Health Human Services has issued guidelines for ensuring the privacy of patients' health information as part of the Health Insurance Portability and Accountability Act of 1996 (HIPAA).[9] Breech of patients' confidentiality for any type of healthcare information, not just for genetic information can result in litigation against the individual or the institution.

Cancer

Hematopoietic Neoplasms

Molecular hematopathology refers to molecular testing for leukemias and lymphomas to identify somatic DNA alterations.[10] These DNA changes are present only in the affected population of hematolymphoid cells and are not a part of the genetic makeup of the individual. The recent development in polymerase chain reaction (PCR)-based techniques has provided greater diagnostic sensitivity and specificity for diagnostic testing, as well as for monitoring of disease during and after therapy. Likewise, the ability to test very small amounts of nucleic acid has enabled the use of a wide range of sample types, including paraffin-embedded tissues (PETs). For molecular testing, however, formalin fixation is optimal. In contrast to genetic tests, which need be performed only once for an individual, molecular oncology tests often are performed repeatedly for initial diagnosis and during and after treatment. If testing is not performed at the time of diagnosis, later testing of the diagnostic specimen may be useful to confirm a molecular marker for minimal residual disease (MRD) testing. Molecular hematopathology results, whenever possible, are interpreted in the context of histopathology, flow cytometry, and clinical findings.

Solid Tumors

As the molecular rearrangements and mutations that cause specific solid tumor types are identified, the detection of these mutations is used for diagnosis and MRD assessment.[11] Many of these molecular rearrangements were identified through molecular characterization of specific chromosomal translocations identified by cytogenetic analysis of solid tumors, including bone and soft tissue sarcomas, such as alveolar rhabdomyosarcoma, Ewing sarcoma, and synovial sarcoma. The expression of the fusion transcripts for the more common translocations can be detected by reverse transcription–polymerase chain reaction (RT-PCR) testing. The preferred specimen for sarcoma testing is frozen tissue because the optimal analyte is RNA.

Specific somatic changes in solid tumors are used to predict response to treatment. For example, genetic analysis for loss of heterozygosity on chromosomes 1p and 19q is used to predict response to treatment for oligodendroglioma tumors, the presence of which correlates with a better outcome.[12]

Laboratory issues specific to solid tumor testing include the use fresh or frozen tissue for RNA-based testing, the use of microdissection to reduce the nonmalignant cell population in the specimen, and working with small tissue specimens such as needle biopsies.

Infectious Diseases

As molecular techniques have become routine, more and more microorganisms are detected or characterized by molecular testing.[13] Molecular tests are especially suitable for infectious agents which are difficult to culture and for drug-resistance testing. Molecular methods also are useful for viral quantitation as a part of monitoring response to therapy, such as human immunodeficiency virus (HIV) and hepatitis C virus (HCV) infections. Because some infectious disease tests are high volume, and automated instrumentation is being developed, commercial test kits for many pathogen types are available. However, these commercial infectious disease test kits are expensive. Although improved patient outcome due to early diagnosis and treatment may outweigh laboratory expenses, such savings have not been fully demonstrated. Finally, the need for confidentiality protections for patients having HIV testing has prompted specific federal and state legislation to ensure the protection of patients and their families.

Identity Testing and HLA Typing

Identity tests use polymorphic DNA markers to establish the identity of an individual or to determine an inheritance pattern.[14] Identity testing is used for several clinical applications including analysis of bone marrow engraftment following bone marrow transplantation, maternal cell contamination (MCC) studies for prenatal genetic testing, paternity testing, and forensic identity testing. Paternity and forensic identity testing, in particular, require additional considerations, including special accreditation (reviewed in Table 49-2), director qualifications, chain-of-custody documentation, verification of identity, familiarity

with complex probability calculations, reporting of results, and legal proceedings.

HLA typing is time-sensitive, requires 24-hour laboratory staffing, and involves complex anlysis of the test results. Laboratories performing HLA typing require special accreditation (see Table 49-2).

Regulatory Agencies

Laboratories performing clinical testing must comply with numerous regulations. With the advent of the Clinical Laboratory Improvement Amendments of 1988 (CLIA),[15] all laboratories are required to implement minimum quality standards. Proof of meeting these standards (accreditation), which may be done through voluntary accreditation programs offered by professional organizations, is critical to providing high-quality laboratory services. Molecular diagnostic laboratories are faced with special challenges due to the complex technical issues and ethical considerations involved in DNA-based testing. Several regulatory agencies are currently developing guidelines and standards for molecular pathology tests (Table 49-2). Some of these regulatory agencies and regulations are discussed.

CLIA and CLIA '03

All clinical laboratory testing is regulated by CLIA. The Congress passed CLIA to establish quality standards for all clinical laboratory testing to ensure the accuracy, reliability, and timeliness of patient test results.[16] CLIA is "test-site neutral," meaning that the same regulations apply to any location performing testing. Every laboratory examining "material derived from the human body for the purpose of providing information for the diagnosis, prevention, or treatment of any disease" is subject to CLIA. Final regulations to meet the law were published in 1992 in the *Federal Register*.[17] In 2003, new, extensively revised CLIA regulations were published and nicknamed CLIA '03. The changes include major reorganization and consolidation of the regulations by basing requirements on the flow of a patient sample through the laboratory, and updating the requirements to accommodate new technologies.[18]

The Centers for Medicare and Medicaid Services (CMS), formerly known as the Health Care Financing Administration (HCFA), in conjunction with the Centers for Disease Control and Prevention (CDC), were originally charged with developing and enforcing the CLIA regulations. CMS continues to oversee much of these regulatory activities, including laboratory registration, on-site inspection, training, and accreditation.

CLIA certificates are based on the complexity of the test method and are divided into three main categories: waived, moderate complexity, and high complexity. The level of complexity is determined by assigning a numerical score for each test or methodology based on numerous criteria including the knowledge, training, and experience required to perform the test, complexity of reagent and material preparation, characteristics of the operational steps, availability of calibrators, controls and proficiency testing, equipment and test system troubleshooting, and interpretation of results. Under the new CLIA '03 regulations, however, all quality control and quality assessment requirements now apply equally to high- and moderate-complexity laboratories, although some of the personnel qualifications remain more stringent for high-complexity laboratories.[18] These criteria are considered to be key elements in performing clinical testing and can be found on the Food and Drug Administration (FDA) Web site (http://www.fda.gov/).

Molecular pathology tests are considered high-complexity tests and as such must comply with CLIA requirements. Some of the requirements for high-complexity tests include qualifications of personnel performing and overseeing the testing, procedure manual specifications, method verification of performance specifications, proficiency testing, quality assurance, patient test management, and inspection. CLIA, however, does not provide specific guidelines for molecular testing, and, therefore, each molecular pathology laboratory is responsible for the development of a test management program according to CLIA criteria.

Voluntary Accreditation Organizations

CLIA regulations allow CMS to approve nonprofit, professional organizations that have laboratory testing and inspection standards equivalent to or more stringent than CLIA to inspect clinical testing laboratories in place of CLIA inspection. The two major organizations providing CLIA inspections are the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), which accredits more than 80% of the US healthcare organizations, and the Laboratory Accreditation Program of the College of American Pathologists (LAP-CAP). The majority of molecular pathology laboratories are inspected and accredited by LAP-CAP. When a test site meets the accrediting agency's requirements, as assessed by inspection, the laboratory, in essence, is meeting CLIA requirements and receives a CLIA license.

The College of American Pathologists

LAP-CAP accredits only laboratory test sites and not entire healthcare organizations. LAP-CAP does not base its requirements for accreditation on the different CLIA complexity levels of testing; instead it requires that all testing laboratories adhere to the same guidelines. LAP-CAP lists its requirements and guidelines in checklists, each consisting of a series of laboratory guideline questions. All laboratories must follow the Laboratory General checklist and the specific laboratory checklist, which corresponds to each

specific laboratory section. In addition to general requirements concerning personnel qualifications, quality control, test performance verification, and operation-related guidelines similar to CLIA, the Molecular Pathology checklist also includes test- and application-specific guidelines. LAP-CAP also requires laboratories to have a procedure manual for each test or process, which complies with the Clinical and Laboratory Standards Institute (CLSI; formerly National Committee for Clinical Laboratory Standards, NCCLS) GP2-3A.[19] LAP-CAP-accredited laboratories must also participate in a proficiency testing program when available, or perform internal proficiency checks when an external proficiency testing program is not available, for each test performed.

The Food and Drug Administration

In the United States, the FDA oversees regulation of medical devices under the 1938 Federal Food, Drug, and Cosmetics Act and subsequent amendments. Unless specifically exempted, medical devices must "be properly labeled and packaged, be cleared for marketing by the FDA, meet their labeling claims, and be manufactured under Good Manufacturing Practices (GMP), which is a mandated quality assurance system." These rules concerning the design, manufacturing, marketing, and surveillance of medical devices can be found in the Code for Federal Regulations (CFR). The FDA also oversees some of the functions related to laboratory operation, including classification of tests as required by CLIA and review of clinical laboratory test kits or systems.

The FDA and Genetic Testing

The recent growth of genetic testing, mostly developed and performed by individual laboratories without FDA review, has raised public concerns about the quality and clinical usefulness of these tests. The Secretary's Advisory Committee on Genetic Testing (SACGT) was established in 1998 to advise the secretary of the Department of Health and Human Services (DHHS) on medical, scientific, ethical, legal, and social issues raised by the development and use of genetic testing. The SACGT, in conjunction with the Clinical Laboratory Improvement Advisory Committee (CLIAC), recommended that the FDA provide regulatory oversight of laboratory-developed genetic testing. The vast majority of genetic testing is performed using laboratory-developed assays rather than by commercial FDA-approved test kits, raising concerns about the quality and clinical usefulness of these tests. As the first step in developing a review process for laboratory-developed tests, the FDA developed a "test review template" for gathering data on the use, performance, interpretation, and reporting of laboratory-developed tests.[20] Although this template was well received by SACGT, the FDA determined that laboratory-developed tests were part of medical practice, which is not regulated by the FDA, and did not develop a review mechanism for laboratory-developed tests. A new committee has been formed, called the Secretary's Advisory Committee on Genetics, Health, and Society (SACGHS), and the committee has determined to monitor the activities of CLIA and FDA for regulatory oversight of laboratory-developed genetic tests.

Analyte-Specific Reagents

Although the majority of the FDA regulations for clinical assays target commercially developed in vitro diagnostic test kits intended for clinical use, the FDA issued regulations for commercial reagents used in laboratory-developed tests in 1997. The phrase "home-brew" assay[21,22] has been used to describe a laboratory-developed test used by clinicians for patient management. Although these regulations are not specific for molecular tests, almost all tests in the molecular pathology laboratory are developed, validated, and performed without the use of FDA-approved test kits. These tests are distinct from research assays that have unproven clinical utility and cannot be used for patient care. In 1997, the FDA developed regulations for manufacturers of reagents used in laboratory-developed tests. These reagents, termed "analyte-specific reagents" (ASRs), serve as the key component for laboratory-developed tests including molecular pathology tests, and include primers or probes that hybridize to specific DNA sequences. The manufacturers of ASRs are prohibited from making statements about the analytic or clinical performance of the reagents, and validation of tests using ASRs is the responsibility of the laboratory. The laboratory must include a disclaimer in the patient report stating, "This test was developed and its performance characteristics determined by (laboratory name). It has not been cleared or approved by the U.S. Food and Drug Administration." The result of the ASR regulations has been an increase in the availability of commercial reagents for molecular pathology tests.

Elements of Laboratory Design

Successful laboratory operation requires optimal utilization of space, appropriate equipment, qualified personnel, adequate information system services, and financial management.

Facilities

The power of molecular testing is achieved by the use of PCR as the main diagnostic procedure. A serious problem in the clinical use of PCR is false-positive amplification or PCR contamination. Because the product of PCR amplification serves as the substrate for the generation of additional products, PCR amplification produces a very

large number of amplicons that can potentially contaminate subsequent amplifications of the same target sequence. This kind of contamination has been termed "carryover" to differentiate it from contamination by genomic DNA. This potential for PCR product contamination requires special precautions. Therefore, the single most important step in designing laboratory space for molecular pathology testing is meticulous attention to the physical separation of the pre-PCR and post-PCR work areas to minimize PCR contamination.[23]

The pre-PCR area, also referred to as a "clean room" or "preamplification area," is used for the pre-PCR steps of sample receiving and processing, including extraction of nucleic acids and setting up amplification reactions. The post-PCR area, also considered a "dirty room" or "postamplification area," is used for PCR amplification and analysis of PCR products. The pre-PCR laboratory ideally contains separate work areas for nucleic acid extractions and PCR setup to minimize cross-contamination of patient nucleic acid samples. Additionally, a containment unit, such as a biosafety cabinet, may be used for setting up PCR reactions. These cabinets usually are equipped with ultraviolet light for reducing DNA contamination of the work area at the end of reaction setup.

If possible, the air system for the pre-PCR and post-PCR laboratories should be independent. Otherwise, air filters can be installed in the inflow to the pre-PCR room air system, which can be cleaned routinely. An additional measure to limit PCR contamination is to maintain the pre-PCR and post-PCR rooms at different air pressures. The pre-PCR area is maintained at positive pressure, reducing the entrance of airborne contaminants and PCR products, while the post-PCR room is maintained at negative pressure, reducing contaminants from exiting the room. If installation of a pressure system is not possible, then the two work areas can be located at a relative distance from each other.

To comply with universal blood-borne pathogen precautions, nucleic acid extraction from patient specimens can be performed in a closed safety cabinet. Specimens known to contain infectious agents such as HIV and HCV can be preferentially processed in a separate biological safety hood. If organic solvents such as phenol, chloroform, or xylenes are used, a chemical fume hood should be available in the work area. Within the pre-PCR laboratory, RNA extractions should be physically separated from DNA extraction if DNA extraction uses RNase, which can degrade RNA. All work areas should be cleaned daily with 10% bleach solution. In addition to pre-PCR and post-PCR areas, the molecular pathology laboratory also may have a neutral reagent preparation area (unless reagents are separately prepared in both rooms) and a data management room for the processing, review, and reporting of test results, which also may be performed in the post-PCR area.

To maximize use of space, equipment, technical expertise, and trained technologists, all types of molecular pathology testing (genetics, infectious disease, cancer, and identity testing) can be consolidated into a single laboratory. Implementing a core molecular pathology laboratory can reduce labor costs by approximately 30%[24] as well as reduce equipment costs and space requirements due to duplication of these resources in several laboratories.

PCR Contamination Control

In addition to physical separation of the pre-PCR and post-PCR areas, specific procedures are recommended for use of PCR in the clinical laboratory. To ensure that PCR products are not transferred from the post-PCR to the pre-PCR laboratory, each should contain separate supplies and equipment. Pipettes, instruments, and supplies should not be transferred from the post-PCR to the pre-PCR area. This separation of supplies and equipment also dictates the workflow in the laboratory. PCR samples are first set up in the pre-PCR room and then transferred to the post-PCR room for PCR amplification and post-PCR analysis. Worksheets, once moved to the post-PCR area for post-PCR analysis, can no longer be taken back into the pre-PCR area and must be processed and filed in the post-PCR area or a neutral area.

Reagents and solutions for sample preparation and PCR work should be prepared with type I water only in the designated work areas. All reagents, except primers, deoxynucleotide triphosphates (dNTPs), and enzymes, should be purchased as sterile reagents or sterilized to prevent bacterial growth. Reagents, including autoclaved type I water and 10 mM Tris-1 mM EDTA buffer (TE) used to dissolve DNA, should be stored as small aliquots to minimize the number of samplings from the same aliquot. Similarly, oligonucleotides used for PCR amplification should be synthesized and purified in a clean environment. To minimize the number of sample transfers and the chance of sporadic contamination and errors, PCR reagents, including primers, dNTPs and buffers, can be combined into a "premixture master mix" (MM), divided into aliquots that are appropriately labeled, and stored frozen at –20°C.[23] Before using for clinical testing, the MM aliquot must be tested to ensure that it specifically amplifies the target sequence without producing PCR products in the negative "minus DNA" controls. A "minus DNA" negative control tested last in each run ensures the lack of DNA contamination in PCR reagents.

Although carryover of amplified sequences contributes to the majority of false positives, cross-contamination between samples can also be a factor. Consequently, precautions must be taken not only during the setup of PCR amplification reactions but also in all aspects of sample handling, from sample collection to sample extraction. For good laboratory practice, gloves should be changed frequently, at least when entering or reentering the pre-PCR area and whenever the technologist notes the gloves to have been contaminated. Technologists should wear

protective clothes to prevent operator-borne spread of contaminating DNA products, with separate protective clothes (i.e., laboratory coats) for the pre-PCR and post-PCR areas.

To minimize aerosolization of PCR products, PCR tubes should be pulse-centrifuged before opening. Microcentrifage tube caps should be opened using both hands, and not flicked open with the thumb of the same hand, to prevent aerosilization. Because cross-contamination of pipettors can lead to false-positive results and to minimize aerosolization, aerosol barrier tips are used for all pre-PCR steps. MM and other nonsample components should be added to the reaction tubes before the DNA samples are added; DNA should be added last and each tube capped before the technologist proceeds to the addition of the next sample. Positive and negative control DNA or RNA should be the last reaction tubes to be set up, after the patient sample reactions are set up and closed.

Another approach to minimize carryover of PCR products is to synthesize all PCR products with deoxyuridine triphosphate (dUTP) in place of deoxythymidine triphosphate (dTTP). Prior to amplification, the PCR reactions are treated with uracil-N-glycosylase (UNG), which will degrade uracil-containing amplicons, allowing amplification only from thymidine-containing target DNA.[25] Some commercially available test kits incorporate this process into the kits.

Equipment

The equipment required for molecular pathology testing is rarely manufactured for clinical use. Most equipment is designed for research and adapted for clinical use.[23] Nucleic acid extractions may be performed manually, although for higher-volume testing, automated extraction instruments are available and have demonstrated variable success for the clinical laboratory. PCR can be set up manually in biosafety cabinets; for higher-volume testing, robotic systems are available. Robotic systems, however, have reagent dead volumes that increase reagent wastage. Thermal cyclers are standard equipment in the molecular pathology laboratory, and the number required depends on the work volume. Post-PCR analysis is highly variable, although electrophoresis apparatus is commonly used.

New specialized equipment for post-PCR analysis includes automated sequencers or capillary electrophoresis instruments, chemiluminescent or colorimetric plate readers, real-time PCR instruments, and denaturing high-performance liquid chromatography (DHPLC) instruments for mutation screening. Most of the available viral load tests are performed using commercial kits, which require the use of specific instruments supplied by the test kit manufacturer. Because of the high level of testing complexity and the concern of PCR contamination, researchers or personnel not trained in molecular pathology should not use the clinical molecular pathology equipment.

Because these instruments are used for clinical testing, they require rigorous surveillance and prompt technical support by manufacturers who are not always aware of the critical clinical testing issues. In addition, these instruments require ongoing maintenance. Most maintenance procedures are defined by the manufacturer and include temperature checks of the PCR wells for thermal cyclers and calibration of pipetters. In the absence of manufacturer's guidelines, the laboratory should set up its own maintenance program in consultation with the manufacturer.

Personnel

Personnel Qualifications

Effective laboratory operation requires well-trained staff and a good management team. CLIA regulations specify that for high-complexity testing, individuals for the positions of director, technical supervisor, and testing staff must have specific qualifications.

The director, as listed on the CLIA certificate, in addition to having 2 years' experience supervising a high-complexity laboratory, must be a licensed doctor of medicine, osteopathy, or podiatry or have a doctoral degree in one of the biological, chemical, physical, or medical sciences. The laboratory director is responsible for the overall operation and administration of the laboratory, including the development and implementation of new tests, current procedure manuals, quality control and training programs for staff, and teaching residents.

The technical supervisor establishes the quality standards of the laboratory by selecting and monitoring methods and instrumentation and documenting the competency of laboratory personnel, while the general supervisor provides day-to-day supervision of testing, personnel, and reporting of results. The qualifications for these positions range from a licensed doctor to an individual holding a bachelor's degree in science plus specific training or experience or both. The laboratory also may have clinical consultants, who must hold a medical or doctoral degree in a relevant discipline with appropriate experience in molecular testing. The Genetic Workgroup (GW) of CLIAC, created in 1998, reviewed personnel qualifications and recommended including specific genetic experience and board certification for supervisors who oversee genetic testing.[20]

Testing personnel who perform high-complexity testing must have an associate's degree in laboratory science or medical laboratory technology, or education and training equivalent to an associate's degree. The laboratory technologists are responsible for all tasks associated with the daily operation of the laboratory, including specimen receiving and processing, testing, identifying problems and troubleshooting, maintaining equipment, and documenting quality control procedures. In addition, they are required to work according to established procedures in adherence with the quality control standards implemented in the laboratory to comply with CLIA, CAP or other reg-

ulations. Because there are very few commercially available molecular test kits, molecular pathology technologists perform much of the development of new tests, as well as validation of commercial molecular kits used for clinical testing.

Because of the high level of desired expertise, technologists also need to be encouraged to attend educational courses to obtain the appropriate background in molecular pathology and the skills required for performing these tests. For the same reason, it is also crucial for the senior management of the laboratory to acknowledge the staff, encourage personal growth, and create new mechanisms for promotion. Although some of these recommendations represent the authors' personal views and may present a burden in the stringent financial environment of many healthcare institutions, the cost of training new technologists is higher than that of creating an environment that promotes job satisfaction and employee retention.

Genetic Counselors

The role of the genetic counselor is to assess the need for genetic testing through family history, inform the patient about the specific risks and benefits of specific genetic tests, order the tests, and communicate the results to the patient. Although the genetic counselor is traditionally a part of the clinical medical genetics service, an increasing number of test sites use genetic counselors as a link between the patient and healthcare providers and the laboratory.

Resident and Fellow Training in Molecular Pathology

Pathology residency programs are required to provide molecular pathology training to residents. Because the molecular biology knowledge and practical experience of pathology residents is highly variable, the Association for Molecular Pathology (AMP) Training and Education Committee generated general goals for molecular pathology resident training programs. AMP molecular pathology training recommendations for residents include basic knowledge in human genetics and molecular biology, which are relevant to all aspects of molecular testing, specific technology information, as well as knowledge of specific molecular pathology tests.[26] Because molecular pathology training time can be brief (1 to 3 months), training may need to be accomplished through didactic lectures that encompass the tests performed in the laboratory as well as basic concepts in molecular pathology and technology, rather than more extensive practical experience.

Education of fellows requires broader training, resources, and time. The increasing importance of this field for clinical practice and the desire to undergo formal molecular pathology training led to the development of fellowship training programs and professional certification by several professional boards, specifically the American Board of Medical Genetics (ABMG) and the American Board of Pathology (ABP). Official training programs in clinical molecular genetics by the ABMG and in molecular genetic pathology (jointly by the ABMG and the ABP) are accredited by the Accreditation Committee on Graduate Medical Education (ACMGE). Official accreditation of either of these training programs requires filing of a program application with review and on-site formal inspection by an ACMGE representative.[27] The goal of these programs is to provide structured educational training for qualified individuals seeking to integrate molecular genetic pathology into their clinical practice experience in all current aspects of the field, including basic science, diagnostic laboratory procedures, laboratory management, and consultation. To be eligible for the examination, fellows should have a strong background in molecular biology and molecular genetics as well as practical experience, and should be able to use this information for diagnosis and management of genetic disorders, infectious diseases, malignancies, identity testing, and HLA typing.

Residents and fellows can contribute significantly to the daily operation of the laboratory, including acquisition of clinical information and communication of test results. In addition, they can assist with the development and implementation of new tests, according to the skill level of the trainee and as time permits. In return, they gain new expertise that may be an advantage when seeking employment.

Staff Training and Accreditation Programs

Appropriate technical staff training is vital to the successful operation of a molecular pathology laboratory.[27] Accreditation programs in medical genetics and molecular pathology ensure that members of the laboratory staff are well trained for performing their assigned roles. Individuals who want to practice clinical molecular pathology have several options, depending on their initial education.

Individuals who hold a medical or doctoral degree and are interested in directing a molecular pathology laboratory can be certified by the ABMG. ABMG is a member of the American Board of Medical Specialists (ABMS) that provides certification in all medical specialties. Certification in clinical molecular genetics requires passing of a general examination in medical genetics as well as a subspecialty examination in clinical molecular genetics. Individuals must have a doctoral degree (MD or PhD) and must have completed 2 years in an accredited genetics program to be eligible to take the ABMG examination for certification. In addition, candidates must provide a record (logbook) documenting their involvement in 150 clinical molecular genetic cases, which is approved by the fellowship program director. The examination is offered every 3 years.

In 1999, the ABMS approved a new subspecialty in molecular genetic pathology (MGP) offered jointly by the ABMG and the ABP. Candidates for this certification must hold a medical degree, have board certification in their specific field of medical genetics or pathology, have a valid license to practice medicine in the United States, and have completed a year of training in an accredited MGP fellowship training program. The first MGP examination was given in 2001; the test is given biannually. MGP training programs are accredited based on standards developed jointly by the ABP and the ABMG.

The American Board of Clinical Chemistry (ABCC) has a certification program in molecular diagnosis, which was offered for the first time in 2000. This certification is offered biannually to individuals who hold doctoral degrees and practice in any one of several clinical laboratory specialties.

Certification in molecular diagnosis is offered to medical and molecular biology technologists by the National Credentialing Agency for Laboratory Personnel, Inc (NCA). The examination is given biannually and is useful for clinical molecular laboratory staff, particularly the senior technical staff who become certified laboratory specialists in molecular biology. The American Society of Clinical Pathology certifies technical staff as a technologist in molecular pathology, based on qualification for the examination by several routes.

Financial Management

Molecular pathology testing is a market with growing gross revenue, which requires an understanding of the influence of finances on the technical aspects of the laboratory. Because many clinical molecular pathology laboratories are part of healthcare systems, they are constantly subjected to limiting managed care contracts and reduced reimbursements, and as such, they are often viewed as cost centers. Molecular pathology laboratory management should develop cost-effective business plans that present molecular pathology testing in the context of the clinical programs this testing supports, if the testing is not independently cost-effective, such that hospital administration views the laboratory as a revenue center.

Cost Analysis

In the current restrictive medical economic environment, the decision to perform a specific test must include business (financial) as well as medical considerations. The evaluation process for deciding to implement a new test or deciding to discontinue a current test should include a cost analysis. The charges for a test are determined mostly by the costs of performing the test and the value of the test in the market (demand). A cost analysis should account for all expenses associated with performing a given test. In a structured approach, a standard template can be developed and used to price any new test based on the cost for each step of the testing process.

Cost analysis involves two main types of expenses: direct expenses and indirect expenses.[28] The direct expenses enumerate all costs directly related to the testing, including equipment, reagents, consumables, salary, and benefits. The calculation of reagent costs is based on the known volume of reagent used and the cost per volume, including reagents for controls and dead volumes. For manual, low-volume testing, the largest component of direct cost is labor, which can comprise up to 90% of the total direct test cost. Laboratory labor cost calculations are generally done using the hands-on technologist time needed to perform the test multiplied by an average hourly salary rate plus the appropriate benefit rate. For example, the cost for a PCR-RFLP test should take into consideration the time required for specimen accessioning and storage, DNA extraction, PCR setup, post-PCR analysis, interpretation, and reporting of results, usually an average of 6 to 8 hours of technologist time, which is divided by the number of samples tested at the same time to reach the labor cost per test.

Indirect costs include overhead, such as licensing fees, royalties, supervisory and administrative salaries, equipment maintenance fees, building services (such as electricity, phones, heating, etc.), marketing, and information system requirements. These costs are more difficult to determine and are not usually a part of the routine cost analysis for academic laboratories. Accurate determination of laboratory costs is important for assessing profitability.

Billing

Billing for molecular pathology tests follows the same guidelines as billing for other pathology laboratory services, requiring a Current Procedural Terminology (CPT) code and an International Classification of Diseases (ICD-9-CM) diagnostic code. The ICD-9 code, which is required for payment, should be provided by the requesting physician when a test is ordered, while the pathologist is responsible for using the appropriate CPT codes for billing of test services. CPT coding of molecular pathology tests, other than most molecular infectious disease tests which have single test-specific CPT codes, is based on a combination of several CPT codes for the various procedural steps that are performed as part of a molecular test. For example, nucleic acid extraction, amplification, and electrophoresis all have separate CPT codes. A single test using all of these procedures is billed using a combination of all the appropriate CPT codes, used multiple times as needed, with the addition of a modifier code following a hyphen at the end of the primary CPT code. Most molecular CPT codes describe a molecular procedure rather than a specific method. For example, a nucleic acid extraction can be performed for RNA or DNA, using different procedures, each bearing a different cost. Therefore, the same CPT code can be associated

with different billing fees, although the Medicare reimbursement rate remains fixed. Additionally, each procedure (or CPT code) can be billed more than once within a single test, depending on the number of times the procedure was performed. For example, a PCR-RFLP test for the detection of two mutations associated with hereditary hemochromatosis would be coded using one DNA extraction code, two PCR amplification codes, two restriction enzyme digestion codes, one gel electrophoresis code, and one interpretation and report code. A complete list of the CPT codes can be found in the *Physician's Current Procedural Manual*, published by the American Medical Association.[29]

Pathogen-specific codes are available for molecular infectious disease tests, such as HIV-1, HCV, and hepatitis B virus viral loads, for which all steps of the procedure (extraction, amplification, detection, and reagents) are covered by a single CPT code. When specific CPT codes are available for a pathogen type, three coding options are available based on the testing method used, specifically for probe hybridization, amplification, or quantification. Infectious disease tests without pathogen-specific CPT codes are coded using a combination of method-based CPT codes.

For molecular pathology tests requiring professional interpretation, the professional component is coded using the interpretation and report CPT code (83912) with a modifier, "-26," to differentiate the billing from the technical component, reported with a "TC."

Reimbursement for molecular pathology tests is dictated, to a large extent, by the reimbursement policies of CMS, which are generally followed by third-party payers such as insurance companies. FDA approval is not necessary for billing of molecular pathology tests; however, some third-party payers may not reimburse for tests that are not FDA approved. In addition, billing Medicare for tests using a non-FDA-approved commercial test kit is fraudulent; while billing for tests performed using ASRs is not. Due to the low reimbursement rates for molecular tests, some molecular pathology laboratories will not perform tests for patients from other institutions without billing the sending institution or pre-payment by the patient, which usually guarantees full payment.

Patent

Another consideration prior to the implementation of a new test is the patent status of the test. Patents can cover a specific method of testing for an analyte, any method of testing for an analyte, or a mutation-gene-disease association in general. The most common royalty payments for molecular pathology laboratories are payments for the use of the PCR method and Taq polymerase, licensed by Roche Diagnostics, Inc. Many diagnostic tests use PCR. Royalty payments for use of Taq polymerase range from 9% to 15% of the amount billed or reimbursed for a test, and generally are higher for commercial laboratories than for academic laboratories.

Of special concern are patenting and licensing of patents for mutations in specific genes associated with disease. Examples of such patents include apolipoprotein E genotyping for Alzheimer disease, Canavan disease, and Charcot-Marie-Tooth disease type 1A. Patenting imposes a significant risk to the molecular pathology laboratory because the number of diagnostic patents, particularly in the area of medical genetics, has greatly increased in recent years. Many clinically relevant DNA sequences are patented, and the terms of use offered by the holders vary considerably; in some instances, a laboratory will be unable to perform a test because the patent is exclusively licensed to another laboratory. Thus, a laboratory considering implementation of a new molecular test first may conduct searches of existing patents and patent applications to see whether the new test infringes on existing or pending patents. Typically, patent attorneys employed by the medical center conduct patent searches. Internet resources provided by the US Patent and Trademark Office (http://www.uspto.gov) allow laboratories to perform searches directly. If a patent exists that covers any part of the molecular test to be implemented, the laboratory or medical center must negotiate an agreement for clinical testing with the patent holder or licensee, if the patent holder will grant a license. Licensing agreements may include royalty payments that increase the cost of the test, and limitations on the use of the test or the volume of testing the laboratory can perform. For this reason, many molecular pathology leaders argue that patented genetic tests should be broadly licensed at affordable costs to allow any qualified clinical laboratory to perform the tests.

Laboratory Information Systems

Most molecular pathology laboratories require a computer information system to handle the large volume of data that is both received and reported. Selecting and installing a laboratory information system (LIS) is a lengthy and labor-intensive process, usually performed by the department or institution rather than by the laboratory. The application software allows users to perform tasks that are specific to the laboratory operations. These tasks include registering patients and accessioning specimens, ordering laboratory tests, reporting test results, and tracking quality control data (Figure 49-1). In addition, the LIS facilitates use of appropriate templates for billing and reporting of results. The primary function of the LIS application is management of the database of patients and laboratory test results. The most-used system is a relational database management system, which links all the data in tables related to one another by common elements. This kind of structure allows the technologist, for example, to quickly produce a list of only those patient specimens that are going be tested for a particular test. When two or more tables contain the same primary field (patient identifier), information can be retrieved from multiple tables. For example, to monitor

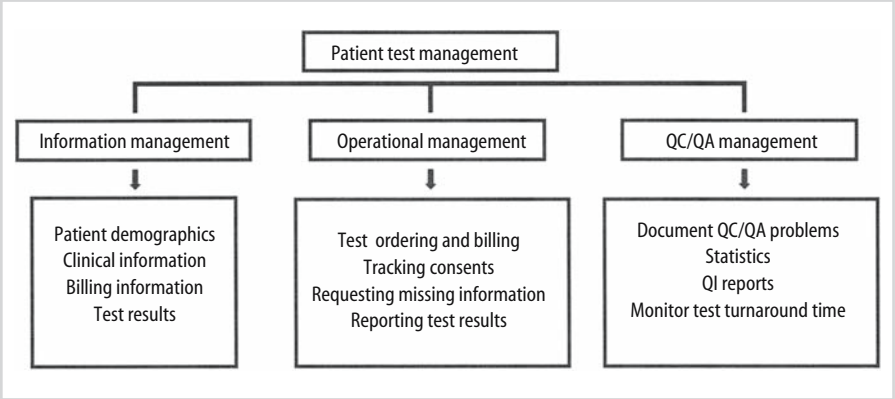


Figure 49-1. Organizational chart for LIS-based molecular pathology testing management.

engraftment after allogeneic bone marrow transplant, the entire list of test results and dates of testing can be obtained for any particular patient. Because departmental LIS systems are not designed to meet the specific information requirements of the molecular pathology laboratory, many molecular pathology laboratories also use networked personal computers (PCs) for data management. These PCs are mostly used as tools for running standard applications, such as word processing, spreadsheets, and other specialized applications, for example, sequencing and genetic risk-assessment applications. Since, however, molecular pathology laboratories have unique data management requirements, a separate patient database may be maintained on the laboratory PCs. To reduce transcription errors and eliminate the labor-intensive work of retyping patient information and test results, laboratory PCs should be interfaced with the LIS system to allow direct transfer of test results for each patient. Furthermore, some molecular pathology instruments, such as the COBAS system (Roche Diagnostics) for HIV-1 and HCV testing, can be interfaced directly to the LIS. The results, collated with specimen identification, are automatically transferred to the LIS, followed by verification by laboratory personnel and report release for clinical use.

To improve laboratory operations, many quality control activities can be performed using the LIS, including monitoring of turnaround time, control results, and statistics. Procedures and other documentation can be maintained and managed by the laboratory computer as well. Security and access control are crucial components of the LIS. Access control (passwords) and patient confidentiality (limiting access to only those who need specific information) are some of the tools required to maintain patient confidentiality.

Test Management

Choice of Test Menu

The menu of tests for molecular pathology laboratories varies significantly and is determined by three principal factors: clinical requirement and usefulness, laboratory competency, and test costs. Other factors that affect the decision about adding a new test include the prevalence of the disease, the mutation(s) causing the disease, the number of different mutations associated with the disease, availability of ASR or FDA-approved test kits, complexity of testing, equipment requirements, and existence of patents that apply to the test. Each of these items requires serious consideration and support on behalf of the institution.

The clinical requirement and usefulness of a test are defined by the significance of the test result for disease diagnosis and clinical management. For example, HIV viral load testing is widely used because it provides immediate information regarding the patient response to treatment and prognosis. Likewise, population-based genetic screening for cystic fibrosis (CF), now offered to all individuals and couples of childbearing age, provides immediate information about a couple's risk for an affected child.[30]

"Competency" refers to the availability of skilled laboratory personnel who can perform, supervise, troubleshoot, and correctly interpret molecular tests. For example, RT-PCR-based tests require knowledge not only of PCR but also of RNA processing in the clinical laboratory.

Test costs are another important factor for selection of tests to offer. Generally, when the cost of a test is low and the volume is high, a test will have a profit margin. A common issue is deciding whether to perform the test or to use a reference laboratory for a specific test. This usually requires a break-even analysis to determine the point (threshold) at which there is no profit or loss for performing the test, taking into consideration the cost of the test, the expected revenue, and the expected volume of testing. Rare or urgent tests may be more expensive since tests are run individually rather than batched analyses, as can be done in a reference laboratory, with combined testing from many sites. In contrast, high-volume tests are likely to be profitable when performed by an individual laboratory.

The prevalence of the disease or mutation in the population affects the testing volume and therefore the number of laboratories that perform a specific test. The laboratory, however, may choose to test for less-common diseases or

mutations if it receives a sufficient specimen volume to make the testing cost-effective. For example, some laboratories offer comprehensive testing for neurological diseases when there is specific clinical expertise at their medical center, or with inclusion of clinical samples referred to the laboratory from other sites in the United States or internationally.

The number of different mutations in one or more genes accounting for a disease is a consideration for testing. In general, the greater the number of mutations in a single gene or the number of different genes that can cause a disease, the more complex and labor-intensive testing becomes. Therefore, for common disorders, such as CF, commercial kits are available that detect 25 of the more common mutations in the *CFTR* gene, requiring substantially less time for test development than laboratory-developed methods or complete *CFTR* gene analysis.

Ease of testing and equipment are two important considerations. Tests requiring lengthy procedures, such as Southern blot analysis, or complex analyses (SMA carrier testing, for example) are less likely to be performed by many laboratories, especially if the test volume is low or commercial test kits are not available. The use of ASRs or FDA-approved tests is preferred by many laboratories because they simplify the laboratory workflow and facilitate quality assurance, even though these tests are often more expensive relative to laboratory-developed tests.

Finally, patents that cover the disease mutations, genes, or testing methods for a test can be a consideration when deciding to implement and perform a test. Patent or license exclusivity may be enforced, licensing or sublicensing fees may be too costly, or other conditions of licensing may not be reasonable for the laboratory, such that the laboratory cannot or chooses not to implement and perform a test covered by one or more patents. This may be even more significant for complex genetic disease tests that require testing of multiple genes, each of which may be covered by patents.

Choice of Test Method

Several factors influence the selection of the testing method, including the mutation type, the degree of mutation heterogeneity, the anticipated test volume, the available resources, such as equipment and expertise, and patent issues. The majority of clinical tests target specific mutations or sequences. The greater the number of sequence variations that need to be detected by a test, the more difficult and labor-intensive the test becomes. The most-frequently used method in molecular pathology laboratories is amplification, usually by PCR. Some of the more common factors affecting the choice of methods from nucleic acid extraction through PCR and post-PCR analysis are discussed.

Nucleic acid extraction is a principal part of most molecular tests. Numerous extraction methods are currently available. Some of the issues to consider in choosing an extraction method for DNA or RNA are yield, quality of the nucleic acid required for the subsequent testing steps, storage, processing speed, and costs. Laboratories may use more than one method for DNA or RNA purification. Some of the more common methods currently utilized are desalting methods and silica-gel membrane columns. Columns are particularly convenient for processing of small volumes of blood for a large number of samples and are available for either DNA or RNA. Organic-based methods can be used but are labor-intensive and have significant safety risks for the clinical laboratory, so are less commonly used now than in the past. Quick lysis methods (alkaline or boiling) generate crude DNA, which may be suitable for some test applications.

The most common method in the molecular pathology laboratory is PCR, which has enabled the implementation of hundreds of clinical molecular tests. The method is particularly suitable for detection of point mutations and other small sequence variations, for detection of chromosomal rearrangements, for detection or quantification of pathogens by amplification, or for sequence analysis. Variations of PCR include RT-PCR for amplification of RNA sequences, multiplex PCR for the simultaneous amplification of two or more sequences, and real-time PCR for quantification of specific sequences (see chapter 2).

While PCR amplification is relatively standard, many methods are used to analyze the resulting PCR products. The simplest method for analysis of PCR products is gel electrophoresis to determine the size of the PCR products that have been synthesized with comparison to a sizing standard that has been run on the same gel. A second common method for analysis of PCR products is digestion with a restriction enzyme followed by analysis of the DNA fragments by gel electrophoresis. PCR products can be hybridized to allele-specific oligonucleotide (ASO) probes complementary for either the normal or mutant allele. A variation on this approach is the reverse dot-blot, in which numerous allelic probes are fixed to specific locations on a solid membrane and hybridized with multiplex PCR products. Methods using hybridization steps are more labor-intensive and expensive than enzymatic digestions and are used less frequently than other post-PCR analysis methods.

Small length changes and genetic rearrangements present in PCR products can be detected by size variations using agarose or acrylamide gel electrophoresis with ethidium-bromide staining, or by a fluorescent-dependent detection method using either a plate- or capillary-based automated sequencer. Automated sequencers are especially useful for the simultaneous analysis of numerous PCR products labeled either with a single or multiple fluorophores, or when quantitation of a PCR product is required.

A significant portion of PCR testing is being transitioned to real-time PCR, which monitors the generation of the PCR product throughout the amplification process. This approach can be used for either point mutation detection or quantitation, eliminating the need for time-consuming post-PCR analysis and decreasing opportunities for PCR contamination.

Table 49-3. Specimen Types and Handling for Molecular Pathology Testing

Specimen Type	Test Type	Requirements	Shipment	Storage
Blood	Genetics, cancer, identity, HLA typing	EDTA (0.5–10 ml)*	RT	4°C
Bone marrow	Cancer, identity testing	EDTA (0.5–1 ml)	RT	4°C
Plasma	Infectious diseases	EDTA (10 ml)	Dry ice	–80°C
Cerebrospinal fluid	Cancer, infectious diseases	None (0.5–2 ml)	Dry ice	–80°C
Sputum	Cancer, infectious diseases	None (0.5–2 ml)	RT	4°C
Amniotic fluid (<15 weeks gestation)	Genetics	None (10–15 ml)	RT	4°C
Chorionic villus sampling (CVS)	Genetics	5–15 mg	RT	4°C
Cultured amniocytes or CVS	Genetics	2 T25 flasks	RT	4°C
Buccal cells	Genetics, cancer, identity	2 swabs	RT	4°C
Fresh tissue	Cancer, identity	50–100 mg	Dry ice	–80°C
Paraffin-embedded tissue	Cancer, identity	2–5 µm sections	RT	RT
DNA	Genetics, cancer, identity, HLA typing	<1–100 µg	RT	4°C
RNA	Genetics, cancer	<1–20 µg	Dry ice	–80°C

*Sample volume is based on application and age of patient.
RT, room temperature.

When the disease gene is known but the mutations are unknown, mutation-scanning methods are used. The most common scanning methods are single-strand conformation polymorphism (SSCP) and DHPLC, both of which detect single base pair variations with high sensitivity.

Automated platforms for molecular tests are becoming available and are more significant for the clinical laboratory as test volume increases. Currently available automated systems include extraction systems, and instruments that combine PCR amplification and detection. Some automated systems are routinely used in many molecular pathology laboratories for viral load testing, such as for HIV-1. Automated instrumentation can reduce turnaround time and human error.

Clinical Information Requirements

Preanalytic clinical information is essential for determining the appropriateness of the test and is critical for interpretation of results. Clinical information usually is obtained from the requisition form or an electronic test order system. A complete requisition contains patient demographic information to allow for identification of the patient, the ordering healthcare provider's name and contact information, the type of sample, the name of the ordered test(s), relevant clinical information, and any other information essential for appropriate interpretation of test ordered results. For some genetic disorders, such as CF, ethnic background is particularly important because disease prevalence can vary significantly with ethnic background, while a pedigree is required for linkage analysis studies and risk assessment. Laboratories should document informed consent for genetic tests (as discussed in the Genetic Testing section), which may be required by state regulation. For cancer or infectious disease testing, indication of the purpose for testing (diagnosis or minimal residual disease assessment) is needed for appropriate interpretation of the test result.

Specimen Requirements

Molecular pathology testing can be performed using a variety of patient sample types. A list of the more common specimen types, applications, handling and storage requirements is presented in Table 49-3. The most common sample for genetic testing is peripheral blood (PB); buccal cells may be acceptable for some tests. Prenatal testing is performed on cultured or direct amniotic cells and chorionic villus sampling (CVS). A maternal sample may be required for prenatal testing to rule out maternal cell contamination (MCC) of the fetal specimen, to ensure that the test result reflects the fetal genotype rather than the maternal genotype, usually by identity testing of the fetal and maternal samples. Samples for molecular oncology and infectious disease testing may include PB, bone marrow, tissues, and other body fluid (cerebral spinal fluid, sputum, etc). Paraffin-embedded tissues (PET) are used predominantly for DNA-based testing such as gene rearrangement studies for lymphomas, while fresh or frozen tissue generally is used for RNA-based testing, such as the detection of specific fusion transcripts in cancers.

Sample requirements and collection and shipment information for molecular tests should be provided by the laboratory to ensure that samples are collected, handled, shipped, and stored appropriately before shipment, during transit, and during storage in the laboratory before sample processing.

Laboratory Operations

Workflow includes the steps of testing from specimen collection to reporting of results. This process can be divided into preanalytic, analytic and postanalytic phases. The preanalytic phase consists of collecting, transporting, accessioning, and storage of the sample(s) prior to testing. The analytic phase is the process of performing the test. The

postanalytic phase consists of all steps after the test is performed, including analysis of the test result, generation of the test report, and communication of the results. Although the preanalytic phase is difficult to control since the steps are performed by individuals outside the laboratory, most errors occur in the analytic phase because of the multistep, predominantly manual nature of molecular testing.

Preanalytic Phase

Samples are usually transported to a central receiving area of the laboratory or directly to the molecular pathology laboratory, either by a local hospital transport system or by overnight carrier if the samples are sent from remote sites. Upon arrival, samples are accessioned into the LIS and assigned a unique identifier or accession number, allowing the sample to be tracked in the laboratory and hospital database. Information entered during accessioning includes patient demographics, test(s) requested, specimen type, and time and date a specimen is collected and received by the laboratory. If available, bar-coded labels can be attached to the requisition form and specimen container.

The clinical information and specimen type are reviewed for appropriateness of the test requested. Missing clinical information can be obtained by contacting the healthcare provider. Genetic test requests are reviewed for ethical considerations.

Some sample types require additional processing prior to nucleic acid extraction. For example, PET blocks are sectioned to allow efficient extraction, and CVS tissue is examined by a qualified cytogeneticist to remove contaminating maternal tissue. Similarly, prenatal cultured cells are examined for confluence; low numbers of cells may not produce sufficient DNA for some analyses, such as Southern blot analysis. For HIV and HCV viral load tests, for example, the plasma must be separated from the blood cells within 4 to 6 hours of collection. The specimen is appropriately stored until processing begins (Table 49-3).

Analytic Phase

For laboratory efficiency and cost-effectiveness, testing is performed on batches of specimens, and the frequency of testing depends on test volume and clinical urgency. The first step in molecular-based testing is nucleic acid isolation. DNA or RNA can be extracted from the patient specimens using a variety of laboratory-developed methods or commercial kits, and the nucleic acid stored appropriately (see Table 49-3). Most tests use PCR, which is set up in the pre-PCR area of the laboratory. To verify sample identity throughout all the steps of testing and prevent possible mix-up and cross-contamination, a worksheet can be used to indicate sample identifiers and numerical locations in the run, as well as reagent information such as lot numbers and expiration dates. The same worksheet can be used later for documentation of test results for patient samples and controls. The patient samples should be set up first, followed by the controls, as discussed in the PCR Contamination Control section. Regardless of the procedure used, certain principles are essential. Specimen and reagent identifications should be checked with every manipulation performed. Expiration dates and identification of reagents should be checked before each use. Test results should be documented, either by photographing the gel or by printing out the data from the instrument. It is very important at this point to label and correctly identify the samples on pictures or data sheets to prevent interpretation errors.

Postanalytic Phase

After the test is performed, the results are analyzed, the test results and interpretation are entered into the laboratory computer manually or by electronic transfer from an instrument, and the results are communicated verbally to the healthcare provider, if appropriate. To reduce errors, results should be interpreted by two independent reviewers, which may include a technologist previously certified to interpret the specific test through the laboratory's training program, a resident or fellow, or the laboratory director or appropriate designee. Results must comply with the established criteria for result analysis and interpretation as defined in the test procedure manual. Control results are reviewed for accuracy before analysis of patient test results. For PCR-based tests analyzed by electrophoresis, the water/no template control should not have bands other than primer dimers, and control samples must demonstrate expected results. For quantitative assays, control results must fall within the established assay ranges. Patient test results are interpreted as positive or negative, or a quantitative result is obtained by comparison with the assay controls and by comparison to size markers, as appropriate for the specific test method. Unexpected results are assessed for errors in the analytic process. The identified problems are corrected and testing of all or selected specimens is repeated. Results should be recorded manually or electronically transferred to the assay worksheet by the technologist who performed the assay, signed, dated, and entered in the computer using a report template, if available. Entered reports should then be reviewed and released by a second interpreter. For efficient workflow, use of report templates for the most common test results and interpretations, with revisions for specific specimens, facilitates test reporting.

The information required for test reports is defined in the CAP Molecular Pathology checklist and includes patient demographics, methods, mutation(s) tested, and a clinical interpretation in an easy-to-interpret format. Although the CAP, in compliance with CLIA, does not require supervisory review of all test results, it is

recommended that all manual assays be reviewed and approved by the laboratory director or qualified designee before reporting. A final report is generated only after results are verified and electronically signed by the laboratory director or a qualified designee. When laboratory director review is not deemed necessary, such as for tests performed using automated equipment, two separate technologists trained to perform the assay, or the same technologist in duplicate, should review the results after entry into the LIS to ensure accurate entry into the LIS prior to release of the reports. Results most often are communicated by printed report or an electronic information system, but direct discussion of urgent or complex test results may be useful for the clinician. Verbal communication or faxing of test results should be performed in compliance with HIPAA regulations to protect patient confidentiality.

Quality Control and Quality Assessment Programs

The molecular pathology laboratory must establish and maintain quality control (QC) and quality assessment also known as quality assurance (QA) programs that are adequate and appropriate for the validity of the procedures performed. The next few sections focus on the principal aspects of these programs.

Quality Control Program

The QC program focuses on the analytic validity of the testing process with the purpose of assessing and ensuring the reliability of patient results. According to CLIA, "The laboratory must establish and follow written QC procedures for monitoring and evaluating the quality of the analytic testing process of each method to assure the accuracy and reliability of patient test results and reports."[17] Specific requirements include adherence to the manufacturer's instructions for instrument operation and test performance, having a current procedure manual, performing and documenting QC test results, maintaining records of all QC activities, and verifying performance specifications. Test sites, especially high-complexity laboratories, may choose to impose more, but not less, stringent QC requirements for the testing process. Molecular pathology tests are largely manual, laboratory-developed, multistep assays, requiring a variety of commercial or laboratory-produced reagents. QC procedures should be designed to detect and reduce errors in the entire testing process. Special attention should be given to the most common types of errors that occur during testing, such as use of expired reagents, improper preparation and storage of reagents and controls, and clerical errors. Some of the key components of the QC program are discussed below.

Test Validation (Verification of Test Performance Characteristics)

There are major differences, as defined by CLIA, between the implementation of an FDA-approved test and the implementation of a laboratory-developed test. If the test is performed using an FDA-approved in vitro diagnostic test kit, the laboratory needs only to verify the performance characteristics of the test for the population of patients seen at that specific laboratory site. Implementation of a laboratory-developed test requires collection of more extensive analytic data to establish the performance characteristics of the test. The performance characteristics that need to be evaluated include accuracy, precision, sensitivity, specificity, reportable range, reference range, and any other applicable characteristics. Laboratories can use the manufacturer's criteria as long as the director determines that these are appropriate for the laboratory's population of patients. To assist in this process, CLIA also recommends that performance characteristics of a new test be compared to a known "gold standard." This is not always practical, since molecular methods often are developed because of the inadequacy of the currently available test methods. The test method also should be validated for all types of specimens that will be accepted for testing.

Establishing the test accuracy requires determining that the test yields appropriate positive, negative, or quantitative results, based on samples with known characteristics as determined by an independent test method. Samples used to determine accuracy can be obtained from several sources, including proficiency surveys or other laboratories performing the test by the same or a different test method. For example, validation results for new RT-PCR procedures for the diagnosis of certain translocations in leukemias can be supported by cytogenetic results for the same specimens.

Establishing the precision of an assay requires demonstrating that the assay results are reproducible. Testing the control specimens and a variety of patient specimens multiple times in separate runs and comparing the results can demonstrate the precision of a test. For quantitative tests, precision is tested over the range of reportable values.

Analytic sensitivity refers to the lowest amount of analyte (nucleic acid) detected or quantified by a test. Analytic sensitivity is determined as the proportion of patient samples that contain the target nucleic acid sequence that tests positive using the test. In quantitative tests, this represents the smallest amount of analyte that is reliably detected by the test. Sensitivity should be determined for qualitative assays for pathogen detection or for MRD monitoring to establish the utility and limitations of the test.

Analytic specificity measures the degree to which the test reacts with nucleic acids other than the intended sequence. It is measured as the proportion of patient samples that do not contain the target nucleic acid that test negative by the test. The analytic specificity of a test can be

determined by analyzing a series of positive and negative samples and calculating the detection rate of the true-negative samples, as defined by samples that do not contain the target of the test. Distinguishing among closely related pathogens, however, requires a precise design of primers specific for the target microorganism. Nonspecific signals can be produced by nonoptimal amplification conditions or inhibitors, which can be overcome by high-quality DNA extraction procedures and optimization of assay conditions.

The development of a new test usually is driven by new molecular information combined with a clinical need for a better test. Once the laboratory director decides that a new molecular test will be implemented, the methodology for the assay, controls, specimen types, costs, and other factors are determined. The method for the assay is chosen based on the type of test information required, clinical utility, current use of the test platform in the laboratory, ease of use, and clinically required turnaround time. Once the test method is chosen, the specific test parameters are designed. Each step of the assay is optimized, including nucleic acid extraction, amplification, detection, and interpretation of results. For laboratories that develop many assays, standard starting conditions and optimization criteria facilitate test implementation. For example, PCR assays can initially use standard reagent concentrations and amplification conditions, followed by fine optimization of the PCR conditions as needed. Once the optimal assay conditions are established, the analytic validation of the assay is performed on clinical specimens and/or controls using those testing conditions.

Initial optimization and performance of the assay must include establishing the controls to be used for testing. Appropriate controls are often commercially available from organizations such as Coriell (Camden, NJ) and the American Type Culture Collection (ATCC; Manassas, VA) as frozen immortalized cell lines or DNA containing the pathogen genome or mutation of interest. With the exception of HIV and HCV, for which standardized reference panels are now available, standardized controls are not available for most of the molecular diagnostic tests currently performed. In the absence of purchased controls, patient samples tested by a different method or provided by another laboratory may be used.

The last step of the test evaluation requires clinical validation of the new test. Clinical validation determines the probability that a sample from a patient with the disease or disease risk will test positive (clinical sensitivity), and that a sample from a patient without the disease or disease risk will test negative (clinical specificity). A final step of the clinical validation may be to obtain a panel of mixed positive and negative patient samples without the known test results from another laboratory that performs the same test. After the testing, results are compared to the results obtained by the other laboratory and documented. Any discrepancies should be resolved by a different test method or a third laboratory.

Test Procedure Manual

Test procedure manuals are essential for molecular pathology laboratory function and are required by CLIA. The manual should be simple, easy to follow, and functional, but should provide sufficient detail to act as a reference for all aspects of testing for laboratory personnel and for training purposes. The manual, in a written or electronic format, must be current and available to all laboratory personnel. Manuals for high-complexity tests include written policies for sample collection, equipment performance evaluation, QC program, and standard operating procedures (SOP) for tests. An example of the contents of a test procedure manual is shown in Table 49-4. The procedures are usually written by laboratory technologists or a supervisor, and

Table 49-4. Contents of a Standard Test Procedure (an example of a PCR-RFLP test)

Item	Content
Background	Theoretical information concerning the disorder and the test
Principle of the assay	Description of how the restriction enzyme detects the specific mutation
Summary of the assay	Characteristics of the restriction-based assay used
Clinical significance	Use of the test for patient care
Strategy for analysis	Overview of the testing steps (i.e., extraction, PCR amplification, digestion, gel electrophoresis)
Specimen collection or acquisition	Description of acceptable specimen type, storage, shipment, and required clinical information
Criteria for rejected samples	Description of unacceptable samples (i.e., specific specimen types, improper labeling, damage)
Equipment and supplies	List of equipment necessary for testing (i.e., PCR machine, gel electrophoresis)
Reagents	List of reagents used for each test step (e.g., PCR, restriction enzyme digestion, gel electrophoresis equipment)
Assay procedure	Description of the testing steps, including general considerations, sample preparation, pre-PCR preparation and setup, thermal cycling, PCR cleanup, PCR product digestion, gel preparation and electrophoresis, gel documentation
Analysis of results	Description of expected results, control results, criteria for gel analysis and interpretation, documentation of results in the LIS and laboratory data base, if used
Turnaround time and schedule for retaining specimen	Expected turnaround time for the specific test, specimen and DNA storage, and schedule for retaining the specimen and DNA
References	Literature or commercial-based information
Addenda	E.g., the worksheet for the specific test, the manufacturers' product insert

Table 49-5. Complete Molecular Pathology Laboratory Procedure Manual: Examples of Required Information, Guidelines, and Protocols

General Information

Staff and contact information
Test information
Requisition forms (genetics, oncology, infectious diseases, identity testing)
Information for clients (shipping, sample type, billing information, consent form, test-specific fax coversheets)
Licenses and permits (CLIA certification and CAP accreditation form)
Reagent handling for molecular testing guidelines
Probe and primer documentation

General Laboratory Operating Procedures

Quality control, quality assurance, and quality improvement programs
Specimen-receiving procedure
Prevention of PCR contamination
Logging-of-primers procedure
Autoclaving of solutions procedure
Use of PCR biosafety chambers
Programming of PCR thermal cyclers
Sequencing using a capillary electrophoresis instrument
Equipment-maintenance procedures
Centrifuges, balance, pH meter, and spectrophotometer: checks and cleanup
Water check procedure
Timer check procedure
Thermometer validation
Temperature-check procedure
Maintenance and validation of PCR thermal cyclers

Nucleic Acid Extraction Procedures

Large-scale DNA extraction from blood by desalting method
Small-scale DNA extraction from blood by column
Fresh- and frozen-tissue DNA extraction
Paraffin-embedded-tissue DNA extraction
Prenatal (amniocytes and CVS) DNA extraction
RNA extraction from blood by column
Nucleic acid quantitation

Test-Specific Procedure Manuals

Genetics

Factor V Leiden analysis
Myotonic dystrophy PCR and Southern analysis
Cystic fibrosis mutation analysis
Spinal muscular atrophy direct testing
Spinal muscular atrophy carrier testing (dosage analysis)
Spinal muscular atrophy linkage analysis

Oncology

IGH gene PCR for detection of B-cell clonality
T-cell receptor PCR for detection of T-cell clonality
BCL2 gene PCR for detection of M-bcr and m-bcr breakpoints
PML-RARA RT-PCR for detection of t(15;17) in APL
BCR-ABL RT-PCR for detection of t(9;22) in CML and ALL
RT-PCR for detection of translocations in sarcomas

Identity Testing

Bone marrow engraftment analysis by genotyping of STR markers
Parentage and identity testing
Maternal cell contamination analysis

Infectious Diseases

HIV-1 viral load assay
HCV viral load assay
HBV viral load assay
Microbial molecular identification analysis

IGH, immunoglobulin heavy chain; *BCL2*, B-cell lymphoma protein 2; M-bcr, major breakpoint cluster region; m-bcr, minor breakpoint cluster region; *PML-RARA*, promyelocytic leukemia-retinoic acid receptor alpha; APL, acute promyelocytic leukemia; CML, chronic myelogenous leukemia; ALL, acute lymphoblastic leukemia; STR, short tandem repeat; HCV, hepatitis C virus; HBV, hepatitis B virus.

then reviewed, edited, and approved by the laboratory director or designee. Any changes to an established procedure manual are documented in writing or electronically and approved by the director. Any laboratory staff members that perform a specific test must review the procedure manual annually, with documentation of the review.

The necessary elements of a test procedure manual are detailed in the CLIA guidelines, which follow closely the items described in CLIS GP2-A3, Clinical Laboratory Technical Procedure Manuals.[19] An example of a complete laboratory procedure manual is presented in Table 49-5. Manufacturer's product inserts or operation manuals can be used as a component of a test procedure manual, but additional information specific to the testing as performed by the laboratory must be provided.

Controls

Every clinical test requires use of appropriate controls. Controls are used to document reproducibility and to ensure that the test is working properly and results are reliable. Several types of controls are used for molecular tests: positive, negative, sensitivity, inhibition, water, and QC controls. Combinations of these controls are used as appropriate based on the method and purpose of the test. A positive control is defined as a sample that contains the target sequence and produces a positive result when tested. Sources of positive controls include positive patient samples and commercially available cell lines or nucleic acids. A negative control is a sample that does not contain the target sequence and produces a negative result when

tested. Sensitivity controls are needed for assays that require a specified level of analytical sensitivity. This control is particularly important for MRD tests for cancer and viral load tests to define the lower limit of detection of the specific test. In general, a sensitivity control is included in each run of the test, since the sensitivity of each test run should be documented and not be assumed to be the same as during validation of the test. A no-template control or water control is a control reaction containing all reagents but no template nucleic acid. This control tests for contamination of the PCR reagents and should not produce any PCR products except for primer dimers. An inhibition control (addition of external target nucleic acid or amplification of another template sequence to check for PCR amplification) is used only when the lack of production of a PCR product is interpreted as a negative result for the test. A QC control is a patient sample that was previously tested and is included for repeat testing on a new run. These samples function to assess the accuracy of the reported results as well as to monitor for interassay variability.

The controls should be processed in the same manner as the patient samples during testing; however, their use and storage should be well defined in the test procedure manual to avoid degradation. Controls should be validated for acceptability and approved by the laboratory director or designee before use in clinical testing. Failure of controls to perform as expected should be investigated, with findings and corrective action documented. Some QC results, for example, in HIV viral load testing, should be charted and analyzed on standard Levi-Jennings plots. Deviations and adverse trends indicate that something is changing in the analytic system, and require investigation and corrective action.

Preventive Equipment Maintenance

All instruments used in the molecular pathology laboratory require written standard procedures for normal operation, performance evaluation, and preventive maintenance. Function checks should be performed on a regular schedule to check critical characteristics and detect malfunctions before test results are affected. In the absence of manufacturer's guidelines, the laboratory should establish a maintenance procedure that appropriately reflects the use of the instrument. Of particular importance are the thermal cyclers, because any change in their performance will have a direct impact on many tests. Thermal cyclers should be routinely monitored for cycle time reproducibility, verification of temperature accuracy, and efficiency of heating and cooling rates. These diagnostic checks are usually incorporated into the instrument's software and are a part of the manufacturer's recommendations for instrument maintenance. Temperature logs for each temperature-dependent piece of equipment, including an acceptable temperature range, are required by CLIA. Temperatures of refrigerators, freezers, and incubators must be documented daily if used for reagent storage. Other instruments, such as water baths, should be documented each day of use. An equipment maintenance and performance evaluation log should be maintained for every laboratory instrument and should be kept for the duration of the clinical use of the instrument.

Quality Assessment Program

CLIA requires each laboratory to have an active QA program to assess various aspects of the technical and nontechnical performance of the molecular pathology laboratory. CLIA requires establishment of guidelines and procedures designated not only "to evaluate the ongoing and overall quality of the testing process," but also to evaluate "the effectiveness of its policies and procedures," followed by requirements to "identify and correct problems, assure reliable and prompt reporting of test reports and assure the adequacy and competence of the staff."[17] In the CLIA '03 version, "quality assurance" was renamed "quality assessment," recognizing that quality cannot always be assured but it can be evaluated or assessed.[18] The laboratory QA program must monitor and evaluate both the ongoing and the overall quality of the total testing process through error detection, corrective actions and their review, and the integration of improvements in procedures. To meet the CLIA requirements, the molecular pathology laboratory must have a QA program that includes numerous laboratory practices, including personnel training and performance evaluation, proficiency testing, inspections, correlation of molecular results with clinical data, and a QC program. Elements monitored by the QA program, such as turnaround time, rejected specimens, and various indicators of test quality, do not relate directly to the analytic validity of the testing process and thus are not a part of the QC program. A complete list of these indicators can be found in the LAP-CAP Molecular Pathology checklist. Table 49-6 lists some of the more important components for which the LAP-CAP requires ongoing surveillance and documentation.

Although CLIA requires that clinical laboratories continually evaluate and improve their work in all areas, some requirements are particularly important from an institutional perspective. CLIA mandates that test results obtained from different methodologies performed under the same certificate be compared, documented, and reviewed for the detection of errors that can affect patient care. This means that molecular oncology test results, for example, should be compared with the results obtained by other laboratories, such as the flow cytometry and cytogenetics laboratories, for the same patient.

Training

Personnel performing high-complexity molecular pathology testing must meet established qualifications in

Table 49-6. Examples of QC and QA Topics That Require Documentation and Corrective Action

Proficiency Testing
Ongoing evaluation of proficiency test results
QC and QA of Testing
Detection and correction of clerical and analytic errors
Recording of failed nucleic acid isolations
Recording of failed hybridization reactions
Monitoring of test turnaround time
Discrepancies between preliminary and final reported results
Discrepancies between the molecular pathology laboratory and other laboratory findings
Performance of statistics and appropriate comparative studies on all molecular pathology tests
Logging of unusual, difficult, and instructive cases
Laboratory-Developed Assay Validation
Performing validation to confirm analytic test performance characteristics
Documenting validation studies for establishing assay performance characteristics
Documenting analytic characteristics including accuracy, sensitivity, specificity, and precision
Demonstrating and documenting clinical validity
Procedure Manual
Annual review of all policies and procedures by the laboratory director or designee
Review of all new policies, procedures, and changes to existing procedures
Knowledge of testing procedures for the tests performed by each technologist in the laboratory
Information for all DNA probes, PCR primers, and other nucleic acid reagents used for testing
Paternity and Forensic Identity
Chain-of-custody (proof of identity)
Genetic characteristics (heterogeneity, recombination rates, etc.) of the DNA markers used
Release of information
Specimen Handling
Disposition of unacceptable specimens
Sample condition, inadequate volumes, and evidence of tampering
Reagents
Validation of reagents used for tests
Characteristics of the specific reagents used in each assay
Controls
Failed control results or when results exceed defined tolerance limits
Instruments
Ongoing evaluation of results of instrument maintenance and function (temperatures, instrument characteristics, cleanliness)
Function checks for detection of problems and malfunctions
Repairs and services
Temperature checks for temperature-dependent equipment
Personnel
Continuing medical laboratory education
Safety
Radioactive area surveys and wipe tests
Radioactive waste disposal

training, experience, job performance, and competency. Training includes competency in performance of the test methodology, calculations, independent judgment, quality control of the method, and knowledge of the instruments. These training requirements cannot be based simply on academic requirements or experience. The on-site training must be sufficient to ensure that staff members have the skills necessary for performing every step of the analytic process, including assay method, preventive maintenance, and quality control measures.

To this end, a training program ensuring that technologists are properly trained to perform laboratory operations must be used. As the first step, technologists can be trained by observing a technologist who knows how to perform the procedure and by reading the procedure manual. Then the technologist should test two to three blind sets of specimens in addition to known controls, with or without observation depending upon the level of experience of the technologist, to ensure that the technologist understands the test and performs it correctly. Monitoring of the training testing ensures the accuracy of the technologist's performance. Finally, the results of training runs must be documented and approved by the laboratory director or designee.

A part of ongoing training is annual competency evaluation. This allows the laboratory director or designee to review different aspects of the technologist's performance, including technical skills, control of PCR contamination, equipment operation, procedure knowledge, and laboratory safety. The overall performance of each technologist is documented, and corrective or additional training measures are implemented as necessary.

Training also includes continuing education programs, which assist the technologist in attaining some of the technical skills and knowledge required to keep up with changes in procedures, instruments, and new tests, as well as improve their ability to perform new assays, troubleshoot, and handle new problems.

Proficiency Testing

Proficiency testing plays a key role in assessing the internal quality of work in CLIA-regulated high-complexity testing laboratories. CLIA requires the laboratory to successfully participate in a CMS-approved proficiency testing program for all analytes tested. CAP provides approved proficiency testing programs in several relevant fields, including genetics, infectious diseases, oncology and identity testing (Table 49-7). Proficiency testing samples are sent at least once per year. In the absence of an external proficiency testing program for a specific test, the laboratory can set up a sample exchange proficiency testing program with another laboratory performing the same test. Proficiency samples must be tested in the same manner as patient samples, by the same technologists performing the clinical testing, using the same interpretation procedures. Laboratories

Table 49-7. CAP Proficiency Surveys for Molecular Pathology Testing

Category/Name	Survey	Type of Testing
Genetics		
MGL	Molecular genetics	Molecular testing for various genetic diseases or genes
Oncology		
MO	Molecular oncology	Molecular analysis of leukemia and lymphomas
MSI	Microsatellite instability	Molecular analysis of PET colorectal carcinoma by DNA amplification of microsatellites
SARC	Sarcoma translocation	Molecular analysis of various sarcoma translocations by RT-PCR
CYH	Fluorescence in situ hybridization	Analysis of FISH for *HER2/NEU* amplification using chromosome-specific DNA probes
Identity		
ME	Monitoring engraftment	Monitoring for bone marrow and stem cell engraftment
PAR, PARF	Parentage testing	DNA testing by PCR and RFLP
ML, DL, DML	HLA molecular typing	Molecular typing of class I and class II HLA markers
FID, FIDM	Forensic sciences—nuclear and mitochondrial DNA	Nuclear and mitochondrial DNA analysis for numerous polymorphic loci
DNA, DNAF	Forensic sciences—DNA database	Analysis of polymorphic loci for the DNA database program
Infectious Disease		
ID	Nucleic acid amplification identification	Analysis of infectious pathogens analyzed by advanced amplification methods (NASBA, bDNA, LCR, PCR)
HIV, HV2	HIV viral load	Quantitative analysis of HIV by nucleic acid amplification and genotyping, respectively
HCVN, HCV2	Hepatitis C viral load	Quantitative analysis of HCV by nucleic acid amplification and genotyping, respectively
HPV	Human papillomavirus	Qualitative analysis of HPV by nucleic acid testing
HC	*C. trachomatis,* herpes, and *N. gonorrhoeae*	Pathogen analysis by nucleic acid amplification and nucleic acid probe methods
NAT	Nucleic acid testing	Qualitative analysis of infectious pathogens designated for blood donor centers
ISH	In situ hybridization	In situ hybridization testing for various viral nucleic acid targets

PET, paraffin-embedded tissue; RT-PCR, reverse transcription–polymerase chain reaction; FISH, fluorescence in situ hybridization; RFLP, restriction fragment length polymorphism; HLA, human leukocyte antigen; NASBA, nucleic acid sequence–based amplification; bDNA, branched DNA technology; LCR, ligase chain reaction; PCR, polymerase chain reaction; HCV, hepatitis C virus; HPV, human papillomavirus.

failing two of three consecutive testing events for any test must submit a plan to the CAP or other proficiency testing agency for and implementation of corrective action.

Laboratory Inspection

Molecular pathology laboratories must be inspected by an external inspection team every 2 years for compliance with CLIA regulations. LAP-CAP inspections are performed by a team of inspectors from an institution other than that of the laboratories being inspected, and the inspection team assesses the quality indicators for continued quality performance as listed in the LAP-CAP checklist.[31] The checklist addresses laboratory procedures including specimen processing and testing, reagents, controls, instrumentation, personnel, physical facilities, and laboratory safety. Inspection also assesses for PCR contamination control, participation in proficiency testing programs, and documentation of the QC and QA program activities (Table 49-6). Failure to comply is documented as a "Phase I" or "Phase II" deficiency. Most items on the molecular pathology checklist are Phase II, with a serious potential to affect patient care. Phase II deficiencies require immediate attention and correction, documented in writing to the CAP. Phase I deficiencies are less serious but also require documentation of corrective action. On alternate years, the laboratory is required to perform a self-inspection.

Recently, as a part of ongoing efforts to strengthen the laboratory accreditation process, to promote public confidence in the LAP-CAP, and to help assure regulators that CAP-accredited laboratories meet or exceed CLIA requirements, the CAP has made significant changes in its LAP, including unannounced routine inspections beginning in January 2006. This means that laboratories will not know the exact date of the inspection and the identity of the inspection team members prior to the inspection. Instead, the laboratories will be provided with a 90-day window of the anniversary of their last inspection, 45 days prior to and 45 days after, pending CMS approval, in which the inspection will occur. By implementing this process for routine inspections, CAP expects to help ensure that laboratories are in continuous compliance and that they are providing quality patient testing at all times.

Alternatively, laboratories can be inspected directly by CLIA or by JCAHO as an accrediting organization for clinical laboratories under the CLIA program. The CLIA inspection regulations are found in Subpart Q of the *Code of Federal Regulations*, which addresses both basic and specific inspection requirements.[32] CLIA inspections under JCAHO are conducted by CMS. The CMS policy for laboratory inspections includes announced initial and biennial recertification inspections and unannounced complaint and follow-up inspections. The CMS inspection focuses more on outcomes than on laboratory process. The goals of this outcome-oriented survey are not only to determine the laboratory's regulatory compliance but also to assist laboratories in improving patient care by emphasizing those aspects that have a direct impact on the laboratory's overall test performance.

Accreditation by JCAHO meets the conditions required by federal law and regulations. Consequently, laboratories that are accredited by JCAHO in lieu of direct federal oversight and that meet the CLIA requirements for laboratories are not subject to routine inspection by state survey agencies to determine their compliance with federal requirements.[33] To be eligible, JCAHO must apply standards and criteria that are equal to or more stringent than the requirements established by DHHS. The JCAHO accreditation process is more stringent than federal certification requirements in that JCAHO accredits a healthcare organization (for example, a hospital) as a whole, not just the laboratory. As such, an organization that loses JCAHO accreditation for reasons other than poor laboratory performance also would lose the ability to legally test human specimens under the laboratory's existing certificate of accreditation, if JCAHO accreditation is being used to meet the CLIA requirements. The JCAHO also began conducting unannounced inspections in 2006.

Summary

In this chapter, we have summarized many of the issues and considerations that are important for the successful management of a molecular pathology laboratory. Although these issues are addressed separately, many are interconnected. If even one function is weak, the success of testing is compromised. Molecular pathology is a rapidly growing field of anatomic pathology and laboratory medicine. The most significant challenges facing the management team of a molecular pathology laboratory are continuous acquaintance with newly developed technologies, new approaches for data analysis (bioinformatics), changing regulations, and evolving testing considerations for a variety of clinical applications, as well as understanding basic business concepts in marketing and finance. Keeping pace with these challenges ensures that the molecular pathology laboratory will be an even greater driving force of medical practice in the 21st century and beyond.

References

1. Klein M. Marketing research products to molecular pathologists. *BioTactics in Action.* 1998;1:1–5; Online Newsletter at www.biotactis.com/newsletter/v1i5/molbio.htm.
2. Grody WW, Matteson KJ, Murphy PD, et al. Molecular diagnostic methods for genetic diseases; approved guidelines. CLSI, MM1-A2. Vol 26. No. 27. Wayne PA, 2000.
3. Casey DK. Genes, dreams and reality. The promise and risks of the new genetics. *Judicature.* 1999;83:105–111.
4. Genetic Information Nondiscrimination Act of 2005 (S 306). Available at: http://www.govtrack.us/congress/bill.xpd?bill=s109-306.
5. Grody WW, Noll WW. Molecular diagnosis of genetic diseases. In: Henry JB, ed. *Clinical Diagnosis and Management by Laboratory Methods.* Philadelphia: WB Saunders; 2001:1372–1381.
6. Giardiello FM, Bresinger JD, et al. The use and interpretation of commercial APC gene testing for familial adenomatous polyposis. *N Engl J Med.* 1997;336:823–827.
7. National Human Genome Research Institute. *Promoting Safe and Effective Genetic Testing: Final Report of the Task Force on Genetic Testing.* Bethesda, MD: National Human Genome Research Institute; 1997.
8. US Department of Health and Human Services, Centers for Disease Control and Prevention. Notice of intent: genetic testing under the Clinical Laboratory Improvement Amendments. *Fed Regist.* 2000;65:25928–25934.
9. *Health Insurance Portability and Accountability Act of 1996.* Public Law 104–191. 104th Congress (August 21, 1996). Available at: http://aspe.hhs.gov/admnsimp/pl104191.htm.
10. Gulley ML. Leukemias and lymphomas. In: Leonard DGB, ed. *Diagnostic Molecular Pathology.* Philadelphia: WB Saunders; 2003:113–127.
11. Barr FG, Ladanyi M. Sarcomas. In: Leonard DGB, ed. *Diagnostic Molecular Pathology.* Philadelphia: WB Saunders; 2003:53–76.
12. Fallon KB, Palmer CA, Roth KA, et al. Prognostic value of 1p, 19q, 9p, 10q, and EGFR-FISH analyses in recurrent oligodendrogliomas. *J Neuropathol Exp Neurol.* 2004;63:314–322.
13. Unger ER, Guarner J. Infectious diseases. In: Leonard DGB, ed. *Diagnostic Molecular Pathology.* Philadelphia: WB Saunders; 2003:129–141.
14. Van Deerlin V. Identity testing: applications of DNA typing in anatomic pathology. In: Leonard DGB, ed. *Diagnostic Molecular Pathology.* Philadelphia: WB Saunders; 2003:163–188.
15. *Clinical Laboratory Improvement Amendments (CLIA) of 1988.* Public Law 100–578. 100th Congress (October 31, 1988). Available at: http://www.cms.hhs.gov/clia.
16. Ehrmeyer SS, Laessig HR. Compliance issues—the regulations. In: Harmening DM, ed. *Laboratory Management Principles and Processes.* Upper Saddle River, NJ: Prentice-Hall; 2003:225–243.
17. US Department of Health and Human Services. Medicare and CLIA programs. Regulations implementing the Clinical Laboratory Improvement of 1988 (CLIA). *Fed Regist.* 1992;57:7002–7186.
18. US Department of Health and Human Services. Medicare and CLIA Programs. Regulations implementing the Clinical Laboratory Improvement of 1988 (CLIA). *Fed Regist.* 2003;68:3639–3714.
19. Hoeltge GA, Dynek DA, Delaharty DC, et al. Clinical laboratory technical procedure manuals approved guidelines. Wayne, PA: National Committee for Clinical Laboratory Standards; 1996. Document GP-2A3-P.
20. Kaul KL, Leonard DG, Gonzales A, Garrett CT. Oversitght of genetic testing: an update. *J Mol Diag.* 2001;3:85–91.
21. Food and Drug Administration. Medical devices: classification/reclassification: restricted devices: analyte specific reagents. *Fed Regist.* 1997;62:62243–62260.

22. Takes PA, Murray G. The analyte specific reagent rule: a new set of regulations for labs, manufacturers. *Clin Lab News.* 1998;24:10.
23. Ferreira-Gonzales A. Establishing a molecular diagnostics laboratory. In: Henry JB, ed. Clinical Diagnosis and Management by Laboratory Methods. Philadelphia: WB Saunders. 2001:1333–1339.
24. Dadoun R. Impact on human resources: Core laboratory versus laboratory information system versus modular robotics. *Clin Lab Manage Rev.* 1998;12:248–255.
25. Udaykumar E, Epstein JS, Hewlett IK. A novel method employing UNG to avoid carry-over contamination in RNA-PCR. *Nucleic Acids Res.* 1993;21:3917–3918.
26. Dumler S, Fink LM, Frank TS, et al. Goals and objectives for molecular pathology education in residency programs. *J Mol Diag.* 1999;1:5–15.
27. Killeen AA, Leung WC, Payne D, et al. Certification in molecular pathology in the United States (Training and Education Committee, The Association for Molecular Pathology). *J Mol Diagn.* 2002;4:181–184.
28. Griffey PA. Fundamentals of financial management. In: Harmening DM, ed. *Laboratory Management: Principles and Processes.* Upper Saddle River, NJ: Prentice-Hall; 2003:163–179.
29. American Medical Association. 2006 CPT. Current Pnocedural Terminology (Professional Edition). Pathology and Laboratory Medicine Guidelines. Chicago: Americian Medical Association; 2005.
30. Grody WW, Cutting GR, Klinger KW, Richards CS, Watson MS, Desnick RJ. Laboratory standards and guidelines for population-based cystic fibrosis screening. *Genet Med.* 2001;3:149–154.
31. Laboratory Accreditation Program—College of American Pathologists. Molecular Pathology checklist. Northfield, IL: College of American Pathologists; 2005.
32. Department of Health and Human Services, Centers for Disease Control and Prevention. Subpart Q—Inspection. *Fed Regist.* 1998;63:986–987.
33. Department of Health and Human Services. CLIA Program: approval of the Joint Commission on Accreditation of Healthcare Organizations as an accrediting organization. *Fed Regist.* 1995;60:130–132.

Chapter 50

From Research to Clinical Practice

Karen Snow-Bailey

Molecular pathology cannot be discussed without including the translational nature of the clinical practice. While the number of Food and Drug Administration (FDA)-cleared or FDA-approved molecular pathology in vitro diagnostic test kits and analyte-specific reagents is increasing, the development of commercial test kits and reagents does not keep pace with the clinically useful genomic information generated by basic, translational, and clinical research. In the genomic in vitro diagnostic testing arena, the molecular pathology laboratory stands at the interface between science and medicine. With existing molecular biology methods, the vast majority of genomic discoveries linking genomic sequences or variations to disease states or risk can be translated into clinical tests in the molecular pathology laboratory. The laboratory director and personnel must understand the regulations and standards that govern this translational process to ensure that the high quality of clinical practice is assured.

Regulations for Clinical Laboratories

Laboratory Accreditation

State and federal regulations stipulate that only laboratories certified under the Clinical Laboratory Improvement Amendments (CLIA) may perform testing when test results are released to a patient or healthcare provider and results are used for medical management.[1] Many US laboratories participate in the Laboratory Accreditation Program offered by the College of American Pathologists, which has CLIA deemed status. New York State also has CLIA deemed status due to the high stringency of the state standards. The New York Clinical Laboratory Evaluation Program monitors the quality of New York State laboratories and laboratories in other states that perform testing on samples from New York.

Accreditation agencies in other countries include Clinical Pathology Accreditation Ltd (CPA; United Kingdom), the National Association of Testing Authorities (NATA; Australia), International Accreditation New Zealand (IANZ), the Dutch Accreditation Council (RvA; the Netherlands), and the European Molecular Genetics Quality Network (EMQN). Standards and checklists used for accreditation vary internationally but include CLIA, CAP checklists, and International Standards Organization (ISO) standards. Unfortunately, the lack of global standardization creates somewhat of a barrier to using offshore laboratories for clinical testing, because equivalency of accreditation status has not been determined. This is of particular relevance in testing for very rare disorders, for which only one or two clinical laboratories worldwide may be performing the testing.

Proficiency Testing

A requirement for laboratory accreditation is participation in an external proficiency testing program or external quality assurance (EQA) program, where one exists. The CAP offers surveys for numerous analytes (covering inherited disorders, infectious diseases, and somatic mutations). Programs also are provided by the EMQN, the United Kingdom External Quality Assurance System, the Human Genetics Society of Australasia (HGSA), and the Royal College of Pathologists of Australasia (RCPA). At an international level, the Organisation for Economic Co-operation and Development (OECD) has surveyed laboratories in member countries regarding quality assurance (QA) and proficiency testing practices.[2] If external proficiency testing is not available for a particular analyte or test, then interlaboratory sample exchange, retesting of previously analyzed samples, or both, are options to demonstrate proficiency. Recommendations have been made for expansion of proficiency testing programs and development of a wider range of QA materials.[3]

Personnel Certification

CLIA regulations outline minimum requirements for laboratory directors. Some states require certification of

technologists working in clinical laboratories. Certification processes and options have been reviewed.[4]

Regulation of Assays and Reagents

The FDA regulates commercial kits for clinical testing. Additionally, reagents making up assays developed by individual laboratories are regulated by the FDA as analytespecific reagents (ASRs). This means that the reagents are manufactured under current good manufacturing practices (CGMP) and are subject to restrictions on distribution, use, and labeling.[5] Laboratories using FDA-approved kits, ASRs or laboratory-developed methods are required to validate the tests under CLIA regulations prior to clinical implementation.

Participation in Research Studies

Regulatory Standards

Title 45, Part 46 (Protection of Human Subjects) of the *US Code of Federal Regulations* describes the policy of the US Department of Health and Human Services for the protection of human research subjects. The document covers functions of the Internal Review Board (IRB) and general requirements for informed consent. The relationship between study participants and researchers is addressed, including complex issues related to participants' access to test results from their own specimen(s).[6] Genetic counselors can play a significant role in facilitating access to research studies and ensuring appropriate informed consent.[7] Recommendations for consent form design and content are available.[8,9] Clinical laboratories that participate in research studies should comply with the research regulations, including assuring IRB review and approval of the research protocols, informed consent, and proper protection of the research information.

Academic Interest

Development of a clinical test may stem from a primary research interest within the laboratory or may develop through collaborations with basic science investigators, clinical investigators, or clinicians with a focused medical practice requiring molecular testing. Data gathered during research can suggest that a test will have clinical application. Ongoing investigation using clinical samples under an IRB-approved protocol may then be used to determine the clinical sensitivity, clinical specificity, and clinical utility of the test.[10] Even after a test becomes offered as a clinical service, ongoing academic interest can be very valuable in making further test improvements or furthering the molecular understanding of the disease (e.g., adding methodologies to increase the clinical sensitivity, elucidating genotype-phenotype correlations to enhance result interpretation, or identifying additional genetic loci or variants associated with the disorder).

Definition of a Clinical Test

For the consumer, a clinical test is a test used for patient management that is performed within the scope of practice of an accredited laboratory. For the laboratory, requirements are analytical and clinical validation of the test, with internal, peer, or regulatory review, or some combination of these, of the validation data, as required by regulation standards applicable to the specific laboratory, usually based on location. All aspects of the clinical test must be determined, including test conditions, controls, procedure manual, personnel training, and all other aspects of clinical testing, as described in chapter 49. Mechanisms and materials for pre- and postanalytical steps also must be implemented, including test information, instructions for specimen collection, result reporting, and billing.

Considerations in Introducing New Clinical Tests

Oversight of Molecular Tests

In 1995, the Task Force on Genetic Testing was convened by the National Institutes of Health–Department of Energy Working Group on Ethical, Legal, and Social Implications of Human Genome Research. The task force was charged to review genetic testing in the United States and to make recommendations where necessary to promote safe and effective testing practices.[11] As a result of the task force recommendations, the Secretary's Advisory Committee on Genetic Testing (SACGT) was formed to address issues associated with new genetic test development, including analytical and clinical validity, clinical utility, laboratory quality, and healthcare provider education. Functions of the SACGT have subsequently been assumed by the Secretary's Advisory Committee on Genetics, Health, and Society (SACGHS), whose charter is to explore, discuss, and make recommendations regarding various issues (e.g., medical, ethical, legal, and social issues) associated with the development and application of genetic technologies. The Office of Genomics and Disease Prevention (OGDP) at the Centers for Disease Control and Prevention (CDC) also plays an active role in genetic test development and monitoring as related to public health. Similar processes have been carried out, or are occurring, in other countries,[12] for example, by the Human Genetics Commission in the United Kingdom, the Institute for Prospective Technological Studies (IPTS) and the European Society of Human Genetics in Europe,[13] and through the Australian Health Ministers' Advisory Council (AHMAC).

Economic Considerations

The cost-effectiveness of genetic testing requires consideration of the prevalence of the disease (and associated genetic defects), burden of suffering, availability and efficacy of treatment, and cost and accuracy of genetic testing. As illustrative examples, a recent article evaluates cost-effectiveness of genetic tests for hereditary nonpolyposis colorectal cancer, familial breast cancer, and periodontal disease; CYP2C9 pharmacogenetic testing for warfarin therapy; and thiopurine S-methyltransferase for 6-mercaptopurine therapy.[14]

Medicare payments for clinical tests are determined by the Current Procedural Terminology (CPT) codes that apply to the specific test. Each CPT code has a monetary value. For many tests, the value of CPT codes is insufficient to cover the cost of performing the test.[15,16] SACGHS has developed recommendations to address the inadequacy of reimbursement for genetic testing.

Predicted Test Volume

Test volume will affect how frequently a test is performed and the batch size of each test run. Test costs should be calculated, incorporating fixed and variable costs accounting for the batch size of each test run, with adjustment for control costs. Tests run in small batches will be more costly than tests with high batch sizes, predominantly because the labor costs are distributed over fewer tests with smaller batch sizes. The test volume should be high enough for the laboratory to maintain competency in performing the test and interpreting the results.

Resource Requirements

Staffing resources are required for test development, implementation, and ongoing provision of new clinical tests. The adequacy of staffing and equipment for the additional testing must be assessed and business plans developed and implemented to obtain any additional resources required for the new testing. New reagents will be required, including control materials (e.g., from repositories such as Coriell Cell Repositories, American Type Culture Collection (ATCC), European Collection of Cell Cultures, and the National Institute of Standards and Technology). Recent research initiatives have been successful in developing additional QA materials for genetic testing.[17,18] Novel testing platforms and technologies are being introduced into the molecular marketplace at an astonishing rate. Unfortunately, there is no one-size-fits-all technology in molecular pathology practice. The nature of the genetic defect, predicted test numbers, and instrument, reagent, and royalty costs should be considered in determining the most appropriate method and instrument for each test.[19,20]

Intellectual Property

Genes, mutations, methods, and technologies may be associated with intellectual property rights and patents. Many academic organizations now have intellectual property offices that work with researchers to identify "inventions" that can be patented. An overview of the patent process, a "patent primer," has been developed by the American College of Medical Genetics (ACMG).[21,22] The Genetics and Molecular Medicine section of the American Medical Association Web site contains additional overview material and useful links.[22] Clinical laboratories wishing to implement a new test should be aware that patents may limit the utilization of specific inventions pertinent to that test. Indeed, the patents and licensing practices for genetic sequences associated with disease processes have a significant impact on molecular pathology laboratories and have recently caused concern, particularly when licenses are not broadly available for clinical practice.[23] In response to this concern, the ACMG has prepared an educational document for consideration in preparing licenses with guidelines that are compatible with wide access to affordable genetic testing.[24]

The Test Development Process

ACMG Standards and Guidelines for Clinical Genetics Laboratories describe three levels in the development of a diagnostic test: research and development, investigational studies, and accepted clinical test.[20] In the United States, CLIA regulations require clinical tests to have been validated by the testing laboratory. As outlined in the ACMG standards and guidelines, test validation requires attention to pretest components (e.g., sample and information collection), assay methodology, analytical validity, clinical validity, clinical utility, and ethical, legal, and social issues. Many of these aspects and additional items such as cost analysis and billing have been discussed.[25]

Analytical Validity

The development of assay protocols should utilize appropriately maintained and calibrated equipment, testing standards, quality controls, and defined interpretation criteria. For numerical results, reference ranges need to be determined and sources of assay variability need to be defined and quantified. Analytical validity requires assessment of the analytical sensitivity and analytical specificity. Although publications have addressed analytical validity of specific tests among cohorts of clinical laboratories, each testing laboratory must determine the analytical sensitivity and specificity of its own test. The test procedure must be documented, for example, by following guidelines established by the Clinical and Laboratory Standards Institute (CLSI; formerly NCCLS).[26] Other CLSI documents, available on the CLSI Web site (http://www.clsi.org/source/

custom/sortby.cfm?category-category) under the section heading "Molecular Methods", provide guidelines for the application of molecular methods for genetic diseases, infectious diseases, hematopathology, and fluorescence in situ hybridization (FISH).

Clinical Validity

Clinical validity is an assessment of the frequency with which the test identifies individuals who have the disorder or clinical phenotype identified by the test. Components of clinical validity are clinical sensitivity and clinical specificity. Data may be derived from the laboratory's own studies or by assessment of published data, or both.

Clinical Utility

Clinical utility incorporates the impact and usefulness of the test weighed against the cost of testing and any adverse health or psychosocial consequences. Factors to consider include available interventions following positive or negative test results, economic or psychosocial benefits or harms of testing, and implications of results for other family members.

Informational Materials and Reports

In addition to development of a written test procedure, the laboratory is required to provide information for the healthcare provider on specimen requirements and transportation, test requisition requirements including informed consent, test turnaround time, and other information relevant to the ordering and use of testing.

The report format and content need to be established. Reports should be interpretable by nongeneticist healthcare providers. Studies undertaken during recent years highlight some of the difficulties encountered by primary care physicians and other nongeneticists in correctly interpreting genetic risk information and in appreciating the limitations of genetic testing.[27,28] These studies are instructive for laboratory directors in pointing out the need to avoid jargon as much as possible and to consider the importance of the report format. Standards and guidelines, best practice guidelines, and checklists for reporting are available.[20,29,30] Disease-specific guidelines also are helpful in report formatting and content. For example, both the EMQN guidelines for Prader Willi-Angelman syndromes and for Friedreich ataxia[31] and the ACMG guidelines for cystic fibrosis and for fragile X syndrome[20] discuss key elements to include in the reports and provide suggestions for interpretive comments.

Mutation databases are invaluable resources in the interpretation of sequence alterations detected in many genes and disorders. However, such databases are not peer-reviewed or quality-controlled resources. Therefore, peer-reviewed publications for a particular mutation are essential for providing the most appropriate report interpretation. A challenging issue is interpretation of missense alterations when no previous data exist to substantiate the change as either a pathogenic mutation or a benign polymorphism. Methods to characterize variants, using BRCA1 mutation interpretation as a model, have been described.[32]

Considerations for molecular pathology reports for tests other than inherited disorders are available.[33] Discussions include consideration of creating "unified" reports, for example, incorporating molecular, flow cytometry, and surgical pathology into a single hematopathology report.

Special Considerations for Rare Disorders

Conferences held in 2004 and 2005 evaluated the status of testing for rare disorders and considered issues related to quality, availability, access, and resources for rare disease testing.[34] Several recommendations for promoting quality testing were made, and issues requiring further discussion were identified. Subsequently, the ACMG published Technical Standards and Guidelines for Molecular Genetic Testing for Ultra-Rare Disorders.[35] These guidelines acknowledge unique challenges associated with rare disorder testing that affect test validation (analytical and clinical), quality assurance, and result interpretation.

A number of clinical laboratories offer testing to confirm mutations identified in research laboratories. These services range from provision of testing for a limited number of disorders to custom mutation analysis for perhaps any rare disorder for which a research laboratory has identified a pathogenic mutation. Laboratories that offer this clinical confirmation can be located in the GeneTests directory available at www.genetests.org/servlet/ under the heading "Custom Molecular Genetic Testing."

Resources for Locating Testing Laboratories

Association for Molecular Pathology Test Directory

The Association for Molecular Pathology (AMP; see the AMP Test Directory, available at: http://www.amptestdirectory.org) maintains a directory of laboratories that provide molecular testing for infectious diseases, solid tumors, and hematopathology. The directory is searchable by various fields, including name of the disorder, molecular abnormality, disease agent, and laboratory location.

GeneTests

The GeneTests laboratory directory (http://www.genetests.org) includes international laboratories that perform testing for inherited disorders. Listed tests include mole-

cular, biochemical, and FISH tests. Searchable fields include disease name, gene, clinical feature, OMIM number, and laboratory location. The GeneTests resource also includes GeneReviews, a clinic directory, and educational materials.

The European Directory of DNA Diagnostic Laboratories

The European Directory of DNA Diagnostic Laboratories (EDDNAL; http://www.eddnal.com) provides information on molecular testing services for heritable syndromes and disorders offered by laboratories throughout the European Union countries and Norway, Poland, and Switzerland. Searchable fields include disease name, OMIM number, and laboratory name or location.

Orphanet

The Orphanet Web site (http://www.orpha.net/) provides information on clinical- and research-based testing for rare diseases, as well as information on support groups and clinical trials. Search terms include disease name, OMIM number, clinical signs, and laboratory name or location.

GENDIA

GENDIA (for genetic diagnostics; http://www.gendia.net/) is an international network of more than 50 laboratories located in the United States, Europe, and Australia. Samples submitted for testing are sent to a single laboratory for distribution to the testing laboratories. Available tests can be located by disease or gene.

Summary

Our knowledge of the molecular pathology of inherited and infectious diseases and of numerous cancers has vastly increased over the past several years. Additionally, the practice of molecular pathology has been transformed by the availability of new technologies, testing platforms, and automation. Alongside the growth in knowledge and available technologies has been the introduction of numerous standards, guidelines, and regulations that are designed to ensure the provision of safe and accurate molecular tests in the clinical setting. These documents provide guidance to laboratories that are introducing new clinical tests. Ongoing challenges such as the limited availability of QA materials and proficiency testing programs are being actively addressed, as are the economic, ethical, legal, and social issues associated with molecular testing. The continued growth of molecular pathology practice holds great promise for the future practice of medicine.

References

1. Schwartz MK. Genetic testing and the clinical laboratory improvement amendments of 1988: present and future. *Clin Chem.* 1999;45:739–745.
2. Organisation for Economic Co-operation and Development. Quality assurance and proficiency testing for molecular genetic testing. Summary results of a survey of 18 OECD countries. Available at: http://www.oecd.org/dataoecd/25/12/34779945.pdf.
3. Williams LO, Cole EC, Lubin IM, Iglesias NI, Jordan RL, Elliott LE. Quality assurance in human molecular genetics testing: status and recommendations. *Arch Pathol Lab Med.* 2003;127:1353–1358.
4. Killeen AA, Leung WC, Payne D, et al. Certification in molecular pathology in the United States (Training and Education Committee, The Association for Molecular Pathology). *J Mol Diagn*. 2002;4:181–184.
5. Gutman S. The role of Food and Drug Administration regulation of in vitro diagnostic devices—applications to genetics testing. *Clin Chem*. 1999;45:746–749.
6. American Society of Human Genetics Ad Hoc Committee on Consumer Issues. Genetic lay advocacy groups: significant others in the conduct of human genetics research. October 1999 (revised November 10, 2000 and October 18, 2001). Available at: http://genetics.faseb.org/genetics/ashg/pubs/policy/pol-46.pdf.
7. Markel DS, Yashar BM. The interface between the practice of medical genetics and human genetic research: what every genetic counselor needs to know. *J Genet Couns*. 2004;13:351–368.
8. The American Society of Human Genetics Statement on Informed Consent for Genetic Research. *Am J Hum Genet*. 1996;59:471–474.
9. Hull SC, Gooding H, Klein AP, Warshauer-Baker E, Metosky S, Wilfond BS. Genetic research involving human biological materials: a need to tailor current consent forms. *IRB*. 2004;26:1–7.
10. American College of Medical Genetics Standards and Guidelines for Clinical Genetics Laboratories. 2005 ed. Available at: http://www.acmg.net/Pages/ACMG_Activities/stds-2002/stdsmenu-n.htm.
11. Holtzman NA. Promoting safe and effective genetic tests in the United States: work of the task force on genetic testing. *Clin Chem*. 1999;45:732–738.
12. Cox SM, Faucett WA, Chen B, et al. International genetic testing. *Genet Med.* 2003;5:176–182.
13. Godard B, Kaariainen H, Kristoffersson U, Tranebjaerg L, Coviello D, Ayme S. Provision of genetic services in Europe: current practices and issues. *Eur J Hum Genet*. 2003;11(suppl 2):S13–S48.
14. Higashi MK, Veenstra DL. Managed care in the genomics era: assessing the cost effectiveness of genetic tests. *Am J Manag Care*. 2003;9:493–500.
15. Root CB. Medicare coding and reimbursement for clinical laboratory services. *Clin Chem*. 1998;44:1713–1727.
16. Logue LJ. Genetic testing coverage and reimbursement: a provider's dilemma. *Clin Leadersh Manag Rev*. 2003;17:346–350.
17. Bernacki SH, Stankovic AK, Williams LO, et al. Establishment of stably EBV-transformed cell lines from residual clinical blood samples for use in performance evaluation and quality assurance in molecular genetic testing. *J Mol Diagn*. 2003;5:227–230.
18. Jarvis M, Iyer RK, Williams LO, et al. A novel method for creating artificial mutant samples for performance evaluation and quality control in clinical molecular genetics. *J Mol Diagn*. 2005;7:247–251.
19. Amos J, Patnaik M. Commercial molecular diagnostics in the U.S.: The Human Genome Project to the clinical laboratory. *Hum Mutat*. 2002;19:324–333.
20. Fortina P, Surrey S, Kricka LJ. Molecular diagnostics: hurdles for clinical implementation. *Trends Mol Med*. 2002;8:264–266.
21. Murphy PD. American College of Medical Genetics Intellectual Property Subcommittee of the Committee on the Economics of Genetic Services. Patent Primer. October 2003. Available at: http://www.acmg.net.
22. American Medical Association. Gene patenting. Available at: http://www.ama-assn.org/ama/pub/category/2314.html. Updated December 10, 2004.

23. Cho MK, Illangasekare S, Weaver MA, Leonard DG, Merz JF. Effects of patents and licenses on the provision of clinical genetic testing services. *J Mol Diagn*. 2003;5:3–8.
24. Tait JF, Intellectual Property Subcommittee of the Economics Committee of the ACMG. Points to consider in preparing license agreements for patented genetic tests. November 2004. Available at: http://www.acmg.net.
25. Association for Molecular Pathology. Recommendations for in-house development and operation of molecular diagnostic tests [statement]. *Am J Clin Pathol*. 1999;111:449–463.
26. Clinical and Laboratory Standards Institute. Available at: http://www.clsi.org.
27. Giardiello FM, Brensinger JD, Petersen GM, et al. The use and interpretation of commercial APC gene testing for familial adenomatous polyposis. *N Engl J Med*. 1997;336:823–827.
28. Sandhaus LM, Singer ME, Dawson NV, Wiesner GL. Reporting BRCA test results to primary care physicians. *Genet Med*. 2001;3:327–334.
29. Clinical Molecular Genetics Society. Best Practice Guidelines. Available at: http://www.cmgs.org/BPG/Default.htm.
30. Swiss Society of Medical Genetics. Best Practice Guidelines. Available at: http://www.ssgm.ch/sections/Documents/Statements/publications.htm.
31. European Molecular Genetics Quality Network. Best Practice Guidelines. Available at: http://www.emqn.org/bpguidelines.php.
32. Judkins T, Hendrickson BC, Deffenbaugh AM, Scholl T. Single nucleotide polymorphisms in clinical genetic testing: the characterization of the clinical significance of genetic variants and their application in clinical research for BRCA1. *Mutat Res*. 2005;573:168–179.
33. Check W. Fleshing out molecular pathology reports [and related articles by Check W and Gulley M]. CAP Today. February 2004. Available at: http://www.cap.org.
34. Available at: https://www.phppo.cdc.gov/dls/genetics/RareDiseaseConf.aspx.
35. Maddalena A, Bale S, Das S, Grody W, Richards S, ACMG Laboratory Quality Assurance Committee. Technical standards and guidelines: molecular genetic testing for ultra-rare disorders. *Genet Med*. 2005;7:571–583.

Index