Sameek Roychowdhury
Eliezer M. Van Allen *Editors*

Precision Cancer Medicine

Challenges and Opportunities

Springer

Precision Cancer Medicine

Sameek Roychowdhury • Eliezer M. Van Allen
Editors

Precision Cancer Medicine

Challenges and Opportunities

Springer

Editors
Sameek Roychowdhury
Department of Internal Medicine
Division of Medical Oncology
Comprehensive Cancer Center &
The Arthur G. James Cancer Hospital at
The Ohio State University
Columbus, OH, USA

Eliezer M. Van Allen
Dana-Farber Cancer Institute
Boston, MA, USA

ISBN 978-3-030-23636-6 ISBN 978-3-030-23637-3 (eBook)
https://doi.org/10.1007/978-3-030-23637-3

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Chapter 1
Between Hype and Hope, on the Cutting Edge of Precision Cancer Medicine

Sameek Roychowdhury and Eliezer M. Van Allen

The concept of precision cancer medicine emerged with a National Academics of Science report in 2011 describing the need and opportunity to classify human disease through a precision taxonomy. This new taxonomy would be based on the availability of genomics data and other sources of big data to enable a more precise diagnosis and management of human disease. Since cancer is characterized by somatic and germline genetic changes, precision medicine has taken a natural direction that has included clinical oncology. This text includes contributions from leading researchers in the emerging and changing field of precision cancer medicine.

The discovery and characterization of biomarkers and their inherent biology has been dependent on clinical observations and concurrent technologies to study cancer biology. While next generation sequencing (NGS) technologies has been the main accelerator for the identification of biomarker targets or vulnerabilities and the development of matching therapies, there were early examples that predated NGS. This includes the estrogen receptor in breast cancer where patients with metastatic breast cancer serendipitously benefited from hormone deprivation with oophorectomy. Subsequent biomarkers were discovered through early approaches for chromosome analysis including cytogenetics to identify and classify hematologic malignancies such as acute promyelocytic leukemia (t,15;17 translocation involving the retinoic acid receptor) and chronic myeloid leukemia (BCR-ABL1 translocation or Philadelphia chromosome). The development of polymerase chain reaction, FISH, and microarray were concurrent with the discoveries of the

S. Roychowdhury (✉)
Department of Internal Medicine, Division of Medical Oncology
Comprehensive Cancer Center & The Arthur G. James Cancer Hospital
at The Ohio State University, Columbus, OH, USA
e-mail: Sameek.roychowdhury@osumc.edu

E. M. Van Allen
Dana Farber Cancer Institute, Harvard Medical School, Boston, MA, USA
e-mail: eliezerm_vanallen@dfci.harvard.edu

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_1

Epidermal Growth Factor Receptor (EGFR) mutations in lung cancer and HER2 (ERBB2) amplification in breast cancer, which are sensitive to corresponding targeted therapies. However, what these technologies in the laboratory lacked was the scalability, affordability, and speed needed to characterize thousands of patients in the clinic for individualized or precision cancer care.

Around 2010, next generation sequencing technologies accelerated large-scale projects to study the most common cancer types through international efforts including The Cancer Genome Atlas and International Cancer Genome Consortium. These efforts helped to identify the landscape of genomic, epigenetic, and transcriptomic alterations in cancer. Advances in technology and drug development have accelerated the time between "target discovery" to "first patient treated in a clinical trial" (Fig. 1.1). Further, these studies described the vast inter-patient genomic heterogeneity that exists in each cancer type. This illustrated the need for individualized characterization of each patient's cancer in the clinic and that there is no routine cancer.

In 2011, there were early efforts to bring NGS approaches to patient care in real-time and this has led to new questions being addressed through clinical and research efforts:

- How do we characterize one person's cancer?
- Should we get new biopsies? Does cancer change over time?
- Should we sequence the whole genome, the exome, targeted exome, or RNA?

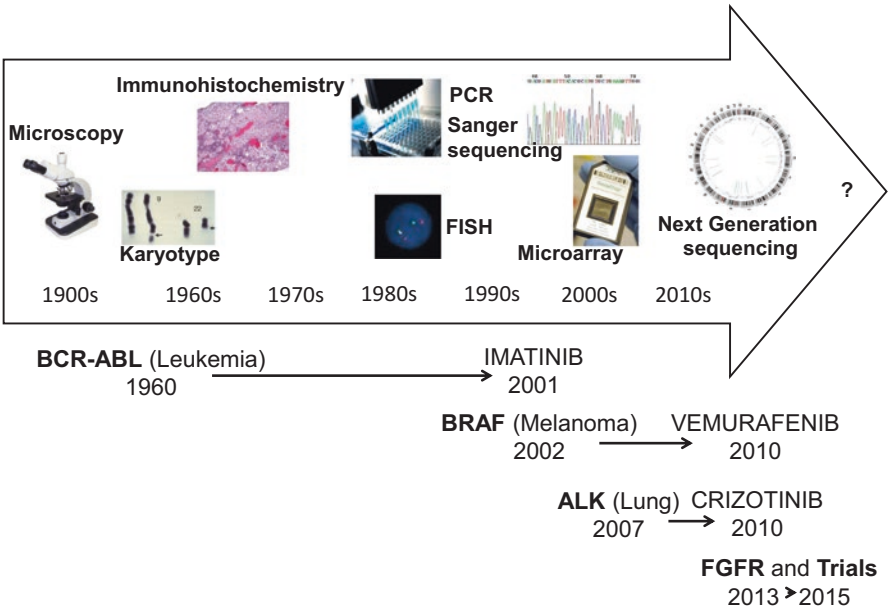

Fig. 1.1 Technology advances precision oncology. This timeline compares advances in genomic technologies and drug development timelines for novel targets in cancer

- What is the best way to analyze this data? Which algorithm is better?
- How do we interpret somatic findings of unknown significance?
- How do we develop and deliver clinical grade diagnostics? What are ideal positive and negative controls? How many?
- What constitutes a proper analytic validation of an assay?
- Are all commercial and academic tests the same quality for point mutations? Amplifications? Fusions?
- Where do store this data? How do we share this data?
- How do we report these results to physicians? What should reports look like?
- How do we connect genomic results to eligible therapies or clinical trials?
- How should we design clinical trials for patients with rare mutations?
- How do we develop and approve new therapies for patients with rare mutations?
- How can we learn from rare patients with an exceptional response to therapy?
- How do connect genomics data to the electronic medical record?
- How do interpret and manage germline findings?
- How do connect genomics data to clinical outcomes?
- How does genomics influence immunotherapy?
- How can we use genomics to study drug resistance?
- How do we speed up genomic testing in the clinic?
- How do we educate our physician workforce?
- Can we connect genomics data to prognosis for indolent and aggressive cancer subtypes?
- Can we use liquid biopsy from the blood instead of tumor tissue?
- What is the concordance of somatic alterations in blood and tumor tissue?
- How do we study clonal hematopoiesis?
- How do get more patients on clinical trials? Or
- How do we characterize patients with no driver mutations or quite genomes?
- How do we bring genomics to health care sites where resources are scarce?
- Why are only 10% of patients who undergo genomic testing able to receiving matching targeted therapies?

There are many answers and ongoing efforts to address these questions. Importantly, there are patients living longer, benefiting from therapy, and even cured thanks to precision cancer medicine approaches. This textbook covers a range of topics from basic science to clinical application for patient care to help describe these solutions and the new problems we need to solve to deliver precision cancer medicine.

Chapter 2
Molecular Diagnostics in Cancer: A Fundamental Component of Precision Oncology

Wendy Yang and Michael F. Berger

Keywords Pathology · Molecular Diagnostics · Precision Oncology · Biomarkers · Diagnosis · Prognosis · Disease Monitoring · Next-Generation Sequencing · Liquid Biopsy

Precision oncology embodies the targeted and rational treatment of cancer according to the specific molecular alterations underlying an individual patient's disease. It is now well-established that cancer is primarily a genetic disease caused by inherited and acquired genomic aberrations [1]. Each cancer carries a unique set of "driver" genomic aberrations that work together to promote cancer initiation and maintenance. It is also understood that cancer genomes are inherently unstable, and the accumulation of new driver genomic aberrations can lead to cancer progression and drug resistance [2]. The promise of precision oncology postulates that each cancer can be treated more effectively, and even possibly cured, through understanding and targeting key driver genomic aberrations. This requires first identifying key oncogenic genomic aberrations in cancer cells, enabled via molecular diagnostics, to guide the rational selection of molecular therapeutic agents specifically targeting these alterations. Thus, cancer molecular diagnostics are a fundamental and integral component of precision oncology.

2.1 History of Molecular Diagnostics in Cancer

At the beginning of the twentieth century, following remarkable advancements in physics and chemistry, biology became the "new frontier" in searching for a fundamental understanding of the empirical world. The term "molecular biology" was coined in the 1930's [3] only as a theoretical concept reflecting the hope that understanding life at its most fundamental level could be explained in physical and

W. Yang (✉) · M. F. Berger
Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA
e-mail: wendy.yang@moffitt.org; bergerm1@mskcc.org

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_2

chemical terms. The birth of molecular biology as a coherent discipline occurred in the 1950's, when Francis Crick and James Watson proposed the duplex model of DNA biopolymers [4, 5] leading to the central dogma of molecular biology describing the flow of genetic information among the three biopolymers from DNA to RNA to protein. The genes or genetic codes stored in the sequences of DNAs, are expressed by first transcribing DNAs to RNAs, and then translating RNAs to the final functional products of the genes: proteins [6–9]. By definition, molecular diagnostics represent the application of molecular biology to medical testing and encompass the analysis of biomarkers in any of the three biopolymers by using a collection of molecular biology techniques. However, despite the biological significance of proteins, progress of protein-based molecular technologies has significantly lagged behind that of nucleic acid-based methodologies due to the technical difficulties of working with proteins that are structurally and functionally far more complex. Thus, current routine molecular diagnostics are mostly limited to nucleic-acid based testing, which will be the focus of this chapter. Historically, the field of molecular diagnostics was born out of molecular genetic research laboratories investigating human hereditary disease in the 1980's, limited initially to the diagnosis of a few uncommon genetic diseases such as thalassemias [10–13]. In the 1990's, the molecular diagnostic field was dominated largely by high-volume molecular diagnostics of infectious diseases such as Chlamydia tracomatis and HIV [14]. Although investigations into cancer have always been a substantial component of molecular biology research since the 1950's, molecular diagnostics of cancer did not rise into prominence until decades later.

The beginning of precision oncology was marked by the emergence of two molecularly targeted therapies in the late 1990's. Trastuzumab (Herceptin), a monoclonal antibody targeting the receptor tyrosine kinase HER2 (ErbB2) protein, was approved by the US Food and Drug Administration (FDA) in 1998 to treat patients with HER2 gene amplification-positive metastatic breast cancer [15]. Imatinib (Gleevec), a small molecule tyrosine kinase inhibitor (TKI), was approved by the FDA in 2001 to treat patients with chronic myelogenous leukemia (CML) in 2001 [16]. The landmark "bench-to-bedside" story of imatinib as a treatment for CML actually began in the 1960's and the 1970's, when a small aberrant chromosome resulting from translocation between chromosomes 9 and 22, coined the Philadelphia chromosome, was discovered to be consistently linked to CML [17, 18]. This rearrangement produced the BCR-ABL fusion gene, homologous in sequence to a mouse retroviral oncogene ABL, with tyrosine kinase activity and cancer transforming capability [19–23]. Molecular functional studies at that time also confirmed that the BCR-ABL fusion protein had tyrosine kinase activity and was oncogenic [24–31]. In the mid 1990's, imatinib was found to suppress the oncogenic tyrosine kinase activity of the BCR-ABL fusion protein in the laboratory [32]. Prior to imatinib, chronic phase CML (CP-CML) was a deadly fatal malignancy with inevitable progression to the acute leukemia phase (blast phase CML or BP-CML), yet it is now a manageable chronic condition in the majority of patients with long term sustained remission or possible cure. The remarkable success of imatinib in the treatment of

CML validated the promise of molecularly targeted therapies and brought the hope of curing all cancers through the precision oncology approach.

A key feature of cancer genomes, and a central challenge of molecular diagnostics, is the multitude of different classes of genomic alterations that may occur and lead to dysregulation of cellular processes [2, 33]. Tumor genomic alterations may represent sequence variants and structural variants. Sequence variants include single nucleotide variants (SNV) and small deletions or insertions (indels), spanning fewer than 50 base pairs. SNVs are the most common type of genetic alteration, typically representing >70% of all mutations in a given tumor [2]. Structural variants (SVs) include both copy number variants (CNVs) in which genomic regions are gained or lost, and copy neutral SVs in which the total copy number remains diploid. CNVs may represent broad chromosome-level gains and losses or more focal amplifications and deletions targeting individual genes. Copy neutral SVs include balanced inversions (a segment of DNA with its two ends flipped on a chromosome), balanced translocations (a segment of DNA joined to a distant locus on another chromosome), and copy neutral-loss of heterozygosity (CN-LOH; duplication of one parental chromosome or chromosomal region and concurrent loss of the other allele) (Fig. 2.1). Unbalanced translocations, inversions, or LOH may also occur with additional loss or gain of genetic loci within the affected DNA region. Oncogenic genomic alterations not only involve protein coding regions but may also involve non protein-coding regions including cis-regulatory elements (i.e., promoters or enhancers), genes for regulatory non-coding RNAs (e.g., miRNA, lnRNA) and some pseudogenes [33]. Additionally, epigenetic aberrations including aberrant gene silencing through DNA methylation or aberrant gene activation/silencing through histone modifications, may lead to dysregulation of cell signaling pathways that promote oncogenesis and tumor progression [2, 33].

The major focus of current cancer molecular diagnostics is the detection of key acquired (somatic) cancer driver mutations that arise during the lifetime of individuals. In contrast to inherited (germline) genomic alterations that are usually uniformly present in every cell of the body, cancer somatic mutations are only present in cancer cells and their detection is thus technically more challenging. First, while germline mutation testing has access to abundant fresh cells from peripheral blood or buccal swabs for extraction of high quality nucleic acids, tumor somatic mutation testing is often challenged by limited quantities of tumor tissue due to the small volume of biopsy samples, driven by the increased utilization of minimally invasive tissue sampling techniques. Somatic mutation testing also faces the challenge of suboptimal quality of tumor nucleic acids, which are usually extracted from tumor tissue routinely embedded in formalin fixed, paraffin embedded (FFPE) blocks. Formalin fixation is known to cause fragmentation and chemical modification of nucleic acids, leading to difficulty in amplifying DNA molecules and potential downstream sequencing artifacts [34, 35]. Although optimization of tumor nucleic acid extraction techniques have since enabled routine clinical DNA sequencing using tumor DNA from FFPE tumor specimens [36], clinical DNA methylation sequencing and RNA-seq using tumor nucleic acids extracted from FFPE samples have not yet been robustly established [37, 38]. Second, cancer cells containing

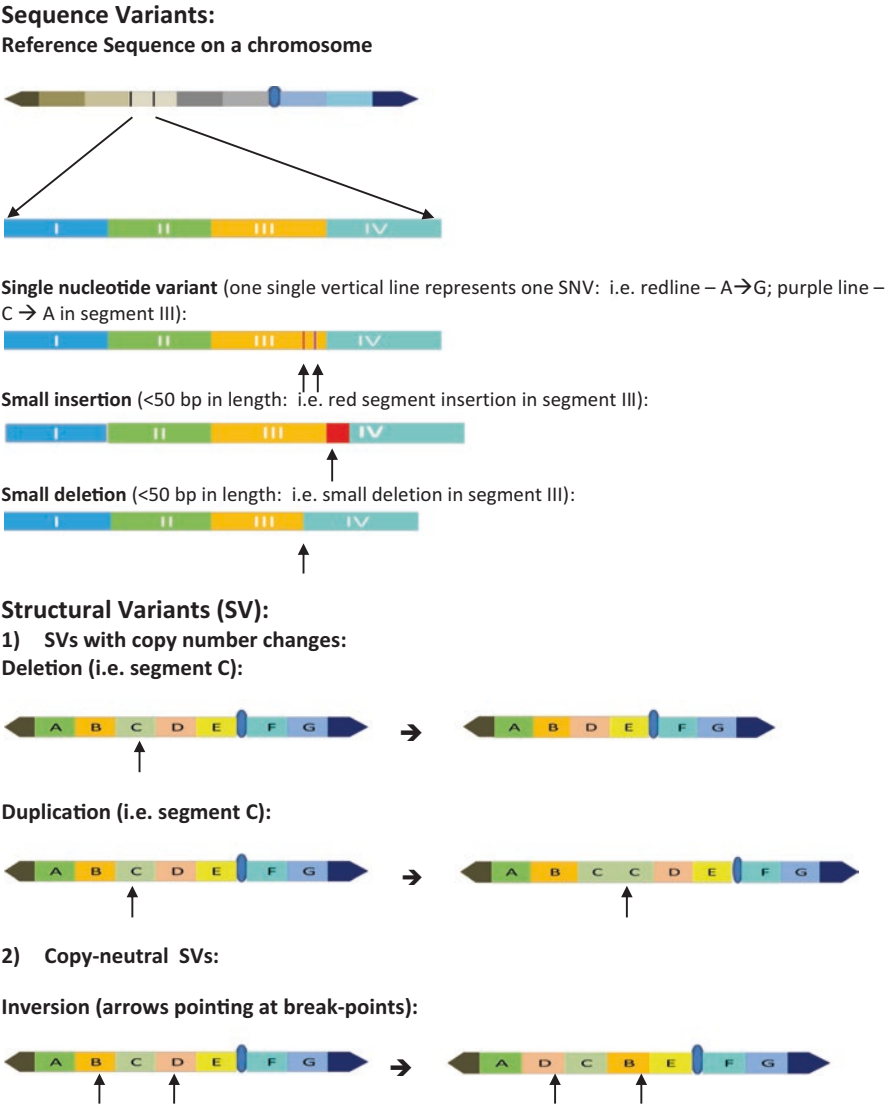

Fig. 2.1 Types of genomic alterations

somatic mutations are intermixed with background benign cells that do not harbor the mutations, leading to reduced mutant allele fractions (MAFs). The MAF of a cancer somatic mutation is equal to the number of somatic mutant alleles divided by the total number of both somatic mutant alleles and wild-type alleles in a mixed population of cancer cells and background benign cells. For a pure and genetically homogeneous diploid cancer cell population, somatic MAFs can either be 50% (heterozygous mutations) or 100% (homozygous mutations). However, low tumor

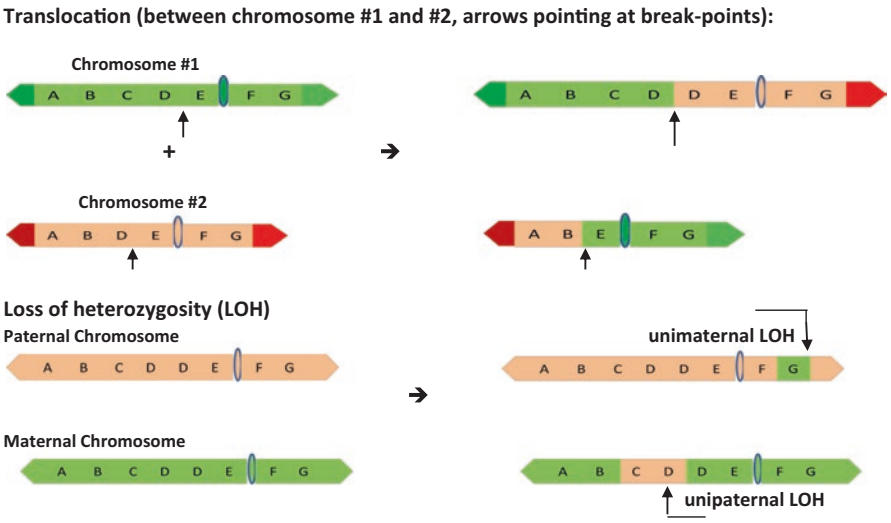

Fig. 2.1 (continued)

purity, which is not uncommon in tumor specimens, can result in substantially lower somatic MAFs. In addition, somatic MAFs may be further reduced due to intra-tumor genetic heterogeneity secondary to inherent cancer genomic instability and clonal evolution, with certain somatic mutations present in only a subset of cancer cells. For example, a sub-clonal somatic heterozygous mutation with 10% tumor purity and presence in only 40% of cancer cells will result in a low MAF of approximately 2%. Facing the challenges of limited tumor quantity and quality and low somatic MAFs due to low tumor purity and/or intra-tumoral genetic heterogeneity, tumor somatic mutation screening thus demands higher analytical sensitivity than that of germline mutation tests.

Prior to the implementation of a molecular test, thorough test validation is required to establish its performance characteristics such as analytical sensitivity, specificity, accuracy, and precision for detecting the intended alteration or class of alterations, compliant with the regulations and guidelines of regulatory agencies governing clinical laboratories such as Clinical Laboratory Improvement Amendments (CLIA) and College of American Pathologists (CAP) [39]. In addition, its clinical validity (the ability of the test to predict the presence or absence of the clinical condition that the test is developed to detect) and clinical utility (the value of the test to determine treatment or patient management) are also important considerations.

Various nucleic acid-based technologies have been used to detect genetic, epigenetic and gene expression alterations in molecular diagnostic labs, many of which were initially developed as molecular biology research tools. Until very recently, most molecular technologies used for routine clinical molecular oncology testing were low-throughput, owing to both their technical limitations and the clinical

necessity to test only a small number of biomarkers [40]. Some of these technologies are suitable for detecting structural variants (e.g., fluorescence in-situ hybridization (FISH) and reverse transcription-PCR (RT-PCR), while the others are suitable for detecting mutations (e.g., Sanger sequencing, quantitative polymerase chain reaction (PCR), and conventional PCR with or without using additional technologies (e.g., allele-specific hybridization, (capillary) gel electrophoresis) for signal detection [41]. Major milestones for the development of molecular technologies prior to the millennium are listed below in chronological order:

1960's: Nucleic acid hybridization was first developed [10, 12]. Nucleic acid hybridization is one of the most basic molecular techniques in molecular biology research and diagnostics alike. Numerous methodologies have been developed based on hybridization of complementary strands of DNA and/or RNA.

1975–1977: The first generation Sanger DNA sequencing method was invented [42, 43], which was further optimized and automated over the following decades [10, 12]. It is based on selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication. Automated Sanger sequencing has long served as the gold standard in molecular diagnostics for detecting SNVs or small indels with single nucleotide resolution. While the read length of Sanger sequencing is relatively long (up to 1000 base pairs) it is only semi-quantitative with a limit of detection of MAF near 20%. Bisulfite sequencing, a variation of Sanger sequencing, can be used to identify DNA methylation [44].

1982: FISH was developed as a hybridization-based cytogenetic technique that uses fluorescent probes binding specifically to chromosomal regions with a high degree of sequence complementarity [45]. FISH was developed to significantly improve the resolution over that of classic banding karyotype, from millions of base pairs to thousands of base pairs. FISH is semi-quantitative and can be used to detect all SV forms.

1985–1986: PCR was developed, representing one of the most important and influential scientific advances in the twentieth century. It has revolutionized both molecular biology research and molecular diagnostics. PCR is able to rapidly generate millions of copies of a target DNA fragment from a very small amount of DNA, bringing dramatically improved sensitivity, specificity and speed to DNA-based molecular analysis. PCR is still the most widely used molecular technique in molecular diagnostics, with broad applications in DNA sequencing, allele-specific mutation detection, gene expression and DNA methylation detection. Further improvements and variations of PCR include:

- Quantitative PCR: Real-time PCR [46–48] and digital PCR [49–52], invented in the 1990's, have extremely high sensitivity on the order of one mutant in 10^5 copies and wide dynamic range. The high sensitivity and specificity of real-time PCR has made it the gold standard for gene detection with several US FDA-approved tests [53].

- RT-PCR: RT-PCR uses reverse transcriptase to convert RNA to cDNA and thus enables the study of RNA expression using DNA-based methodology, because RNA is less stable than DNA. In addition, RT-PCR can be used to detect fusion gene transcripts (e.g., BCR-ABL in CML) [53].

Mid-1990's: The DNA microarray was invented as a high throughput and quantitative technology for measuring the abundance of genomic DNA or RNA transcripts by hybridization to complementary DNA probes affixed to the surface of a slide [54]. Genome-wide copy number variants may be measured using chromosomal microarrays (CMA) [55], including SNP microarrays [56, 57] and array comparative genomic hybridization (aCGH) [55, 58, 59]. Gene expression microarrays, in contrast, can be used to derive genome-wide expression profiles across all transcripts and identify gene expression signatures associated with tumor prognosis [60–65] or tumor sub-classification [66, 67].

2.2 Clinical Applications of Molecular Diagnostics in Cancer Care

Cancer research has significantly advanced our understanding of the molecular pathogenesis of many cancer types. Consequently, a multitude of clinically meaningful molecular alterations have been identified and catalogued, which not only dictate tumors' innate aggressiveness and their clinical courses but are also potential therapeutic targets for molecularly targeted cancer drugs. For the last decade-and-a-half, molecular diagnostics have experienced an unprecedented growth to become an integral part of current oncology practice, revolutionizing virtually all aspects of cancer care including the incorporation of molecular biomarkers into traditional cancer taxonomy for diagnosis and prognosis, selection of targeted therapies, post-treatment disease monitoring, cancer prevention and early detection.

2.2.1 Taxonomy and Molecular Biomarkers for Prediction of Therapy, Diagnosis, and Prognosis

The majority of cancer diagnoses are still currently classified using the traditional cancer taxonomy that is based on tumors' sites of origin, as well as their histopathology including tumor cell lineage origins and other microscopic morphologic features. However, there is a paradigm shift toward a clinical oncogenic-molecular-biomarker based cancer taxonomy, accompanying the rapidly increasing importance of molecular diagnostics in the practice of oncology. The optimal clinical cancer taxonomy for the precision oncology practice should consist of distinctive tumor diagnoses, defined by specific profiles of key molecular cancer drivers, which not only can guide therapies by predicting efficacy of the drugs targeting the

specific cancer drivers, but also enable the best prognostic stratifications to formulate the most appropriate cancer treatment strategies. The most salient examples of this paradigm shift are the new clinical molecular classifications and molecular reclassifications in several general tumor types, still within the traditional cancer taxonomy's framework of tumor site of origin, but conferring robust diagnostic, predictive and prognostic superiority.

For example, among hematological malignancies with their common tissue origin within the blood/lymphoid circulation system, CML was the first malignancy defined by its key molecular oncogenic driver (BCR-ABL) by the World Health Organization (WHO) in 2001. Except for the requirement for evidence of clinical chronicity and the cancer cell lineage origin restriction to myeloids, the diagnosis is otherwise independent of histopathology. The BCR-ABL fusion gene is not only the diagnostic biomarker of CML, but it also predicts the efficacy of TKIs as well as the prognosis of a well defined disease course. Further, in 2008, the WHO also created a new hematologic malignancy classification, including three myeloid-and-lymphoid neoplasms with eosinophilia and rearrangements of PDGFa, PDGFb or FGFR1 respectively [18]. This new classification is so far the most radical example of the departure from the traditional cancer taxonomy, with each of its three malignancies unrestricted to a cell lineage origin and potentially including neoplastic proliferations of either a lymphoid lineage or a myeloid lineage. Nevertheless, these histopathologically divergent disease entities within each malignancy are united by their similar underlying tyrosine kinase-driven oncogenic mechanisms, which may be targeted by imatinib or other TKIs and carry similar prognosis. The WHO has also re-classified acute myeloid leukemia (AML) with a large set of subtypes defined by recurrent cytogenetic abnormalities and recurrent somatic mutations [68], leading to significantly improved prognostic risk stratification and well defined treatment strategies to facilitate clinical decision making such as bone marrow transplant.

Among solid tumors, there is a strong growing consensus to re-classify gliomas, which are a group of tumors derived from glial cells located in the central nervous system, based on mutation status in the genes IDH1 and IDH2 rather than the existing histopathologic grading and diagnostic criteria. IDH-mutation-driven gliomas exhibit epigenetic dysregulation and genomic hypermethylation, which are biologically very different from IDH-wild-type gliomas. This re-classification has shown robust superiority in prognostic stratification, with IDH-mutant gliomas characterized by favorable prognosis with extended overall survival relative to the IDH-wide-type counterparts, transcending the current WHO grading classification of gliomas [69].

Much more commonly, the incorporation of molecular biomarkers have led to the refinement of traditional tumor classifications, resulting in improved prediction of drug efficacy, better diagnostic algorithms, and superior prognostic stratification:

2.2.1.1 Predictive Biomarkers

In contrast to the molecular diagnostics of hereditary diseases, which relate largely to disease diagnosis, a major focus of cancer molecular diagnostics is on its predictive value for guiding the selection of therapies. Since the characterization of the BCR-ABL fusion gene as a predictive biomarker for response to imatinib in CML, there has been a proliferation of genomic alterations that may predict response to molecularly targeted therapies that are either themselves the molecular targets or closely related to the molecular targets in the associated oncogenic pathways. Some common positive predictive biomarkers for targeted therapies include certain activating EGFR mutations for EGFR tyrosine kinase inhibitors (TKIs) and the EML4-ALK fusion gene for ALK-targeted inhibitors in lung adenocarcinomas, oncogenic BRAF V600E/K mutations for BRAF inhibitors in melanomas and hairy cell leukemia (HCL) [70], inactivating BRCA1 and BRCA2 mutations for PARP inhibitors in ovarian cancers, and a subset of activating mutations in certain exons of KIT and PDGFRA for imatinib and/or other TKIs in gastrointestinal stromal tumors (GISTs) [71]. Some common negative predictive biomarkers that confer resistance to targeted therapies include oncogenic KRAS mutations for EGFR TKIs in lung adenocarcinomas and EGFR monoclonal antibodies in colon cancers [72], as well as a different subset of activating mutations in different exons of KIT gene or PDGFRA gene for imatinib and/or other TKIs in GISTs [71]. In addition, a few predictive biomarkers can be both a negative predictor of drug resistance for certain targeted therapy drugs, but also a positive predictor of drug responsiveness for other target therapy drugs. For example, the T315I mutation of the ABL kinase domain is predictive of drug resistance for all the first/second generation CML TKIs; but it is also predictive of drug responsiveness for ponatinib, a third generation TKI [73]. Similarly, in lung adenocarcinomas, T790 M mutation of EGFR is predictive of drug resistance to all first/second generation EGFR TKIs, but it is also predictive of drug responsiveness to osimertinib, a third generation EGFR TKI [74].

The impact of precision oncology has also further expanded into the arena of cancer molecularly targeted therapy testing with the advent of “basket” clinical trials. This innovative clinical trial design enrolls patients with cancers across different anatomic sites and histopathology, on the basis of shared key oncogenic genomic alterations that are putative predictive biomarkers [70, 75]. Molecular diagnostics play a critical role in selecting the patient populations for basket clinical trials by screening the presence of specific putative predictive biomarkers.

In addition, some molecular biomarkers can predict responses to non-targeted therapy. Mismatch repair-deficiency confirmed by microsatellite instability testing in colorectal cancer is a predictive marker for resistance to 5-fluorouracil (5-FU)-based chemotherapy and positive responsiveness to immunotherapy such as PD-1 inhibitors [76].

2.2.1.2 Diagnostic Biomarkers

Incorporation of molecular biomarkers has contributed to the improvement of the diagnostic algorithms of many cancers. Molecular biomarkers have aided confident grouping of various histopathologically heterogeneous tumors with the same underlying tumor biology and clinical course into a single tumor type; and have also enabled confident differentiation between different cancer types. In addition, molecular markers have been used to distinguish between a cancer recurrence and a new primary tumor, between a malignant tumor and a benign tumor, as well as between a malignancy and a non-neoplastic reactive condition, when there are overlapping histopathology precluding definitive diagnoses.

As an example in hematological malignancies, the BRAF V600E mutation can serve as a diagnostic biomarker to reliably differentiate HCL from other morphologically similar mature B cell leukemias and lymphomas, including splenic marginal zone lymphoma and hairy cell leukemia variant [77]. In solid tumors, genomic alterations have brought much needed improvement on the diagnostic algorithms for the differential diagnoses of many sarcomas and salivary gland tumors that are well known to be diagnostically challenging based on histopathologic evaluations alone. The challenges include significant histopathologic heterogeneity within one single tumor type as well as significant histopathologic overlap between different tumor types that exhibit completely different biological and clinical behavior. For example, MDM2 and CDK4 amplifications are a hallmark for dedifferentiated liposarcoma, which is highly valuable for its diagnosis, as the sarcoma often presents with diverse variant histopathology or is poorly differentiated with no specific morphologic features for a definitive diagnosis [71]. Detection of fusion events involving the MAML gene on chromosome 11q21 can be used to confirm the diagnosis of mucoepidermoid carcinoma of salivary gland tumors with diverse variant histopathology, or to differentiate high grade mucoepidermoid carcinoma from other more aggressive high grade salivary gland tumors with overlapping histopathology [78].

In the clinical settings of peripheral blood myeloid cytosis or cytopenia, differential diagnosis of chronic myeloid neoplasms versus benign reactive conditions can be challenging due to significant histopathologic overlap between these entities. Since the WHO included particular recurrent somatic mutations in JAK2, MPL and KIT genes as part of the diagnostic criteria for a few myeloproliferative neoplasms in 2008, many additional somatic mutations in multiple cancer driver genes have been associated with various chronic myeloid neoplasms [79, 80]. Detection of more than one recurrent mutation in the appropriate clinical settings of peripheral cytopenia or cytosis, particularly at more than 10% MAF, can help to arrive at diagnoses of myeloid neoplasms with higher confidence in challenging cases. In the clinical setting of peripheral lymphocytosis or enlarged lymph nodes, B cell or T cell clonality testing for the presence of clonal rearrangement of immunoglobulin genes and the T cell receptor gene may also help to differentiate between lymphoid neoplasms versus benign reactive conditions [81].

For pancreatic cysts with only cyst fluid aspirate and no tissue section for histopathologic evaluation, differentiation between benign pancreatic cystic lesions that

only need monitoring versus cystic mucinous neoplasms with the potential to progress to pancreatic adenocarcinomas that may require surgery can be very challenging in the absence of molecular analysis. Mutations of KRAS, GNAS and RNF43 are specific for mucinous neoplasms and can serve as effective diagnostic biomarkers. Incorporation of molecular analysis of mutations in TP53, PIK3CA and PTEN genes can further reveal high grade mucinous neoplasms [82].

Finally, molecular testing including tumor gene expression profiling (GEP) often serves as the last resort to predict the tissue site origins of cancers of unknown primary (CUPs), based on gene expression patterns retained from the normal tissues, when the conventional histopathologic evaluation and immunohistochemical workup of CUPs fail to do so [83, 84].

2.2.1.3 Prognostic Biomarkers

Numerous molecular biomarkers with or without diagnostic or predictive values may provide additional prognostic information to improve existing cancer risk stratification, resulting in the formulation of more tailored treatment strategies. For example, recurrent somatic mutations in ASXL1 and EZH2 in myelodysplastic syndromes (MDSs) are valuable prognostic markers associated with shorter overall survival independent of the MD Anderson Lower-Risk Prognostic Scoring System (LR-PSS) [85, 86]. Inactivating mutations in *B2M* in Classical Hodgkin Lymphoma (CHL) is associated with lower stage of disease, younger age at diagnosis, and better overall and progression-free survival [87]. While inactivating mutations of p53 are unfavorable prognostic markers in many cancer types including breast cancer, osteosarcoma and leiomyosarcoma [71, 88, 89], mismatch repair-deficiency in colorectal cancer is a favorable prognostic biomarker. In addition to DNA alterations, gene expression signatures have also been developed for risk stratification to inform decisions among different therapeutic interventions such as chemotherapy and radiation therapy. For example, the gene expression-based Oncotype DX test helps to determine whether a patient with early stage ER-positive breast cancer or ductal carcinoma in-situ (DCIS) would benefit from chemotherapy or radiation therapy, respectively [64, 65].

It is worth noting that predictive or prognostic biomarkers are not limited to somatic mutations, as some germline mutations can impact cancer therapy and prognosis similar to their somatic mutation counterparts. For example, germline inactivating BRCA mutations may also predict responsiveness to PARP inhibitors. Similarly, germline mutations of mismatch repair genes in colorectal cancer (i.e. Lynch syndrome) may also indicate favorable prognosis and predict drug resistance to 5-FU chemotherapy. In addition, there is some evidence that germline variants with pharmacogenomic significance may impact targeted cancer drug efficacy including imatinib for CML and tamoxifen for breast cancer [90–92].

2.2.2 Disease Monitoring

Various molecular analyses have also been deployed for post-treatment monitoring of leukemias to follow treatment response, acquired drug resistance, disease progression, minimal residual disease (MRD), and disease recurrence. For example, highly sensitive and quantitative RT-PCR to detect BCR-ABL fusion transcripts in peripheral blood is routinely used for post-treatment monitoring of CML. Treatment response is well defined by a grading system based not only on the absolute tumor load reduction, but also the dynamic speed of the reduction [73]. After 3 log reduction or major molecular response (MR3 or MMR), the molecular test is used to monitor potential acquired drug resistance or disease progression. Persistent rise in BCR-ABL fusion transcript level with loss of MMR will trigger sequencing analysis to investigate acquired drug resistance mutations, and cytogenetic karyotyping to investigate new cytogenetic abnormalities indicative of disease progression. When a drug resistance mutation is confirmed by sequencing, the patient will be switched to a different TKI with known efficacy toward the specific mutation [93]. The molecular monitoring may continue indefinitely for potential relapse, even for some CML patients who may have achieved potential cure with sustained treatment-free remission [94]. Studies involving MRD detection of acute promyelocytic leukemia (APL) have also shown that pre-emptive chemotherapy at the time of molecular relapse improves survival compared to treatment at the point of hematological relapse with abnormal peripheral blood count [95, 96].

2.2.3 Cancer Prevention and Early Detection

A significant fraction of cancers have a hereditary component, and genetic risk assessment for hereditary cancer is integral in the comprehensive care of today's cancer patients. Detection of pathogenic germline variants in cancer predisposition genes such as BRCA1, BRAC2, ATM, TP53 and mismatch repair genes in cancer patients may initiate testing of blood-related family members at risk for developing cancer. Detection of germline cancer predisposition gene mutations in patients with no prior cancer history may dictate chemoprevention for cancer prevention or suppression, more frequent surveillance for early cancer detection, as well as risk-reduction prophylactic surgeries such as colectomy or mastectomy [97–99].

Cervical cancer has become the most preventable cancer due to cervical cancer screening in the past 50 plus years, using cervical cytology or Papanicolaou test [100–102]. With the more recent elucidation of causal agent for almost all cervical cancers being persistent infection of high risk human Papillomaviruses (HR-HPVs) [103], nucleic acid-based HR-HPV testing of cervical liquid cytology samples has become an integral part of cervical cancer screening in addition to cervical cytology test, or even in lieu of cervical cytology test, with higher sensitivity and lower cost [104–106].

Sequencing analysis of peripheral blood, which is often performed as part of the paired analysis of tumors and patient-matched normal samples to enable unambiguous detection of somatic mutations, may detect clonal hematopoiesis of indeterminate potential (CHIP). CHIP is a clinical entity defined only recently, referring to detection of clonal somatic mutations in peripheral blood that are frequently associated with hematologic malignancies when there is no evidence of hematologic disease. It is associated with aging and occurs more often in older patients. In most cases, CHIP does not progress to hematologic malignancies. However, it is associated with an increased, albeit still low risk of progression to hematological malignancies. With such uncertain clinical implications, closer monitoring of complete blood count with differential, particularly in CHIP patients with cytopenias, may be prudent [107, 108].

2.3 Frontiers in Molecular Diagnostics of Cancer

2.3.1 Current State-of-the-Art Nucleic-Acid-Based Analysis: Next-Generation Sequencing

The major technological advancement in the new millennium for both cancer research and cancer molecular diagnostics has been massively parallel "next generation" sequencing (NGS), which followed the completion of Human Genome Project in the early 2000's. Using the then state-of-the-art automated Sanger sequencing technology, the sequencing of the first human genome spanned more than 10 years and cost more than a billion dollars. Stimulated by the unmet need for more powerful sequencing technologies, novel NGS platforms based on different chemistries were developed throughout the decade that followed [52, 109, 110]. Compared to Sanger sequencing technology, all NGS platforms exhibit ultra-high throughput and substantially lower cost per base (i.e., ~$0.02 per million bases for the Illumina Hiseq in 2016) [52], enabling a variety of nucleic acid-based applications including genome sequencing, gene expression profiling and transcriptome analysis via RNA-seq, and epigenetic analysis via methyl-seq, ChIP-seq or ATAC-seq [52].

For the last 10 years, NGS has had a profound impact on cancer genomic research by enabling the complete characterization of entire cancer genomes at the single nucleotide resolution. After the first cancer genomes were reported in 2008 [111, 112], tens of thousands of tumors have been comprehensively analyzed at the whole-genome or whole-exome scale, either by individual research groups or by large consortiums such as TCGA (The Cancer Genome Atlas) and ICGC (International Cancer Genome Consortium). The effort to map and analyze cancer genomes in a large-scale and a systematic fashion has led to the identification of significantly recurrently mutated genes and biological pathways, prompting novel hypotheses regarding cellular processes governing tumor initiation and progression

[2]. Precision oncology has thus experienced its most rapid growth in the past decade, accompanied by a deeper understanding of the complexity of cancer genomes with rapid proliferation of biological targets for novel therapies as well as new predictive, prognostic and diagnostic biomarkers for molecular diagnostics. There currently exist 57 FDA approved molecularly targeted cancer drugs (FDA website, last accessed 10/2016) [113], with many more in the pipeline, either at the clinical trial stage or the development stage.

By enabling simultaneous sequencing of multiple genomic regions in multiple samples in a single NGS run, the high-throughput sequencing capability of NGS is an advantageous technical attribute to not only minimize turn-around time and significantly lower cost, but also to address the challenge of limited tumor sample volumes and ever expanding list of biomarkers for routine somatic tumor testing. Only a single input of relatively low quantity of tumor nucleic acids is required for screening all biomarkers in multiple genomic regions in one NGS run, in contrast to traditional Sanger sequencing requiring cumulatively larger amounts of input nucleic acids [40]. NGS is capable of detecting all classes of genomic alterations, including CNVs and SVs. Further, owing to its digitally quantitative nature, individual DNA molecules are analyzed independently instead of as an ensemble, leading to high sensitivity for detecting mutations with low MAFs. The high sensitivity and the digitally quantitative nature of NGS are technical attributes particularly suitable for interrogation of cancer somatic mutations, as low MAFs due to low tumor purity and/or intra-tumoral genetic heterogeneity can often occur. RNA molecules can also be analyzed by RNA-Seq, following conversion of RNA to cDNA. Compared to cDNA microarrays, RNA-Seq enables digitally quantitative assessment of gene expression profiling with broader dynamic range and higher sensitivity, and offers the additional capabilities of detecting allele-specific expression, novel RNAs including splicing variants and gene fusion events and phasing mutations across fusion transcripts in highly complex cancer genomes [114–117].

The sequencing and signal detection platforms of NGS can generally be divided into two main categories: platforms that require clonal amplification of templates prior to sequencing and platforms enabling real-time sequencing of single DNA molecules. The two NGS platforms most widely used in clinical laboratories, manufactured by Illumina and Ion Torrent (Life Technologies), belong to the first category and are based on "sequencing by synthesis" (SBS) chemistry [52, 109]. In the SBS approach, single-stranded DNA is used as a template for synthesizing a complementary strand in a manner allowing the identity of the incorporated nucleotide to be detected. Because it is necessary to first clonally amplify each template molecule in order to enhance the signal to facilitate detection, the signal represents a consensus readout across multiple molecules, which may lead to errors due to dephasing as the reaction progresses. Thus, these methods typically only produce up to 200–400 bp short reads [52, 109, 118]. Furthermore, different platforms exhibit different characteristic errors: Illumina sequencing is predominated by substitution errors with an error rate on the order of 0.1%, while the Ion Torrent platform is predominated by small insertion and deletion errors particularly at homopolymer sites, with an overall error rate closer to 1% [52, 109].

Due to existing limitations in throughput, cost and accuracy, single molecule sequencing has not yet been widely adopted for molecular diagnostics. However, these platforms hold appealing technical advantages provided that the technology will continue to improve. Single molecule sequencing does not require any clonal amplification, thereby reducing sample preparation time and PCR amplification errors or bias. Much longer sequence reads can be achieved (~10,000 to ~350,000 bp) [52, 109, 117], which enable the detection of structural variants in complex cancer genomes and the phasing of mutations across multiple protein-coding exons [52, 114, 119]. Examples of single molecule sequencing include the SMRT platform of Pacific Biosciences and the Nanopore platform of Oxford Nanopore Technologies. While the SMRT platform also utilizes SBS, the Nanopore technology directly detects the DNA composition of single-stranded DNA as it crosses through a narrow pore. Both technologies have shown abilities to directly detect DNA methylation. Both technologies currently harbor high error rates with single pass, although error rates may be reduced by sequencing to higher coverage (i.e., the number of DNA molecules sequenced from a given genomic region). For example, the SMRT platform can achieve lower error rates of 1% via higher coverage through circular consensus reads. Through further development and optimization, real time single molecule sequencing may significantly expand the clinical applications of NGS [52].

Efficient utilization of NGS technology is necessary to maximize throughput and minimize turnaround time in clinical laboratories. Two key innovations have enabled laboratories to leverage the ever-expanding data output of NGS instruments and increase the number of samples that can be analyzed in a single sequencing run. First, the development of sample-specific barcodes (i.e., 6–8 bp DNA sequences) enabled sequence libraries derived from different samples to be uniquely tagged and pooled together on the same sequencing instrument. Second, the development of target enrichment strategies allowed specific sub-genomic regions of interest to be captured and sequenced as an alternative to whole genome sequencing. There are two general categories of target enrichment: "hybridization capture" by annealing to complementary synthetic DNA or RNA probes, and multiplex-PCR-based "amplicon capture". While amplicon capture approaches typically incur a shorter turnaround-time and require less input genomic DNA, hybridization capture can be extended to a larger fraction of the genome and enables the detection of CNVs and SVs in addition to sequence mutations. In general, hybridization capture also provides higher uniformity of sequence coverage than amplicon capture methods [120].

The technical innovations of NGS have considerably expanded the variety of options for clinical NGS applications. NGS-based diagnostic tests can be designed to target the genome at different scales including targeted gene panels, whole exome sequencing (WES) and whole genome sequencing (WGS). Gene panel-based targeted sequencing includes both small amplicon capture-based "hotspot" panels targeting only frequently mutated regions in 1–50 well-established cancer genes, and larger hybridization capture-based cancer gene panels that can target all exons of up to hundreds of cancer-associated genes as well as clinically relevant non-coding regions. WES and WGS are capable of revealing a far greater number of alterations

across a much larger genomic territory, though these primarily represent variants of uncertain significance (VUS) with no immediate clinical applications. Compared to smaller targeted gene panels, WES and WGS also incur higher cost and greater computational requirements for data storage and processing, which can negatively impact sequencing throughput for molecular diagnostics testing. Gene panels may also exhibit increased sensitivity for detecting clinically actionable alterations compared to WES and WGS, owing to deeper sequence coverage in these most important genomic regions [121]. Nevertheless, WES and WGS do significantly increase the breadth of genomic analysis and offer the greatest discovery potential, with WGS further facilitating the detection of translocations, inversions, and other structural alterations [40].

As NGS sequencing reactions are capable of generating hundreds of millions to billions of sequence reads, innovative computational tools and bioinformatic pipelines are required for rapid and accurate NGS data processing to deliver timely reports to oncologists for cancer therapy guidance. The major steps in NGS bioinformatic pipelines include base calling to covert raw sequencing signals to sequence reads, alignment of sequence reads to the reference human genome, variant calling to detect candidate genomic alterations (including mutations, copy number alterations, and structural rearrangements), variant filtering to eliminate spurious variant calls, and finally variant annotation according to the annotation guidelines from the Human Genome Variation Society (HGVS) [122] including standardized gene name (per HUGO nomenclature guidelines) [123], variant class, transcript isoform, and amino acid change if in a coding sequence. In addition to variant annotation, it is also recommended that the functional and clinical significance of variants, including prognostic and therapeutic implications, should be interpreted. This may involve database and literature review for information on prevalence, biological significance, and clinical predictive or prognostic implications. Cancer databases capturing mutation recurrence and clinical annotations include COSMIC [124, 125], the cBioPortal for Cancer Genomics [126, 127], My Cancer Genome/DIRECT [128, 129], Personalized Cancer Therapy [130, 131], TARGET [132], the Human Gene Mutation Database (HGMD) [133] and OncoKB [134].

2.3.2 Liquid Biopsy

Traditional oncogenic mutation detection is based on the direct analysis of cancer tissue obtained by surgical resection or biopsy. However, body fluids of cancer patients, including blood, saliva, cerebrospinal fluid and urine, have been shown to contain cell-free tumor-derived nucleic acids as well as circulating tumor cells (CTC) harboring detectable cancer biomarkers. The term “liquid biopsy” refers to molecular diagnostics using body-fluid samples. For some cancer patients, surgical resection or biopsy of a tumor may not be feasible, either due to the location of the tumor or the inability of a patient to tolerate an invasive procedure. For example, in approximately 30% of lung cancer patients, clinicians are not able to obtain tumor

tissue by surgical biopsy or resection [135]. CTC and circulating tumor DNA (ctDNA) based assays represent easily accessible, minimally invasive means to not only detect but also follow the dynamic molecular makeup of a patient's tumor longitudinally. These approaches also have the potential to more completely represent the full genetic heterogeneity of a patient's disease compared to a single biopsy site [135]. Analyses of CTCs in blood and ctDNA in plasma are promising avenues for molecular diagnostics in cancer.

CTCs, which are extremely rare in individuals without malignancy, are present at a wide range of frequencies in patients with various metastatic carcinomas [136]. The detection and molecular characterization of circulating tumor cells (CTCs) are one of the most active areas of translational cancer research [137]. The enumeration of CTCs has been shown to be a key indicator of cancer diagnosis, prognosis, and response to therapy [138–140]. An FDA approved clinical CTC counting assay has been used for cancer prognosis and monitoring treatment response since 2004 [141]. However, a combination of CTC enumeration, molecular profiling at the DNA, RNA and protein level [137], and establishment of individual patient–based cell lines and xenograft models for drug sensitivity testing [142] may be critical to exploit the full potential of CTCs. There has been intense research in the field of CTCs to develop novel protocols for CTC capture and sequencing and to demonstrate their clinical utility in cancer molecular diagnostics [137]. Due to the rarity of CTCs, limited purity of the isolated CTC populations and intra-tumoral genetic and epigenetic heterogeneity, molecular profiling of CTCs at single-cell resolution is required. Various technologies using a combination of whole genome amplification and NGS are able to detect CNVs and SNVs in a single cell [143]. These approaches have demonstrated the potential of CTC-based molecular analysis in molecular diagnosis [144, 145] and monitoring of cancer clonal evolution and intra-tumoral heterogeneity [146, 147]. However, CTC molecular profiling is currently largely limited to preclinical and academic studies.

For ctDNA, only a small amount of circulating cell-free DNA (cfDNA) is typically present in plasma (average ~17 ng/ml), of which only a tiny fraction may actually be derived from tumor [135]. This poses a significant challenge for mutation detection in cfDNA when the fraction of tumor-derived DNA in plasma may begin to approach the error rate of NGS. Technical advances have been made to increase the detection sensitivity by decreasing the background error rate though the use of unique molecular barcodes, including Safe-Seq and Duplex Sequencing [148, 149] as well as the newly developed CAPP-Seq, reaching a sensitivity of ~0.02% [150, 151]. ctDNA analysis holds great promise as a less invasive alternative to tissue sequencing that may capture a wider spectrum of sub-clonal mutations. Furthermore, the minimally invasive nature of the ctDNA liquid biopsy allows repeated sampling and sequencing over time to facilitate the monitoring of response to therapy and the acquisition of drug resistance mutations. Repeated sampling and sequencing may help increase the sensitivity of ctDNA NGS by differentiating persistent rare mutations versus random sequencing artifacts [152]. Studies have shown that analyses of ctDNA can be used to stratify and guide cancer target therapy, to determine tumor

prognosis, and to monitor treatment response, acquired drug resistance, minimal residual disease, and disease recurrence in real time.

Compared to ctDNA, CTCs may have the advantage of being more representative of the current tumor state, as they are viable tumor cells while ctDNA often originates from necrotic cells. CTCs may also provide information at the DNA, RNA and protein levels, while ctDNA can only capture genomic alterations [152]. However, some studies have shown that ctDNA is a biomarker with higher sensitivity than that of CTCs for detecting new mutations and disease progression [153]. ctDNA has also been proposed as a mechanism for screening and early detection of cancer, though the sensitivity and specificity may be affected by the presence of circulating non-tumor DNA containing somatic mutations arising from non-malignant tissue or indicative of CHIP associated with aging [154]. While circulating RNA is too unstable to be used for clinical diagnostics, stable extracellular exosomes released by tumor cells contain mainly RNA, which may be analyzed to provide information about the tumor [135, 155, 156].

The integration of liquid biopsy-based molecular diagnostics to precision oncology practice has begun, as the FDA approved the Cobas EGFR ctDNA mutation test in 2016 [157]. This ctDNA test is indicated as an alternative companion diagnostic to tumor tissue-based sequencing, aiming to inform the selection of molecularly targeted drugs for patients with non-small cell lung cancer. Liquid biopsy-based mutation profiling has thus officially entered the cancer molecular diagnostic arena and holds great potential in precision oncology.

2.3.3 Current Challenges and Near Term Solutions

A trend observed in in-vitro diagnostics in general, and in cancer molecular diagnostics in particular, is the natural evolution of technology platforms toward greater levels of sensitivity, versatility as well as automation [10, 12]. This trend has characterized the new era of clinical NGS with its gradual replacement of the traditional low-throughput nucleic acid-based methodologies with high-throughput technologies capable of interrogating cancer genomes at single base resolution from a limited amount of tumor DNA of variable quality that is extracted from either tissue or body fluid. NGS has become routine for clinical tumor genomic mutation screening and has been adapted for nucleic acid-based clinical interrogation of cancer transcriptomes and epigenetics as well. The high complexity of NGS workflows and result-reporting has made its validation and implementation in clinical diagnostic labs challenging. Guidelines from regulatory agencies including College of American Pathologists (CAP), American College of Medical Genetics (ACMG) and the New York State Department of Health, as well as publications from clinical laboratories, have recently provided much needed guidance [158, 159]. There are a variety of target NGS panels implemented in different molecular diagnostic labs, in terms of the numbers and the identities of the selected genes or gene regions, the indicated tumor types and the sample throughput. In addition to developing custom

target NGS panels, due to the significant investment requirement for the operational and bioinformatics infrastructure, many diagnostic labs have chosen to validate ready-made commercial target NGS panel solutions with or without modification to expedite the implementation of NGS testing [121].

Bioinformatics considerations continue to pose a significant challenge to clinical laboratories that have historically operated lower-throughput technologies requiring minimal computational support. Bioinformatics staff capable of managing and analyzing large, complex data sets are in high demand. Computational pipelines must be established during assay validation, including variant filtering algorithms that account for strand bias and minimum cut-offs for quality scores, depth of coverage, number of variant reads and MAF [40]. Manual review using tools such as Integrated Genome Viewer (IGV) is an important final step of variant filtering for excluding spurious variant calls. For variant interpretation and reporting, further improvement on data sharing of cancer somatic mutations is needed, due to the lack of a single centralized repository of cancer somatic mutations with de-identified patient information [152]. The storage of large amounts of NGS data is another technical challenge and should follow the ACMG and AMP guidelines to retain at least raw data files so that the entire data analysis process can be repeated [158, 159].

Sequencing paired normal samples is essential for large scale tumor analysis, particularly WES and WGS, in order to make unambiguous tumor somatic mutation calls by filtering out germline variants. Sequencing both tumors and paired normal samples may reveal germline variants with clinical implications other than those directly related to the tumors. The disclosure of these "incidental" findings to the patients and its legal and ethical implications are an area of controversy. Recent guidelines from the ACMG and AMP have provided some guidance on dealing with the ethical challenge of reporting incidental findings, stressing the importance of informed consent from patients before testing [160, 161]. On the other hand, as discussed previously, some germline variants may have significant predictive and prognostic implications for cancer treatment, in the context of cancer predisposition genes as well as germline pharmacogenomics. The controversy on sequencing and reporting these germline variants may be clarified in the near future by continued active research in the area [152, 162].

The implementation of clinical NGS testing has also been hampered by the reimbursement challenges. The current reimbursement policies of Medicare and private insurance companies were written mostly for single-gene or biomarker molecular testing, prior to the advent of multi-gene high throughput NGS testing in the clinical labs [163]. However, progress on NGS reimbursement, albeit slow, has been made with some private insurance companies showing more rapid adaptation of their reimbursement policies to the NGS technology platforms. For example, insurance coverage for NantHealth's Cancer GPS test, which includes tumor WGS, by Blue Cross was announced in 2016 [164, 165]. Despite all the challenges, NGS technology is destined in the near future to revolutionize molecular diagnostics with an impact comparable to that of PCR [52].

2.3.4 Future Perspective

The remarkable success of imatinib in CML validated the promise of precision oncology early on; however, the vast majority of molecularly targeted therapies to date have only shown modest success by prolongation of life for months rather than years, and only in a subset of cancer patients. This reveals that our ability to effectively treat patients with complex cancer genomes may be limited compared to CP-CML, which has only one single key oncogenic driver - the BCR-ABL fusion gene [166]. In order to enhance the sensitivity and specificity of identifying key molecular cancer drivers in complex cancer cells, a systems biology approach involving multi-omic interrogation of cancer genomes, epigenomes, transcripsomes, proteomes and metabolomes, has emerged in cancer research [167–169] and is starting to expand to the clinical testing arena [164, 165]. Recent technical advances in proteomics, including mass spectrometry and protein microarrays, have further enabled and propelled this systems biology approach in cancer research [170, 171]. The multi-modal interrogation of cancer cells can identify not only genomic alterations that are mainly static and often cryptic, but also molecular aberrations at the dynamic and complex functional genomic levels of both gene expression (epigenetic/RNA) and protein expression. The integration of multi-omic data has enabled identification of key molecular cancer drivers with more robust diagnostic, predictive and prognostic implications [172, 173], thus providing more complete and precise molecular portraits of underlying cancer biology. We may no longer be naive about the daunting challenges to cure all cancers through precision oncology, but the promise and hope are still very real.

Acknowledgements We gratefully acknowledge Jonathan Reichel, Jiabin Tang, and Fanli Meng for critical review of the manuscript and Eric Lin for advice regarding molecular diagnostic reimbursement policies. Funding was provided by the Empire Clinical Research Investigator Program and the MSK Cancer Center Support Grant (P30 CA008748).

References

1. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70. [PubMed:10647931]
2. Vogelstein B et al (2013) Cancer genome landscapes. Science 339(6127):1546–1558. [PubMed: 23539594]
3. Weaver W (1970) Molecular Biology: Origin of the term. Science 170(3958):581–582. [PubMed: 4919180]
4. Watson JD (1953) *Crick F.* A structure for deoxyribose nucleic acid. Nature 171(4356):737–738. [PubMed: 13054692]
5. Crick F, Watson JD. The complementary structure of deoxyribonucleic acid. Proc R Soc Lond 1954*;* 223*(*Series A*):*80–96
6. Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 3:318–356. [PubMed: 13718526]

7. Grunberg-Manago M, Ortiz PJ, Ochoa S (1955) Enzymatic synthesis of nucleic acid like polynucleotides. Science 122(3176):907–910. [PubMed: 13274047]
8. Rich A, Davies DR (1956) A new, two-stranded helical structure: polyadenylic acid and polyuridylic acid. J Am Chem Soc 78(14):3548–3549
9. Felsenfeld G, Davies DR, Rich A (1957) Formation of a three-stranded polynucleotide molecule. J Am Chem Soc 79(8):2023–2024
10. Tsongalis GJ, Silverman LM (2006) Molecular diagnostics: a historical perspective. Clin Chim Acta 369(2):188–192. [PubMed: 16701603]
11. Gingeras TR et al (2005) Fifty years of molecular (DNA/RNA) diagnostics. Clin Chem 51(3):661–671. [PubMed: 15650028]
12. Demidov VVDNA (2003) Diagnostics in the fifty-year retrospect. Expert Review Mol Diagn 3(2):121–124. [PubMed: 12647990]
13. Kan YW et al (1980) Polymorphism of DNA sequence in the β globin gene region. N Engl J Med 302(4):185–188. [PubMed: 6927915]
14. Emmadi R et al (2011) Molecular methods and platforms for infectious diseases testing - a review of FDA-approved and cleared assays. J Mol Diagn 13(6):583–604. [PubMed: 21871973]
15. Hudziak RM et al (1989) p185HER2 monoclonal antibody has antiproliferative effects in vitro and sensitizes human breast tumor cells to tumor necrosis factor. Mol Cell Biol 9:1165–1172. [PubMed: 2566907]
16. Druker BJ et al (2001) Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N Engl J Med 344(14):1031–1037. [PubMed: 11287972]
17. Nowell PC, Hungerford DA (1960) A minute chromosome in chronic granulocytic leukemia [abstract]. Science 132:1497
18. Rowley JD (1973) Letter: a new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature 243(5405):290–293. [PubMed: 4126434]
19. Abelson HT, Rabstein LS (1970) Lymphosarcoma: virus-induced thymic-independent disease in mice. Cancer Res 30:2213–2224. [PubMed: 4318922]
20. Siegler R, Zajdel S (1972) Pathogenesis of Abelson-virus-induced murine leukemia. J Natl Cancer Inst 48:189–218. [PubMed: 4347027]
21. Rosenberg N, Baltimore D, Scher CD (1975) In vitro transformation of lymphoid cells by Abelson murine leukemia virus. Proc Natl Acad Sci U S A 72:1932–1936. [PubMed: 168584]
22. Witte ON et al (1981) Phosphorylation of the Ableson murine leukemia virus transforming protein. J Virol 39:870–878. [PubMed: 6270355]
23. Ponticelli AS et al (1982) In vivo tyrosine phosphorylations of the Abelson virus transforming protein are absent in its normal cellular homolog. Cell 29:953–960. [PubMed: 6185233]
24. Heisterkamp N et al (1982) Chromosomal localization of human cellular homologues of two viral oncogenes. Nature 299:747–749. [PubMed: 7121606]
25. de Klein A et al (1982) A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. Nature 300:765–767. [PubMed: 6960256]
26. Bartram CR et al (1983) Translocation of c-abl oncogene correlates with the presence of a Philadelphia chromosome in chronic myelocytic leukaemia. Nature 306:277–280. [PubMed: 6580527]
27. Canaani E et al (1984) Altered transcription of an oncogene in chronic myeloid leukaemia. Lancet 1:593–595. [PubMed: 6142307]
28. Groffen J et al (1984) Philadelphia chromosomal breakpoints are clustered within a limited region, bcr, on chromosome 22. Cell 36(1):93–99. [PubMed: 6319012]
29. Heisterkamp N et al (1985) Structural organization of the bcr gene and its role in the Ph' translocation. Nature 315(6022):758–761. [PubMed: 2989703]
30. Shtivelman E et al (1985) Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature 315(6020):550–554. [PubMed: 298962]

31. Stam K et al (1987) Evidence that the *phl* gene encodes a 160,000 dalton phosphoprotein with associated kinase activity. Mol. Cell. Biol 1955–60:7. [PubMed: 3299055]
32. Buchdunger E et al (1996) Inhibition of the Abl protein tyrosine kinase in vitro and in vivo by a 2-phenylaminopyrimidine derivative. Cancer Res 56(1):100–104. [PubMed: 8548747]
33. Khurana E et al (2016) Role of non-coding sequence variants in cancer. Nat Rev Genet 17(2):93–108. [PubMed: 26781813]
34. Do H, Dobrovic A (2015) Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. Clin Chem 61(1):64–71. [PubMed: 25421801]
35. Williams C et al (1999) A high frequency of sequence alterations is due to formalin fixation of archival specimens. Am J Pathol 155(5):1467–1471. [PubMed: 10550302]
36. Van Allen EM et al (2014) Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. Nat Med 20(6):682–688. [PubMed: 24836576]
37. Hedegaard J et al (2014) Next-generation sequencing of RNA and DNA isolated from paired fresh-frozen and formalin-fixed paraffin-embedded samples of human cancer and normal tissue. PLoS One 9(5):e98187. [PubMed: 24878701]
38. Dumenil TD et al (2014) Genome-wide DNA methylation analysis of formalin-fixed paraffin embedded colorectal cancer tissue. Genes Chromosomes Cancer 53(7):537–548. [PubMed: 24677610]
39. Jennings L, van Deerlin VM, Gulley ML (2009) Recommended principles and practices for validating clinical molecular pathology tests. Arch Pathol Lab Med 133(5):743–755. [PubMed: 19415949]
40. Luthra R et al (2015) Next-generation sequencing in clinical Molecular diagnostics of Cancer: advantages and challenges. Cancers 7(4):2023–2036. [PubMed: 26473927]
41. Katsanis SH, Katsanis N (2013) Molecular genetic testing and the future of clinical genomics. Nat Rev Genet 14(6):415–426. [PubMed: 23681062]
42. Sanger F (1975) Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol 94(3):441–448. [PubMed: 1100841]
43. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74(12):5463–5467. [PubMed: 71968]
44. Frommer M et al (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A 89(5):1827–1831. [PubMed: 48546]
45. Langer-Safer PR, Levine M, Ward DC (1982) Immunological method for mapping genes on Drosophila polytene chromosomes. Proc Natl Acad Sci 79(14):4381–4385. [PubMed: 6812046]
46. Gibson UE, Heid CA (1996) Williams PM. A novel method for real time quantitative RT-PCR. Genome Res 6(10):951–1001. [PubMed: 8908519]
47. Rasmussen R, Meuer S, Wittwer C, Nakagawara K (eds) (2001) Rapid cycle real-time PCR, 1st edn. Springer, Berlin, pp 21–34
48. Bustin SA et al (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55(4):611–622. [PubMed: 19246619]
49. Milbury CA et al (2014) Determining lower limits of detection of digital PCR assays for cancer-related gene mutations. Biomolecular Detect Quantification 1(1):8–22. [PubMed: 27920993]
50. Vogelstein B, Kinzler KW (1999) Digital PCR. Proc Natl Acda Sci 96(16):9236–9241. [PubMed: 10430926]
51. Dong L et al (2015) Comparison of four digital PCR platforms for accurate quantification of DNA copy number of a certified plasmid DNA reference material. Sci Rep 5:13174. [PubMed: 26302947]
52. Goodwin S, McPherson JD, McCombie WR (2016) Coming of age: ten years of next generation sequencing technologies. Nat Rev Genet 17(6):333–351. [PubMed: 27184599]

53. Freeman WM, Walker SJ, Vrana KE (1999) Quantitative RT-PCR: pitfalls and potential. BioTechniques 26(1):112–122. 124–5. [PubMed: 9894600]
54. Pinkel D et al (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20(2):207–211. [PubMed: 9771718]
55. Miller DT et al (2010) Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet 6(5):749–764. [PubMed: 20466091]
56. Affymetrix Genome-Wide Human SNP Array 6.0 data sheet. Affymetrix, Santa Clara, California; 2007. (https://cancergenome.nih.gov/abouttcga/aboutdata/platformdesign/affymetrixSNP6array)
57. Illumina Genome-Wide DNA Analysis Bead Chips Data Sheet. Illumina, San Diego, California; 2009 (https://www.illumina.com/content/dam/illumina-marketing/documents/products/brochures/datasheet_omni_whole-genome_arrays.pdf)
58. Ylstra B et al (2005) BAC to the future! or oligonucleotides: a perspective for microarray comparative genomic hybridization (array CGH). Nucleic Acid Res 34(2):445–450. [PubMed: 16439806]
59. Li W, Olivier M (2013) Current analysis platforms and methods for detecting copy number variation. Physiol Genomics 45(1):1–16. [PubMed: 23132758]
60. Perez-Diez A, Morgun A, Shulzhenko N. Microarrays for cancer diagnosis and classification. Madame Curie Bioscience Database Austin, Texas: Landes Bioscience; 2000–2013. (https://www.ncbi.nlm.nih.gov/books/NBK6624/)
61. Bullinger L et al (2004) Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med 350:1605–1616. [PubMed: 15084693]
62. Bittner M et al (2000) Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 406(6795):536–540. [PubMed: 10952317]
63. Alizadeh AA et al (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403(6769):503–511. [PubMed: 10676951]
64. Paik S et al (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351(27):2817–2826. [PubMed: 15591335]
65. Albain KS et al (2010) Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomized trial. Lancet Oncol 11(1):55–65. [PubMed: 20005174]
66. Kumar R, Sharma A, Tiwari RK (2012) Application of microarray in breast cancer: an overview. J Pharm Bioallied Sci 4(1):21–26. [PubMed: 22368395]
67. Buyse M et al (2006) Validation and clinical utility of a 70-gene prognostic signature in women with node-negative breast cancer. J Natl Cancer Inst 98:1183–1192. [PubMed: 16954471]
68. Swerdlow SH et al (2008) WHO classification of Tumours of Haematopoietic and lymphoid tissues, 4th edn. IARC Press, Lyon
69. Huse JT (2016) Establishing a robust molecular taxonomy for diffuse gliomas of adulthood. Surg Pathol Clin 9(3):379–390. [PubMed: 27523967]
70. Hyman DM et al (2015) Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. N Engl J Med 373(8):726–736. [PubMed: 26287849]
71. Mariño-Enríquez A, Bovée J (2016) Molecular pathogenesis and diagnostic, prognostic and predictive molecular markers in sarcoma. Surg Pathol Clin 9(3):457–473. [PubMed: 27523972]
72. Ong FS et al (2012) Personalized medicine and pharmacogenetic biomarkers: progress in molecular oncology testing. Expert Rev Mol Diagn 12(6):593–602. [PubMed: 22845480]
73. Baccarani M, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. Blood. 2013; 122(6):872–884. [PubMed: 23803709]
74. Greig SL (2016) Osimertinib: first global approval. Drugs 76(2):263–273. [PubMed: 26729184]

75. Hyman DM et al (2015) Precision medicine at Memorial Sloan Kettering Cancer Center: clinical next-generation sequencing enabling next-generation targeted therapy trials. Drug Discov Today 20(12):1422–1428. [PubMed: 26320725]
76. Nowak JA, Hornick JL (2016) Molecular evaluation of colorectal adenocarcinoma: current practice and emerging concepts. Surg Pathol Clin 9(3):427–439. [PubMed: 27523970]
77. Ho C, Kluck MJ (2016) Molecular pathology: predictive, prognostic, and diagnostic markers in lymphoid neoplasms. Surg Pathol Clin 9(3):489–521. [PubMed: 27523974]
78. Seethala RR, Griffith CC (2016) Molecular pathology: predictive, prognostic, and diagnostic markers in salivary gland tumors. Surg Pathol Clin. 9(3):339–352. [PubMed: 27523965]
79. Vainchenker W et al (2011) New mutations and pathogenesis of myeloproliferative neoplasms. Blood 118:1723–1735. [PubMed: 21653328]
80. Kuo FC (2016) Molecular pathology: prognostic and diagnostic genomic markers for myeloid neoplasms. Surg Pathol Clin 9(3):475–488. [PubMed: 27523973]
81. Langerak AW et al (2012) EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. Leukemia 26(10):2159–2171. [PubMed: 22918122]
82. Theisen BK, Wald AI, Singhi AD (2016) Molecular diagnostics in the evaluation of pancreatic cysts. Surg Pathol Clin. 9(3):441–456. [PubMed: 27523971]
83. Greco FA (2013) Molecular diagnosis of the tissue of origin in cancer of unknown primary site: useful in patient management. Curr Treat Options in Oncol 14(4):634–642. [PubMed: 23990214]
84. Hainsworth JD, Greco FA (2014) Gene expression profiling in patients with carcinoma of unknown primary site: from translational research to standard of care. Virchows Arch 464(4):393–402. [PubMed: 24487792]
85. Schlegelberger B et al (2012) Update on cytogenetic and molecular changes in myelodysplastic syndromes. Leuk Lymphoma 53:525–536. [PubMed: 21877899]
86. Bejar R et al (2012) Validation of a prognostic model and the impact of mutations in patients with lower-risk myelodysplastic syndromes. J Clin Oncol 30:3376–3382. [PubMed: 22869879]
87. Reichel J et al (2015) Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. Blood 125(7):1061–1072. [PubMed: 25488972]
88. Olivier M et al (2006) The clinical value of somatic TP53 gene mutations in 1,794 patients with breast cancer. Clin Cancer Res 12:1157–1167. [PubMed: 16489069]
89. Freed-Pastor WA, Prives C (2012) Mutant p53: one name. many proteins Genes Dev 26(12):1268–1286. [PubMed: 22713868]
90. Clark RE et al (2008) Pharmacologic markers and predictors of responses to imatinib therapy in patients with chronic myeloid leukemia. Leuk Lymphoma 49(4):639–642. [PubMed: 18398725]
91. Giannoudis A et al (2013) The hOCT1 SNPs M420del and M408V alter imatinib uptake and M420del modifies clinical outcome in imatinib-treated chronic myeloid leukemia. Blood 121(4):628–637. [PubMed: 23223357]
92. Thurner P, Nanoff C (2008) Limitations in adjuvant breast cancer therapy: the predictive potential of pharmacogenetics and pharmacogenomics. Breast Care (Basel) 3(6):401–406. [PubMed: 21048911]
93. Soverini S et al (2011) BCR-ABL kinase domain mutation analysis in chronic myeloid leukemia patients treated with tyrosine kinase inhibitors: recommendations from an expert panel on behalf of European LeukemiaNet. Blood 118(5):1208–1215. [PubMed: 21562040]
94. Richter J, et al (2016) Stopping Tyrosine kinase inhibitors in a very large cohort of European chronic myeloid leukemia patients: results of the Euro-SKI Trial [abstract]. European Hematology Association (EHA) 2016 Congress. Abstract S145
95. Lo-Coco F, Cicconi L (2011) History of acute promyelocytic leukemia: a tale of endless revolution. Mediterr J Hematol Infect Dis 3(1):e2011067. [PubMed: 22220264]
96. Grimwade D, Lo CF (2002) Acute promyelocytic leukemia: a model for the role of molecular diagnosis and residual disease monitoring in directing treatment approach in acute myeloid leukemia. Leukemia 16(10):1959–1973. [PubMed: 12357347]

97. Syngal S et al (2015) ACG clinical guideline: genetic testing and management of hereditary gastrointestinal cancer syndromes. Am J Gastroenterol 110(2):223–262. [PubMed: 25645574]
98. Bayraktar S (2017) Arun B. BRCA mutation genetic testing implications in the United States. Breast 31:224–232. [PubMed: 27931006]
99. Couch FJ, Nathanson KL, Offit K (2014) Two decades after BRCA: setting paradigms in personalized cancer care and prevention. Science 343(6178):1466–1470. [PubMed: 24675953]
100. Gustafsson L et al (1997) International incidence rates of invasive cervical cancer before cytological screening. Int J Cancer 71(2):159–165. [PubMed: 9139836]
101. Gustafsson L et al (1997) International incidence rates of invasive cervical cancer after introduction of cytological screening. Cancer Causes Control 8(5):755–763. [PubMed: 9328198]
102. Saslow D et al (2002) American Cancer Society guideline for the early detection of cervical neoplasia and cancer. CA Cancer J Clin 52(6):342–362. [PubMed: 12469763]
103. Schiffman M et al (2007) Human papillomavirus and cervical cancer. Lancet 370(9590):890–907. [PubMed: 17826171]
104. Schiffman M et al (2011) Human papillomavirus testing in the prevention of cervical cancer. J Natl Cancer Inst 103(5):368–383. [PubMed: 21282563]
105. https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/InVitroDiagnostics/ucm330711.htm#microbial
106. http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P100020S008
107. Heuser M et al (2016) Clonal hematopoiesis of indeterminate potential: a risk factor for hematologic neoplasms. Dtsch Arztebl Int 113(18):317–322. [PubMed: 27215596]
108. Steensma DP et al (2015) Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood 126(1):9–16. [PubMed: 25931582]
109. Metzker ML (2010) Sequencing technologies - the next generation. Nat Rev Genet 11(1):31–46. [PubMed: 19997069]
110. Liu L et al (2012) Comparison of next-generation sequencing systems. J Biomed Biotechnol 251364([PubMed: 22829749]):2012
111. Ley TJ et al (2008) DNA sequencing of a cytogenetically normal acute myeloid leukemia genome. Nature 456(7218):66–72. [PubMed: 18987736]
112. Mardis ER et al (2009) Recurring mutations found by sequencing an acute myeloid leukemia genome. New Eng. J Med 361(11):1058–1066. [PubMed: 19657110]
113. http://www.fda.gov/drugs/scienceresearch/researchareas/Pharmacogenetics/ucm083378.htm
114. Byron SA et al (2016) Translating RNA sequencing into clinical diagnosis: opportunities and challenges. Nat Rev Genet 17(5):257–271. [PubMed: 26996076]
115. Zhao S et al (2014) RNA-Seq and microarray in transcriptome profiling of activated T cells. PLoS One 9(1):e78644. [PubMed 24454679]
116. Fumagalli D et al (2014) Transfer of clinically relevant gene expression signatures in breast cancer: from Affymetrix microarray to Illumina RNA-sequencing technology. BMC Genomics 15:1008. [PubMed: 25412710]
117. Zhang W et al (2015) Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. Genome Biol 16:133. [PubMed: 26109056]
118. Erlich Y et al (2008) Alta-cyclic: a self-optimizing base caller for next-generation sequencing. Nat Methods 5(8):679–682. [PubMed: 18604217]
119. Lee H, et al (2016) Third-generation sequencing and the future of genomics. http://www.biorxiv.org/content/biorxiv/early/2016/04/13/048603.full.pdf
120. Samorodnitsky E et al (2015) Evaluation of hybridization capture versus amplicon-based methods for whole-exome sequencing. Hum Mutat 36(9):903–914. [PubMed: 26110913]
121. Cheng DT et al (2015) Memorial Sloan Kettering-integrated mutation profiling of actionable Cancer targets (MSK-IMPACT) a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. J Mol Diagn 17(3):251–264. [PubMed: 25801821]
122. den Dunnen JT, Antonarakis SE (2000) Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat 15(1):7–12. [PubMed: 10612815]

123. Wain HM et al (2002) Guidelines for human gene nomenclature. Genomics 79(4):464–470. [PubMed: 11944974]
124. Forbes SA et al (2010) COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. Nucleic Acids Res 38(Database issue):D652–D657. [PubMed: 19906727]
125. COSMIC: http://cancer.sanger.ac.uk/cosmic/
126. Cerami E et al (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2(5):401–404. [PubMed: 22588877]
127. cBioPortal: http://www.cbioportal.org/
128. Yeh P et al (2013) DNA-mutation inventory to refine and enhance cancer treatment (DIRECT): a catalog of clinically relevant cancer mutations to enable genome-directed anticancer therapy. Clin Cancer Res 19(7):1894–1901. [PubMed: 23344264]
129. My Cancer Genome: http://www.mycancergenome.org/
130. Meric-Bernstam F et al (2015) A decision support framework for genomically informed investigational cancer therapy. J Natl Cancer Inst 107(7). [PubMed: 25863335]
131. Personalized Cancer Therapy: https://pct.mdanderson.org/
132. TARGET: http://archive.broadinstitute.org/cancer/cga/target
133. The Human Gene Mutation Database (HGMD): http://www.hgmd.cf.ac.uk/ac/index.php
134. http://oncokb.org/
135. Brock G et al (2015) Liquid biopsy for cancer screening, patient stratification and monitoring. Transl Cancer Res 4:3
136. Allard WJ et al (2004) Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res 10(20):6897–6904. [PubMed:15501967]
137. Alix-Panabières C, Pantel K (2016) Clinical applications of circulating tumor cells and circulating tumor DNA as liquid biopsy. Cancer Discov 6(5):479–491. [PubMed:26969689]
138. Beerepoot LV et al (2004) Increased levels of viable circulating endothelial cells are an indicator of progressive disease in cancer patients. Ann Oncol 15(1):139–145. [PubMed: 14679134]
139. Cohen SJ et al (2006) Isolation and characterization of circulating tumor cells in patients with metastatic colorectal cancer. Clin Colorectal Cancer 6(2):125–132. [PubMed: 16945168]
140. Joy JE, Penhoet EE, Petitti DB (eds) (2005) Saving women's lives: strategies for improving breast cancer detection and diagnosis. National Academies Press, Washington DC. [PubMed: 20669447]
141. http://www.accessdata.fda.gov/cdrh_docs/pdf3/k031588.pdf
142. Yu M et al (2014) Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. Science 345(6193):216–220. [PubMed: 25013076]
143. Gawad C, Koh W, Quake SR (2016) Single-cell genome sequencing: current state of the science. Nat Rev Genet 17(3):175–188. [PubMed: 26806412]
144. Xu X et al (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148(5):886–895. [PubMed: 22385958]
145. Navin NE (2014) Cancer genomics: one cell at a time. Genome Biol 15(8):452. [PubMed: 25222669]
146. Navin N et al (2011) Tumour evolution inferred by single-cell sequencing. Nature 472(7341):90–94. [PubMed: 21399628]
147. Hou Y et al (2012) Single-cell exome sequencing and monoclonal evolution of a *JAK2*-negative myeloproliferative neoplasm. Cell 148(5):873–885. [PubMed: 22385957]
148. Kinde I et al (2011) Detection and quantification of rare mutations with massively parallel sequencing. Proc Natl Acad Sci U S A 108(23):9530–9535. [PubMed: 21586637]
149. Schmitt MW et al (2012) Detection of ultra-rare mutations by next-generation sequencing. Proc Natl Acad Sci U S A 109(36):14508–14513. [PubMed: 22853953]
150. Newman AM et al (2014) An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med 20(5):548–554. [PubMed: 24705333]

151. Newman AM et al (2016) Integrated digital error suppression for improved detection of circulating tumor DNA. Nat Biotechnol 34(5):547–555. [PubMed: 27018799]
152. Gagan J, Van Allen EM (2015) Next-generation sequencing to guide cancer therapy. Genome Med 7(1):80. [PubMed: 26221189]
153. Dawson SJ et al (2013) Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 368(13):1199–1209. [PubMed: 23484797]
154. Genovese G et al (2014) Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med 371(26):2477–2487. [PubMed: 25426838]
155. Redzic JS et al (2014) Extracellular RNA mediates and marks cancer progression. Semin Cancer Biol 28:14–23. [PubMed: 24783980]
156. SEQC/MAQC-III Consortium (2014) A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the sequencing quality control consortium. Nat Biotechnol 32(9):903–914. [PubMed: 25150838]
157. http://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs/ucm504540.htm
158. Aziz N et al (2015) College of American Pathologists' laboratory standards for next-generation sequencing clinical tests. Arch Pathol Lab Med 139(4):481–493. [PubMed: 25152313]
159. Rehm HL et al (2013) ACMG clinical laboratory standards for next-generation sequencing. Genet Med 15(9):733–747. [PubMed: 23887774]
160. Green RC et al (2013) ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med 15(7):565–574. [PubMed: 23788249]
161. Hegde M et al (2015) Reporting incidental findings in genomic scale clinical sequencing—a clinical laboratory perspective: a report of the Association for Molecular Pathology. J Mol Diagn 17:107–117. [PubMed: 25684262]
162. Heather EW et al (2013) Cancer pharmacogenomics: strategies and challenges. Nat Rev Genet 14(1):23–34. [PubMed: 23183705]
163. Deverka PA, Dreyfus JC (2014) Clinical integration of next generation sequencing: coverage and reimbursement challenges. J Law Med Ethics 42(S1):22–41. [PubMed: 25298289]
164. http://nanthealth.com/nations-first-insurance-coverage-of-next-generation-whole-genome-sequencing-and-proteomic-diagnostic-platform-announced-by-independence-blue-cross-and-nanthealth-in-the-war-against-cancer/
165. http://finance.yahoo.com/news/nanthealth-expands-gps-cancer-coverage-170600097.html
166. Prasad V, Fojo T, Brada M (2016) Precision oncology: origins, optimism, and potential. Lancet Oncol 17(2):e81–e86. [PubMed: 26868357]
167. Yu KH, Snyder M (2016) Omics profiling in precision oncology. Mol Cell Proteomics 15(8):2525–2536. [PubMed: 27099341]
168. Waldron D (2016) Cancer genomics: a multi-layer omics approach to cancer. Nat Rev Genet 17(8):436–437. [PubMed: 27418154]
169. Vucic EA et al (2012) Translating cancer 'omics' to improved outcomes. Genome Res 22(2):188–195. [PubMed: 22301133]
170. Hanash S, Taguchi A (2010) The grand challenge to decipher the cancer proteome. Nat Rev Cancer 10(9):652–660. [PubMed: 20733593]
171. Zhang H et al (2016) Integrated proteogenomic characterization of human high-grade serous ovarian cancer. Cell 166(3):755–765. [PubMed: 27372738]
172. Hoadley KA et al (2014) Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell 158(4):929–944. [PubMed: 25109877]
173. Ritchie MD et al (2015) Methods of integrating data to uncover genotype-phenotype interactions. Nat Rev Genet 16(2):85–97. [PubMed:25582081]

Chapter 3
Clinical Interpretation

Nathanael D. Moore, Parastou Ghazi, and Eliezer M. Van Allen

Keywords Clinical interpretation · Computational biology · Algorithms · Knowledge base · Variant classification · Machine learning

3.1 Introduction

Applied to cancer, the precision medicine model can dramatically improve patient care by matching the unique constellation of alterations seen in an individual cancer patient to therapies that have the best chance of successfully treating that patient. Executing this model requires that researchers successfully bridge knowledge gained in the laboratory with the patients seen in the clinic: Physicians must be able to rapidly and accurately perform clinical interpretation of patient genetic data in order to optimize the care they provide.

With the advent of next-generation sequencing technology, the amount of genetic data available to both physicians and researchers has rapidly increased. The first human genome was sequenced over a 13 year period through the efforts of a multi-institutional project that cost over $3 billion dollars [1]. As of 2015, the cost of sequencing a single genome has dropped to $1245 [2], a cost comparable to many routine diagnostic tests, and the procedure can take as little as 26 hours from data acquisition to preliminary interpretation [3]. The use of more narrowly focused sequencing techniques, including whole-exome and targeted panel sequencing, allow for rapid acquisition of patient genetic data at even lower costs (Fig. 3.1).

As a result of these advances, patient genetic data are increasingly becoming available at the point of care, whether through targeted panels of commonly altered genomic locations, or by complete sequencing of the exome or genome. The broadened availability of genetic information has potential implications for the treatment

N. D. Moore · P. Ghazi · E. M. Van Allen (✉)
Dana Farber Cancer Institute, Harvard Medical School, Boston, MA, USA
e-mail: eliezerm_vanallen@dfci.harvard.edu

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_3

Fig. 3.1 The decreasing cost of whole-genome sequencing, from 2001 to 2015 (cost per genome on vertical axis) [2]. Cost reductions accelerated with the introduction of next-generation sequencing (NGS), a family of technologies that enable parallel sequencing

of cancer patients, giving clinicians the ability to analyze the mutations and oncogenic drivers that underlie a patient's cancer and select therapeutics personalized to the set of mutations observed in a given patient.

However, the vast quantities of data produced per patient introduce new challenges. The genetic abnormalities present in cancer patients often result in dramatically increased mutation rates, further obscuring the genetic landscape; in order to understand the fundamental determinants of a patient's cancer, causal "driver" mutations must be identified and distinguished from "passenger" mutations that appear alongside driver mutations, but are not biologically relevant [4–6]. Once mutations of interest are identified, they must be linked to potentially effective therapeutic options (which can vary dramatically depending on the cancer type and the specific mutations involved), prognostic information, and indicators of clinical trial eligibility in order to provide clinical benefit to the patient.

Achieving these goals requires a multi-pronged approach, combining manual review of patient tumor data with automated analyses that take advantage of cutting-edge oncology research. The remainder of this chapter will explore the recent history of precision cancer developments from the perspective of the clinic, describe the current state of available technologies and techniques for bringing robust interpretation of these data into the clinic, and provide an overview of the challenges faced both by clinicians and researchers in performing clinical interpretation of patient-specific cancer genomic data.

3.2 History

Far prior to the introduction of next-generation sequencing technologies, genetic information has been used to improve treatment of cancer patients. In 1988, an association was discovered between the presence of translocations in the *RARα* retinoic acid receptor gene and increased rates of remission after treatment with all-trans retinoic acid (ATRA) [7]. The discovery in 1990 that *BRCA1* and *BRCA2* alterations are associated with heritable breast and ovarian cancer [8] was quickly used to tailor how clinicians approach patients with these abnormalities, whether by encouraging increased screening and/or prophylactic mastectomy and salpo-oophorectomy in high-risk patients [9]. Soon after, one of the first drugs directly targeting a genetic abnormality was developed; the connection between the "Philadelphia chromosome," a genomic translocation creating the *BCR-ABL* fusion gene, and several forms of leukemia led to the creation of imatinib, a small molecule inhibitor that targets the mutated tyrosine kinase signaling molecule produced by the *BCR-ABL* gene. Imatinib vastly improves outcomes in patients with this abnormality; the five-year survival rate of chronic myelogenous leukemia (CML) patients, the primary indication for imatinib, has increased from 31% in the early 1990s to 63% in patients diagnosed between 2005 to 2011 [10]. Additional targeted therapies soon followed, including treatments for patients with BRAF-mutant metastatic melanoma [11], and non-small-cell lung cancers with mutations in either EGFR [12] or ALK [13].

Each of these relationships between a genetic abnormality and a clinical action is an application of precision cancer medicine techniques; a discriminating feature was identified in a subset of patients with a certain cancer type, which was then used by clinicians to tailor their treatment plans to individual patients. The massive increase in genetic data enabled by modern sequencing technology holds vast potential for discovering additional genetic abnormalities linked to clinical actions. The availability of routine clinical whole-exome and whole-genome sequencing, deployed at multiple institutions [5, 14, 15], allow for more holistic approaches to genetic analysis. Rather than limit analyses to the presence of alterations at individual genetic loci, global features that weigh exome- or genome-wide effects may also be interpreted. For example, the overall mutational burden of a patient's tumor has been linked to clinically actionable information in the settings of bladder cancer [16], non-small cell lung cancer [17], and metastatic melanoma [18]. Analyses of the differing patterns of somatic mutations observed in cancer genomes have led to the identification of "mutational signatures" that reflect the mutational processes observed in a cancer [19]; the presence of these signatures has been linked to actionable information in breast cancer [20], non-small cell lung cancer [12], bladder cancer [21], and esophageal adenocarcinoma [22] among others.

However, research into targeted therapeutics and prognostic associations is not sufficient; clinical interpretation efforts need to be made to match these discoveries with the patients that can benefit most from them. The field of oncology research moves far too rapidly for clinicians to feasibly be aware of the current state of

research at any given time; not only do new therapies frequently receive approval, but new indications are applied to existing therapies as discoveries are made, new clinical trials are created that may benefit subsets of patients who fulfil specific eligibility criteria, and contemporary pre-clinical research may help guide treatment in patients who have failed all standard approved therapies. Molecular tumor boards have been the traditional approach to this problem, in which experts within various areas of clinical oncology at a hospital meet to discuss complex patient cases and recent research advances. While useful, tumor boards are limited by the areas of expertise and fields of research performed at the home institution, cannot provide a comprehensive landscape of the current body of literature applicable to each patient, and are often unable to meet clinically acceptable turnaround times [23]. In response, several approaches have been pursued towards automatic clinical interpretation. While clinicians are limited to manual review of a limited number of well-known alterations that have been previously validated, automated computational approaches are able to quickly review the entirety of a patient's genetic landscape and draw from the vast body of current literature to make clinical suggestions. When provided with additional knowledge about the biological processes underpinning oncogenesis and progression, these tools can make more remote inferences, potentially suggesting off-label or novel uses of therapies compatible with the patient's genetic status.

3.3 Current Developments

Current research into improved clinical interpretation involves three separate aspects: Acquisition, Analysis, and Action (Fig. 3.2).

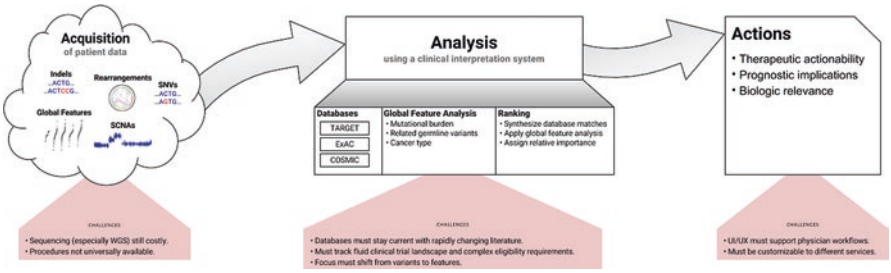

Fig. 3.2 The acquisition-analyses-action workflow. Patient data are acquired via tissue sample sequencing, and input into clinical interpretation systems for analysis. Recommendations are made based on direct analysis of variants and more complex global feature analysis, and synthesized into a report of potential actions for review by clinicians

3.3.1 Acquisition

Acquisition involves the capture of tissue samples from a cancer patient, and the processing of these samples to produce genetic sequence data. In general, two kinds of tissue samples may be acquired: tumor samples, which are taken from the site of the neoplasm caused by the patient's disease, and normal samples, which are taken from a patient's healthy tissue. The decision to acquire a tumor sample, normal sample, or both kinds of samples from a patient determines the kinds of analyses that can be performed on the mutations present in the resulting sequence data.

Mutations may be divided into *somatic* and *germline* alterations. *Somatic* alterations are those that occur during the lifetime of the patient. They are not passed down from a parent, and cannot be passed on to their descendants. In contrast, *germline* alterations are inherited from a parent, and can be passed on to progeny. Both types of alterations may lead to cancer, either by activating or increasing the expression of a *proto-oncogene* (a gene that encourages growth) to become an *oncogene* (a dysregulated pro-growth gene), or by disabling or reducing the expression of a *tumor suppressor* gene (a gene that normally restricts growth) (Fig. 3.3).

The acquisition of a tumor sample provides information on both the somatic alterations that are involved in the patient's cancer and the patient's germline alterations. A normal sample primarily enables identification of germline alterations. Acquiring a tumor sample is of usually of primary importance, since it will reveal cancer-specific alterations that may suggest use of a targeted therapy or provide prognostic insight. However, acquiring a matched normal sample may reveal that some mutations identified in the tumor sample are in fact germline alterations, with separate implications for treatment selection and heritability counseling [24].

See the "Molecular Pathology" and "Liquid Biopsies" chapters for more information on the acquisition of patient tissue samples.

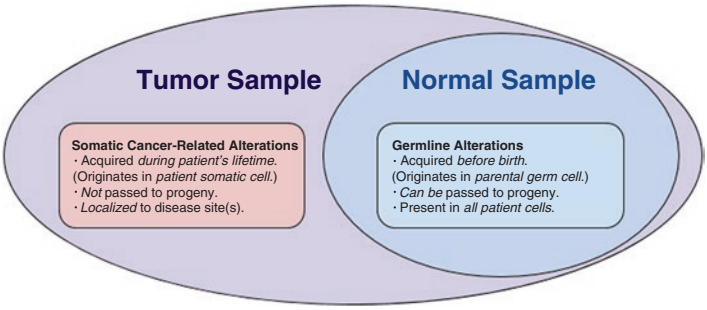

Fig. 3.3 Tumor samples contain both germline alterations and the somatic alterations that contribute to (or have occurred because of) a patient's cancer. Normal samples primarily contain germline alterations that are present in nearly all of the patient's cells. Sampling both tumor and normal tissue allows the determination of which alterations are somatic and thus more likely to contribute to the cancer state

3.3.2 Analysis

Once a patient's genetic information has been obtained, the next step towards clinical interpretation is analysis of the data. Usually, a pipeline of software is first used to identify specific kinds of alterations within the raw genetic data provided (including point mutations [25], insertions/deletions [indels] [26], fusions [27, 28], and others). These data are often further annotated by algorithms that match additional information to the identified alterations, such as the predicted impact of an alteration, information about the altered protein resulting from an alteration, and population-level statistics among others [29–31].

Clinical interpretation systems attempt to make therapeutic suggestions and prognostic estimates that build upon previously annotated alteration information. These systems annotate alteration data with information about clinical actionability, often extracted from cancer-specific databases (with a focus on clinically actionable assertions) or by performing independent analyses on sequencing data. Clinical interpretation systems range from taking disease-specific approaches (e.g., OncotypeDX [32], shown below, produces products specifically for analysis of breast, colon, and prostate cancer) to pan-cancer approaches (e.g., FoundationOne [33] and PHIAL [5], shown below). See "Challenges" below for a case study of PHIAL, a clinical interpretation system currently under development, and an examination of how it attempts to address several obstacles faced by clinical interpretation approaches in general.

3.3.3 Action

"Clinical actionability" is vaguely defined, but generally refers to the potential of an assertion to change a physician's treatment plan, or to provide prognostic information about the severity or likely outcome of the disease. The primary intention of most clinical interpretation systems is to recommend a therapy targeted to the mutations involved in an individual's cancer, usually after non-targeted first-line therapies have failed. These suggestions may involve the off-label use of a therapy (e.g., recommending a drug approved for a different cancer, but which blocks a pathway altered in several cancers), or the enrollment of the patient in an ongoing clinical trial. Information about the state of the patient (e.g., what treatments the patient has failed, or the stage of the cancer) may improve the ability of an interpretation system to recommend better clinical actions tailored to the disease progress in a given patient (Table 3.1).

Table 3.1 Examples of current techniques applied to fulfil several clinical interpretation unmet needs across the spectrum of precision cancer medicine

Aspect	Examples	
Acquisition	Illumina HiSeq	*Mass-sequencing of many samples.*
	Illumina TruGenome	*Sequencing service for small numbers of samples.*
Analyses	FoundationOne	*Reports putative actionable alterations within 315 genes. Estimates immunotherapy response and suggests clinical trial eligibility.*
	OncotypeDX	*Reports putative actionable alterations within small sets of genes optimized for specific cancer types (breast, colon, and prostate cancer).*
	PHIAL	*Reports putative actionable alterations with 307 genes. Analyzes global features and attempts to identify novel therapeutic applications.*
Action	Therapeutics	E.g., *imatinib, vemurafenib, trametinib, sorafenib, gefitinib, etc.*
	Heritability Counseling	E.g., BRCA1/BRCA2 *(breast and ovarian cancer)*, FANCA *(Fanconi anemia).*

3.3.4 Case Study: PHIAL

PHIAL (Precision Heuristics for Interpreting the Alteration Landscape) is a computational algorithm developed to help bridge the gap between clinical research and the clinic, and serves as an example of a modern approach to performing rapid clinical interpretation [5]. This algorithm operates on patient genetic sequence data that has been thoroughly investigated to detect several types of abnormalities (such as point mutations, indels, and fusions, among others). PHIAL performs multiple analyses on the collected abnormalities in an attempt to discover new information that could influence how a clinician decides to treat a patient, or provide additional prognostic information. It then ranks the information it has collected according to its relative importance and creates a report intended for use by physicians at the point-of-care.

A key component of PHIAL is its ability to search several oncology databases for additional information about genetic abnormalities and synthesize this information to provide targeted treatment suggestions. One such database, TARGET (Tumor Alterations Relevant for GEnomics-driven Therapy), is co-developed alongside PHIAL to directly relate genetic alterations with potential treatment options according to published research discoveries and clinical trial findings (Fig. 3.4).

Notably, TARGET curates not only clinically validated discoveries but also preclinical findings of various likelihood to have biological relevance in cancer patients. TARGET uses a series of "predictive implication levels" to categorize assertions, allowing PHIAL to rank matching assertions by the comparative rigor of the aggregated existing research into their potential effects (see Fig. 3.2). For example, the

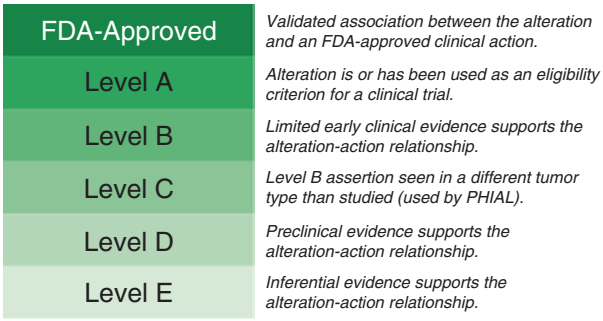

Fig. 3.4 The predictive implication levels used by TARGET to classify the likelihood of an alteration-action relationship's clinical impact

use of imatinib in a CML patient harboring the BCR-ABL fusion gene would be given an "FDA-Approved" predictive implication level due to imatinib's status as an FDA-approved drug for patients in this setting, while the use of a molecule that has only been validated in mouse models would be given a "Level D" predictive implication level since it has not been tested in the clinic.

PHIAL also performs global feature analyses in order to make suggestions based on the overall "cancer state" of the patient, rather than individual alterations. One such global feature is the overall mutational burden of a given patient, which has been associated with differences in patient responses in several tumor types (including metastatic melanoma [17], non-small cell lung cancer [34], and urothelial carcinoma [35]); PHIAL and TARGET record available information on how mutational burden relates to actionability in various cancers, and PHIAL also reports when the observed mutational burden lies outside the norm for a patient's type of cancer. Additionally, TARGET contains information about how the presence of certain mutational signatures, discussed above, relate to potential clinical actions; work is underway to enable PHIAL to derive and interpret mutational signatures using genetic data from a single patient.

This case study demonstrates the paradigm shift occurring in the field of clinical interpretation away from using single alterations to make clinical predictions and towards the synthetic use of multiple features to obtain a global view of a patient's disease, leading to an ability to make more accurate predictions and suggest more precisely targeted therapies. As PHIAL moves from a heuristic system based on the detection of damaging alterations towards the use of global features such as mutational burden and mutational signatures, it gains the ability to better model the underlying mutational processes occurring in a patient. In the future, PHIAL and similar systems could combine biological knowledge of the interplay observed in pathways associated with cancer with an accurate modeling of how these pathways are disrupted in individual patients to suggest novel uses of pre-existing drugs that could have a biological effect based on their mechanisms of action, or help guide cancer researchers by highlighting how pathways are disrupted across cohorts of patients within a cancer type. As the ability of these systems to more finely discriminate between disease processes improves, the categorization of cancer types may

move from the current tissue-based ontology (e.g., breast cancer, bladder cancer, prostate cancer, etc.) to a feature-based ontology predicated upon recurring mutational patterns.

3.4 Challenges

Improving clinical interpretation promises many potential benefits; as cancer treatment moves towards a precision medical model and drugs become more specifically targeted, it will become increasingly more important that patients are correctly assigned appropriate treatments. Achieving this goal requires that researchers and clinicians surmount several obstacles.

3.4.1 Tissue Acquisition

In order to sequence patient genetic data, it is necessary to extract DNA and RNA from patient tissue samples. The quality of these samples correlates with sequencing accuracy; contaminated samples will result in high sequencing error and an inability to derive useful information from a sample.

The most common method of sample acquisition creates archival formalin-fixed paraffin-embedded (FFPE) tissue, which performs well at maintaining cell morphology over long periods of time. However, the FFPE process severely degrades DNA and RNA, resulting in low purity samples and high rates of sequencing error. Fresh frozen samples are far better at maintaining the integrity of DNA and RNA, but have not achieved the same popularity as FFPE tissue; frozen samples are more difficult to acquire, and incur higher costs to archive compared to FFPE, which can be stored at room temperature.

To perform accurate clinical interpretation, it will be necessary for institutions to standardize their methods for acquiring tissue samples intended for genetic analysis, and for researchers to optimize methods for extracting genetic information from impure samples. FFPE tissue has become more useful with the introduction of NGS; parallel sequencing allows higher coverage of targets of interest, providing more power to detect alterations despite the low quality of individual reads [22]. Alternative methods of sample acquisition may also further increase the number of patients who are able to undergo genetic testing (see “Liquid Biopsies”).

3.4.2 Sequencing Approaches and Infrastructure

Generally, three approaches are taken in genetic sequencing: whole-genome sequencing (WGS), which determines the entire genetic sequence of an individual; whole-exome sequencing, which only provides the sequence of an individual’s

coding regions of DNA; and targeted panels, which only sequence certain predetermined genetic loci (generally of areas known to be frequently mutated in cancer and associated with clinical actions). More comprehensive sequencing methods (such as WGS) provide more information that may lead to clinical actions, particularly when searching for large-scale alterations such as rearrangements and large-scale deletions, but also demand higher costs. The sequencing of additional genetic information can provide additional important genetic information; for example, RNA sequencing can help identify expression level changes and increase confidence in alteration detection. It is thus important that patients receive a level of sequencing that allows accurate clinical interpretation at a non-prohibitive cost that allows widespread adoption [36].

Access to sequencing infrastructure also factors into patient access to precision cancer medicine. The relatively high cost and specialized equipment required to perform high-quality genetic sequencing generally limits its use to large academic centers, limiting the ability of patients in medically underserved areas to receive targeted therapies. It is also vital that physicians receive accurate and contemporary education on how to use and interpret modern genetic tests. Oncology remains one of the most rapidly advancing fields in medicine, and it is challenging for clinicians to remain abreast of ongoing research and clinical trials that could benefit their patients. Additionally, the over-representation in genetic studies of individuals of European descent has resulted in a dearth of research on potential genetic differences that could alter drug response rates and outcome prognostics in patients of different ethnicities [37]. Researchers must seek to include a greater diversity of patient populations in future studies to ensure that as many patients as possible can receive targeted treatments.

3.4.3 Clinical Adoption

No matter how accurate its predictions, a clinical interpretation system will not provide benefit to patients unless it successfully integrates itself into the clinic. Multiple factors can slow clinical adoption. Cost remains a potential deterrent, especially at institutions not located within major academic centers; while the price of capturing a genome, exome, or targeted gene panel from a patient has rapidly decreased in recent years, performing these procedures requires the use of specialized sequencing machines with high reagent and maintenance costs, reducing the likelihood that a non-academic center will have easy access to them. Clinician education also presents a potential barrier. Although a major advantage of clinical interpretation is its ability to rapidly provide information to clinicians about the most recent research relevant to their patients, it also requires that clinicians understand how to interpret the information generated by clinical interpretation systems and apply it in their practice. Doing so will require that researchers develop methods for meaningfully ranking actionability suggestions and emphasizing what is unique about a patient's disease state and how that knowledge can be leveraged to provide clinical benefit.

3.5 Future Approaches

The field of clinical interpretation moves rapidly alongside oncology research, incorporating new discoveries and techniques for analyzing patients as they become available. New and less costly methods of acquiring patient samples allow for greater exploration of sequence data in more patients. Improved clinical interpretation systems allow clinicians to make greater use of patient data and make novel insights into potential actions they may take. New methods for curating alteration-action relationship databases increase the potential impact of discoveries and the audience of patients that could benefit from them. Improved reporting of action suggestions increases the likelihood that clinicians will be able to act on the suggestions made by interpretation systems and bring precision medicine into the clinic.

3.5.1 Acquisition Improvements

Currently, FFPE sample acquisition is the most-used technique in the clinic. As described above, the FFPE process is not optimal for use in sequencing; techniques such as fresh tissue freezing prevent degradation of genetic information, and result in more accurate sequencing results. It is possible that wider acquisition of fresh frozen tissue will be encouraged among institutions as precision medicine techniques supplant the exclusive use of histology in managing cancer patients. However, many advances have been made in adopting FFPE samples for use in genetic profiling [5, 38, 39]; this approach not only allows institutions to continue to use standard FFPE procedures, but also enables researchers to make use of archived samples acquired before modern sequencing techniques existed [40]. Future developments will likely include research into new tissue sampling methods optimized for genetic profiling, as well as improved protocols for making the most use out of existing sampling methods.

3.5.2 Analysis Improvements

Reduced sequencing costs will allow researchers and clinicians to use greater amounts of data to inform clinical decisions. Initially, only targeted panels composed of genes well-known to be involved with cancer were sequenced and analyzed; while useful, these panels do not capture the full alteration landscape of each cancer patient, and thus limit their ability to make clinical suggestions. Panels also impair future cancer research; even if a patient's data does not implicate a known therapeutic option at the time of sequencing, it may contribute to research that leads to the development of a future therapy. As sequencing costs continue to decrease, greater numbers of patients will have access to more comprehensive forms of

sequencing, such as WES and WGS. One approach to utilizing this increased quantity of data is to develop pan-cancer clinical interpretation systems that draw from frequently updated databases of alteration-action associations (e.g., the pairing of the PHIAL clinical interpretation system with the TARGET database; see PHIAL case study above). However, it will soon become possible for interpretation systems to move beyond alteration-based prediction to feature-based prediction, in which exome- or genome-wide information is synthesized to make prediction. An early example of this approach lies in the use of mutational signatures [18] to classify cancers by the specific mutational processes that are taking place as they progress, rather than by the tissue type they appear in.

The acquisition of greater quantities of data per patient, combined with the greater number of patients from which samples may be acquired, will provide new opportunities for applying artificial intelligence approaches to clinical interpretation, particularly in the field of machine learning (ML) [41]. In ML, large quantities of data samples are collected and used to train algorithms to better classify novel data. In general, ML methods are divided into *supervised* and *unsupervised* techniques. In supervised learning, large amounts of input training data are pre-classified by a researcher (e.g., a genetic sequence sample might be labeled as belonging to a patient with metastatic prostate cancer resistant to androgen deprivation therapy). An ML algorithm may then use this set of training data to "learn" a generalized approach to classifying new information not contained in the training set. In contrast, unsupervised learning involves the classification of data without the benefit of learning from pre-labeled data (e.g., classifying the cancer types corresponding to several genetic sequence samples with no accompanying clinical information). Several ML techniques have already been applied to cancer research and clinical interpretation [41–43].

Increased use of cancer-related databases will also improve clinical interpretation by providing clinical interpretation systems with additional information they may relate to patients. It is apparent that several kinds of databases will be needed to optimize precision medicine delivery.

Databases capturing population-level statistics on alteration prevalence will assist interpretation systems by helping determine which alterations are potentially deleterious and which are likely benign by filtering out variants that are relatively prevalent in normal populations [44, 45]. One such database, the Exome Aggregation Consortium (ExAC) [46], pursues this goal via a collection of over 60,000 exomes aggregated from multiple studies; its planned successor, the Genome Aggregation Database (gnomAD) [47], seeks to create an even larger database that includes genomic data.

There is a perhaps even greater need for databases that track the associations that have been discovered between genetic alterations and clinical actionability. TARGET, described above in the PHIAL Case Study, is one example of such a database; it chooses to focus on including a wide spectrum of alteration-action discoveries, whether or not they have been clinically validated. Other databases, such as the Precision Oncology Knowledge Base (OncoKB) [48] and My Cancer Genome [49], also capture clinically actionable alteration information with more focus on

clinically validated information. All of these databases must stay current with the rapidly changing field of oncology and the massive quantity of new information that is released daily in newly published literature, a process termed *biocuration*. Currently, most databases pursue a manual curation approach, in which researchers and physicians periodically review current literature and add new findings to their database. This approach results in high-quality additions that are highly relevant to cancer and have a high likelihood of clinical applicability, but also inevitably creates a delay between new developments and the ability for clinical interpretation. Several research projects involve the development of methods to automatically curate new literature and extract information relevant to researchers. One such method, PubTator [50], assists biocuration efforts by attempting to automatically detect "bio-entities" (names of genes, diseases, species, chemicals [such as drugs], and mutations) in article abstracts. Methods that focus more specifically on cancer include EMU (extractor of mutations) [51], MutationFinder [52], and tmVar [53], all of which focus on identifying references to various genetic alterations within journal articles.

3.5.3 Action Improvements

As precision oncology research develops, greater numbers of drugs targeted to specific genetic alterations or cancer stated will be developed. As the clinician's armamentarium grows, it is important that clinical interpretation systems accurately represent how precision therapies can help their patients. Thus, researchers developing clinical interpretation systems must focus not only on producing accurate results for patients, but also explore how to design intuitive user interfaces that display their results in a manner that enables accurate physician understanding.

This will be a particular important area to explore with regard to clinical trial matching. A cornerstone of precision medicine is the execution of trials with carefully defined inclusion and exclusion requirements that enable researchers to tease apart how a candidate drug affects patients with different alteration states. However, the rapid proliferation of these trials, combined with increasingly complex eligibility requirements, has resulted in an incomplete understanding on the behalf of individual physicians about which of their patients could qualify for which trials. This is a particular important problem for patients that have failed all available standard therapies. Clinical interpretation systems could provide significant improvements in patient treatment by matching patient genetic information with clinical trials that could provide a benefit, and by providing clinicians an interface that allows them to quickly evaluate their patient's suitability for the trial and enroll the patient.

Electronic medical records (EMR) systems offer a potential avenue through which clinical interpretation systems can present their findings in the context of a patient's clinical history. They also have the opportunity to increase the amount of information available to researchers, allowing clinical interpretation systems to both provide insights to patients and bring new data back to the lab. The developments of

targeted therapies requires not only the acquisition of patient genetic data but also the ability to match that data to the phenotypes observed in patients. The additional clinic information stored in EMR systems could be a source of this information; however, innovators must be careful to preserve patient privacy as data are shared between individuals and institutions.

3.6 Conclusion

Advances in computational biology techniques, alongside the acquisition of vast amounts of knowledge on cancer pathways and increased availability of genetic sequencing, have enabled the field of precision medicine to grow dramatically. Research insights hold new promise for patients who have few or no remaining standard treatment options, and new analysis techniques have the potential to massively increase the number of targeted therapies available to patients. However, rapid, accurate, and comprehensive clinical interpretation techniques are required to fulfill the promise of precision medicine to patients. Novel targeted therapies cannot make an impact unless physicians can unambiguously identify who will respond to them. As researchers make new discoveries in the biology of cancer, clinical interpretation will be critical to giving their discoveries meaning in the clinic.

References

1. Venter JC, Adams MD, Myers EW et al (2001) The sequence of the human genome. Science 291(5507):1304–1351
2. Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcostsdata. Accessed March 20, 2017
3. Miller NA, Farrow EG, Gibson M et al (2015) A 26-hour system of highly sensitive whole genome sequencing for emergency management of genetic diseases. Genome Med 7:100
4. Lawrence MS, Stojanov P, Polak P et al (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 499(7457):214–218
5. Van Allen EM, Wagle N, Stojanov P et al (2014) Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. Nat Med 20(6):682–688
6. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144(5):646–674
7. Huang ME, Ye YC, Chen SR et al (1988) Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. Blood 72(2):567–572
8. Hall JM, Lee MK, Newman B et al (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250(4988):1684–1689
9. Bayraktar S, Arun B (2017) BRCA mutation genetic testing implications in the United States. Breast 31:224–232
10. Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. CA Cancer J Clin 66(1):7–30
11. Flaherty KT, Infante JR, Daud A et al (2012) Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. N Engl J Med 367(18):1694–1703

12. Maemondo M, Inoue A, Kobayashi K et al (2010) Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med 362(25):2380–2388
13. Shaw AT, Kim DW, Nakagawa K et al (2013) Crizotinib versus chemotherapy in advanced ALK-positive lung cancer. N Engl J Med 368(25):2385–2394
14. Roychowdhury S, Iyer MK, Robinson DR et al (2011) Personalized oncology through integrative high-throughput sequencing: a pilot study. Sci Transl Med 3(111):111–121
15. Beltran H, Eng K, Mosquera JM et al (2015) Whole-exome sequencing of metastatic cancer and biomarkers of treatment response. JAMA Oncol 1(4):466–474
16. Cazier JB, Rao SR, McLean CM et al (2014) Whole-genome sequencing of bladder cancers reveals somatic CDKN1A mutations and clinicopathological associations with mutation burden. Nat Commun 5:3756
17. Rizvi NA, Hellmann MD, Snyder A et al (2015) Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science 348(6230):124–128
18. Van Allen EM, Miao D, Schilling B et al (2015) Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science 350(6257):207–211
19. Alexandrov LB, Nik-Zainal S, Wedge DC et al (2013) Signatures of mutational processes in human cancer. Nature 500(7463):415–421
20. Nik-Zainal S, Alexandrov LB, Wedge DC et al (2012) Mutational processes molding the genomes of 21 breast cancers. Cell 149(5):979–993
21. Kim J, Mouw KW, Polak P et al (2016) Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. Nat Genet 48(6):600–606
22. Secrier M, Li X, De Silva N et al (2016) Mutational signatures in esophageal adenocarcinoma define etiologically distinct subgroups with therapeutic relevance. Nat Genet 48(10):1131–1141
23. Ghazani AA, Oliver NM, St Pierre JP et al (2017) Assigning clinical meaning to somatic and germ-line whole-exome sequencing data in a prospective cancer precision medicine study. Genet Med
24. Jones S, Anagnostou V, Lytle K et al (2015) Personalized genomic analyses for cancer mutation discovery and interpretation. Sci Transl Med 7(283):283ra53
25. Cibulskis K, Lawrence MS, Carter SL et al (2013) Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 31(3):213–219
26. Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ, Cheetham RK (2012) Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics 28(14):1811–1817
27. Iyer MK, Chinnaiyan AM, Maher CA (2011) ChimeraScan: a tool for identifying chimeric transcription in sequencing data. Bioinformatics 27(20):2903–2904
28. Hass B, Dobin A, Stransky N (2017) STAR-fusion: fast and accurate fusion transcript detection from RNA-Seq. bioRxiv. https://doi.org/10.1101/120295
29. Ramos AH, Lichtenstein L, Gupta M et al (2015) Oncotator: cancer variant annotation tool. Hum Mutat 36(4):E2423–E2429
30. Ng SB, Turner EH, Robertson PD et al (2009) Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461(7261):272–276
31. Mclaren W, Gil L, Hunt SE et al (2016) The Ensembl variant effect predictor. Genome Biol 17(1):122
32. OncotypeDX. Available at: http://www.oncotypedx.com/. Accessed 23 Mar 2017
33. FoundationOne. Available at: http://www.foundationone.com/. Accessed 23 Mar 2017
34. Rizvi NA, Hellmann MD, Snyder A et al (2015) Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science 348(6230):124–128
35. Chang SS (2016) Re: Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. J Urol 196(6):1637–1638

36. Garofalo A, Sholl L, Reardon B et al (2016) The impact of tumor profiling approaches and genomic data strategies for cancer precision medicine. Genome Med 8(1):79
37. Popejoy AB, Fullerton SM (2016) Genomics is failing on diversity. Nature 538(7624):161–164
38. Beltran H, Yelensky R, Frampton GM et al (2013) Targeted next-generation sequencing of advanced prostate cancer identifies potential therapeutic targets and disease heterogeneity. Eur Urol 63(5):920–926
39. Wagle N, Berger MF, Davis MJ et al (2012) High-throughput detection of actionable genomic alterations in clinical tumor samples by targeted, massively parallel sequencing. Cancer Discov 2(1):82–93
40. Carrick DM, Mehaffey MG, Sachs MC et al (2015) Robustness of next generation Sequencing on older formalin-fixed paraffin-embedded tissue. PLoS One 10(7):e0127353
41. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI (2015) Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 13:8–17
42. Polley MY, Freidlin B, Korn EL, Conley BA, Abrams JS, McShane LM (2013) Statistical and practical considerations for clinical evaluation of predictive biomarkers. J Natl Cancer Inst 105(22):1677–1683
43. Cruz JA, Wishart DS (2007) Applications of machine learning in cancer prediction and prognosis. Cancer Inform 2:59–77
44. Bamshad MJ, Ng SB, Bigham AW et al (2011) Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 12(11):745–755
45. MacArthur DG, Manolio TA, Dimmock DP et al (2014) Guidelines for investigating causality of sequence variants in human disease. Nature 508(7497):469–476
46. Lek M, Karczewski KJ, Minikel EV et al (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536(7616):285–291
47. The genome Aggregation Database. Available at: http://gnomad.broadinstitute.org/. Accessed 31 Mar 2017
48. OncoKB. Available at: http://oncokb.org. Accessed 31 Mar 2017
49. My Cancer Genome. Available at: https://www.mycancergenome.org. Accessed 31 Mar 2017
50. Wei CH, Kao HY, Lu Z (2013) PubTator: a web-based text mining tool for assisting biocuration. Nucleic Acids Res 41(Web Server issue):W518–W522
51. Doughty E, Kertesz-Farkas A, Bodenreider O et al (2011) Toward an automatic method for extracting cancer- and other disease-related point mutations from the biomedical literature. Bioinformatics 27(3):408–415
52. Caporaso JG, Baumgartner WA, Randolph DA, Cohen KB, Hunter L (2007) MutationFinder: a high-performance system for extracting point mutation mentions from text. Bioinformatics 23(14):1862–1865
53. Wei CH, Harris BR, Kao HY, Lu Z (2013) tmVar: a text mining approach for extracting sequence variants in biomedical literature. Bioinformatics 29(11):1433–1439

Chapter 4
Precision Cancer Medicine and Clinical Trial Design

Senthil Damodaran, Jason Willis, and Sameek Roychowdhury

Keywords Basket · Genomics · Precision · Targeted therapy

4.1 Introduction

Over the past decade, large-scale genomic surveys have uncovered a wide landscape of somatic alterations in cancers, both across different types and within histopathologically identical sub-types. The contribution of this molecular heterogeneity towards clinical outcome (e.g. overall survival, response, toxicity) has become the focus of intense study and is undoubtedly an important component of precision cancer medicine. Importantly, as our understanding of cancer as a heterogeneous and complex disease continues to evolve, so do the approaches towards therapeutic discovery [1, 2]. Currently, multiple targeted therapies to matching genomic alterations have been approved by the Food and Drug Administration (FDA) while many others are being evaluated in clinical trials. While the availability of cancer genomic testing in the clinic has led to opportunities in oncology such as drug target discovery, it has also led to challenges including how to develop targeted therapies for small populations of patients.

Since many of the targetable alterations typically occur at low frequencies within a single tumor type, accrual of patients in clinical trials becomes a challenge. In this chapter, we outline the opportunities and challenges presented by precision medicine and clinical trial implementation.

S. Damodaran · J. Willis
UT MD Anderson Cancer Center, Houston, TX, USA

S. Roychowdhury (✉)
Department of Internal Medicine, Division of Medical Oncology
Comprehensive Cancer Center & The Arthur G. James Cancer Hospital
at The Ohio State University, Columbus, OH, USA
e-mail: Sameek.roychowdhury@osumc.edu

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_4

4.2 Target Discovery and Validation

Clinical implementation of precision medicine involves two broad approaches: genotype to phenotype strategy where trials select for genomic alterations that may predict for response to therapy, and phenotype to genotype strategy where patients who had an exceptional response are retrospectively analyzed to identify molecular features associated with response. While cancer genomic testing has facilitated the identification of novel targets, it has also raised challenges for clinical implementation. The effectiveness of targeted therapy depends on matching with the right target; targets that are differentially expressed in tumor cells and provide growth and survival advantage as a driver mutation. While the mutational rates vary across different tumors, only a few represent driver mutations whereas most are passenger mutations that do not provide selective growth advantage. However, it is often challenging to ascertain the clinical significance of driver mutations that are potentially actionable through matching targeted therapies. While potential activation and actionability is often inferred based on the location of alterations (e.g. kinase domains), activation through mutations in non-kinase domains can also be observed (pleckstrin homology in AKT1) [3, 4]. Thus, for actionability, functional characterization of the mutation with alteration at the level of protein is important [5]. However, this information is often not available. Also, presence of concurrent mutations can affect the response to targeted agents; this variability is often not captured in preclinical models. Thus, enrollment into clinical trials with matching targeted therapy is often predicated on limited preclinical data that the specific biomarker or variant predicts for response.

Currently, while there are numerous approved therapies that target gain-of-function oncogenic mutations and are approved for clinical use, there is dearth of therapies that are effective against tumor suppressors. Thus, *TP53* the most commonly mutated cancer gene is currently not actionable. Further, it has been challenging to develop drugs for certain oncogenes due to their structure. The RAS proteins (HRAS, KRAS, NRAS) belonging to the family of small GTPases are among the most commonly mutated oncogenes across multiple tumors [6]. However, efforts to develop direct inhibitors of have not been successful thus far.

To support precision medicine initiatives, the American Society of Clinical Oncology (ASCO) has launched the TAPUR (The Targeted Agent and Profiling Utilization Registry) study to evaluate the safety and efficacy of FDA approved targeted agents in patients with advanced cancers with matching actionable genomic alterations. Through collaborating pharmaceutical entities, the study provides molecularly targeted drugs to discern benefits beyond the approved indication and collects data on outcomes to develop hypotheses for future clinical trials [7].

4.3 Implications of Tumor Heterogeneity

Intratumor heterogeneity is a phenomenon in which somatic alterations vary across different clonal cell populations within the same tumor or across tumors from different foci in the same patient. Recent high-throughput sequencing studies have demonstrated have demonstrated molecular heterogeneity in several tumor types including clear cell renal cell carcinoma, lung adenocarcinoma, and acute myeloid leukemia [8–10]. Gerlinger et al. performed sequencing analysis to highlight the degree of molecular heterogeneity and the limitations of a single representative biopsy [9]. Multiple tissue biopsies from primary and metastatic sites revealed varied tumor sub-clones and mutations suggesting that a biopsy and genomic profile from a single lesion or site may not be representative. The contributions of intratumor heterogeneity towards the natural history of cancers, response to therapy, and risk of relapse / recurrence are areas of on-going investigation. Nonetheless, there are clear hypothetical implications for modern clinical trial design. For example, in cases where a particular biomarker varies spatially within a tumor, clinical studies based on that maker can be susceptible to cross-contamination as a result of tumor sampling biases. That is, whether an individual subject is labeled as biomarker-negative or positive may be influenced by the ability of a biopsy to capture the relevant clonal population within the tumor. Furthermore, the clinical relevance of a genomic alteration (and by proxy the treatment efficacy) can theoretically depend on the abundance of the alteration within the same tumor. Also, pretreatment of tumors with cytotoxic therapies can often lead increased genomic instability and alterations making interpretation of tumor evolution and heterogeneity an arduous task [11]. Despite the prevalence of mutations restricted to specific sites of metastasis, it is expected that targeting the "truncal" mutations would likely lead to significant tumor responses.

Molecular heterogeneity across patients with histopathologically identical tumor types raises important challenges for clinical trial design and targeted therapeutic development. Foremost among these is the observation that when analyzing unselected patient groups, the degree to which a treatment benefit can be detected is highly dependent on the prevalence of the presumed molecular target(s) within the study population. Sleiffer et al. modeled this phenomenon as it pertains to the anti-human epidermal growth factor receptor 2 (*HER2*) agent trastuzumab and its use in breast cancer [2]. In a phase III study of 469 women with advanced *HER2*-positive breast cancer, treatment with trastuzumab plus chemotherapy combination was associated with improved progression-free survival compared to chemotherapy alone (hazard ratio, 0.51, 95% CI, 0.41 to 0.63) leading to FDA approval [12]. Interestingly, it was projected that the clinical study conducted in an unselected population (i.e., both *HER2*-positive and negative tumors) would require at least 2500 patients in order to achieve 90% power with reduction in hazard ratio to 0.9.

Furthermore, at a fixed effect size of HR 0.4, the estimated number of patients required to detect benefit at 90% power varied substantially depending on the prevalence of *HER2*-positivity: from 7540 patients at 10% prevalence to 114 patients at 80% prevalence. Thus, without biomarker selection approval of trastuzumab would not have been feasible. Altogether, clinical trials of anti-*HER2*-directed therapy in breast cancer exemplify a best-case scenario in which the drug mechanism is well known and a biomarker (e.g. *HER2* amplification) exists to reliably identify a patient subpopulation that is predicted to benefit from treatment. As an additional consequence, other subgroups of patients are able to avoid unnecessary treatment and/or harms. However, as the landscape of somatic tumor alterations grows ever complex, there is still a relative deficit of studies that examine the functional consequences of these alterations. Therefore, it seems more likely that investigators will encounter a scenario in which the mechanism of action is only partially defined and therefore the assumption of treatment benefit in various molecular subgroups cannot be made.

Intertumor molecular heterogeneity seems relevant not only to the development of targeted therapies, but also of novel immunotherapies. In a phase II study of 945 patients with unresectable stage III or stage IV melanoma, treatment with the anti-programmed death 1 (PD-1) check point inhibitor nivolumab was associated with improved progression free survival compared to treatment with anti-cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4) inhibitor alone [13]. However, subgroup analyses showed that survival benefit was significantly higher in patients with tumors that had detectable expression of programmed cell death 1 ligand (PD-L1) as compared to those that did not (14.0 months vs. 5.3 months).

4.4 Clinical Trial Paradigm

Advances in tumor genome analysis have created a fundamental shift in clinical trial paradigm. The classical clinical trial eligibility based on histopathology is being phased out in favor of enrollment based on molecular alteration. Due to increasing emphasis on biomarker and patient identification, clinical trials that integrate genomic and proteomic approaches have increased noticeably. Molecular enrichment, in addition to identifying potential responders, helps avoid treatment of patients who are unlikely to be benefit from therapy. While, target enrichment can potentially limit accessible patient population and accrual, due to biomarker selection and matched targeted therapies studies can be designed with large treatment effects to compensate for reduction in numbers enhancing the probability of success of the clinical trial and drug approval.

Recognition of tumor heterogeneity and clonal tumor evolution has led to widespread adoption of enrichment strategies and tumor tissue collection to assess correlatives in clinical trials. Acquired resistance to targeted therapies with prolonged exposure invariably develops in most patients. Although mechanisms of resistance tend to vary across different targeted therapies there are often shared principles. Consequently, most clinical trials mandate pre-treatment biopsies to identify predic-

tive biomarkers and post-progression biopsies after response to identify secondary mechanisms of resistance. This has enabled identification of secondary alterations in the target leading to the development of next generation of agents that are effective against gatekeeper mutations in alterations such as *EGFR* and *ALK* [14, 15]. Alternatively, characterization of alternative pathways for survival that bypasses the target (e.g. *MET* amplification and resistance to gefitinib), have led to efforts on identification of rational targeted therapy combinations to delay or overcome resistance.

Due to biomarker selection and notable tumor responses to matching therapies in susceptible populations, phase-1 trials often provide unequivocal evidence of clinical activity and benefit. Consequently, a large confirmatory randomized clinical may not be obligatory for regulatory approval. Lately, this has led to the accelerated approval of novel targeted therapies based on activity observed in early phase trials [15–17].

4.5 Enrichment and Adaptive Strategies

Genomic analyses have corroborated that cancer is a heterogeneous disease characterized by varying degree of sensitivity and responses to treatment. Thus, it is essential to identify or enrich for subsets of patients who are likely to benefit from treatment. Enrichment is a biomarker-driven strategy that utilizes *a priori* knowledge about the putative target of a study drug to select a subgroup of patients that is hypothesized to benefit (Fig. 4.1a). By definition, this approach depends upon the presence of a biomarker assay that can reliably discriminate between patient populations. Patients who are putatively treatment-sensitive are selected for enrollment, whereas treatment-resistant patients are excluded. Enrichment design offers some advantages. The sample size required for a well-powered study decreases as the sample is enriched for putative responders. Also it avoids exposure to the drug in populations who are unlikely to benefit. Target enrichment has been successfully implemented in several recent trials leading to approval of novel targeted therapies (e.g. vemurafenib for BRAF V600E mutated melanomas, trastuzumab for HER2 amplified breast and gastric cancers). Interestingly, the initial evaluation of endocrine therapy was carried out in an unselected breast cancer population. However, due to the high prevalence of estrogen receptor positivity in breast cancer, clinical activity was discernable even in a non-enriched patient population. On the other hand, use of trastuzumab without enrichment for HER2 positivity, would not have been able to identify significant clinical efficacy.

Adaptive design strategies have become particularly useful in settings where *a priori* knowledge of predictive biomarker(s) for a given agent is either lacking or insufficient at the outset of a study (Fig. 4.1b). For example, a study may begin with an unselected patient population that is randomized into one or more treatment arms. Efficacy data and prospective biopsy data from this stage of analysis can be used to discover novel predictive biomarkers. An enriched phase III cohort can

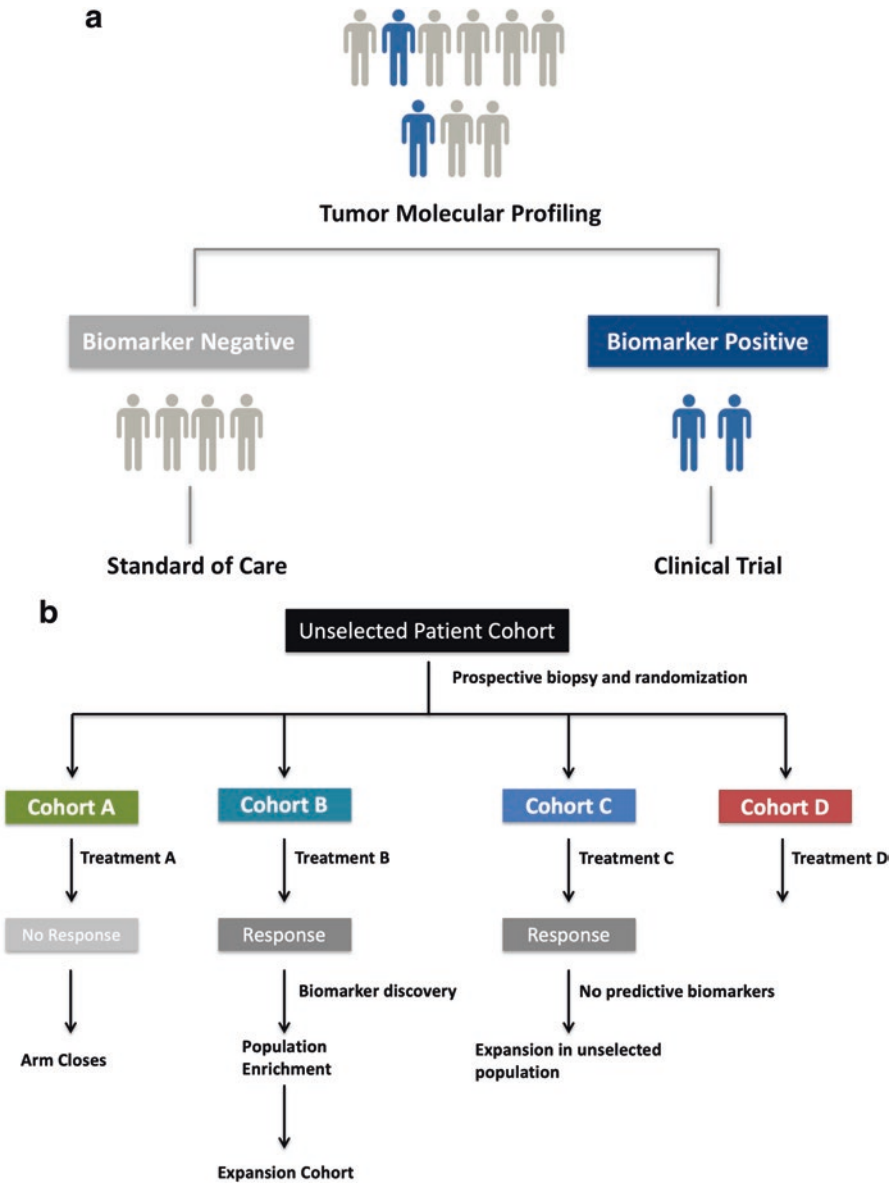

Fig. 4.1 Enrichment and adaptive strategies for precision medicine trials. (**a**) Enrichment based on a prospectively identified biomarker to compare biomarker-guided therapy with conventional therapy (**b**) Biomarker identification and population enrichment for expansion in larger cohort

thereafter be developed for additional testing in the treatment arm. Concurrently, treatment arms that lack promising efficacy data can be closed. Enrichment and adaptive strategies lend an appealing degree of agility to modern drug development

by enabling focused and hypothesis-driven investigation. However, this agility brings some important limitations-namely, that iterative biomarker-based patient selection decreases sample sizes and makes generalizability even more challenging. For common cancers, these limitations are readily overcome by expanding recruitment. For rare cancers, these limitations may make enrichment / adaptive design less appealing.

4.6 Umbrella or Master Trials

As discussed above, enrichment and adaptive-design can help accelerate investigation of individual therapeutic targets. Building upon these advancements, additional strategies such as umbrella and basket trials have emerged to facilitate systematic and large-scale investigation in an increasingly complex genomic landscape.

Umbrella studies provide a biomarker-integrated, histology-dependent framework for clinical investigation of multiple therapeutic targets in a population of patients with identical tumor types. In the typical umbrella study design, standardized biomarker profiling (e.g. multiplex gene panel sequencing) is used to define molecular subgroups of patients within the study population, which are then matched to particular arms of the overall study (Fig. 4.2a). Table 4.1 lists representative examples of the umbrella trial design.

The BATTLE-2 trial (BATTLE-2 Program: A Biomarker-Integrated Targeted Therapy Study in Previously Treated Patients With Advanced Non-Small Cell Lung Cancer) is a phase-2, multi-stage umbrella trial designed to evaluate the effects of targeted therapies in *KRAS*-mutated metastatic non small cell lung cancer (NSCLC). This study highlights some important advantages and limitations with the umbrella design [8]. Patients underwent tumor genome profiling and were randomly assigned to four hypothesis-driven treatment groups: (1) erlotinib, (2) erlotinib and MK-2206 (Akt inhibitor), (3) MK-2206 plus AZD6244 (MEK inhibitor) or (4) sorafenib. In the first stage of the study, the primary end-point (disease control rate (DCR) at 8 weeks) was compared across all four arms, with the first arm (erlotinib alone) serving as the reference group. A second stage was conceived whereby the efficacy data from the first stage would be used to identify predictive biomarkers for the four arms of treatment and therefore enrich their respective patient populations. While the intent of the first half was to perform prospective testing of biomarkers to identify predictive biomarkers to guide patient assignment for the second part, results from the first stage revealed overall modest activity without significant survival difference between the treatment arms yielding new predictive biomarkers. Consequently, stage two of the study was not pursued. Though, patients with mesenchymal tumors had improved survival when treated with MK-2206 and AZD6244 when compared to those with epithelial tumors [18].

I-SPY2 (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis) is an adaptive, phase-2 study of neoadjuvant therapy for high- risk clinical stage II or III breast cancer to identify novel targeted

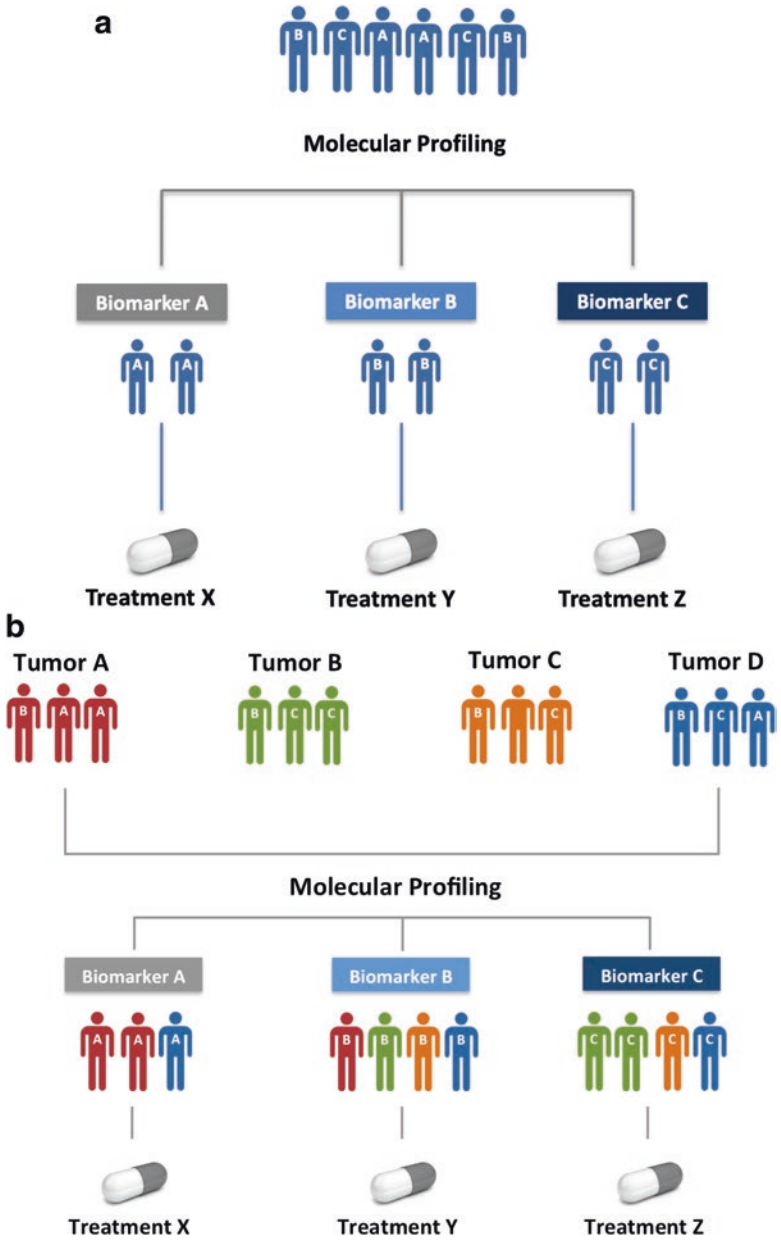

Fig. 4.2 Umbrella and basket trial designs. (**a**) Umbrella or master trial – patients of similar tumor type are screened for biomarkers and allocated to matching targeted therapies. (**b**) Basket trial – patients across multiple tumor types are screened for molecular alteration of interest(s) and treated with matching targeted therapy (tumor agnostic)

Table 4.1 Umbrella trials

Trial	Tumor type	Design	Clinical trial ID	Reference(s)
BATTLE	NSCLC	Phase II/III	NCT00409968	[34]
BATTLE-2	NSCLC	Phase II/III	NCT01248247	[18]
ALCHEMIST	Adenocarcinoma lung	Phase III	NCT02194738	[35]
I-SPY2	Breast	Phase II	NCT01042379	[36]
Lung-MAP	Lung	Phase II	NCT02154490	[21]

therapies in combination with standard chemotherapy to improve rates of pathological complete response. Patients underwent adaptive randomization to either standard chemotherapy with an experimental regimen or standard chemotherapy alone. Ten clinically relevant biomarker signatures, based on HER2 status, hormone-receptor status, and 70-gene profile (MammaPrint, Agendia) were used to assess efficacy. Therapies that meet pre-specified thresholds of efficacy are said to "graduate" for further evaluation. This study has identified two agents, veliparib in combination with carboplatin in triple negative, and neratinib in HER2 positive tumors as promising agents to be studied in phase-3 study [19, 20].

ALCHEMIST (Adjuvant Lung Cancer Enrichment Marker Identification and Sequencing Trial) trial is planned to screen nearly 8000 patients with early stage non-small cell lung cancer to identify patients with alterations in EGFR or ALK genes. Patients with EGFR mutations would be enrolled on a randomized trial evaluating erlotinib, and patients with ALK fusions would be enrolled on a randomized trial evaluating crizotinib. The trial involves genomic analysis of primary tumor and at the time of relapse along with circulating tumor DNA.

The Lung-MAP (Lung Cancer Master Protocol) will screen approximately 1000 patients with advanced squamous lung cancer who have progressed after one line of chemotherapy to identify actionable genomic alterations using a commercial targeted gene panel (Foundation Medicine, Cambridge, MA) [21]. Subsequently, patients would be randomized to one of several phase-2/3 trials evaluating matching targeted therapies with a control arm of standard therapy. Patients without any genomic alterations of interest will be enrolled in the arm testing immunotherapy.

As exemplified by BATTLE-2 or I-SPY2, umbrella trial designs can provide the advantage of exploring multiple lines of novel therapy in parallel across a well-characterized patient population. The design is agile, insofar as individual arms of the study can be selected or halted based upon interim efficacy data, thereby promoting efficient investigation. These features can be particularly useful in uncommon or rare cancer types. However, the logistical requirements of umbrella designs are non-trivial, including large sample size, centralized patient tracking, robust and high-throughput biomarker profiling.

4.7 Basket Trials

Basket trials offer a similar framework for investigating multiple therapeutic targets in parallel, but differ from umbrella design in that it is histology-independent (Fig. 4.2b). Despite the increasingly complex spectrum of genomic alterations, clear patterns are emerging across different cancer types. Recognition of this has given rise to the hypothesis that molecularly defined tumors will respond to targeted therapy against a common vulnerability regardless of tissue of origin. Since many of the actionable alterations typically occur at low frequencies within a single tumor type, accrual of patients in clinical trials becomes a challenge. For instance, in contrast to *ERBB2* amplification, observed in approximately 20% of breast cancers, activating point mutations in *ERBB2* are observed in only 1–2% of breast cancer patients [22]. Preclinical studies showed that these mutations while conferring resistance to lapatinib were sensitive to neratinib, an irreversible *ERBB2* inhibitor [22]. Consequently, a phase 2 study of neratinib, restricted to metastatic *ERBB2*-mutant breast cancer is currently underway [23]. While this standard approach is amenable for a common tumor type like breast cancer, it is not pragmatic for less common tumor types with rare genomic alterations. Thus, basket trials, where accrual is open to patients across multiple tumor types based on a specific genomic alteration (e.g *BRAF, FGFR*) instead of confining to a single tumor type, have been explored. While basket trials facilitate patient accrual, importantly they also help to assess whether different tumor subtypes harboring specific genomic alterations would indeed respond to targeted inhibition. The presence of an actionable alteration does not guarantee response to a matching targeted therapy. For instance, *BRAF* mutant colon cancers do not show sensitivity to BRAF inhibition as melanomas with similar mutation [24]. This was demonstrated to be due to EGFR-mediated reactivation of MAP kinase pathway, and combined EGFR and BRAF inhibition was able to improve response [25]. Representative basket trials that are complete or in progress are listed in Table 4.2.

One of the largest basket trial endeavors launched to date is the NCI-MATCH (National Cancer Institute-Molecular Analysis for Therapy Choice) [26]. The trial was opened in August 2015 with the intention of screening 3000 patients harboring refractory solid tumors or lymphomas for enrollment into one of ten treatment arms based on molecular matching. Each treatment arm focuses on a molecular target (e.g. ALK, BRAF, PIK3CA, ERBB2) for which there was either an FDA-approved agent available or a novel agent identified in early-phase clinical studies (Table 4.3).

Table 4.2 Basket trials

Trial	Design	Clinical trial ID	Reference(s)
NCI-MATCH	Phase II	NCT02465060	[26]
NCI-Pediatric match	Phase II	NA	[26]
NCI-MPACT	Phase II	NCT01827384	[28]
CUSTOM	Phase II	NA	[29]
VE-BASKET	Phase II	NA	[31]

Table 4.3 FDA approved targeted therapies

Gene	Alteration	Cancer	Agents	Biomarker selection required
BCL2	Deletion 17p	CLL	Venetoclax	Yes
BCR-ABL	Rearrangement	CML, ALL	Imatinib, Nilotinib, Dasatinib, Bosutinib, Ponatinib	Yes
BRAF	Mutation	Melanoma	Vemurafenib, Dabrafenib, Trametinib, Cobimetinib	Yes
BRCA1/2	Mutation	Ovary	Olaparib	Yes
CD20	Expression	CLL, FL, NHL	Rituximab, Obinutuzumab, Ofatumumab, Tositumomab, Ibritumomab tiuxetan	No
CD38	Expression	Multiple myeloma	Daratumumab	No
CD52	Expression	CLL	Alemtuzumab	No
CDK4/6	Amplification	Breast	Palbociclib	No
cKIT	Mutation	GIST, Mastocytosis	Imatinib, Sunitinib, Regorafenib	Yes
COL1A1-PDGFB	Rearrangement	DFSP	Imatinib	Yes
EML4-ALK	Rearrangement	Lung	Crizotinib, Ceritinib, Alectinib	Yes
EGFR	Mutation, deletion	Lung	Erlotinib, Gefitinib, Afatinib, Osimertinib	Yes
	Expression	Colon, head and neck, lung	Cetuximab, Panitumumab, Necitumumab	No
ERBB2	Amplification	Breast, Gastric[a]	Trastuzumab, Lapatinib, Pertuzumab, Ado-trastuzumab emtansine	Yes
JAK1/2	Mutation	Myelofibrosis, Polycythemia Vera	Ruxolitinib	No
mTOR	Mutation	Breast, renal	Everolimus, Temsirolimus	No
PDGFRA	Mutation	GIST	Imatinib, Sunitinib	Yes
PML-RARa	Rearrangement	APL	ATRA, arsenic trioxide	Yes
RET	Mutation	Thyroid, lung	Vandetanib, Cabozantinib, Lenvatinib	No
SMO and PTCH1	Mutation	Basal cell	Vismodegib, Sonidegib	No
VEGF/ VEGFR	Expression	Kidney, Colon, lung, gastric, cervix, ovary	Bevacizumab, Ramucirumab, Regorafenib, Ziv-aflibercept, Axitinib, Pazopanib, Sunitinib, Sorafenib	No

ALL acute lymphoblastic leukemia, *APL* Acute promyelocytic leukemia, *ATRA* all-trans retinoic acid, *CDK4/6* cyclin dependent kinase, *CLL* chronic lymphocytic leukemia, *CML* Chronic myelogenous leukemia, *COL1A1* collagen type I alpha 1, *DFSP* dermatofibrosarcoma protuberans, *EGFR* epidermal growth factor receptor, *ERBB2* also known as HER2, *FL* follicular lymphoma, *GIST* gastrointestinal stromal tumor, *mTOR* mammalian target of rapamycin, *NHL* non-hodgkin's lymphoma, *PDGFB* platelet-derived growth factor B, *PDGFR* platelet-derived growth factor receptor, *PTCH1* patched 1, *RET* rearranged during transfection, *SMO* smoothened, *VEGFR* vascular endothelial growth factor

[a]Only trastuzumab is approved for use in gastric cancers

After disease progression, patients who have more than one actionable alteration can potentially be eligible for another study that evaluates matching therapy to the second genomic abnormality.

The trial incorporates pre-treatment and post-treatment biopsies to evaluate mechanisms of acquired resistance. The study was briefly paused in November 2015 for a pre-planned interim analysis, the results of which were reported in April 2016 [27]. Of the 795 patients screened, 87% underwent successful biopsy and molecular profiling, which included colorectal, breast, pancreatic, and neuroendocrine cancers. It is notable that only 9% were found to have an actionable mutation that permitted matching into the one of the ten treatments arms. Nonetheless, the study showed that large-scale, high-throughput molecular profiling was feasible in such a way as to benefit both common and rare cancer types. The study has since re-opened with an expansion to include 14 additional treatment arms. Complementarily, the NCI-Pediatric MATCH will enroll children with advanced cancers who have progressed on standard therapy. As with the NCI-MATCH study, tumor genome sequencing will be used to identify children whose tumors harbor an actionable genomic abnormality and target with approved or investigational agent. This trial is planned to open for accrual in early 2017.

NCI-MPACT (Molecular Profiling-Based Assignment of Cancer Therapy) on the other hand seeks to determine if patients with a mutation in a certain pathway are more likely to benefit from a treatment that targets that pathway, as opposed to another treatment not targeting that pathway [28]. After tumor tissue analysis, eligible patients are randomized 2:1 to a drug that targets the mutated pathway versus a drug that is not known to target the pathway. Patients who received treatment on non-targeted therapy arm will be allowed to cross over to a drug targeting the mutation at progression. Currently agents that target the RAS, PI3K, and DNA repair pathways are being evaluated.

The CUSTOM (Molecular Profiling and Targeted Therapies in Advanced Thoracic Malignancies) trial was designed to evaluate multiple targeted therapies against matching molecular aberrations in NSCLC, small cell lung cancer and thymic tumors [29]. Patients were enrolled into a standard-of-care treatment arm or one of the following five biomarker-enriched treatment groups: erlotinib for EGFR mutations; selumetinib for KRAS, NRAS, HRAS, or BRAF mutations; MK2206 for PIK3CA, AKT, or PTEN alterations; lapatinib for ERBB2 alterations; and sunitinib for KIT or PDGFRA alterations. However, the study was not feasible due to limited patient accrual and lack of adaptive design.

Recently, a phase-2 basket trial of vemurafenib in BRAF V600 mutated non-melanoma cancers, with the exception of papillary thyroid carcinoma and hairy cell leukemia, was reported [30]. Though the study enrolled 122 patients across various tumor cohorts clinically significant activity was observed in NSCLC, Langerhan's cell histiocytosis, and Erdheim-Chester disease, suggesting that *BRAF* is an actionable oncogene only in selected tumors defined by histopathology. Basket trials for genomic alterations in FGFR, ALK, ROS1, and PI3K pathway activation are in progress.

Though basket trials have the potential to accelerate the discovery of targeted therapies it is unlikely that they would lead to direct regulatory drug approval. Due

to genomic variability across tumors, not all mutations are likely to be actionable across all tumors. While, they can help identify tumor type(s) with significant clinical responses to pursue a subsequent biomarker enriched tumor specific clinical trial, interpreting treatment activity based on data from a small number of patients is a major challenge. While the observed responses might be notable, particularly for rare tumors with limited therapeutic options, it is difficult to make definitive inferences on clinical efficacy without a larger study, which may not be possible due to logistic issues. Nevertheless, the adaptive nature of basket trials allows to make rapid adjustments to incorporate emerging preclinical data to explore novel therapeutic combinations [30, 31].

4.8 N of 1 Trial

While this method has been more frequently explored in the treatment of respiratory and musculoskeletal ailments, in oncology, such an approach could be utilized for rare cancers or genomic alterations with very low prevalence when randomized trials are not feasible. The N of 1 trial approach allows patients to serve as their own controls. Patients with tumor genomic analysis are randomly assigned to matching targeted therapies or placebo in alternating sequence with an intervening washout period. Von Hoff et al. evaluated this approach in 86 patients with refractory metastatic cancers who had molecular profiling performed [32]. Twenty seven percent of patients experienced longer progression free survival (PFS) on a regimen based on molecular profile compared to PFS with prior systemic therapy meeting the primary endpoint for the study (PFS ratio > 1.3 in 15% or more patients). Considering that meta-analysis involving over 30,000 patients has demonstrated that a precision medicine guided treatment strategy improves response rates and survival with less toxicity, N of 1 trial approach is rationale and practicable for select tumors [33].

4.9 Conclusion

High-throughput genomic and proteomic techniques have enabled the identification of novel targets and matching therapies. While targeted therapies have transformed the approach to cancer care, their overall utility has been limited due to lack of sustained responses and emergence of acquired resistance. Accordingly, intense efforts have been focused to identify rational targeted therapy combinations to delay or overcome resistance. With improvements in cost, time, and technology, a wider utilization of unbiased tumor profiling methods to increasingly define treatment approaches in future is anticipated. Therapy would be based more on genomic alterations in a subset of patients with defined histopathology. Novel trial designs incorporating innovative statistical approaches will be required to enable treatment decisions based on small patient populations.

References

1. Biankin AV, Piantadosi S, Hollingsworth SJ (2015) Patient-centric trials for therapeutic development in precision oncology. Nature 526:361–370
2. Sleijfer S, Bogaerts J, Siu LL (2013) Designing transformative clinical trials in the cancer genome era. J Clin Oncol 31:1834–1841
3. Carpten JD, Faber AL, Horn C et al (2007) A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. Nature 448:439–444
4. Grabiner BC, Nardi V, Birsoy K et al (2014) A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. Cancer Discov 4:554–563
5. Damodaran S, Miya J, Kautto E et al (2015) Cancer driver log (CanDL): catalog of potentially actionable cancer mutations. J Mol Diagn 17:554–559
6. Downward J (2003) Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer 3:11–22
7. NCT02693535. TAPUR: testing the use of Food and Drug Administration (FDA) approved drugs that target a specific abnormality in a tumor gene in people with advanced stage cancer (TAPUR). https://clinicaltrials.gov/ct2/show/NCT02693535
8. Ding L, Ley TJ, Larson DE et al (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 481:506–510
9. Gerlinger M, Rowan AJ, Horswell S et al (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366:883–892
10. Zhang J, Fujimoto J, Zhang J et al (2014) Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. Science 346:256–259
11. Shyr D, Liu Q (2013) Next generation sequencing in cancer research and clinical application. Biol Proced Online 15:4
12. Slamon DJ, Leyland-Jones B, Shak S et al (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 344:783–792
13. Larkin J, Chiarion-Sileni V, Gonzalez R et al (2015) Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. N Engl J Med 373:23–34
14. Janne PA, Yang JC, Kim DW et al (2015) AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. N Engl J Med 372:1689–1699
15. Shaw AT, Kim DW, Mehra R et al (2014) Ceritinib in ALK-rearranged non-small-cell lung cancer. N Engl J Med 370:1189–1197
16. Kwak EL, Bang YJ, Camidge DR et al (2010) Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. N Engl J Med 363:1693–1703
17. Ledermann J, Harter P, Gourley C et al (2012) Olaparib maintenance therapy in platinum-sensitive relapsed ovarian cancer. N Engl J Med 366:1382–1392
18. Papadimitrakopoulou V, Lee JJ, Wistuba II et al (2016) The BATTLE-2 study: a biomarker-integrated targeted therapy study in previously treated patients with advanced non-small-cell lung cancer. J Clin Oncol 34:3638–3647
19. Park JW, Liu MC, Yee D et al (2016) Adaptive randomization of neratinib in early breast cancer. N Engl J Med 375:11–22
20. Rugo HS, Olopade OI, DeMichele A et al (2016) Adaptive randomization of veliparib-carboplatin treatment in breast cancer. N Engl J Med 375:23–34
21. Herbst RS, Gandara DR, Hirsch FR et al (2015) Lung Master Protocol (Lung-MAP)-a biomarker-driven protocol for accelerating development of therapies for squamous cell lung cancer: SWOG S1400. Clin Cancer Res 21:1514–1524
22. Bose R, Kavuri SM, Searleman AC et al (2013) Activating HER2 mutations in HER2 gene amplification negative breast cancer. Cancer Discov 3:224–237
23. https://clinicaltrials.gov/ct2/show/NCT01670877
24. Prahallad A, Sun C, Huang S et al (2012) Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. Nature 483:100–103

25. Corcoran RB, Ebi H, Turke AB et al (2012) EGFR-mediated re-activation of MAPK signaling contributes to insensitivity of BRAF mutant colorectal cancers to RAF inhibition with vemurafenib. Cancer Discov 2:227–235
26. NCT02465060. NCI-MATCH: Targeted therapy directed by genetic testing in treating patients with advanced refractory solid tumors or lymphomas. https://clinicaltrials.gov/ct2/show/NCT02465060
27. Conley BA, Gray R, Chen A, O'Dwyer P, Arteaga C, Coffey B, Patton D, Li S, McShane LM, Rubinstein L, Comis R, Abrams J, Williams PM, Lih C-J, Hamilton S, Mitchell E, Zwiebel J, Flaherty K, NCI MATCH team (2016) NCI-molecular analysis for therapy choice (NCI-MATCH) clinical trial: interim analysis. In: Proceedings of the 107th Annual Meeting of the American Association for Cancer Research; 2016 Apr 16–20; New Orleans, LA Philadelphia (PA): AACR. Cancer Res 76(14 Suppl):CT101
28. Shivaani Kummar MW, Lih C-J, Chen AP, Rubinstein L, Antony R, Polley E, Zhao Y, Conley BA, Simon R, Doroshow JH (2014) NCI mpact: National Cancer Institute molecular profiling-based assignment of cancer therapy. J Clin Oncol 32(5s suppl):TPS2642
29. Lopez-Chavez A, Thomas A, Rajan A et al (2015) Molecular profiling and targeted therapy for advanced thoracic malignancies: a biomarker-derived, multiarm, multihistology phase II basket trial. J Clin Oncol 33:1000–1007
30. Hyman DM, Puzanov I, Subbiah V et al (2015) Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. N Engl J Med 373:726–736
31. Josep Tabernero EC, Baselga J, Blay J-Y, Chau I, Hyman DM, Raje NS, Wolf J, Sirzen F, Veronese ML, Mitchell L, Hidalgo M (2014) VE-BASKET, a Simon 2-stage adaptive design, phase II, histology-independent study in nonmelanoma solid tumors harboring BRAF V600 mutations (V600m): activity of vemurafenib (VEM) with or without cetuximab (CTX) in colorectal cancer (CRC). J Clin Oncol 32(5s, suppl):3518
32. Von Hoff DD, Stephenson JJ Jr, Rosen P et al (2010) Pilot study using molecular profiling of patients' tumors to find potential targets and select treatments for their refractory cancers. J Clin Oncol 28:4877–4883
33. Schwaederle M, Zhao M, Lee JJ et al (2015) Impact of precision medicine in diverse cancers: a meta-analysis of phase II clinical trials. J Clin Oncol 33:3817–3825
34. Kim ES, Herbst RS, Wistuba II et al (2011) The BATTLE trial: personalizing therapy for lung cancer. Cancer Discov 1:44–53
35. Govindan R, Mandrekar SJ, Gerber DE et al (2015) ALCHEMIST trials: a golden opportunity to transform outcomes in early-stage non-small cell lung cancer. Clin Cancer Res 21:5439–5444
36. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ (2009) I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. Clin Pharmacol Ther 86:97–100

Chapter 5
Resistance to Anti-Cancer Therapeutics

Jennifer A. Woyach

Keywords Drug resistance · Chemotherapy · Targeted therapy · Precision medicine · Mutation · Tumor heterogeneity · DNA damage · Drug efflux

5.1 Introduction

Resistance to anti-cancer agents is an issue that has plagued physicians and scientists since the development of effective anti-cancer drugs. The scope of the problem of resistance is reflected in the consistently high death rate for advanced cancers, even in the era of molecularly targeted agents. Resistance can be differentiated into broad categories in a number of different ways. First, resistance can be either intrinsic to the cancer and present prior to the receipt of therapy, or it can be acquired as an adaptation to a previously effective therapy. While intrinsic resistance is often due to incomplete drug penetration or lack of reliance on the pathway being targeted and often affects multiple classes of drugs, acquired resistance is generally more complex and involves acquisition of specific mutations or specific adaptations of the cancer or the microenvironment to limit the effectiveness of one or a few classes of agents. As our tools to more effectively detect mechanisms of resistance expands, so does the complexity of the mechanisms underlying this issue.

In this review, we will discuss resistance in the context of specific therapies, and divide this broad topic into resistance to chemotherapeutic agents, and resistance to targeted therapies. While many resistance mechanisms affect both chemotherapeutic and targeted agents, as the field moves toward more molecularly targeted therapies, currently understood resistance mechanisms can help anticipate resistance toward new classes of agents and lead to the evolution of rational combination therapies that may prevent or treat resistant disease.

J. A. Woyach (✉)
Division of Hematology, Department of Internal Medicine, The Ohio State University, Columbus, OH, USA
e-mail: Jennifer.woyach@osumc.edu

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_5

5.2 Chemotherapy Resistance

5.2.1 Introduction

Chemotherapy resistance can be intrinsic to tumor cells or can be acquired after exposure to therapy. Unfortunately, numerous mechanisms exist by which cancer cells can become resistant to therapy, and multiple mechanisms can exist in a single patient. Especially in solid tumors, tumor heterogeneity represents a mechanism of resistance in itself, but also can also lead to the evolution of multiple disparate mechanisms that can exert multi-drug resistance. Here I will describe some of the most common resistance mechanisms.

5.2.2 Modulation of Drug Efflux from Cells

The classic mechanism of intrinsic chemotherapy resistance, primary to oral drugs, is through upregulation of drug efflux proteins. These proteins are members of the ATP-binding cassette (ABC) transporter family, and are responsible for drug efflux in normal cells, where they are present in the lining of the GI tract and blood brain barrier (Table 5.1). Multidrug resistance protein 1 (MDR1; also known as permeability glycoprotein [P-gp]), is a member of this family, encoded by the ABCB1 gene. This protein is often upregulated in cancer as an intrinsic abnormality, which can lead to reduced oral bioavailability of anti-cancer agents through excessive drug efflux [1–3]. Upregulation of MDR1 has also been shown to induce CYP3A4 expression which can deactivate some anticancer drugs [4]. Although MDR1 is the primary drug efflux protein that has been found to be upregulated in cancer, other members of the ABC family have been shown to be present in cells that do not express MDR1, including multidrug resistance-associated protein 1 (MRP1), encoded by ABCC1 [5], and mitoxantrone resistance protein (MXR) [6].

In addition to intrinsic upregulation, MDR1 can be upregulated in response to therapy. A variety of drugs used throughout medicine can modulate MDR1, notably to cancer, both tamoxifen [7, 8] and retinoic acid [9] have been shown to upregulate MDR1, which can have negative repercussions for agents given concurrently. Interestingly, this phenomenon can be cell specific, such as in the case of doxorubicin

Table 5.1 ABC family proteins known to be involved in drug resistance

Gene Symbol	Alias	Chromosomal Location
ABCA2	ABC2	9q34
ABCB1	MDR1, PgP	7p21
ABCC1	MRP1	16.13.1
ABCC3	MRP3	17q21.3
ABCG2	ABCP, MXR, BCRP	4q22

in sarcoma, where the drug was shown to induce upregulation of MDR1 in lung tumor cells specifically, but not in normal adjacent lung tissue [10]. Multiple mechanisms likely exist for the activation of MDR1 post-therapy. It has been shown in ovarian cancer that the ABCB1 gene can be upregulated after therapy through acquisition of mutations resulting in gene fusions or translocations in addition to simple point mutations [11]. As well, demethylation of the ABCB1 promoter has been shown to upregulate MDR1, leading to a multi-drug resistant phenotype [12, 13]. This diversity in mechanism of upregulation, and ubiquitous phenomenon of upregulation of MDR1 in response to therapy, has made predicting and overcoming this resistance mechanism extremely difficult.

5.2.3 Intratumoral Heterogeneity and Cancer Stem Cells

All cancers, both solid and liquid, are heterogeneous to some degree, and in general, as cancers are treated, tumor heterogeneity increases. This is due both to natural genetic drift within the unstable tumor genome, and also specific abnormalities that are introduced as the result of agents which damage DNA. Within a tumor, therefore, there will likely exist some cells that have intrinsic drug resistance, and when treated, as the bulk of the tumor dies, the resistant clone can become dominant, leading to secondary drug resistance (Fig. 5.1) [14]. An increasing amount of data suggest that intrinsically resistant cells possess stem cell-like features (termed cancer stem cells) which may aid in the propagation of these resistant clones [15–17]. Through asymmetric division, these stem cells continue to propagate both self-renewing and also more differentiated progeny, which leads to increased heterogeneity with multiple differentiation stages present in a single tumor. After therapy, it has been shown that

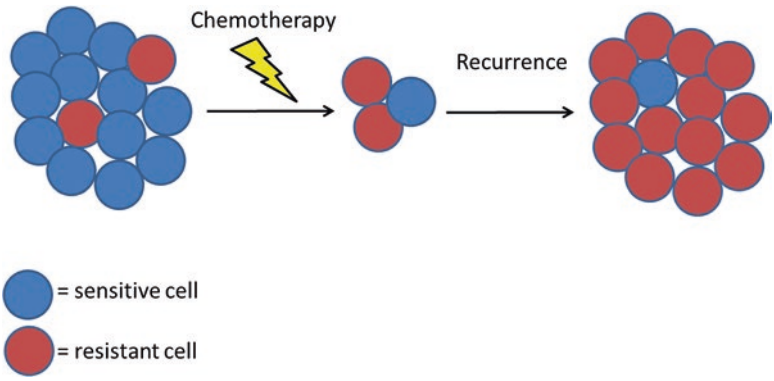

Fig. 5.1 Tumor Heterogeneity Leads to Resistance. Due to tumor heterogeneity, intrinsically chemoresistant cells can be present within a chemosensitive tumor. Following chemotherapy, resistant cells remain, and can lead to relapse

cancer stem cell concentration increases, which both demonstrates the intrinsic resistance of these cells and accounts for rapid growth in relapse [18, 19].

It has been suggested that the difference between cancer stem cells and non-stem cells is attributable to the process of the epithelial to mesenchymal transition (EMT) [20] that defines metastasis in solid tumors. During this transition, epigenetic and later genetic modifications occur, which results in the loss of epithelial characteristics and gain of mesenchymal characteristics, including increased capacity for migration and invasion. This process, which is biased toward cells with stem-cell like characteristics, imparts a high degree of heterogeneity into the tumor, as certain cells are singled out for transition [21]. It has also been shown that cells resistant to therapy tend to have upregulation of genes common in stroma, which suggests that cells undergoing EMT are preferentially resistant to therapy [22].

5.2.4 DNA Damage Checkpoint and Repair Mechanisms

Both normal and neoplastic cells have a complex mechanism to protect against DNA damage, which involves both detection of damage and response to damage including cell cycle arrest, DNA repair, or apoptosis, collectively referred to as DNA damage response (DDR) mechanisms. Mutations of genes involved in detection of DNA damage and DDR response, including ATM, are very common in cancer. When defects in specific DNA repair pathways are present such as this, the cell can be completely dependent upon a backup pathway, which can lead to sensitivity to specific classes of agents, but complete resistance against others. A notable example of this is the cancer susceptibility genes BRCA 1 and 2 are representative genes. Although conferring susceptibility to cancer, by inactivating homologous recombination repair, BRCA 1 and 2 mutations confer relative sensitivity to DNA damaging chemotherapy agents. Conversely, reversion of BRCA1 or 2 to wild type through acquired mutation can be seen in cases of chemotherapy resistance [23–25].

5.2.5 Genomic Complexity and Acquisition of Specific Mutations

Genomic instability, like tumor heterogeneity, is a hallmark of cancer itself, and becomes more prevalent with administration of chemotherapy. Chemotherapy resistance can be mediated by the acquisition of specific mutations, and also by the general acquisition of mutations leading to a more complex genome.

General genomic evolution in cancer cells has been shown in some cases to be driven by chromosomal rearrangements. Unlike in a normal diploid cell, where two alleles would have to be mutated for many events to be physiologically relevant, in cancer cells the rates of mutation and consequences of mutation are much higher,

leading to the theory that large scale genetic changes are catalyzed by aneuploidy. Aneuploidy both can account for the high frequency of non-silent mutations that occurs in cancer cells, and also leads to more genetic changes by creating an unstable spindle apparatus during mitosis [26, 27]. This then can lead to widespread chromosomal changes that have been shown to lead to increasing drug resistance by activating alternative biochemical pathways of survival [28].

Further evidence that widespread genomic changes as the result of catastrophic chromosomic events can lead to drug-resistant phenotypes comes from disease specific examples. The mechanisms of these structural changes include chromothripsis, a one-time catastrophic DNA rearrangement [11, 29, 30], chromoplexy [31], where multiple DNA translocations and deletions arise in an interdependent manner, or variations of these with a less catastrophic phenotype. These mechanisms have been observed in multiple different tumors, and although each is fairly infrequent, together they are a significant mechanism of clonal evolution. In a variety of tumor types, it has been shown that higher genomic complexity is related to drug resistance, and can explain the progression from drug-sensitive to drug-resistant phenotypes in some cases [11, 32].

5.2.6 Specific Gene Alterations

As well as widespread genomic alterations, acquisition of specific mutations can lead to resistance against single or multiple classes of drugs. I will outline some of the best described mutations here.

The tumor suppressor TP53 is the most common mutation in cancer. This gene regulates cell cycle, apoptosis, and metabolism-associated genes. Many mutations are dominant negative loss of function, which eliminates the tumor suppressor phenotype. However, some TP53 mutations can convert p53 from tumor suppressor to oncogene (reviewed in [33]). In a variety of malignancies, including ovarian, breast, and biliary cancers, this has been shown to be associated with resistance to chemotherapy, including microtubule stabilizers [34], platinum agents [35–37], and topoisomerase inhibitors [35, 36]. In some circumstances, resistance is associated with increased transcription of genes commonly associated with chemotherapy resistance such as c-myc or MDR1.

Common antimetabolite agents that target the folate pathway include methotrexate, 5-fluorouricil, and pemetrexed. Methotrexate exerts cytotoxicity via inhibition of dihydrofolate reductase (DHFR), a co-enzyme in DNA methylation. In a variety of cancers, DHFR can be upregulated in the presence of methotrexate to lead to resistance [38]. As well, inactivating mutations of the human reduced folate carrier (RFC) [39–43] and quantitative decrease in gene expression without known inactivating mutations are alternative mechanisms [44]. Similarly, mutations of folylpolygamma-gluatamate synthetase (FPGS) as well as aberrant splicing can decrease the intracellular retention of antifolates and lead to resistance [45–47]. The other antifolate metabolites can share mechanisms of resistance with methotrexate, however,

5-fluorouricil and pemetrexed act primarily through inhibition of thymidylate synthetase (TS) which can also be altered through acquisition of mutations. The most common mechanism of resistance to 5-fluorouricil appears to be amplification of TS [48–52] which can be either an innate or acquired characteristic.

Drugs that inhibit DNA topoisomerase I and II are commonly used drugs in cancer therapy. In vitro, resistance to topoisomerase I inhibitors including topotecan and irinotecan has been linked to decreased expression of topoisomerase I [53] as well as mutations that lead to alteration in structure that disrupts the binding to the drug or alters the linker [54–56]. Similarly, topoisomerase II inhibitors, including doxorubicin and etoposide, can induce resistance through targeted gene alterations that either decrease topoisomerase II levels [57] or mutate binding sites [58, 59]. As well, alterations of genes that regulate topoisomerase II-induced DNA break formation or subsequent DNA repair have been shown to induce resistance in vitro [60].

The taxanes, which include paclitaxel and docetaxel, are microtubule-stabilizing drugs which are the backbone of therapy in a variety of cancers. Mechanisms of resistance to the taxanes have not been fully elucidated, but a variety of mechanisms have been proposed. From a genomic standpoint, alterations in protein kinesins, which work with microtubule associated proteins to direct cytokinesis, have been shown to lead to taxane resistance [61–63]. As well, mutations in beta tubulin have been shown to lead to resistance in a variety of cancers [64–66].

5.2.7 Other Mechanisms of Chemotherapy Resistance

Numerous other mechanisms of chemoresistance have been reported, but are less common than the ones focused on previously. These include upregulation of anti-apoptotic proteins [67, 68], tumor desmoplasia [11, 69] tissue hypoxia [70, 71], and autophagy (reviewed in [72]).

5.3 Resistance to Targeted Therapies

Over the past decade, targeted therapies have moved to the forefront of treatment for many malignancies. While many of these therapies have been paradigm changing for their respective diseases, most cancers are able to eventually become resistant, especially to single agent therapies. This can either be due to mutations acquired during the course of therapy, minor resistant subpopulations that become dominant in the presence of drug, or intrinsic resistance mechanisms that induce primary resistance. Unlike chemotherapy resistance where mechanisms often confer resistance to multiple drugs, with targeted therapies the mechanisms of resistance are often very specific. I will therefore discuss resistance to common individual classes of drugs, noting some mechanisms in common with chemotherapy resistance mechanisms.

5.3.1 BCR-ABL Inhibitors

The inhibitors of BCR-ABL, with the prototypical drug imatinib, were the first targeted agents in wide use, and changed the natural history of chronic myelogenous leukemia (CML). While resistance to imatinib, and the second generation inhibitors nilotinib and dasatinib, are relatively uncommon, the most notable mechanism of resistance to these agents is the T315I mutation in ABL1, which is the gatekeeper residue for the fusion protein that disrupts drug binding [73]. Alternative mutations that alter drug binding through direct inhibition or alteration of the ATP-binding site or the catalytic domain, as well as upregulation of BCR-ABL can also lead to resistance with imatinib [73–77]. Most mutations that confer resistance to imatinib can be overcome by the second generation inhibitors, except the T315I mutation [78, 79]. A third generation inhibitor, ponatinib, is effective even in the presence of T315I [80].

In addition to CML, imatinib is also commonly used to treat GIST tumors and mastocytosis due to inhibition of KIT, and chronic eosinophilic leukemia due to inhibition of PGFRα. Similar to CML, mutation of these targets is the most common mechanism of resistance [81, 82].

Besides mechanisms specific to BCR-ABL, upregulation of the MDR1 gene also confers resistance to imatinib in CML [83, 84].

5.3.2 EGFR Inhibitors

Inhibitors of EGFR have been widely adopted in the treatment of patients with non-small-cell lung cancer with activating mutations in EGFR. Resistance can appear relatively quickly, and the most common mechanism is acquisition of a gatekeeper mutation at T790M [85–87]. This mutation in most cases is a pre-existing minor clone that quickly develops dominance, and has been found in up to 80% of patients pre-treatment [88]. This mutation has even been found in germline, however that is relatively rare [89–91].

Amplification of alternative oncogenic pathways can also drive resistance to EGFR inhibitors. Amplification of the MET proto-oncogene has been observed in about 20% of resistant cases [92–94], and its ligand HGF can promote resistance as well [95]. Both of these alterations lead to maintenance of signaling through ERBB3 and the PI3K pathway.

5.3.3 BRAF-V600E Inhibitors

V600E mutations in BRAF are common in a variety of cancers, including melanoma and hairy cell leukemia. Small molecule inhibitors targeting this abnormality have shown great promise in these diseases, with about an 80% response rate in

melanoma, however resistance develops invariably. There are numerous mechanisms for resistance, most commonly involving re-expression of signaling through the MAPK pathway [96]. BRAF specific abnormalities include amplification of BRAF [97] and aberrant splicing of RAF leading to dimerization in a RAS-independent manner [98]. Non-RAF specific mechanisms are prevalent as well and include MEK1 activation [99], PDGFRB upregulation, N-RAS activating mutations [100], and loss of PTEN which leads to suppression of apoptosis via BIM down-regulation [101].

Mechanisms of resistance to BRAF inhibitors in hairy cell leukemia have not been as clearly elucidated, but likely also involve reactivation of MAPK [102]. As well, in vitro, upregulation of MDR1 has been seen with BRAF inhibition in B cells, which may represent another mechanism [103].

5.3.4 HER2 (ERBB2) Targeted Therapies

Therapies targeting HER2 include antibodies and tyrosine kinase inhibitors and are used in the approximately 25% of breast cancers that overexpress HER2. HER2 inhibition in these cases has been shown to prolong survival, but resistance can develop during the course of initial therapy or in relapsed disease.

The antibody trastuzumab was the first HER2 targeting therapy developed (Fig. 5.2). While responses with combination therapy are excellent, especially in the neoadjuvant and adjuvant setting, response durations to single agent therapy tend to be short. There are several mechanisms of resistance to this drug that have been discovered in the preclinical setting and in many cases clinically as well. One mechanism involves over-expression of other members of the ERbB family, including EGFR and HER3/4. [104–107] As well, activation of PI3K/Akt through loss of PTEN [108, 109] or amplification of other signaling pathways that converge downstream with the HER2 pathway to upregulate PI3K/Akt, such as the IGF-I receptor [110–112] or MET [113], have been shown to be potential mechanisms of resistance. Finally, mutations in HER2 itself which prevent antibody binding have been shown to be potential resistance mechanisms [114].

TKIs which target HER2 have been shown to be effective in patients who are resistant to trastuzumab, and thus have some differential mechanisms of resistance from trastuzumab, although the end result of upregulation of PI3K/AKT tends to be similar in most cases. In general, both proposed and proven mechanisms of resistance include activation of compensatory pathways, HER2 amplification, and HER2 gene mutations. HER2 inhibition by the TKI lapatinib induces upregulation of HER3 in vitro, which as mentioned previously can lead to resistance [115]. As well, similar to trastuzumab, MET can lead to resistance, as well as AXL, which has the potential to activate PI3K and bypass either trastuzumab or lapatinib [116]. Inhibition of IAPs, either directly or via mTOR upregulation has also been shown in vitro to lead to lapatinib resistance [117]. Finally, general kinome reprogramming has been observed in lapatinib resistance, leading to overexpression of multiple compensatory pathways [118].

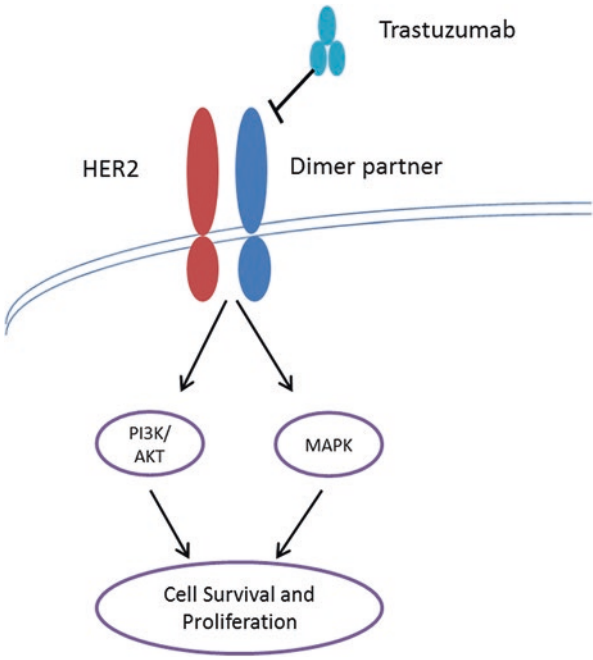

Fig. 5.2 Mechanism of Trastuzumab

5.3.5 ALK Inhibitors

Amplification of and activating mutations in ALK have been found in a number of cancers, and ALK inhibitors have been very effective for these types of tumors. Resistance tends to be mediated through ALK fusion gene amplifications or secondary mutations in ALK [119–121]. As well, activation of bypass signaling pathways including EGFR and c-KIT can be seen [120, 122]. The second generation ALK inhibitor ceritinib has been shown to be effective in patients resistant to crizotinib [123], and in vitro is effective in the presence of most, but not all, crizotinib-resistance mutations [124].

5.3.6 Proteasome Inhibitors

Inhibitors of the proteasome are utilized in a number of malignancies, most notably in multiple myeloma, where bortezomib and carfilzomib are integral components of most treatment regimens. Bortezomib specifically binds to the β5 subunit of the proteasome. Mutations within the PSMβ5 gene which encodes this protein, that impair drug binding, is a common mechanism of resistance in preclinical models [125], however, this has yet to be confirmed in patients. As well, upregulation of PSMβ5 has been associated with resistance in both cellular models and in patients [126]. Similar to other targeted therapies and chemotherapies, IAPs are also

implicated in resistance, where MCL1 overexpression has been linked to resistance to bortezomib [127]. Second generation inhibitors such as carfilzomib are effective in some patients resistant to bortezomib, suggesting that it can overcome some but likely not all resistance mechanisms.

5.3.7 VEGF Inhibitors

Like other targeted therapies, VEGF inhibitors often promote excellent responses but then rapid development of resistance via multiple mechanisms. One common mechanism is upregulation of compensatory angiogenic pathways including epidermal growth factor [128], platelet-derived growth factor [129], and fibroblast growth factors 1 and 2 [130]. As well, the tumor microenvironment can promote both intrinsic and acquired resistance. One mechanism that has been seen in glioblastoma is induction of autophagy that is triggered by the tissue hypoxia that results from anti-angiogenic therapy, and leads to adaptive autophagy through HIF-1α/AMPK pathway signaling [131]. As well, in mouse models, hypoxia after anti-angiogenic therapy was shown to lead to infiltration by myeloid cells and acquisition of stem cell features [132].

5.3.8 BTK Inhibitors

Inhibitors of Bruton's Tyrosine Kinase (BTK) have revolutionized therapy for a number of B cell malignancies including CLL and mantle cell lymphoma. The mechanism of relapse appears to be disease-specific, with relapse in CLL and Waldenstrom's macroglobulinemia are primarily mediated through acquisition of mutations in BTK or its immediate downstream target PLCG2 [133–136]. BTK C481S, the most common acquired mutation in BTK, reduces the binding affinity of ibrutinib for BTK and changes ibrutinib from an irreversible to a reversible inhibitor [133, 137]. The mutations identified in PLCG2 in CLL have all been demonstrated to be potentially gain-of-function, allowing activation through the BCR even in the presence of inactive BTK [133, 138]. Clonal evolution has also been shown to be a hallmark of ibrutinib resistance in CLL. [139] In mantle cell lymphoma, these specific mutations are seen only rarely, and upregulation of PI3 kinase through kinome reprogramming appears to play a major role. [140]

5.4 Conclusions

Drug resistance in cancer therapy is a very complex topic, and there is unfortunately not a simple solution. Understanding resistance mechanisms, however, may lead to development of rational next-line therapies. For example, the use of demethylating

agents in the case of methylation-induced 5-FU resistance, or BCL2 inhibitors after therapies which induce IAPs may selectively target resistant cells. As the importance of repeat tumor sampling becomes more widely accepted, and molecular techniques for detection of resistance mechanisms improve, we have a greater ability to understand resistance mechanisms and even consider pre-emptive combination therapies. Although advanced heterogeneous cancers will likely be incurable for the forseeable future, as the armamentarium of targeted therapies and immune activating therapies increases, it becomes more of a realistic possibility that we can successfully treat advanced cancers as a chronic disease and greatly extend quality life for our patients.

References

1. Ueda K, Cardarelli C, Gottesman MM, Pastan I (1987) Expression of a full-length cDNA for the human "MDR1" gene confers resistance to colchicine, doxorubicin, and vinblastine. Proc Natl Acad Sci U S A 84(9):3004–3008
2. Fojo AT, Ueda K, Slamon DJ, Poplack DG, Gottesman MM, Pastan I (1987) Expression of a multidrug-resistance gene in human tumors and tissues. Proc Natl Acad Sci U S A 84(1):265–269
3. Cordon-Cardo C, O'Brien JP, Casals D, Rittman-Grauer L, Biedler JL, Melamed MR et al (1989) Multidrug-resistance gene (P-glycoprotein) is expressed by endothelial cells at blood-brain barrier sites. Proc Natl Acad Sci U S A 86(2):695–698
4. Schuetz EG, Schinkel AH, Relling MV, Schuetz JD (1996) P-glycoprotein: a major determinant of rifampicin-inducible expression of cytochrome P4503A in mice and humans. Proc Natl Acad Sci U S A 93(9):4001–4005
5. Cole SP, Bhardwaj G, Gerlach JH, Mackie JE, Grant CE, Almquist KC et al (1992) Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. Science 258(5088):1650–1654
6. Sauna ZE, Smith MM, Muller M, Kerr KM, Ambudkar SV (2001) The mechanism of action of multidrug-resistance-linked P-glycoprotein. J Bioenerg Biomembr 33(6):481–491
7. Callaghan R, Higgins CF (1995) Interaction of tamoxifen with the multidrug resistance P-glycoprotein. Br J Cancer 71(2):294–299
8. Chen J, Balmaceda C, Bruce JN, Sisti MB, Huang M, Cheung YK et al (2006) Tamoxifen paradoxically decreases paclitaxel deposition into cerebrospinal fluid of brain tumor patients. J Neuro-Oncol 76(1):85–92
9. El Hafny B, Chappey O, Piciotti M, Debray M, Boval B, Roux F (1997) Modulation of P-glycoprotein activity by glial factors and retinoic acid in an immortalized rat brain microvessel endothelial cell line. Neurosci Lett 236(2):107–111
10. Abolhoda A, Wilson AE, Ross H, Danenberg PV, Burt M, Scotto KW (1999) Rapid activation of MDR1 gene expression in human metastatic sarcoma after in vivo exposure to doxorubicin. Clin Cancer Res 5(11):3352–3356
11. Patch AM, Christie EL, Etemadmoghadam D, Garsed DW, George J, Fereday S et al (2015) Whole-genome characterization of chemoresistant ovarian cancer. Nature 521(7553):489–494
12. Baker EK, El-Osta A (2003) The rise of DNA methylation and the importance of chromatin on multidrug resistance in cancer. Exp Cell Res 290(2):177–194
13. Kantharidis P, El-Osta A, de Silva M, Wall DM, Hu XF, Slater A et al (1997) Altered methylation of the human MDR1 promoter is associated with acquired multidrug resistance. Clin Cancer Res 3(11):2025–2032

14. Navin N, Krasnitz A, Rodgers L, Cook K, Meth J, Kendall J et al (2010) Inferring tumor progression from genomic heterogeneity. Genome Res 20(1):68–80
15. Chaffer CL, Brueckmann I, Scheel C, Kaestli AJ, Wiggins PA, Rodrigues LO et al (2011) Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. Proc Natl Acad Sci U S A 108(19):7950–7955
16. Alison MR, Islam S, Wright NA (2010) Stem cells in cancer: instigators and propagators? J Cell Sci 123(Pt 14):2357–2368
17. Ailles LE, Weissman IL (2007) Cancer stem cells in solid tumors. Curr Opin Biotechnol 18(5):460–466
18. Dallas NA, Xia L, Fan F, Gray MJ, Gaur P, van Buren G 2nd, et al. (2009) Chemoresistant colorectal cancer cells, the cancer stem cell phenotype, and increased sensitivity to insulin-like growth factor-I receptor inhibition. Cancer Res 69(5):1951–1957
19. Levina V, Marrangoni AM, DeMarco R, Gorelik E, Lokshin AE (2008) Drug-selected human lung cancer stem cells: cytokine network, tumorigenic and metastatic properties. PLoS One 3(8):e3077
20. Polyak K, Weinberg RA (2009) Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat Rev Cancer 9(4):265–273
21. Singh A, Settleman J (2010) EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. Oncogene 29(34):4741–4751
22. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V et al (2009) A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Med 15(1):68–74
23. Edwards SL, Brough R, Lord CJ, Natrajan R, Vatcheva R, Levine DA et al (2008) Resistance to therapy caused by intragenic deletion in BRCA2. Nature 451(7182):1111–1115
24. Sakai W, Swisher EM, Karlan BY, Agarwal MK, Higgins J, Friedman C et al (2008) Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. Nature 451(7182):1116–1120
25. Norquist B, Wurz KA, Pennil CC, Garcia R, Gross J, Sakai W et al (2011) Secondary somatic mutations restoring BRCA1/2 predict chemotherapy resistance in hereditary ovarian carcinomas. J Clin Oncol 29(22):3008–3015
26. Li R, Yerganian G, Duesberg P, Kraemer A, Willer A, Rausch C et al (1997) Aneuploidy correlated 100% with chemical transformation of Chinese hamster cells. Proc Natl Acad Sci U S A 94(26):14506–14511
27. Duesberg P, Rausch C, Rasnick D, Hehlmann R (1998) Genetic instability of cancer cells is proportional to their degree of aneuploidy. Proc Natl Acad Sci U S A 95(23):13692–13697
28. Duesberg P, Stindl R, Hehlmann R (2000) Explaining the high mutation rates of cancer cells to drug and multidrug resistance by chromosome reassortments that are catalyzed by aneuploidy. Proc Natl Acad Sci U S A 97(26):14295–14300
29. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ et al (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144(1):27–40
30. Rausch T, Jones DT, Zapatka M, Stutz AM, Zichner T, Weischenfeldt J et al (2012) Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell 148(1–2):59–71
31. Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y et al (2013) Punctuated evolution of prostate cancer genomes. Cell 153(3):666–677
32. Campbell PJ, Yachida S, Mudie LJ, Stephens PJ, Pleasance ED, Stebbings LA et al (2010) The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature 467(7319):1109–1113
33. Brachova P, Thiel KW, Leslie KK (2013) The consequence of oncomorphic TP53 mutations in ovarian cancer. Int J Mol Sci 14(9):19257–19275
34. Seagle BL, Yang CP, Eng KH, Dandapani M, Odunsi-Akanji O, Goldberg GL et al (2015) TP53 hot spot mutations in ovarian cancer: selective resistance to microtubule stabilizers in vitro and differential survival outcomes from the cancer genome atlas. Gynecol Oncol 138(1):159–164

35. Di Agostino S, Cortese G, Monti O, Dell'Orso S, Sacchi A, Eisenstein M et al (2008) The disruption of the protein complex mutantp53/p73 increases selectively the response of tumor cells to anticancer drugs. Cell Cycle 7(21):3440–3447
36. Chrisanthar R, Knappskog S, Lokkevik E, Anker G, Ostenstad B, Lundgren S et al (2011) Predictive and prognostic impact of TP53 mutations and MDM2 promoter genotype in primary breast cancer patients treated with epirubicin or paclitaxel. PLoS One 6(4):e19249
37. Ahn DH, Javle M, Ahn CW, Jain A, Mikhail S, Noonan AM et al (2016) Next-generation sequencing survey of biliary tract cancer reveals the association between tumor somatic variants and chemotherapy resistance. Cancer 122(23):3657–3666
38. Curt GA, Carney DN, Cowan KH, Jolivet J, Bailey BD, Drake JC et al (1983) Unstable methotrexate resistance in human small-cell carcinoma associated with double minute chromosomes. N Engl J Med 308(4):199–202
39. Gong M, Yess J, Connolly T, Ivy SP, Ohnuma T, Cowan KH et al (1997) Molecular mechanism of antifolate transport-deficiency in a methotrexate-resistant MOLT-3 human leukemia cell line. Blood 89(7):2494–2499
40. Jansen G, Mauritz R, Drori S, Sprecher H, Kathmann I, Bunni M et al (1998) A structurally altered human reduced folate carrier with increased folic acid transport mediates a novel mechanism of antifolate resistance. J Biol Chem 273(46):30189–30198
41. Wong SC, Zhang L, Witt TL, Proefke SA, Bhushan A, Matherly LH (1999) Impaired membrane transport in methotrexate-resistant CCRF-CEM cells involves early translation termination and increased turnover of a mutant reduced folate carrier. J Biol Chem 274(15):10388–10394
42. Drori S, Jansen G, Mauritz R, Peters GJ, Assaraf YG (2000) Clustering of mutations in the first transmembrane domain of the human reduced folate carrier in GW1843U89-resistant leukemia cells with impaired antifolate transport and augmented folate uptake. J Biol Chem 275(40):30855–30863
43. Rothem L, Ifergan I, Kaufman Y, Priest DG, Jansen G, Assaraf YG (2002) Resistance to multiple novel antifolates is mediated via defective drug transport resulting from clustered mutations in the reduced folate carrier gene in human leukaemia cell lines. Biochem J 367(Pt 3):741–750
44. Rothem L, Aronheim A, Assaraf YG (2003) Alterations in the expression of transcription factors and the reduced folate carrier as a novel mechanism of antifolate resistance in human leukemia cells. J Biol Chem 278(11):8935–8941
45. Stark M, Wichman C, Avivi I, Assaraf YG (2009) Aberrant splicing of folylpolyglutamate synthetase as a novel mechanism of antifolate resistance in leukemia. Blood 113(18):4362–4369
46. Liani E, Rothem L, Bunni MA, Smith CA, Jansen G, Assaraf YG (2003) Loss of folylpoly-gamma-glutamate synthetase activity is a dominant mechanism of resistance to polyglutamylation-dependent novel antifolates in multiple human leukemia sublines. Int J Cancer 103(5):587–599
47. Mauritz R, Peters GJ, Priest DG, Assaraf YG, Drori S, Kathmann I et al (2002) Multiple mechanisms of resistance to methotrexate and novel antifolates in human CCRF-CEM leukemia cells and their implications for folate homeostasis. Biochem Pharmacol 63(2):105–115
48. Johnston PG, Lenz HJ, Leichman CG, Danenberg KD, Allegra CJ, Danenberg PV et al (1995) Thymidylate synthase gene and protein expression correlate and are associated with response to 5-fluorouracil in human colorectal and gastric tumors. Cancer Res 55(7):1407–1412
49. Rooney PH, Stevenson DA, Marsh S, Johnston PG, Haites NE, Cassidy J et al (1998) Comparative genomic hybridization analysis of chromosomal alterations induced by the development of resistance to thymidylate synthase inhibitors. Cancer Res 58(22):5042–5045
50. Ooyama A, Okayama Y, Takechi T, Sugimoto Y, Oka T, Fukushima M (2007) Genome-wide screening of loci associated with drug resistance to 5-fluorouracil-based drugs. Cancer Sci 98(4):577–583
51. Wang TL, Diaz LA Jr, Romans K, Bardelli A, Saha S, Galizia G et al (2004) Digital karyotyping identifies thymidylate synthase amplification as a mechanism of resistance to 5-fluorouracil in metastatic colorectal cancer patients. Proc Natl Acad Sci U S A 101(9):3089–3094

52. Watson RG, Muhale F, Thorne LB, Yu J, O'Neil BH, Hoskins JM et al (2010) Amplification of thymidylate synthetase in metastatic colorectal cancer patients pretreated with 5-fluorouracil-based chemotherapy. Eur J Cancer 46(18):3358–3364
53. Sugimoto Y, Tsukahara S, Oh-hara T, Isoe T, Tsuruo T (1990) Decreased expression of DNA topoisomerase I in camptothecin-resistant tumor cell lines as determined by a monoclonal antibody. Cancer Res 50(21):6925–6930
54. Fiorani P, Bruselles A, Falconi M, Chillemi G, Desideri A, Benedetti P (2003) Single mutation in the linker domain confers protein flexibility and camptothecin resistance to human topoisomerase I. J Biol Chem 278(44):43268–43275
55. Gongora C, Vezzio-Vie N, Tuduri S, Denis V, Causse A, Auzanneau C et al (2011) New topoisomerase I mutations are associated with resistance to camptothecin. Mol Cancer 10:64
56. Tesauro C, della Rocca BM, Ottaviani A, Coletta A, Zuccaro L, Arno B, et al. Molecular mechanism of the camptothecin resistance of Glu710Gly topoisomerase IB mutant analyzed in vitro and in silico. *Mol Cancer.* 2013;12(1):100
57. Burgess DJ, Doles J, Zender L, Xue W, Ma B, McCombie WR et al (2008) Topoisomerase levels determine chemotherapy response in vitro and in vivo. Proc Natl Acad Sci U S A 105(26):9053–9058
58. Zwelling LA, Hinds M, Chan D, Mayes J, Sie KL, Parker E et al (1989) Characterization of an amsacrine-resistant line of human leukemia cells. Evidence for a drug-resistant form of topoisomerase II. J Biol Chem 264(28):16411–16420
59. Hinds M, Deisseroth K, Mayes J, Altschuler E, Jansen R, Ledley FD et al (1991) Identification of a point mutation in the topoisomerase II gene from a human leukemia cell line containing an amsacrine-resistant form of topoisomerase II. Cancer Res 51(17):4729–4731
60. Wijdeven RH, Pang B, van der Zanden SY, Qiao X, Blomen V, Hoogstraat M et al (2015) Genome-wide identification and characterization of novel factors conferring resistance to topoisomerase II poisons in cancer. Cancer Res 75(19):4176–4187
61. Theriault BL, Pajovic S, Bernardini MQ, Shaw PA, Gallie BL (2012) Kinesin family member 14: an independent prognostic marker and potential therapeutic target for ovarian cancer. Int J Cancer 130(8):1844–1854
62. Ehrlichova M, Mohelnikova-Duchonova B, Hrdy J, Brynychova V, Mrhalova M, Kodet R et al (2013) The association of taxane resistance genes with the clinical course of ovarian carcinoma. Genomics 102(2):96–101
63. Tan MH, De S, Bebek G, Orloff MS, Wesolowski R, Downs-Kelly E et al (2012) Specific kinesin expression profiles associated with taxane resistance in basal-like breast cancer. Breast Cancer Res Treat 131(3):849–858
64. Cabral F (2001) Factors determining cellular mechanisms of resistance to antimitotic drugs. Drug Resist Updat 4(1):3–8
65. Huzil JT, Chen K, Kurgan L, Tuszynski JA (2007) The roles of beta-tubulin mutations and isotype expression in acquired drug resistance. Cancer Informat 3:159–181
66. Yin S, Zeng C, Hari M, Cabral F (2012) Random mutagenesis of beta-tubulin defines a set of dispersed mutations that confer paclitaxel resistance. Pharm Res 29(11):2994–3006
67. Liu G, Yuan X, Zeng Z, Tunici P, Ng H, Abdulkadir IR et al (2006) Analysis of gene expression and chemoresistance of CD133+ cancer stem cells in glioblastoma. Mol Cancer 5:67
68. Kitada S, Andersen J, Akar S, Zapata JM, Takayama S, Krajewski S et al (1998) Expression of apoptosis-regulating proteins in chronic lymphocytic leukemia: correlations with in vitro and in vivo chemoresponses. Blood 91(9):3379–3389
69. Olive KP, Jacobetz MA, Davidson CJ, Gopinathan A, McIntyre D, Honess D et al (2009) Inhibition of hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. Science 324(5933):1457–1461
70. Sakata K, Kwok TT, Murphy BJ, Laderoute KR, Gordon GR, Sutherland RM (1991) Hypoxia-induced drug resistance: comparison to P-glycoprotein-associated drug resistance. Br J Cancer 64(5):809–814

71. Pouyssegur J, Dayan F, Mazure NM (2006) Hypoxia signalling in cancer and approaches to enforce tumour regression. Nature 441(7092):437–443
72. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG (2013) Cancer drug resistance: an evolving paradigm. Nat Rev Cancer 13(10):714–726
73. Gorre ME, Mohammed M, Ellwood K, Hsu N, Paquette R, Rao PN et al (2001) Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science 293(5531):876–880
74. le Coutre P, Tassi E, Varella-Garcia M, Barni R, Mologni L, Cabrita G et al (2000) Induction of resistance to the Abelson inhibitor STI571 in human leukemic cells through gene amplification. Blood 95(5):1758–1766
75. Branford S, Rudzki Z, Walsh S, Grigg A, Arthur C, Taylor K et al (2002) High frequency of point mutations clustered within the adenosine triphosphate-binding region of BCR/ABL in patients with chronic myeloid leukemia or Ph-positive acute lymphoblastic leukemia who develop imatinib (STI571) resistance. Blood 99(9):3472–3475
76. Roche-Lestienne C, Soenen-Cornu V, Grardel-Duflos N, Lai JL, Philippe N, Facon T et al (2002) Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. Blood 100(3):1014–1018
77. Roumiantsev S, Shah NP, Gorre ME, Nicoll J, Brasher BB, Sawyers CL et al (2002) Clinical resistance to the kinase inhibitor STI-571 in chronic myeloid leukemia by mutation of Tyr-253 in the Abl kinase domain P-loop. Proc Natl Acad Sci U S A 99(16):10700–10705
78. Weisberg E, Manley PW, Breitenstein W, Bruggen J, Cowan-Jacob SW, Ray A et al (2005) Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. Cancer Cell 7(2):129–141
79. Shah NP, Tran C, Lee FY, Chen P, Norris D, Sawyers CL (2004) Overriding imatinib resistance with a novel ABL kinase inhibitor. Science 305(5682):399–401
80. O'Hare T, Shakespeare WC, Zhu X, Eide CA, Rivera VM, Wang F et al (2009) AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance. Cancer Cell 16(5):401–412
81. Heinrich MC, Corless CL, Blanke CD, Demetri GD, Joensuu H, Roberts PJ et al (2006) Molecular correlates of imatinib resistance in gastrointestinal stromal tumors. J Clin Oncol 24(29):4764–4774
82. Wardelmann E, Merkelbach-Bruse S, Pauls K, Thomas N, Schildhaus HU, Heinicke T et al (2006) Polyclonal evolution of multiple secondary KIT mutations in gastrointestinal stromal tumors under treatment with imatinib mesylate. Clin Cancer Res 12(6):1743–1749
83. Mahon FX, Belloc F, Lagarde V, Chollet C, Moreau-Gaudry F, Reiffers J et al (2003) MDR1 gene overexpression confers resistance to imatinib mesylate in leukemia cell line models. Blood 101(6):2368–2373
84. Burger H, van Tol H, Brok M, Wiemer EA, de Bruijn EA, Guetens G et al (2005) Chronic imatinib mesylate exposure leads to reduced intracellular drug accumulation by induction of the ABCG2 (BCRP) and ABCB1 (MDR1) drug transport pumps. Cancer Biol Ther 4(7):747–752
85. Kobayashi S, Boggon TJ, Dayaram T, Janne PA, Kocher O, Meyerson M et al (2005) EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. N Engl J Med 352(8):786–792
86. Pao W, Miller VA, Politi KA, Riely GJ, Somwar R, Zakowski MF et al (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. PLoS Med 2(3):e73
87. Inukai M, Toyooka S, Ito S, Asano H, Ichihara S, Soh J et al (2006) Presence of epidermal growth factor receptor gene T790M mutation as a minor clone in non-small cell lung cancer. Cancer Res 66(16):7854–7858

88. Watanabe M, Kawaguchi T, Isa S, Ando M, Tamiya A, Kubo A et al (2015) Ultra-sensitive detection of the pretreatment EGFR T790M mutation in non-small cell lung cancer patients with an EGFR-activating mutation using droplet digital PCR. Clin Cancer Res 21(15):3552–3560
89. Bell DW, Gore I, Okimoto RA, Godin-Heymann N, Sordella R, Mulloy R et al (2005) Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. Nat Genet 37(12):1315–1316
90. Oxnard GR, Miller VA, Robson ME, Azzoli CG, Pao W, Ladanyi M et al (2012) Screening for germline EGFR T790M mutations through lung cancer genotyping. J Thorac Oncol 7(6):1049–1052
91. Yu HA, Arcila ME, Harlan Fleischut M, Stadler Z, Ladanyi M, Berger MF et al (2014) Germline EGFR T790M mutation found in multiple members of a familial cohort. J Thorac Oncol 9(4):554–558
92. Bean J, Brennan C, Shih JY, Riely G, Viale A, Wang L et al (2007) MET amplification occurs with or without T790M mutations in EGFR mutant lung tumors with acquired resistance to gefitinib or erlotinib. Proc Natl Acad Sci U S A 104(52):20932–20937
93. Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO et al (2007) MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. Science 316(5827):1039–1043
94. Turke AB, Zejnullahu K, Wu YL, Song Y, Dias-Santagata D, Lifshits E et al (2010) Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer Cell 17(1):77–88
95. Yano S, Wang W, Li Q, Matsumoto K, Sakurama H, Nakamura T et al (2008) Hepatocyte growth factor induces gefitinib resistance of lung adenocarcinoma with epidermal growth factor receptor-activating mutations. Cancer Res 68(22):9479–9487
96. Shi H, Hugo W, Kong X, Hong A, Koya RC, Moriceau G et al (2014) Acquired resistance and clonal evolution in melanoma during BRAF inhibitor therapy. Cancer Discov 4(1):80–93
97. Shi H, Moriceau G, Kong X, Lee MK, Lee H, Koya RC et al (2012) Melanoma whole-exome sequencing identifies (V600E)B-RAF amplification-mediated acquired B-RAF inhibitor resistance. Nat Commun 3:724
98. Poulikakos PI, Persaud Y, Janakiraman M, Kong X, Ng C, Moriceau G et al (2011) RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E). Nature 480(7377):387–390
99. Wagle N, Emery C, Berger MF, Davis MJ, Sawyer A, Pochanard P et al (2011) Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. J Clin Oncol 29(22):3085–3096
100. Nazarian R, Shi H, Wang Q, Kong X, Koya RC, Lee H et al (2010) Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. Nature 468(7326):973–977
101. Paraiso KH, Xiang Y, Rebecca VW, Abel EV, Chen YA, Munko AC et al (2011) PTEN loss confers BRAF inhibitor resistance to melanoma cells through the suppression of BIM expression. Cancer Res 71(7):2750–2760
102. Tiacci E, Park JH, De Carolis L, Chung SS, Broccoli A, Scott S et al (2015) Targeting mutant BRAF in relapsed or refractory hairy-cell leukemia. N Engl J Med 373(18):1733–1747
103. Tsai YT, Lozanski G, Lehman A, Sass EJ, Hertlein E, Salunke SB et al (2015) BRAFV600E induces ABCB1/P-glycoprotein expression and drug resistance in B-cells via AP-1 activation. Leuk Res 39:1270–1277
104. Wang SE, Xiang B, Guix M, Olivares MG, Parker J, Chung CH et al (2008) Transforming growth factor beta engages TACE and ErbB3 to activate phosphatidylinositol-3 kinase/Akt in ErbB2-overexpressing breast cancer and desensitizes cells to trastuzumab. Mol Cell Biol 28(18):5605–5620
105. Ritter CA, Perez-Torres M, Rinehart C, Guix M, Dugger T, Engelman JA et al (2007) Human breast cancer cells selected for resistance to trastuzumab in vivo overexpress epidermal growth factor receptor and ErbB ligands and remain dependent on the ErbB receptor network. Clin Cancer Res 13(16):4909–4919

106. Motoyama AB, Hynes NE, Lane HA (2002) The efficacy of ErbB receptor-targeted anticancer therapeutics is influenced by the availability of epidermal growth factor-related peptides. Cancer Res 62(11):3151–3158
107. Moulder SL, Yakes FM, Muthuswamy SK, Bianco R, Simpson JF, Arteaga CL (2001) Epidermal growth factor receptor (HER1) tyrosine kinase inhibitor ZD1839 (Iressa) inhibits HER2/neu (erbB2)-overexpressing breast cancer cells in vitro and in vivo. Cancer Res 61(24):8887–8895
108. Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hijmans EM, Beelen K et al (2007) A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. Cancer Cell 12(4):395–402
109. Nagata Y, Lan KH, Zhou X, Tan M, Esteva FJ, Sahin AA et al (2004) PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. Cancer Cell 6(2):117–127
110. Harris LN, You F, Schnitt SJ, Witkiewicz A, Lu X, Sgroi D et al (2007) Predictors of resistance to preoperative trastuzumab and vinorelbine for HER2-positive early breast cancer. Clin Cancer Res 13(4):1198–1207
111. Lu Y, Zi X, Zhao Y, Mascarenhas D, Pollak M (2001) Insulin-like growth factor-I receptor signaling and resistance to trastuzumab (Herceptin). J Natl Cancer Inst 93(24):1852–1857
112. Nahta R, Yuan LX, Zhang B, Kobayashi R, Esteva FJ (2005) Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells. Cancer Res 65(23):11118–11128
113. Shattuck DL, Miller JK, Carraway KL 3rd, Sweeney C (2008) Met receptor contributes to trastuzumab resistance of Her2-overexpressing breast cancer cells. Cancer Res 68(5):1471–1477
114. Scaltriti M, Rojo F, Ocana A, Anido J, Guzman M, Cortes J et al (2007) Expression of p95HER2, a truncated form of the HER2 receptor, and response to anti-HER2 therapies in breast cancer. J Natl Cancer Inst 99(8):628–638
115. Garrett JT, Olivares MG, Rinehart C, Granja-Ingram ND, Sanchez V, Chakrabarty A et al (2011) Transcriptional and posttranslational up-regulation of HER3 (ErbB3) compensates for inhibition of the HER2 tyrosine kinase. Proc Natl Acad Sci U S A 108(12):5021–5026
116. Liu L, Greger J, Shi H, Liu Y, Greshock J, Annan R et al (2009) Novel mechanism of lapatinib resistance in HER2-positive breast tumor cells: activation of AXL. Cancer Res 69(17):6871–6878
117. Brady SW, Zhang J, Tsai MH, Yu D (2015) PI3K-independent mTOR activation promotes lapatinib resistance and IAP expression that can be effectively reversed by mTOR and Hsp90 inhibition. Cancer Biol Ther 16(3):402–411
118. Stuhlmiller TJ, Miller SM, Zawistowski JS, Nakamura K, Beltran AS, Duncan JS et al (2015) Inhibition of Lapatinib-induced Kinome reprogramming in ERBB2-positive breast cancer by targeting BET family Bromodomains. Cell Rep 11(3):390–404
119. Doebele RC, Pilling AB, Aisner DL, Kutateladze TG, Le AT, Weickhardt AJ et al (2012) Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer. Clin Cancer Res 18(5):1472–1482
120. Katayama R, Shaw AT, Khan TM, Mino-Kenudson M, Solomon BJ, Halmos B et al (2012) Mechanisms of acquired crizotinib resistance in ALK-rearranged lung Cancers. *Sci Transl Med* 4(120):120ra117
121. Choi YL, Soda M, Yamashita Y, Ueno T, Takashima J, Nakajima T et al (2010) EML4-ALK mutations in lung cancer that confer resistance to ALK inhibitors. N Engl J Med 363(18):1734–1739
122. Sasaki T, Koivunen J, Ogino A, Yanagita M, Nikiforow S, Zheng W et al (2011) A novel ALK secondary mutation and EGFR signaling cause resistance to ALK kinase inhibitors. Cancer Res 71(18):6051–6060
123. Shaw AT, Kim DW, Mehra R, Tan DS, Felip E, Chow LQ et al (2014) Ceritinib in ALK-rearranged non-small-cell lung cancer. N Engl J Med 370(13):1189–1197

124. Friboulet L, Li N, Katayama R, Lee CC, Gainor JF, Crystal AS et al (2014) The ALK inhibitor ceritinib overcomes crizotinib resistance in non-small cell lung cancer. Cancer Discov 4(6):662–673
125. Franke NE, Niewerth D, Assaraf YG, van Meerloo J, Vojtekova K, van Zantwijk CH et al (2012) Impaired bortezomib binding to mutant beta5 subunit of the proteasome is the underlying basis for bortezomib resistance in leukemia cells. Leukemia 26(4):757–768
126. Shuqing L, Jianmin Y, Chongmei H, Hui C, Wang J (2011) Upregulated expression of the PSMB5 gene may contribute to drug resistance in patient with multiple myeloma when treated with bortezomib-based regimen. Exp Hematol 39(12):1117–1118
127. Nencioni A, Hua F, Dillon CP, Yokoo R, Scheiermann C, Cardone MH et al (2005) Evidence for a protective role of Mcl-1 in proteasome inhibitor-induced apoptosis. Blood 105(8):3255–3262
128. Cascone T, Herynk MH, Xu L, Du Z, Kadara H, Nilsson MB et al (2011) Upregulated stromal EGFR and vascular remodeling in mouse xenograft models of angiogenesis inhibitor-resistant human lung adenocarcinoma. J Clin Invest 121(4):1313–1328
129. Crawford Y, Kasman I, Yu L, Zhong C, Wu X, Modrusan Z et al (2009) PDGF-C mediates the angiogenic and tumorigenic properties of fibroblasts associated with tumors refractory to anti-VEGF treatment. Cancer Cell 15(1):21–34
130. Casanovas O, Hicklin DJ, Bergers G, Hanahan D (2005) Drug resistance by evasion of antiangiogenic targeting of VEGF signaling in late-stage pancreatic islet tumors. Cancer Cell 8(4):299–309
131. Hu YL, DeLay M, Jahangiri A, Molinaro AM, Rose SD, Carbonell WS et al (2012) Hypoxia-induced autophagy promotes tumor cell survival and adaptation to antiangiogenic treatment in glioblastoma. Cancer Res 72(7):1773–1783
132. Piao Y, Liang J, Holmes L, Zurita AJ, Henry V, Heymach JV et al (2012) Glioblastoma resistance to anti-VEGF therapy is associated with myeloid cell infiltration, stem cell accumulation, and a mesenchymal phenotype. Neuro-Oncology 14(11):1379–1392
133. Woyach JA, Furman RR, Liu TM, Ozer HG, Zapatka M, Ruppert AS et al (2014) Resistance mechanisms for the Bruton's tyrosine kinase inhibitor ibrutinib. N Engl J Med 370(24):2286–2294
134. Maddocks KJ, Ruppert AS, Lozanski G, Heerema NA, Zhao W, Abruzzo L et al (2015) Etiology of Ibrutinib therapy discontinuation and outcomes in patients with chronic lymphocytic leukemia. JAMA Oncol 1(1):80–87
135. Albitar A, Ma W, De Dios I, Estrella J, Farooqui M, Wiestner A et al (2015) High sensitivity testing shows multiclonal mutations in patients with CLL treated with BTK inhibitor and lack of mutations in Ibrutinib-Naive patients. Blood 126:716
136. Xu L, Tsakmaklis N, Yang G, Chen JG, Liu X, Demos M et al (2017) Acquired mutations associated with ibrutinib resistance in Waldenstrom macroglobulinemia. Blood 129(18):2519–2525
137. Cheng S, Guo A, Lu P, Ma J, Coleman M, Wang YL (2014) Functional characterization of BTK mutation that confers ibrutinib resistance: exploration of alternative kinase inhibitors. Leukemia 29:895–900
138. Zhou Q, Lee GS, Brady J, Datta S, Katan M, Sheikh A et al (2012) A hypermorphic missense mutation in PLCG2, encoding phospholipase Cgamma2, causes a dominantly inherited autoinflammatory disease with immunodeficiency. Am J Hum Genet 91(4):713–720
139. Burger JA, Landau DA, Taylor-Weiner A, Bozic I, Zhang H, Sarosiek K et al (2016) Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. Nat Commun 7:11589
140. Zhao X, Lwin T, Silva A, Shah B, Tao J, Fang B et al (2017) Unification of de novo and acquired ibrutinib resistance in mantle cell lymphoma. Nat Commun 8:14920

Chapter 6
Exceptional Responders

Michael Cheng and Gopa Iyer

Keywords Pathology · Molecular diagnostics · Precision oncology · Biomarkers · Diagnosis · Prognosis · Disease monitoring · Next-generation sequencing · Liquid biopsy

6.1 Introduction

Anecdotes of patients who experience unexpectedly profound and/or durable responses to a specific cancer treatment, inconsistent with the experience of the vast majority of patients who receive that same treatment, are replete within the scientific literature [1]. The relevance of such outlier responses has usually not been well defined; however, in the context of clinical trials of anti-cancer therapies in which only a minority of patients (or even a single patient) achieves a substantial response, durable responses can often have significant repercussions. The efficacy of such drugs is by definition interpreted based upon the response of the overall study population, frequently leading to a failure of further drug development. Before the advent of Next Generation Sequencing (NGS), detailed analysis of patient responders was extremely challenging. Moreover, clinical trials of targeted therapies in large part initially adopted traditional histology-based eligibility criteria, leading to the accrual of patients whose tumors did not harbor the putative target inhibited by targeted agents. The phenotype of enhanced sensitivity to a drug observed within outlier cases or "exceptional responders" can now be interrogated using NGS to define their

M. Cheng
Genitourinary Oncology Service, Department of Medicine,
Memorial Sloan Kettering Cancer Center, New York, NY, USA

G. Iyer (✉)
Genitourinary Oncology Service, Department of Medicine,
Memorial Sloan Kettering Cancer Center, New York, NY, USA

Weill Cornell Medical College, New York, NY, USA
e-mail: iyerg@mskcc.org

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_6

genomic underpinnings and ultimately to use this information to inform the design of future clinical trials. Additionally, outlier responders represent a unique opportunity to identify new predictive biomarkers of exquisite sensitivity to targeted therapy (or traditional chemotherapy), delineating a subset of patients that could significantly benefit from these drugs as compared to the overall population. In contrast to other strategies to identify actionable targets, such as the large-scale analyses of The Cancer Genome Atlas (TCGA) [2], examination of exceptional responders represents a phenotype-to-genotype approach in an individual, outcome-driven context that has the potential to enhance our understanding of tumor biology, inform clinical trial design, identify novel therapeutic targets, salvage drugs deemed ineffective within a genetically undefined patient population, and define more precise subgroups of patients who may benefit from new or existing drugs. In this chapter, we will describe examples of how detailed genomic analysis of extreme responses to a variety of treatments has resulted in improved insight into the molecular pathogenesis of a specific disease subtype or uncovered putative therapeutic targets for further investigation.

6.2 Initial Whole Genome Sequencing (WGS) of an Exceptional Responder in Urothelial Carcinoma

WGS was performed on an exceptional responder identified on a phase II trial of everolimus, an allosteric mTOR complex 1 (mTORC1) inhibitor, for the treatment of metastatic urothelial cancer [3]. This single-arm, non-randomized study enrolled 45 patients with progressive, pre-treated metastatic urothelial cancer to receive everolimus 10 mg orally once daily [4]. The trial did not meet its primary endpoint of 2-month progression free survival rate of 70%, with 51% of patients being progression-free at 2 months. However, one patient experienced a durable near-complete response, with a 94% reduction in size of target lesions. This patient still remains on therapy (6 years out), which is in stark contrast to the median PFS of 2.6 months (95% confidence interval (CI), 1.8–3.5) and median overall survival of 8.3 months (95% CI, 5.5–12.1) of the patients on study. One other patient achieved a partial response and 12 patients achieved minor tumor regression.

Whole-genome sequencing was performed on DNA obtained from the patient's primary tumor and peripheral blood mononuclear cells. The tumor genome was structurally intact, and 140 non-synonymous mutations were identified in protein-coding or non-coding RNA regions. Of these, two notable loss-of-function mutations were identified. The first was a two base-pair deletion in the tuberous sclerosis complex 1 (*TSC1*) gene, which resulted in a frameshift truncation (c.1907_1908del, p.Glu636fs) of the protein product. The second was a nonsense mutation in the neurofibromatosis type 2 (*NF2*) gene, leading to a premature stop codon (c.863C > G, p.Ser288∗) [3]. Alterations in these genes have been shown to result in dysregulated mTORC1 signaling and sensitivity to rapamycin (an mTORC1 inhibitor which is an analog of everolimus) in preclinical models [5, 6]. Moreover, everolimus is FDA approved for the treatment of subependymal giant cell astrocytomas (SEGAs),

which are tumors that develop in individuals with tuberous sclerosis, an inheritable condition characterized by germline loss of function mutations in TSC1 or TSC2. In a single-arm study, 32% of patients experienced ≥50% shrinkage of SEGAs by 6 months of therapy with a median response duration of approximately 8.8 months.

Following these findings, a separate cohort of 96 high-grade bladder cancers was sequenced for mutations in TSC1 and NF2 and identified 5 additional somatic mutations in TSC1, but no additional mutations in NF2. These findings then guided laboratory work to test the hypothesis that this patient's unusual NF2 mutation was the potentiating factor in the setting of the TSC1 frameshift truncation that led to the exceptional response. Short hairpin RNA (shRNA)-knockdown of NF2 was performed in RT-4, a TSC1-null human bladder cancer cell line, and resulted in increased sensitivity to rapamycin [3].

Furthermore, targeted exon capture sequencing was performed on tissue from 13 additional patients from the everolimus trial. Four of the 13 patients had TSC1 mutations, with three nonsense mutations and one missense variant of unknown functional consequence. Two of the patients with nonsense mutations had best overall responses of 17% and 24% tumor regression, and the patient with a missense mutation had 7% tumor regression. In contrast, eight of the nine TSC1 wild type patients had tumor progression (Fig. 6.1). The TSC1-mutant patients remained on treatment significantly longer than the wild-type patients (7.7 versus 2.0 months, $p = 0.004$) and had a significantly longer time to recurrence (4.1 versus 1.8 months; HR 18.5, CI 2.1–162, $P = 0.001$).

Notably, the trial's pre-specified analysis for hotspot mutations in *FGFR3*, *HRAS*, *PIK3CA*, and *BRAF* did not identify any alterations in this patient or the other responder. Also, analysis of IHC for mTOR pathway markers and PTEN expression did not reveal any association with PFS[4]. This case brings up several issues: 1. Use of selective sequencing approaches or protein-based assays may not identify a predictive biomarker of response, 2. Targeted therapies used in the context of unselected patient populations are likely to fail if the target of response is present in only a minority of patients, 3. Such trials may ultimately lead to lack of development of a drug that, in the right genomic context, may result in meaningful clinical responses, and 4. Even in the right genomic context, the panoply of co-existing genomic alterations within solid tumors likely modulates the degree of therapeutic efficacy of single agent targeted therapies. The outlier approach may therefore potentially salvage investigational therapies for use in a subset of molecularly stratified patients even when the drug "fails" in the setting of a larger, unselected trial population (Fig. 6.2).

6.3 Exceptional Response to mTOR Inhibitor Therapy and Pathway Convergence in Clonal Heterogeneity

Everolimus is FDA-approved for the treatment of metastatic renal cell carcinoma (RCC) following progression on first-line tyrosine kinase inhibition based upon a phase III study showing incremental improvements in median progression-free

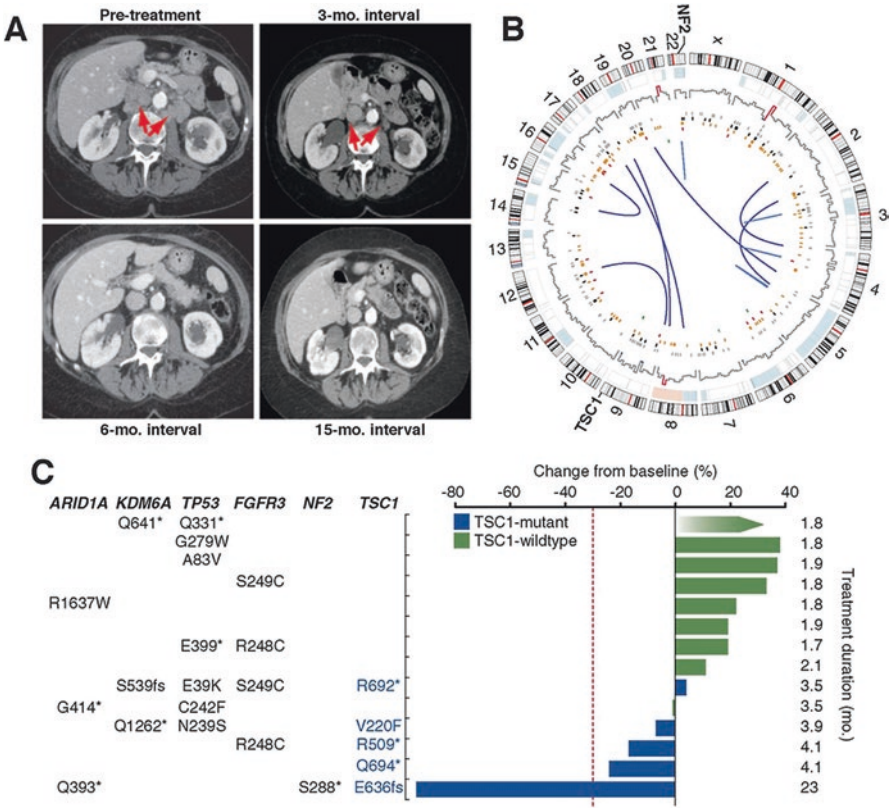

Fig. 6.1 (**a**) Computed tomography images of the index patient demonstrating complete resolution of metastatic disease (arrows). (**b**) Somatic abnormalities in the outlier responder's genome included (from outside to inside) copy number alterations; mutations at ~10-Mb resolution; regulatory, synonymous, missense, nonsense, nonstop, and frameshift insertion and deletion mutations (black, orange, red, green, and dark green); and intra- and interchromosomal rearrangements (light and dark blue). (**c**) Best overall response of 14 sequenced trial patients. Negative values indicate tumor shrinkage (red line, threshold for partial response). Gradient arrow, patient with rapid progression in bone

survival [7, 8]. Long-term disease control with mTOR inhibition has been reported in RCC, even following rapid progression on multi-targeted tyrosine kinase inhibitor therapy, suggesting that these outcomes are driven by sensitivity to drug rather than inherently slow growth kinetics of disease [9]. In order to delineate the molecular causes of these rare yet durable responses, targeted hybridization capture-based sequencing [10] was performed on five outlier metastatic RCC cases identified at a single center. These cases were selected based on extended disease control (ranging from 20 months to >45 months) with either temsirolimus (another mTORC1 inhibitor approved for metastatic RCC) or everolimus. These disease control durations dramatically exceeded prior treatment duration on first-line sunitinib (2–14 months) [11]. The number of prior treatment regimens per patient ranged from 1–3

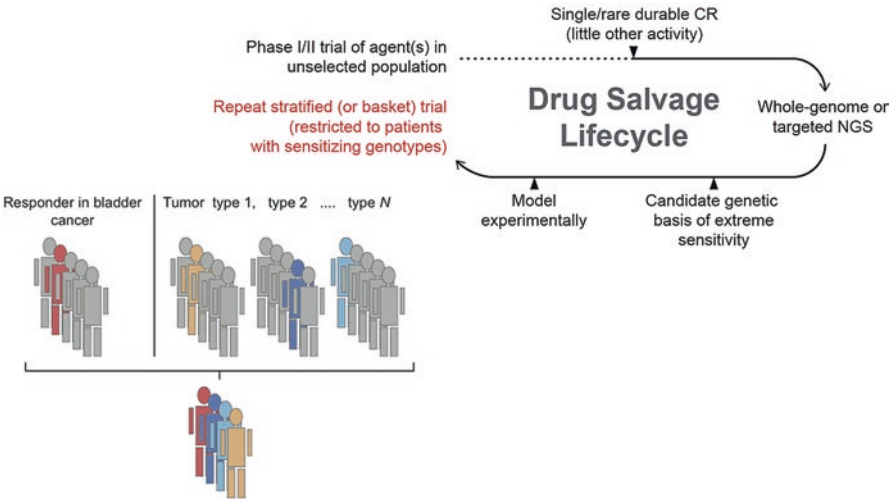

Fig. 6.2 Immunogenomics represents one of many factors that influence response to checkpoint blockade immunotherapy

Table 6.1 Patient characteristics

	Sex	Age	Histologic RCC subtype	MSKCC risk score[a,b]	Number of prior regimens	Treatment duration on prior VEGF-targeted agent (months; agent)	Number of metastatic sites	Rapalog	Treatment duration on rapalog (months)
1	F	58	Clear	Int	1	14 (sunitinib)	≥3	Temsirolimus	27
2	F	73	Clear	Int	1	3 (sunitinib)[c]	1	Temsirolimus	34
3	M	66	Clear	Int	2	5 (sunitinib)	≥3	Everolimus	20
4	F	60	Clear	Fav	3	11 (sunitinib)	≥3	Temsirolimus	28
5	F	50	Unclassified	Fav	1	2 (sunitinib)	≥3	Temsirolimus	45+

Fav favorable, *Int* intermediate, *MSKCC* Memorial Sloan-Kettering Cancer Center
[a]At the time of first rapalog dose
[b]Motzer et al., J Clin Oncol 1999; 17:2530–40
[c]Discontinued due to treatment toxicity

(Table 6.1). Additionally, to explore the potential effects of intratumoral heterogeneity, DNA was extracted from morphologically distinct regions within each primary tumor and sequenced. In three of five patients, a potential genomic etiology for response was identified. In the first patient, analysis of the primary tumor revealed a previously unreported TSC1 frameshift truncation (c.932delC) in two spatially distinct regions of the primary tumor and a TSC1 nonsense mutation (Q527∗) in a third region. Copy number analysis demonstrated loss of heterozygosity in chromosome 9, which contains TSC1, resulting in complete functional loss of the gene. Similarly, the second patient harbored a novel frameshift mutation in TSC1 (c.1738delAT) in all regions of the primary tumor as well as in a metastatic site with concurrent loss of heterozygosity, again resulting in functional loss of TSC1. Loss of the upstream regulator TSC1 results in mTORC1 hyperactivation and is the likely basis for exceptional sensitivity to therapy with temsirolimus in

these two cases. The third patient had a novel missense mutation in mTOR (Q2223K) in two regions of the primary tumor, and a novel TSC1 nonsense mutation (Q781∗) and loss of heterozygosity was noted in two other regions of the primary tumor. The mTOR mutation was located near the ATP-binding site of the catalytic PI3K-related kinase domain, and in vitro work demonstrated that the mutation resulted in hyperactivity.

These findings emphasize the complex issue of intratumoral heterogeneity observed in RCC and other cancer types [12]; moreover, they exemplify the concept of convergence of diverse alterations upon a common phenotype—enhanced mTOR signaling output. These observations also provide a roadmap for the implementation of strategies such as sequential or combinatorial therapies in solid tumors. The evolutionary reasons for the convergence of distinct clones are unknown, and a complex area for future investigation. Furthermore, the inability to identify a genomic basis of durable rapalog sensitivity for the other two patients in this study highlights the limitations of contemporary exceptional responder analysis with next-generation sequencing. Examination at the DNA level alone may not be sufficient, as epigenetic alterations and mechanisms such as post-transcriptional or post-translational gene silencing may influence drug response.

6.4 Concurrent mTOR Mutations and Sensitivity to mTOR Inhibition as a Component of Combination Therapy

Combinatorial treatment with targeted agents is an area of active clinical investigation. Based on suggestive preclinical data supporting the combination of mTOR inhibitors and angiogenesis inhibitors for cancer therapy [13], a phase I trial enrolled nine patients with advanced solid tumors progressing on standard therapies to treatment with pazopanib and everolimus. While no responses were observed in other patients, one patient with platinum- and taxane-refractory urothelial carcinoma achieved a complete response lasting for 14 months. Whole genome sequencing revealed two concurrent mutations in mTOR (E2419K and E2014K), neither of which was previously described in human cancer [14]. Functional analysis demonstrated that both mutations were activating, but do not affect the mTOR activation loop conformation or alter binding to everolimus. While intratumoral heterogeneity was not directly interrogated in this case, each mutation had an allelic fraction of approximately 50%, suggesting that the mutations were heterozygous throughout the tumor sample rather than occurring independently in different tumor subsets. No mutations related to pazopanib sensitivity were identified other than a PDGFRA alteration (p.Y102fs) with an allelic fraction of 2%. Additionally, greater than 100× higher doses of pazopanib (versus rapamycin) were required to overcome the effects of the mTOR mutations in vitro. These findings suggest that the exceptional response was derived from mTOR inhibition with everolimus against a likely clonal mTOR signaling-dependent tumor. Furthermore, this case emphasizes that analysis of

exceptional responses to a combination therapy may reveal exceptional sensitivity to one component agent rather than direct sensitivity to the combination itself.

6.5 Curative Response to Combination Therapy in the Context of a Hypomorphic RAD50 Mutation

Curative outcomes following progression on systemic treatment for metastatic solid tumors are rare. A particularly notable case involved a patient with metastatic invasive small-cell cancer of the ureter, a rare histologic subtype with an especially poor prognosis. The patient achieved a complete response within 5 months of treatment with AZD7762, an ATP-competitive inhibitor of CHK1/2, in combination with weekly irinotecan, a topoisomerase I inhibitor, on a phase I clinical trial [15]. This complete response persisted at least 3 years following discontinuation of all systemic treatments. Whole genome sequencing of pre-treatment tissue identified a clonal RAD50 L1237F mutation and there was concurrent focal heterozygous loss of the wild-type allele. RAD50 is a component of the multi-subunit nuclease Mre11 complex, which is responsible for repair of double strand DNA breaks. Functional modeling demonstrated that this mutation results in an intact Mre11 complex, but reduced Rad50 protein levels and inability to activate the checkpoint kinase ATM, which is responsible for one of two primary signaling cascades that coordinate the DNA-damage response. This results in strong dependence on the remaining ATR pathway in the setting of irinotecan-induced DNA damage. Co-treatment with AZD7762 inhibits CHK1, an effector kinase of ATR, thus leading to a synthetic lethal interaction responsible for the patient's curative outcome. Notably, neither the combination therapy nor the investigational agent AZD7762 as a single drug proceeded in development, as limited activity was observed in the unselected 68 solid tumor patients on this trial. Many examples relate a single mechanism of sensitivity to single-agent targeted therapy while the prior instance demonstrated exceptional response as a result of one component of a combination. However, this case highlights how exceptional responses can occur in the setting of combination treatment as a function of complex interactions between interrelated biological pathways [16].

6.6 Analysis of Exceptional Responses to Treatment with Immunotherapy

Exceptional responders have primarily been studied in the setting of treatment with a targeted agent with activity against a specific molecular lesion or lesions. However, notable recent examples demonstrate how genomic profiling can elucidate the molecular etiology of an exceptional response to immunotherapy. A patient with chemotherapy-treated, recurrent endometrial adenocarcinoma involving

retroperitoneal and supraclavicular lymph nodes was enrolled on a phase I trial of pembrolizumab (an anti-PD-1 antibody) for PD-L1 expressing solid tumors. Shortly after initiating treatment, the patient's lymphadenopathy improved and resultant lower extremity edema resolved. A partial response was documented on imaging at 8 weeks from the start of treatment. This response was maintained for at least 14 months, representing the first known case of endometrial cancer with a durable response to pembrolizumab in the setting of a clinical trial.

Targeted genomic tumor profiling [17] was performed on biopsy tissue from a supraclavicular lymph node metastasis as well as the primary hysterectomy specimen. A remarkably high mutation frequency was identified, with 32 likely pathogenic variants in the primary tumor and 33 in the metastatic specimen. Twenty-eight of these alterations were shared between the two samples, including two mutations in the *POLE* gene, which encodes for the central catalytic subunit of DNA polymerase epsilon and is involved in nuclear DNA repair and replication. The first was a missense mutation in the exonuclease domain responsible for proofreading function (Val411Leu), and the second was a nonsense mutation (Arg114∗) resulting in inactivation of the wild-type allele [18]. Seven percent of endometrial cancers in the TCGA dataset have mutations in the POLE exonuclease domain, and demonstrate an ultramutated phenotype that is an order of magnitude greater than the microsatellite instability (MSI) hypermutated group. Val411Leu represents one of two hotspots present in 76% of the POLE mutants [19]. POLE-mutant endometrial cancers display more antigenic neoepitopes, demonstrate enhanced tumor-infiltrating lymphocytes, have higher intratumoral CD8+ T-cell density, all suggestive of greater immunogenicity compared to non-POLE mutant endometrial cancers [20]. Importantly, these POLE-mutant tumors are largely microsatellite stable, and are not identified using currently approved IHC and PCR-based MSI assays [19].

A subsequent report described durable responses in two heavily pretreated metastatic endometrial cancer patients to off-label nivolumab, another anti-PD-1 antibody [21]. In these cases, treatment selection was informed by genomic analysis performed after progression on salvage chemotherapeutic agents. The first patient had a POLE exonuclease mutation (Pro286Arg, the other hotspot identified by TCGA), and weak membranous PD-L1 expression in 5% of tumor cells. Treatment with nivolumab at 3 mg/kg biweekly resulted in marked clinical improvement and regression of her metastatic pelvic and intra-abdominal tumor deposits, meeting partial response criteria by RECIST. This response was sustained for at least 7 months following treatment initiation. Another patient underwent NGS testing that identified an MSH-6 mutation with a hypermutated phenotype. Tumor cells did not significantly express PD-L1. This patient was treated with nivolumab and achieved a partial response maintained for at least 9 months, with a concomitant improvement in performance status. These results prompted an investigator-initiated phase II study at the same institution evaluating pembrolizumab in chemotherapy-refractory recurrent endometrial cancer with POLE-mediated ultramutator or MMR deficient hypermutator phenotypes (NCT02899793). Ultimately, observations such as the above have led to a landmark approval by the FDA of pembrolizumab for patients with microsatellite unstable tumors. Overall, these experiences are emblem-

atic of how analysis of exceptional responders can contribute to the development of genomic biomarker-selected clinical trials, and synergize with insights gleaned from TCGA and other efforts.

6.7 Analysis of Exceptional Responses to Treatment with Chemotherapy

Unusually deep responses have also been described in the setting of treatment with chemotherapy. Whole-exome sequencing was performed on pretreatment tissue from 50 muscle-invasive urothelial carcinoma patients who received cisplatin-based neoadjuvant chemotherapy prior to cystectomy. The genomic profiles of 25 patients who achieved downstaging to pT0/pTis at cystectomy were compared to 25 patients with residual pT2 or higher stage disease [22]. ERCC2, a nucleotide excision repair (NER) helicase, was identified as the only mutated gene significantly enriched in the pT0/pTis cohort (occurring in 36%, compared to ~12% of unselected muscle-invasive bladder cases from TCGA). All ERCC2 presumed loss-of-function mutations occurred in this group within or adjacent to helicase domains, similar to the clustering of germline mutations responsible for Xerderma pigmentosum (XP) and XP with combined Cockayne syndrome. In vitro work confirmed the sensitivity of ERCC2 mutant cells to cisplatin as well as UV-induced DNA damage, which is also mediated by NER. The allelic fraction of most (78%) ERCC2-mutant cases was <0.5, suggesting a haploinsufficient or dominant-negative effect of a heterozygous mutation in mediating sensitivity to cisplatin via deficiency in NER capacity. These results provide a potential genomic basis for identifying patients most likely to benefit from cisplatin-based chemotherapy in urothelial carcinoma.

6.8 Occult Biomarkers Identified by Outlier Analysis

Extreme outlier analyses can also identify novel biomarkers of response to targeted therapies, since the majority of cancer driver alterations are scattered across cancer types at low frequencies (the so-called “long tail” genes). A phase II trial of selumetinib, a non-ATP competitive MEK1/2 inhibitor, for patients with recurrent low grade serous (LGS) ovarian cancer reported a radiographic response rate of 15% [23]. The only complete response was observed in a patient with initial stage IIIC serous borderline (SB) ovarian cancer who recurred with metastatic LGS ovarian cancer. Despite progressing on multiple lines of prior therapy, she experienced a durable complete response of at least 5 years to treatment with selumetinib on the trial. Capture-based sequencing confirmed wild type BRAF and KRAS but identified a 15-bp in-frame deletion within MAP2K1 that resulted in a 5 amino acid deletion in MEK1 adjacent to the negative regulatory helix [24]. In silico modeling

suggested that this truncation resulted in release of negative regulation of MEK1 kinase activity. In vivo and in vitro work confirmed enhanced tumor growth with the mutation, and sensitivity to selumetinib. This MAP2K1 deletion was previously uncharacterized, and would not have been discovered by hotspot analysis. The results of the index patient led to sequencing of an additional 28 BRAF/RAS wild-type SB/LGS patients, which did not reveal additional MEK1 mutations but identified occult MAPK pathway alterations in all but 5 tumors. These occult alterations included two truncating mutations in the RAS GAP NF1, a HER2 AYVM insertion (previously validated as oncogenic), an NRAS Q61R mutation, and two novel paracentric BRAF fusions (MKRN1:BRAF and CUL1:BRAF). Notably, the patient with a CUL1:BRAF fusion achieved a durable complete response to combination treatment with a MEK inhibitor and paclitaxel. While such alterations may be genetic outliers, additional large scale unbiased sequencing efforts are necessary to define their frequency and relevance within cancer.

6.9 Mechanisms of Acquired Resistance Following Initial Exceptional Response

While some curative outcomes have been reported, the majority of exceptional responders ultimately experience disease progression after an initial durable response. Understanding the mechanisms of this acquired resistance represents an important priority. A patient with metastatic anaplastic thyroid cancer who received everolimus on a phase II study testing the therapy for patients with radioiodine refractory thyroid cancer (NCT00936858) experienced a near complete response that was maintained for 18 months, in contrast to a median survival of 5 months in patients with this disease [25]. Whole exome sequencing was performed on pretreatment tumor (from thyroidectomy and neck dissection) and tumor tissue at the time of progression. An inactivating TSC2 (Q1178∗) mutation was identified in the pretreatment tumor, resulting in mTOR activation and sensitivity to everolimus. A FLCN R17fs mutation may also have contributed to mTOR activation. Sequencing of the resistant tumor revealed persistence of the previously identified mutations, but also identified a novel MTOR F2108 L mutation that was not present in the pretreatment tissue. This mutation was functionally confirmed to occur in the FRB (FKBP-rapamycin binding) domain and likely resulted in acquired resistance to everolimus by preventing allosteric drug binding.

Activation of downstream or bypass pathways can also mediate acquired resistance. A patient with metastatic melanoma harboring a BRAF V600E mutation was treated on a clinical trial of the monomeric mutant BRAF inhibitor vemurafenib (PLX4032) with a profound near complete response by 15 weeks of therapy. However, this response persisted only until 23 weeks, when disease relapse occurred involving most prior sites of disease. Targeted exon sequencing of the resistant tumor revealed a MEK1 C121S, and functional studies confirmed the mutation to

confer increased kinase activity and resistance to RAF inhibition [26]. Additional putative resistance mutations in MEK1 were characterized using a mutagenesis screen. This represents the first report of resistance mediated by an activating mutation acquired downstream of the targeted kinase, and contributes to the framework for understanding relapse following an initial exceptional response.

6.10 NCI Exceptional Responders Initiative

The NCI has embarked on an Exceptional Responders Initiative as part of a broader Precision Medicine effort [27] based upon observations from multiple extraordinary responder analyses. This exploratory study (NCT02243592) aims to molecularly characterize 100 exceptional responder cases with whole exome and/or mRNA sequencing. The NCI reviewed its Cancer Therapy Evaluation Program (CTEP) phase II trial database between 2002–2012, and identified about 100 cases that fit the exceptional responder criteria. While many of these cases lacked biomaterial to further interrogate the outlier response, their presence suggests that obtaining sufficient cases for the Initiative via submission from physicians and researchers is feasible.

The NCI has developed an operational definition for exceptional responders as patients with "a complete response or a partial response that lasted at least 6 months to a systemic treatment that was not expected in more than 10% of patients [28]." This expectation of treatment response is informed either by clinical trial data or extensive historical experience.

6.11 Conclusions and Future Steps

Exceptional responses to targeted agents have been reported with increasing frequency. Application of NGS technologies, including WES, WGS, and targeted sequencing, has deciphered the genomic biomarkers that underpin many of these responses. Similarly, predictors of resistance to therapy have also been defined when these responders progress on treatment.

The ultimate goal of exceptional responder analysis is to generate scientific evidence that informs prospective investigation of specific therapies for genomically defined patient subsets. Notably, many actionable alterations occur in a small fraction of any tumor subtype [29], and a significant fraction are likely to be currently undiscovered or uncharacterized (the "long tail" pattern). Thus, future clinical trials for precision cancer medicine must have the flexibility to enroll and assess patients with specific molecular profiles, even if these account for only a small fraction of the relevant population or even involve only a single patient. While much of the initial experience with exceptional responder analysis employed WGS or whole exome sequencing (WES), targeted exon capture panels available at academic med-

ical centers [10] or commerically [17], can now be utilized to rapidly and prospectively screen patients for the most common genomic alterations.

An important contemporary clinical trial structure in precision cancer medicine is the basket study [30, 31]—one method to prospectively identify exceptional responders. By including baskets of patients with different cancer types, all of which share a common mutation, we can rapidly identify disease groups for whom the mutation is a driver and for whom inhibition of the mutation results in durable, meaningful clinical benefit. The NCI-Molecular Analysis for Therapy Choice (NCI-MATCH) trial (NCT02465060), and the Targeted Agent and Profiling Utilization Registry (TAPUR) trial (NCT02693535) are the largest and most rigorous precision cancer medicine trials to date. NCI-MATCH is a unique phase II trial [32] that will enroll patients with solid tumors, lymphoma, and myeloma that no longer respond to standard therapy for sequencing of tumor tissue to assess a set of pre-specified genes [33]. When sequencing identifies a mutation that fits into one of several substudy arms, the patient is assigned to an appropriate linked therapy. The therapies assigned to each arm are required to have demonstrated clinical activity in at least one molecularly characterized case. NCI-MATCH will evaluate a primary endpoint of objective response rate (ORR), and a secondary endpoint of progression free survival (PFS). Additionally, if patients fail on an initial treatment arm, repeat gene testing can be employed to evaluate for eligibility on a second arm of the trial. The aggregation of patients across tumor types in this and other basket studies enhances the investigation of targeted therapies against specific mutational targets that may be of limited prevalence even in the most common tumor types. TAPUR is similarly designed, but aims to evaluate "real-world" precision medicine by evaluating the efficacy of 15 different FDA approved drugs or combinations matched to potentially actionable genomic variants identified by molecular profiling tests performed in the clinical setting (including genomic tests using cell-free DNA). The trial aims to enroll 1030 patients with advanced solid tumors, myeloma, or Non-Hodgkin lymphoma, and will evaluate ORR as the primary endpoint and OS as the secondary endpoint.

The first published basket trial was a phase II study which tested response rate to vemurafenib for the treatment of BRAF V600 mutant nonmelanoma cancers in 122 patients across multiple disease histologies [34]. The trial enrolled tumor types including: anaplastic thyroid, cholangiocarcinoma, colorectal cancer, Erdheim-Chester disease or Langerhans'-cell histiocytosis, multiple myeloma, non-small cell lung cancer, primary brain tumors, and several others. Notably, work from The Cancer Genome Atlas (TCGA) [2] and other contemporary efforts has demonstrated that the incidence of BRAF V600 mutations in more than half of nonmelanoma cancer types is under 5%. Response to vemurafenib varied between tumor types, with the highest ORRs in the NSCLC cohort (42%) and in the Erdheim-Chester disease or Langerhans'-cell histiocytosis cohort (43%). Response rates were lower in other disease types, including an ORR of 0% in colorectal cancer patients treated with vemurafenib alone. The results of this study highlight that there exists differential sensitivity between tumor types to targeted treatment directed at the same

molecular target, and that the basket trial structure can identify specific tumor types with promising outcomes and thus inform the design of future definitive studies.

The current experience with exceptional responders underscores that these cases are not just statistical outliers, but frequently reveal novel biological interactions and mechanisms of response [35]. Laboratory investigation is frequently required to assess whether putative novel alterations both have functional significance and engender dependency of the tumor upon the alteration. Additional important factors should also be integrated into in the future study of exceptional responders as well as the application of their findings. The presence of co-alterations as well as the clonality of the mutation of interest in the setting of intra- and intertumoral heterogeneity must be considered in evaluating response to a targeted agent. Furthermore, the distance in molecular time from the sequenced tissue to the current treatment must also be considered. Sequencing of archival tissue may not adequately reflect contemporary mutation burden. Biopsies, even if up-to-date, may reflect only the genomic profile of the most anatomically accessible tissue rather than the molecular profile of the entire tumor population. Implementation of cell-free DNA (cfDNA) sequencing [36, 37] and optimization of the turnaround time for NGS assays may help ameliorate these concerns. Also, as more flexible precision medicine trials become more common, should the observation of exceptional responses influence the design or enrollment of the trial? For example, the identification of an unexpected complete response in a particular disease or molecular subset could prompt a dynamic enrichment of that subset within a trial. Such an adaptive trial design would require the close involvement of translational researchers who can biologically validate and extend the findings observed in the clinic to provide insight into why a specific disease subtype responds better than others to a targeted therapy. Overall, innovative clinical investigation strategies to test hypotheses generated by exceptional responder analysis offer tremendous opportunities to advance precision cancer medicine.

References

1. Chang DK, Grimmond SM, Evans TR et al (2014) Mining the genomes of exceptional responders. Nat Rev Cancer 14:291–292
2. Weinstein JN, Collisson EA, Mills GB et al (2013) The cancer genome atlas pan-cancer analysis project. Nat Genet 45:1113–1120
3. Iyer G, Hanrahan AJ, Milowsky MI et al (2012) Genome sequencing identifies a basis for everolimus sensitivity. Science 338:221
4. Milowsky MI, Iyer G, Regazzi AM et al (2013) Phase II study of everolimus in metastatic urothelial cancer. BJU Int 112:462–470
5. Lopez-Lago MA, Okada T, Murillo MM et al (2009) Loss of the tumor suppressor gene NF2, encoding merlin, constitutively activates integrin-dependent mTORC1 signaling. Mol Cell Biol 29:4235–4249
6. James MF, Han S, Polizzano C et al (2009) NF2/merlin is a novel negative regulator of mTOR complex 1, and activation of mTORC1 is associated with meningioma and schwannoma growth. Mol Cell Biol 29:4250–4261

7. Hudes G, Carducci M, Tomczak P et al (2007) Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma. N Engl J Med 356:2271–2281
8. Motzer RJ, Escudier B, Oudard S et al (2010) Phase 3 trial of everolimus for metastatic renal cell carcinoma : final results and analysis of prognostic factors. Cancer 116:4256–4265
9. Molina AM, Ginsberg MS, Motzer RJ (2011) Long-term response with everolimus for metastatic renal cell carcinoma refractory to sunitinib. Med Oncol 28:1527–1529
10. Cheng DT, Mitchell TN, Zehir A et al (2015) Memorial sloan kettering-integrated mutation profiling of actionable cancer targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. J Mol Diagn 17:251–264
11. Voss MH, Hakimi AA, Pham CG et al (2014) Tumor genetic analyses of patients with metastatic renal cell carcinoma and extended benefit from mTOR inhibitor therapy. Clin Cancer Res 20:1955–1964
12. Gerlinger M, Rowan AJ, Horswell S et al (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366:883–892
13. Ikezoe T, Yang Y, Nishioka C et al (2006) Effect of SU11248 on gastrointestinal stromal tumor-T1 cells: enhancement of growth inhibition via inhibition of 3-kinase/Akt/mammalian target of rapamycin signaling. Cancer Sci 97:945–951
14. Wagle N, Grabiner BC, Van Allen EM et al (2014) Activating mTOR mutations in a patient with an extraordinary response on a phase I trial of everolimus and pazopanib. Cancer Discov 4:546–553
15. Al-Ahmadie H, Iyer G, Hohl M et al (2014) Synthetic lethality in ATM-deficient RAD50-mutant tumors underlies outlier response to cancer therapy. Cancer Discov 4:1014–1021
16. Peng G, Woodman SE, Mills GB (2014) RADical response puts an exceptional responder in CHKmate: a synthetic lethal curative response to DNA-damaging chemotherapy? Cancer Discov 4:988–990
17. Frampton GM, Fichtenholtz A, Otto GA et al (2013) Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. Nat Biotechnol 31:1023–1031
18. Mehnert JM, Panda A, Zhong H et al (2016) Immune activation and response to pembrolizumab in POLE-mutant endometrial cancer. J Clin Invest 126:2334–2340
19. Kandoth C, Schultz N, Cherniack AD et al (2013) Integrated genomic characterization of endometrial carcinoma. Nature 497:67–73
20. van Gool IC, Eggink FA, Freeman-Mills L et al (2015) POLE proofreading mutations elicit an antitumor immune response in endometrial cancer. Clin Cancer Res 21:3347–3355
21. Santin AD, Bellone S, Buza N et al (2016) Regression of chemotherapy-resistant polymerase epsilon (POLE) ultra-mutated and MSH6 hyper-mutated endometrial tumors with nivolumab. Clin Cancer Res 22:5682–5687
22. Van Allen EM, Mouw KW, Kim P et al (2014) Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. Cancer Discov 4:1140–1153
23. Farley J, Brady WE, Vathipadiekal V et al (2013) Selumetinib in women with recurrent low-grade serous carcinoma of the ovary or peritoneum: an open-label, single-arm, phase 2 study. Lancet Oncol 14:134–140
24. Grisham RN, Sylvester BE, Won H et al (2015) Extreme outlier analysis identifies occult mitogen-activated protein kinase pathway mutations in patients with low-grade serous ovarian cancer. J Clin Oncol 33:4099–4105
25. Wagle N, Grabiner BC, Van Allen EM et al (2014) Response and acquired resistance to everolimus in anaplastic thyroid cancer. N Engl J Med 371:1426–1433
26. Wagle N, Emery C, Berger MF et al (2011) Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. J Clin Oncol 29:3085–3096
27. Abrams J, Conley B, Mooney M et al (2014) National cancer institute's precision medicine initiatives for the new national clinical trials network. Am Soc Clin Oncol Educ Book 34:71–76

28. Mullard A (2014) Learning from exceptional drug responders. Nat Rev Drug Discov 13:401–402
29. Lawrence MS, Stojanov P, Polak P et al (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 499:214–218
30. Willyard C (2013) 'Basket studies' will hold intricate data for cancer drug approvals. Nat Med 19:655
31. Redig AJ, Janne PA (2015) Basket trials and the evolution of clinical trial design in an era of genomic medicine. J Clin Oncol 33:975–977
32. McNeil C (2015) NCI-MATCH launch highlights new trial design in precision-medicine era. J Natl Cancer Inst 107:djv193
33. Lih CJ, Harrington RD, Sims DJ et al (2017) Analytical validation of the next-generation sequencing assay for a nationwide signal-finding clinical trial: molecular analysis for therapy choice clinical trial. J Mol Diagn 19:313–327
34. Hyman DM, Puzanov I, Subbiah V et al (2015) Vemurafenib in multiple nonmelanoma cancers with BRAF V600 mutations. N Engl J Med 373:726–736
35. Mehra N, Lorente D, de Bono JS (2015) What have we learned from exceptional tumour responses?: review and perspectives. Curr Opin Oncol 27:267–275
36. Forshew T, Murtaza M, Parkinson C et al (2012) Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 4:136ra68
37. Murtaza M, Dawson SJ, Tsui DW et al (2013) Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. Nature 497:108–112

Chapter 7
Immunogenomics

Jonathan J. Havel and Alexandra Snyder

Keywords Neoantigens · Neoantigen prediction · Neoepitopes · Tumor immunology · Tumor antigens · Mouse models · Coley · Major histocompatibility complex (MHC) · T cells · Tetramers

7.1 Introduction and History of Immunogenomics

Checkpoint blockade immunotherapies have changed the face of medical oncology since the approval of ipilimumab to treat metastatic melanoma in 2011. The entry of immunotherapies into standard of care treatment for has heralded many translational studies attempting to identify which patients are most likely benefit from these treatments. Studies describing the importance of the immune system to the development and treatment of cancer have a long history, with moments of success overshadowed by challenges for much of the eighteenth and nineteenth centuries. Perhaps the strongest evidence for the importance of the immune system to control of cancer lies in the epidemiology of cancer in immunosuppressed populations. The incidence of AIDS-defining illnesses associated with immunosuppression (Kaposi sarcoma, non-Hodgkin lymphoma and cervical cancer) have decreased dramatically in the United States since the widespread use of highly active anti-retroviral therapy (HAART), which largely restores immune function in the setting of viral suppression [1].

One of the earliest demonstrations of the capacity of the immune system to be manipulated to control selected cancers was shown by William Coley, a surgeon injected a cocktail of bacterial toxins into tumors and noted regressions of some

J. J. Havel
Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

A. Snyder (✉)
Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Merck & Co, Inc., Kenilworth, NJ, USA
e-mail: alex.snyder@merck.com

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_7

tumors in these patients [2]. This idea was not widely disseminated, however, and the underlying concept of using exogenous means to inflame a tumor by making it appear more foreign to the immune system did not experience resurgence until the development of oncolytic viruses in the 1990's [3]. This strategy has experienced its greatest development in melanoma, where the approval of talimogene laherparepvec (T-VEC) by the Food and Drug Administration (FDA) in 2015 was based on improved overall survival in patients who had had local injection of the oncolytic virus.

Other pioneers of immuno-oncology in the 1940s and beyond approached the issue from a different angle, observing, for example, that sarcomas induced by exogenous toxins experienced rejection in mouse models [4–7]. The idea that mutations in a tumor could generate a novel peptide sequence that could be recognized as non-self was executed using human tumor genomic data for the first time in 2008 [8]. In a study by Segal and colleagues, the authors predicted which neoepitopes, which are also referred to as neoantigens, or mutation-associated neoantigens (MANA), would putatively be generated in a series of colon and breast cancers. In so doing, the authors underscored two points that have become central to the field: first, that different tumor types tend to feature higher or lower ranges of neoantigens depending on their burden of mutations that affect protein coding sequences; and second, that neoantigens are, by and large, patient-specific or "personalized" to a given tumor.

7.2 Preclinical Models Demonstrate Relevance of Immunogenomics to Tumor Fate

Murine and *in vitro* models have been indispensable in the study of tumor immunity. Prior to the advent of high-throughput sequencing technologies and whole genome mapping, the identification of tumor rejection antigens was a daunting task. During the 1980s and 1990s, Boon and colleagues utilized mouse tumor cells treated with a mutagen to derive unique clones readily rejected by the immune system in mice (reviewed in [9]). Parental cells were largely non-immunogenic. T cell clones with specificity for unique tumor clones, but not for parental cells, were generated via a labor-intensive process involving repeated co-cultures of T cells with irradiated tumor cells [10, 11]. Genomic, and later cDNA, plasmid libraries were generated from T cell-sensitive tumor clones and transfected into non-immunogenic cells, i.e. either parental cells or cells that had been immunoselected by specific T cell clones to lack rejection antigens. The transfected cells were sub-cloned and subjected to co-culture with a relevant T cell clone. Plasmids that conferred T cell sensitivity were isolated and sequenced to identify tumor rejection antigens [12]. Notably, two distinct types of tumor antigens were identified using this technique: (a) peptides from ubiquitously expressed proteins that contained a single amino acid substitution mutation (Fig. 7.1) [13, 14], and (b) peptides from non-mutated, but selectively and/

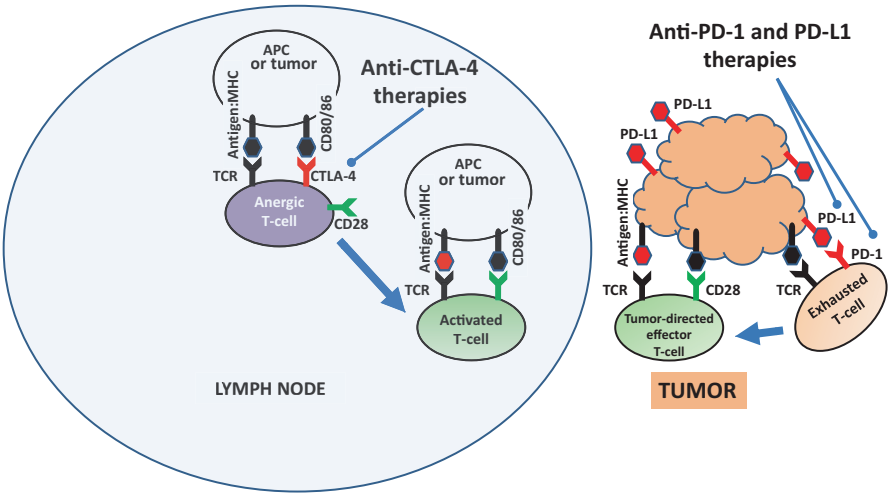

Fig. 7.1 Cartoon demonstrating where in the immune response neoantigens (red hexagon) resulting from tumor mutations are thought to play a role. (Adapted from Intlekofer & Thompson, JLB 2013; Callahan & Wolchok, JLB, 2013). *TCR* T-cell receptor, *MHC* major histocompatibility complex, *APC* antigen presenting cell, *Red hexagon* putative tumor neoantigen, *Black hexagon* co-stimulatory signal

or aberrantly expressed proteins [15]. This approach was used to identify tumor antigens in both mice and humans; however, the process was exceptionally labor-intensive.

As demonstrated above in the Segal et al. study, advances in genomic sequencing technologies and computational biology can now be applied to streamline the identification of mutated tumor antigens, i.e. neoantigens [8]. In 2012, Gubin et al. used a carcinogen-induced murine sarcoma model to test the ability of the immune system to sculpt the mutational landscape of a tumor [16]. Sub-clones were isolated from tumors grown in immune incompetent mice and were subsequently transferred to immune competent mice. Some clones were rejected in the immune competent mice, while others remained tumorigenic. Whole exome sequencing was performed in order to identify the full complement of non-synonymous mutations in each cell line. Mutations unique to clones rejected by immune competent mice were identified as putative neoantigens, and were subsequently validated as immunogenic epitopes in *in vitro* cell culture assays [17].

While the neoantigens identified in the studies above were all restricted by MHC Class I and targeted by CD8+ cytotoxic T lymphocytes (CTLs), other studies have discovered mutation-induced neoantigens restricted by MHC Class II and targeted by CD4+ T cells. Linnemann et al. utilized an innovative co-culture system in which autologous human B cells immortalized by stable transduction of the anti-apoptotic proteins BCL-6 and BCL-XL were used as antigen presenting cells. The immortalized B cells were loaded with long peptides (31 amino acids each) containing somatic mutations identified by whole exome sequencing of a patient's tumor.

Autologous CD4+ T cells were co-cultured with peptide-loaded B cells and supernatant cytokine levels were measured to detect T cell activation. Functional, Class II-restricted neoantigens were identified in tumor infiltrating lymphocytes and in the adoptive cell therapy products used to successfully treat patients with metastatic melanoma [18]. Although this study found that only ~0.5–1% of mutations elicit a CD4+ T cell response in patients, i.e. one or two immunogenic mutations out of 100 or 200 tested, another study utilizing peptide vaccination in syngeneic murine tumor models found the numbers to be much higher. Kreiter et al. found that approximately 20–40% of all mutations tested elicited an immune response from splenocytes of vaccinated mice, and 70–80% of those were due to CD4+ T cell reactivity [19]. These discrepancies may reflect differences in immune responses between mice and humans or between tumor- and vaccination-induced responses. Nevertheless, more work will be required to determine the relative and mechanistic contributions of CD8+ and CD4+ T cells to tumor immune responses in the setting of checkpoint blockade, adoptive T cell therapy, and vaccination.

7.3 Neoantigen Prediction

The prediction of which putative neoantigens will actually induce an immune response and the nature of that response, present computational and biological conundra for which, as of 2017, the field is developing but has not completed solutions. In a process that has been amply and elegantly reviewed elsewhere [20], peptides can be processed intracellularly by an antigen presenting or other cell into antigens that can then be presented by the human immune system by type I or type II major histocompatibility complexes (MHC). The peptide-MHC complex must then be recognized by and engage with a T-cell receptor, which then signals internally to help determine the response of the T-cell to that antigen. In addition to internal signals, external signals both on the T cell and in the tumor immune microenvironment impact this dynamic process.

Algorithms exist to predict which mutations will be translated, processed, presented and then elicit a T-cell response; however, the absence of a gold standard for assessing the output of these steps in this process experimentally in the setting of malignancy (see next section on Technologies to detect a neoantigen-specific T-cell response) limits the feedback process necessary to optimize the computational algorithms. That said, the existing algorithms are undergoing continuous improvements to address these issues. After mutations are called in a tumor sample by comparing the tumor DNA to an individual's matched normal DNA, then the affected areas are "virtually translated," yielding a putative aberrant peptide resulting from the mutation, and a wild type peptide. The prediction algorithms then "scan" the altered peptide to discern whether it would be predicted to be presented by MHC Class I.

Although as stated above there are algorithms to address each step in the process of antigen presentation, the most commonly tools focus on predicting which 9 to 11 amino acid (known as 9- to 11mers) stretches will be presented by MHC Class

I. Initial programs did so using stereotyped motifs, for example assuming anchor residues (positions 2 and 9) would be occupied by certain amino acids. However, current tools are trained on actual HLA binding data which they then use to predict binding in an HLA-specific manner (concepts are discussed further in [21]). In addition, where expression data are available, current strategies take into account expression in order to decrease candidate neoantigens [22, 23].

Significant challenges remain, however. The longer length of peptides loaded onto MHC Class II and the less well-known rules determining peptide position have meant that MHC Class II peptide binding prediction lags behind Class I (reviewed in [24]) Attempts to elute neoantigens from MHC are very labor-intensive and have been successful in only a subset of labs with expertise in that area [16, 25]. Even less is known about the T-cell side of the neoantigen equation. The degeneracy of T-cell receptor (TCR) recognition both enables a large number of peptides to be recognized by a given TCR [26], and simultaneously renders it impossible, with current technology, to identify which peptide(s) a given TCR may recognize. What type of T-cell response is induced also remains a topic of debate, with several studies focusing on the CD8+ T-cell response to neoantigens in preclinical models [16] and patients treated with checkpoint blockade [27], and others demonstrating that a CD4+ T cell anti-neoantigen response can yield immune control in both mouse [19] and human [28] contexts. Data on whether neoantigens can actually induce an immuno-suppressive response has not yet been published, but seems logical given what is known by non tumor-associated antigens [29–31].

7.4 Technologies to Detect a Neoantigen-Specific T Cell Response

In order to improve computational neoantigen prediction algorithms, it will be imperative to gather more empirical data through the experimental identification and validation of neoantigens. Methods for the detection of neoantigen-specific T cell responses can be grouped into two categories, depending on whether they detect T cell activation or the binding of an MHC-presented peptide to specific T cells.

Most functional assays are based upon the following experimental framework – *in vitro* co-culture of peptide-presenting antigen presenting cells (APCs) with autologous T cells, followed by detection of cytokine production from activated T cells. Each element of this experimental framework can be varied based on the scope and purpose of the experiment and on the materials available. Sources of APCs include autologous dendritic cells differentiated *in vitro* from monocytes, dendritic cells and macrophages already present in peripheral blood, immortalized B cells, or artificial antigen presenting cells, in which cells (typically from a different species) are engineered to express a specific human MHC complex and co-stimulatory molecules. The way in which putative neoantigen peptides are introduced to APCs is another source of variation. Synthetic peptides can be loaded directly onto APCs or,

alternatively, RNA encoding longer protein sequences can be electroporated into APCs to allow for natural processing and presentation. Finally, the method of cytokine detection can also vary. The most efficient and high-throughput of these is called ELISPOT. In the ELISPOT assay, cytokine capture antibodies are immobilized on the surface of wells of a microtiter culture plate. Secondary antibodies conjugated to a colorimetric detection agent are used to detect pockets or "spots" of cytokine production. The spots are counted and measured to indicate the number of activated T cells and the strength of the response. Alternatively intracellular cytokine staining can be performed on T cells and analyzed by flow cytometry. This method has the advantage of being able to detect multiple cytokines and assign them to individual cells and it fairly quantitative; however, it is also more labor-intensive and the cells must be killed by fixation and permeabilization during staining, rendering them unavailable for further cell-based assays.

Binding assays are based on peptide-MHC (pMHC) tetramer technology. The fundamental unit of a pMHC tetramer consists of a recombinant MHC molecule (including beta-2 microglobulin) conjugated to biotin and loaded with a specific peptide. The binding affinity of a single pMHC monomer for its cognate TCR is exceptionally low and therefore pMHC monomers are not suitable labeling reagents for antigen-specific T cells. To circumvent this issue, the biotin is then bound to fluorescently-labeled streptavidin, which contains four biotin binding sites per molecule, thus creating fluorescently-labeled tetramers of the pMHC complex. The resulting cooperative binding significantly strengthens TCR: tetramer affinity and these tetramers can be used as a labeling reagent to identify and monitor T cells with a unique antigen specificity [32, 33].

Generation of large libraries of pMHC tetramers is an appealing approach for identifying neoantigen-reactive T cells in patient samples; however, this approach was initially limited by technological challenges. Importantly, the MHC molecule is only stable when bound to a cognate peptide. Initially, this caveat limited the efficiency of generating large pools of pMHC tetramers, as the recombinant MHC molecules could only be purified when already bound to a peptide, requiring each pMHC complex to be purified separately. Schumacher and colleagues devised a solution to this problem by creating ultraviolet (uv) cleavable peptides that could be used as "place-holders" to stabilize the MHC molecule during purification. These peptides could then be easily swapped with other peptides by exposing the tetramers to uv light in the presence of the a peptide of interest [34]. This advance allowed for the rapid generation of large libraries of pMHC tetramer complexes.

Because each tetramer complex can only be labeled with a single fluorophore, the next challenge was to devise a method for uniquely labeling each tetramer so that a patient sample could be screened by a large library of pMHC tetramers in a single well or tube. To this end, a "barcoding" technique was developed in which each unique pMHC tetramer is labeled with one of two fluorophores. Therefore, each pMHC tetramer is encoded by a unique fluorophore signature. More practically speaking, a tetramer is identified by detection in a specific combination of two fluorophore channels on a flow cytometer [35, 36]. This method has been further developed for detection by mass cytometry, in which fluorophores are replaced by

rare earth metals with unique mass [36, 37], or by DNA barcodes than can be identified via sequencing [38], thereby greatly expanding the size limitations of pMHC libraries for screening in a single reaction.

Ultimately, in order to make exponential rather than incremental improvements in neoantigen prediction algorithms, it will be critical to include TCR-binding predictions. Therefore, it will be important to identify both the relevant neoantigen sequences and their cognate T cell receptor sequences in the experimental setting. Although this approach has not yet been documented successfully in a high-throughput manner, it may be possible to use pMHC libraries to label T cells of a certain specificity, followed by sorting and analysis by single-cell RNA sequencing to identify the cognate TCR variable domain sequence for each neoepitope tested [39, 40]. Collection of this type of data will be invaluable to improving neoantigen reactivity in a personalized manner for patients.

7.5 Evidence for T-Cell Reactivity to Neoantigens in Patients Treated with Checkpoint Blockade Immunotherapy

The advent of checkpoint blockade therapies to the treatment of advanced cancers brought immunogenomics to the fore in studies in tumor immunology. The currently FDA-approved therapies target the checkpoints cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), programmed cell death protein 1 (PD-1) and its ligand PD-L1, and are now standard of care in a wide variety of solid tumor malignancies [41]. A study by van Rooij and colleagues examined a single melanoma patient treated with anti-CTLA-4 therapy and noted a neoantigen-specific response in circulating T cells in the peripheral blood [42]. This study utilized the UV-exchange technology described above, and with the sensitivity of that assay, responses to only two neoantigens, one of which was quite weak, were identified out of 448 predicted neoantigens. The authors hypothesized in that study that the higher number of single nucleotide variants (SNVs) in melanomas, particularly those related to ultraviolet radiation exposure, might increase the likelihood of forming one or several neoantigens that generate an anti-tumor T-cell response. Indeed, two follow-up studies on CTLA-4-treated melanoma patients confirmed a correlation between tumor mutation burden (TMB)—specifically coding SNVs leading to an amino acid change—and outcome [27, 43]. This correlation was imperfect in both studies, however, with tumors harboring many mutations lacking benefit, and others with low mutation burdens nonetheless benefiting.

At the time of the initial study [27], it was not known whether this principle would apply to other tumor types or checkpoint targets. However, further studies have demonstrated a correlation—consistently imperfect—between elevated TMB and improved outcome in melanomas, non-small cell lung cancers (NSCLC) [44] and mismatch repair-deficient cancers treated with anti-PD-1 therapy [45], and urothelial cancers treated with anti-PD-L1 therapy [46]. These concepts were con-

firmed retrospectively in a larger dataset: 271 patients with NSCLC were treated with nivolumab 3 mg/kg every 2 weeks and 270 patients were treated with histology-dependent chemotherapy in the front-line setting; subgroups of each underwent whole exome sequencing of tumor and matched normal blood [47]. Those patients in the highest tertile mutation burden (n = 47) of the nivolumab-treated patients demonstrated an improved PFS as compared to medium or low-tertile groups (9.7 months in high versus 3.6 in medium [n = 49] and 4.2 in low [n = 62]); no such distinction was seen in the chemotherapy arm (5.8 [n = 60], 6.5 [n = 53] and 6.9 [n = 41] months PFS in high, medium and low groups respectively).

However, the exceptions to the principle that high TMB correlates with outcome clearly imply that other mechanisms are at play in contributing to a successful or failed anti-tumor immune response. For example, anti-PD-1 therapy has been remarkably successful in Hodgkin's lymphoma [48] and is approved in renal cell carcinoma (RCC) [49], neither of which features a subset of high TMB tumors [50]. Sequencing data from these distinct situations are pending.

Furthermore, as our understanding of neoantigens deepens, the neoantigen features associated with a relevant T-cell response is a topic of intense investigation. One such study that used two of the aforementioned datasets as well as functional *in vitro* data suggested that clonal neoantigens—i.e. those neoantigens resulting from mutations found in every cell of a tumor—comprise the set that are relevant to immune rejection [51]. This study was aided by the enhanced sensitivity of the assay used to identify neoantigens, a technique that uses barcoding rather than combinatorial fluorescence to identify neoantigen-MHC as described above, and can thus identify a larger number of predicted neoantigens using a smaller number of T cells [38].

Furthermore, the relevance of immunogenomics extends beyond putative targets of checkpoint blockade therapies. Several groups have performed cell therapy as immunotherapy and identified neoantigens as targets. This method consists of extracting patients' T cells (either from tumor or peripheral blood), expanding them under stimulated conditions *ex vivo*, then reinfusing them to garner enhanced anti-tumor effects. Indeed, Rosenberg and colleagues identified neoantigens as the targets of the cell therapies pioneered by their group, with published examples of an CD4 anti-neoantigen clinical efficacy in a patient with cholangiocarcinoma [28], and another similar study in patients with gastrointestinal cancers [52].

Finally, preclinical data from the Sahin and Schreiber labs demonstrate that under some conditions, neoantigen vaccines can protect against tumor challenge or even lead to tumor regression, particularly when combined with checkpoint blockade [16, 19]. This effect may be CD4 or CD8 T cell-mediated. Based in part on these data, several neoantigen vaccine studies are recruiting or planned (for example, NCT01970358 and NCT03166254, among others).

7.6 Unanswered Questions and Future Directions

Although remarkable clinical advances have been made in cancer immunotherapy, the field of cancer immunogenomics is still in its infancy. Many intriguing questions about mechanism, biomarkers, and prediction remain. Notably, although neoantigen-specific responses have been detected, mutation or predicted neoantigen load is a relatively poor predictor of clinical benefit. This observation has many implications for future investigation. Perhaps neoantigens are critical determinants of response in some patients but not others. Furthermore, the existence of an experimentally identified neoantigen response does not necessarily imply causality in clinical benefit from that therapy. Ultimately a direct comparison of the relative influences of mutation-derived neoantigens, non-mutated tumor associated antigens, and tumor microenvironment across many patients will be required to better understand the mechanisms of response to cancer immunotherapy (Fig. 7.2).

Another intriguing question is whether neoantigens can ever be detrimental to response. Specifically, the concept of immunodominance, in which reactivity to a single antigen "out-competes" other reactivities, has been documented in infectious disease immunology [53–56]. Whether specific mutation-derived peptide sequences may have a similar effect and can act as "decoy epitopes" has yet to be explored in the setting of tumor immunology. Furthermore, T regulatory cells can be activated

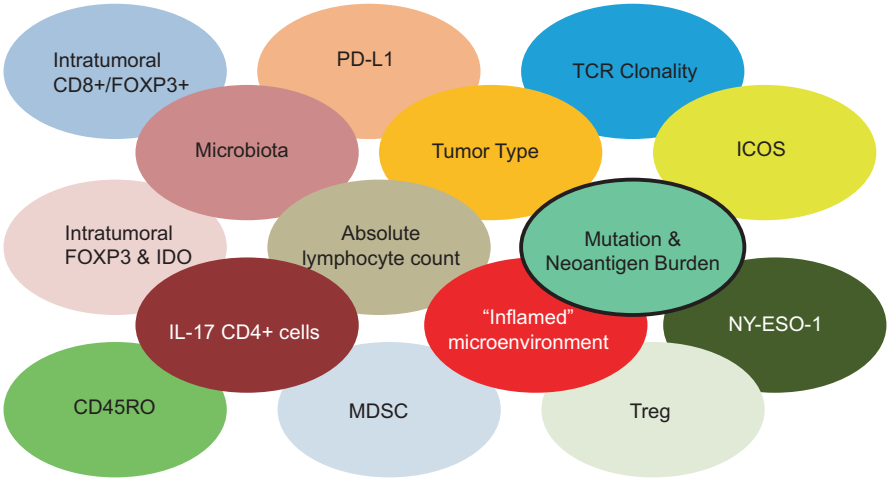

Fig. 7.2 Immunogenomics represents one of many factors that influence response to checkpoint blockade immunotherapy. (Yuan et al., PNAS 2011; DiGiacom lo et al. Cancer Immunol Immunother 2013; Queirolog et al., Cancer Invest 2013; Wolchok et al., Cancer Immun 2010; Tumeh et al. Nature 2105; Snyder et al. NEJM 2014; Rizvi et al. Science 2015; Van Allen et al. Science 2105; Sivan et al. Science 2015; Vetizou et al. Science 2015; Rosenberg et al. Lancet 2016, Ku et al. Cancer 2010; Menard et al. Clin Cancer Res 2008; Weber et al. JCO 2009; Hodi et al. PNAS 2008; Hamid et al. JCO 2009; Ng et al. Cancer Immuno Res 2013; Tarhini et al. PLoS One 2014; Kitano et al. Cancer Immunol Res 2013; Spranger et al. Sci Transl Med 2013; Kitano et al. Cancer Immunol Res 2014; Ji RR et al., Cancer Immunol Immunother 2012)

in an antigen-specific manner [29] and specific peptide sequences, known as Tregitopes, have been identified that trigger T regulatory (Treg) cell-mediated immunosuppression [31]. Therefore, it seems possible that tumor-derived mutations that generate "neo-Tregitope" sequences may have a detrimental effect on patient response to immunotherapy. Detailed experimental work will be required to explore these possibilities.

Ultimately, the ability to predict and improve response to cancer immunotherapies will likely require examination of a number of genetic and environmental factors that influence not only tumor characteristics, but the host immune repertoire and function as well. Therefore, interrogation of the tumor, its surrounding stroma, and the immune system in general at epigenomic, genomic, and proteomic levels will be critical [57]. It is hoped that such an integrative immunogenomics approach will improve the ability to tailor immunotherapies specifically to each patient's tumor and thereby greatly improve survival rates for patients.

References

1. Franceschi S, Lise M, Clifford GM et al (2010) Changing patterns of cancer incidence in the early- and late-HAART periods: the Swiss HIV Cohort Study. Br J Cancer 103:416–422
2. Coley WB (1906) Late results of the treatment of inoperable sarcoma by the mixed toxins of erysipelas and Bacillus prodigiosus. Am J Med Sci 131:375–430
3. Martuza RL, Malick A, Markert JM, Ruffner KL, Coen DM (1991) Experimental therapy of human glioma by means of a genetically engineered virus mutant. Science 252:854–856
4. Srivastava PK, DeLeo AB, Old LJ (1986) Tumor rejection antigens of chemically induced sarcomas of inbred mice. Proc Natl Acad Sci U S A 83:3407–3411
5. DuBois GC, Appella E, Law LW, DeLeo AB, Old LJ (1980) Immunogenic properties of soluble cytosol fractions on Meth A sarcoma cells. Cancer Res 40:4204–4208
6. Foley EJ (1953) Antigenic properties of methylcholanthrene-induced tumors in mice of the strain of origin. Cancer Res 13:835–837
7. Prehn RT, Main JM (1957) Immunity to methylcholanthrene-induced sarcomas. J Natl Cancer Inst 18:769–778
8. Segal NH, Parsons DW, Peggs KS et al (2008) Epitope landscape in breast and colorectal cancer. Cancer Res 68:889–892
9. Coulie PG, Van den Eynde BJ, van der Bruggen P, Boon T (2014) Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. Nat Rev Cancer 14:135–146
10. Maryanski JL, Van Snick J, Cerottini JC, Boon T (1982) Immunogenic variants obtained by mutagenesis of mouse mastocytoma P815. III. Clonal analysis of the syngeneic cytolytic T lymphocyte response. Eur J Immunol 12:401–406
11. Herin M, Lemoine C, Weynants P et al (1987) Production of stable cytolytic T-cell clones directed against autologous human melanoma. Int J Cancer 39:390–396
12. De Plaen E, Lurquin C, Van Pel A et al (1988) Immunogenic (tum-) variants of mouse tumor P815: cloning of the gene of tum- antigen P91A and identification of the tum- mutation. Proc Natl Acad Sci U S A 85:2274–2278
13. Szikora JP, Van Pel A, Brichard V et al (1990) Structure of the gene of tum- transplantation antigen P35B: presence of a point mutation in the antigenic allele. EMBO J 9:1041–1050
14. Sibille C, Chomez P, Wildmann C et al (1990) Structure of the gene of tum- transplantation antigen P198: a point mutation generates a new antigenic peptide. J Exp Med 172:35–45

15. Van den Eynde B, Lethe B, Van Pel A, De Plaen E, Boon T (1991) The gene coding for a major tumor rejection antigen of tumor P815 is identical to the normal gene of syngeneic DBA/2 mice. J Exp Med 173:1373–1384
16. Gubin MM, Zhang X, Schuster H et al (2014) Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens. Nature 515:577–581
17. Matsushita H, Vesely MD, Koboldt DC et al (2012) Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. Nature 482:400–404
18. Linnemann C, van Buuren MM, Bies L et al (2015) High-throughput epitope discovery reveals frequent recognition of neo-antigens by CD4+ T cells in human melanoma. Nat Med 21:81–85
19. Kreiter S, Vormehr M, van de Roemer N et al (2015) Mutant MHC class II epitopes drive therapeutic immune responses to cancer. Nature 520:692–696
20. Janeway CA, Travers P, Walport M, Shlomchik M (2005) Immunobiology: the immune system in health and disease, 6th edn. Garland Science Publishing, New York
21. Snyder A, Chan TA (2015) Immunogenic peptide discovery in cancer genomes. Curr Opin Genet Dev 30:7–16
22. van Rooij N, van Buuren MM, Philips D et al (2013) Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an Ipilimumab-responsive melanoma. J Clin Oncol 31:e439–e442
23. Snyder A, Nathanson T, Funt SA et al (2017) Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: an exploratory multi-omic analysis. PLoS Med 14:e1002309
24. Nielsen M, Lund O, Buus S, Lundegaard C (2010) MHC class II epitope predictive algorithms. Immunology 130:319–328
25. Meyer VS, Drews O, Gunder M, Hennenlotter J, Rammensee HG, Stevanovic S (2009) Identification of natural MHC class II presented phosphopeptides and tumor-derived MHC class I phospholigands. J Proteome Res 8:3666–3674
26. Donermeyer DL, Weber KS, Kranz DM, Allen PM (2006) The study of high-affinity TCRs reveals duality in T cell recognition of antigen: specificity and degeneracy. J Immunol 177:6911–6919
27. Snyder A, Makarov V, Merghoub T et al (2014) Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med 371:2189–2199
28. Tran E, Turcotte S, Gros A et al (2014) Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer. Science 344:641–645
29. Mills KH, McGuirk P (2004) Antigen-specific regulatory T cells--their induction and role in infection. Semin Immunol 16:107–117
30. Sloan-Lancaster J, Allen PM (1996) Altered peptide ligand-induced partial T cell activation: molecular mechanisms and role in T cell biology. Annu Rev Immunol 14:1–27
31. De Groot AS, Moise L, McMurry JA et al (2008) Activation of natural regulatory T cells by IgG Fc-derived peptide "Tregitopes". Blood 112:3303–3311
32. Altman JD, Moss PA, Goulder PJ et al (1996) Phenotypic analysis of antigen-specific T lymphocytes. Science 274:94–96
33. Altman JD, Davis MM (2016) MHC-peptide tetramers to visualize antigen-specific T cells. Curr Protoc Immunol 115:17 3 1–173 44
34. Rodenko B, Toebes M, Hadrup SR et al (2006) Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. Nat Protoc 1:1120–1132
35. Hadrup SR, Bakker AH, Shu CJ et al (2009) Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. Nat Methods 6:520–526
36. Andersen RS, Kvistborg P, Frosig TM et al (2012) Parallel detection of antigen-specific T cell responses by combinatorial encoding of MHC multimers. Nat Protoc 7:891–902
37. Newell EW, Davis MM (2014) Beyond model antigens: high-dimensional methods for the analysis of antigen-specific T cells. Nat Biotechnol 32:149–157
38. Bentzen AK, Marquard AM, Lyngaa R et al (2016) Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. Nat Biotechnol 34:1037–1045

39. Han A, Glanville J, Hansmann L, Davis MM (2014) Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nat Biotechnol 32:684–692
40. Redmond D, Poran A, Elemento O (2016) Single-cell TCRseq: paired recovery of entire T-cell alpha and beta chain transcripts in T-cell receptors from single-cell RNAseq. Genome Med 8:80
41. Balar AV, Weber JS (2017) PD-1 and PD-L1 antibodies in cancer: current status and future directions. Cancer Immunol Immunother 66:551–564
42. van Rooij N, van Buuren MM, Philips D et al (2013) Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. J Clin Oncol 31:e439–e442
43. Van Allen EM, Miao D, Schilling B et al (2015) Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science 350:207–211
44. Rizvi N, Hellmann M, Snyder A et al (2015) Mutational landscape determines sensitivity to programmed cell death-1 blockade in non-small cell lung cancer. Science 348:124–128
45. Le DT, Uram JN, Wang H et al (2015) PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 372:2509–2520
46. Rosenberg JE, Hoffman-Censits J, Powles T et al (2016) Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet 387:1909–1920
47. Peters S, Creelan B, Hellmann MD, Socinski MS, Reck M, Bhagavatheeswaran P, Chang H, Geese WJ, Paz-Ares L, Carbone DP (2017) CT082: impact of tumor mutation burden on the efficacy of first-line nivolumab in stage iv or recurrent non-small cell lung cancer: an exploratory analysis of CheckMate 026. In: AACR. https://doi.org/10.1158/1538-7445.AM2017-CT082
48. Ansell SM, Lesokhin AM, Borrello I et al (2015) PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. N Engl J Med 372:311–319
49. Motzer RJ, Escudier B, McDermott DF et al (2015) Nivolumab versus everolimus in advanced renal-cell carcinoma. N Engl J Med 373:1803–1813
50. Alexandrov LB, Nik-Zainal S, Wedge DC et al (2013) Signatures of mutational processes in human cancer. Nature 500:415–421
51. McGranahan N, Furness AJ, Rosenthal R et al (2016) Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 351:1463–1469
52. Tran E, Ahmadzadeh M, Lu YC et al (2015) Immunogenicity of somatic mutations in human gastrointestinal cancers. Science 350:1387–1390
53. Kelleher AD, Long C, Holmes EC et al (2001) Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. J Exp Med 193:375–386
54. Crowe SR, Miller SC, Shenyo RM, Woodland DL (2005) Vaccination with an acidic polymerase epitope of influenza virus elicits a potent antiviral T cell response but delayed clearance of an influenza virus challenge. J Immunol 174:696–701
55. Weiskopf D, Angelo MA, Sidney J, Peters B, Shresta S, Sette A (2014) Immunodominance changes as a function of the infecting dengue virus serotype and primary versus secondary infection. J Virol 88:11383–11394
56. Welsh RM, Che JW, Brehm MA, Selin LK (2010) Heterologous immunity between viruses. Immunol Rev 235:244–266
57. Weber JS (2017) Biomarkers for checkpoint inhibition. Am Soc Clin Oncol Educ Book 37:205–209

Chapter 8
Managing Germline Findings from Molecular Testing in Precision Oncology

Jessica N. Everett and Victoria M. Raymond

Abstract Molecular genetic testing of the germline for hereditary cancer risk and molecular testing of tumor tissue for therapeutic decision making are no longer clearly distinct. Here we review how changes in sequencing technology have impacted use of germline and tumor molecular testing, implications for identifying incidental pathogenic germline variants through tumor testing, relevance of pathogenic germline variants in cancer care and prevention, and emerging research guiding clinical practice in this area.

Keywords Precision oncology · Genetic counseling · Germline variants · NGS testing · Clinical genomics · Targeted therapy

8.1 Overview of Molecular Testing in Oncology

Molecular genetic and genomic testing has been incorporated into medical oncology practice for two key uses:

1. To identify patients and families at risk for a hereditary cancer syndrome due to pathogenic germline variants (PGV) in cancer risk genes
2. To identify tissue specific (somatic) variants for prognosis and predictive therapeutic decision making in patients with a cancer diagnosis

These two uses of molecular genomic testing emerged independently to meet distinct clinical needs within oncology care, were managed by practitioners with dif-

J. N. Everett (✉) · V. M. Raymond
New York University, New York, NY, USA
e-mail: jessica.everett@nyulangone.org

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_8

ferent clinical expertise, and followed different standards for patient education and consent. Specialist genetics practitioners have primarily provided counseling and molecular testing for hereditary risk. Hereditary cancer risk counseling includes a pre-test discussion of the risks, benefits, and alternatives to testing, and implications of a PGV finding for the medical care of the patient and their family members. Molecular testing for somatic variants has typically been ordered by medical oncologists for prognostic and therapeutic decision making. Traditional pre-test genetic counseling is not included as part of the tumor molecular testing process, since the somatic variants of interest are usually not hereditary. Broader use of these two types of molecular testing has proven that these formerly distinct tests and testing indications are being increasingly intertwined, raising new questions and challenges for clinicians. PGVs identified through testing for hereditary risk may be used to direct patients to targeted therapies (e.g. PGVs in *BRCA1/2* and poly ADP ribose polymerase (PARP) inhibitor therapy). Similarly, expanded use of comprehensive tumor genomic sequencing [1] has increased the potential to identify incidental PGVs, leading to unanticipated diagnoses of hereditary cancer syndromes. Molecular testing of the germline for hereditary cancer risk and molecular testing of tumor tissue for therapeutic decision making are no longer clearly distinct. Here we review how changes in sequencing technology have impacted use of germline and tumor molecular testing, implications for identifying incidental PGVs through tumor testing, relevance of PGVs in cancer care and cancer prevention, and emerging research guiding clinical practice in this area for providers in both genetics and oncology.

8.2 Molecular Testing in Oncology Care – Background

8.2.1 Tumor Molecular Testing for Treatment Planning

Tumor molecular profiling for treatment decision making is a concept that has evolved since the late 1990s, when the Food and Drug Administration (FDA) approved trastuzumab for treatment of breast cancers with overexpression of HER2 (*ERBB2)*. This was followed by approval of imatinib for chronic myelogenous leukemia with *BCR-ABL* gene fusion in 2001, and erlotinib for *EGFR* mutant non-small cell lung cancer in 2004 [2]. Early tumor profiling assays often examined a single variant "hot spot" or expression of a single gene, could be completed as part of routine pathology work flow, and did not require paired normal tissue samples for result interpretation. These early molecular tests had relevance for patients with specific types of cancer. By the early 2000s, rapid changes in the cost and availability of large scale next generation genomic sequencing (NGS) began to greatly expand knowledge of tumor mutational landscapes across cancer types.

Targeted therapies and related clinical trials based on a wide variety of molecular markers continue to grow. Oncologists increasingly order some level of molecular tumor testing in routine clinical practice, ranging from single variant or single gene analysis, to multi-gene panels or whole exome sequencing, for patients with all

types of cancer. The opening of large studies including the National Cancer Institute's Molecular Analysis for Therapy Choice (MATCH), and the American Society of Clinical Oncology's (ASCO) Targeted Agent and Profiling Utilization Registry (TAPUR) suggest that a growing number of patients will undergo tumor sequencing of some kind to guide optimal therapy and clinical trial participation, and the rapid decrease in the price of genomic sequencing will make this testing more accessible for larger patient populations over time.

As tumor molecular testing moved beyond hot spot and single gene assays, it became clear that the presence of the normal germline genome within the tumor was a potentially confounding factor to fully understanding tumor biology. For this reason, research studies began to favor paired sequencing of tumor/germline samples to allow for clearer distinction between somatic and germline variants. As many as one third of actionable variants reported on tumor-only analyses using analytic methods rather than paired germline samples may be incorrectly classified as somatic [3]. Correct distinction of somatic and germline variants has important implications for interpretation of results, as well as downstream impact on treatment decisions and screening for family members. In clinical practice, tumor only sequencing is still favored due to lower cost and ease of sample collection.

8.3 Germline Molecular Testing for Hereditary Cancer Risk

Use of germline genetic testing to identify patients with PGVs causative of hereditary cancer syndromes has become a well-established part of routine oncology care. PGVs in cancer risk genes account for 5–10% of common cancers, with variation depending on cancer type. For patients with cancer, identification of PGVs in cancer risk genes has potential impact on surgical decision making, and increasingly affects systemic treatment decision making as options for targeted therapeutics based on presence of PGVs become available [4, 5]. The two most common hereditary cancer syndromes, Hereditary Breast and Ovarian Cancer (HBOC) caused by PGVs in the *BRCA1* and *BRCA2* genes, and Lynch syndrome, caused by PGVs in the *MLH1*, *MSH2*, *MSH6*, *PMS2*, and *EPCAM* genes, exemplify these points. Women who carry PGVs in *BRCA1/2* face a lifetime risk of breast cancer of up to 72%, and a lifetime risk of ovarian cancer of up to 44% [6]. Long term follow-up studies show that surveillance [7], chemoprevention [8, 9], prophylactic mastectomies [10], and bilateral salpingo-oophorectomies [11] can decrease morbidity and mortality in this population [12]. *BRCA1/2* PGV carriers with ovarian cancer, and possibly those with pancreatic cancer, appear to have improved survival when treated with platinum based therapies [13, 14], and there is evidence suggesting platinum response may be improved in *BRCA2* carriers with prostate cancer [15]. Targeted therapy with PARP inhibitor agents shows efficacy in *BRCA1/2* PGV carriers [16] which can be sustained over months in some patients [17]. Similarly, patients with Lynch syndrome face a lifetime risk of colorectal cancer of up to 85%, which can be dramatically reduced with colonoscopies every 2–3 years beginning at age 20 [18]. Patients with Lynch Syndrome associated stage II

colon cancer can opt out of adjuvant chemotherapy without losing survival benefit [19], and there is evidence supporting use of PD-1 blockade in tumors with defective mismatch repair (MMR) and high rates of somatic mutations, as is found in most Lynch syndrome cancers [4]. In a first of its kind pan-cancer therapy approval, the FDA approved use of pembrolizumab, an anti PD-1 therapy, in patients with microsatellite instable (MSI) or MMR deficient tumors, the classic pathology features of a Lynch Syndrome associated tumor (www.fda.gov).

PGVs in cancer risk genes also have important implications for an affected patient's family members, who may benefit from screening and prevention options to reduce cancer risk. For this reason, pre-test counseling for hereditary cancer risk typically includes discussion of risk to family members, and post-test counseling involves discussion of plans to share test results with relatives [20]. Testing of at-risk family members for known PGVs offers opportunities for intervention, and improves the cost effectiveness of molecular testing for hereditary risk [21].

Genetic testing for hereditary cancer risk entered clinical practice with *BRCA1/2* testing in 1996. At that time, various factors including the high cost of testing, lack of follow up data supporting clinical utility, and fears of genetic discrimination, meant that testing was targeted only to patients at high risk to carry a PGV based on their personal and family history of cancer. Clinical guidelines first established by the ASCO in 1996 [22] and the National Comprehensive Cancer Network (NCCN) in 2010 [23] recommended germline test utilization for patients with specific "red flags" including earlier than average age at cancer diagnosis, suggestive family history, or suggestive tumor characteristics [24]. The National Society of Genetic Counselors (NSGC) published recommended elements of pre-test and post-test counseling for germline cancer genetic testing in 2004 [25]. These basic elements included:

- Collection of 3–4 generation family history, and detailed personal cancer history
- Risk assessment and consideration of genetic testing for patients meeting high risk criteria, where test will influence medical management
- Explicit, written informed consent for testing after discussion of risks, benefits, limitations, and alternatives as well as general education about gene(s), associated cancer risks, and inheritance pattern
- Plan for results disclosure and post-test counseling including discussion of medical and psychosocial implications for patient and family members

8.4 Next Generation Sequencing Technologies – Impact on Clinical Molecular Testing

The arrival of NGS technology in clinical laboratories, and the related decrease in cost of sequencing, led to disruption in standard thinking and practice for use of both germline and tumor testing. Along with adoption of NGS technology, the 2013

Supreme Court decision striking down the patent on *BRCA1/2* sequencing [26] also contributed to a rapid decrease in cost of germline testing. The price of sequencing the *BRCA1* and *BRCA2* genes alone fell from $4000 to $2000 within days of the decision [27]. Multi-gene panels for hereditary risk became clinically available in 2013, and were quickly adopted by clinicians who appreciated the efficiency and cost savings of analyzing multiple cancer risk genes through a single sample from a single clinic visit. By the spring of 2017, multi-gene germline panel tests for 30–79 cancer risk genes, including *BRCA1/2* cost as little as $150–$475 for patients paying out of pocket. Studies of multi-gene panel use reported increased identification of PGVs in patients with suspected hereditary breast and ovarian cancer [28, 29], suspected Lynch syndrome [30], early onset colon cancer [31], and pancreatic cancer [32]. Similarly, the cost of full sequencing of tumors dropped from $5400 to $3600 within 6 months [33], and proved useful in identifying new genes and pathways important to tumor development and progression [34].

Less expensive testing and availability of multi-gene germline and somatic panels led to increased use of testing for a broader range of genes. While this improved diagnostic capabilities for both germline and tumor testing, it came with a side effect: unanticipated or incidental findings. Genetics providers began to see multi-gene panel testing identifying PGVs in patients who did not meet testing criteria for the associated hereditary cancer syndrome, or had atypical phenotypes that may have been missed with a single gene testing approach [30, 35–37]. Patients found to carry more than one cancer predisposing PGV were reported [38]. Cancer genetics specialists noted that new approaches would be needed to address changing needs for pre-test education and informed consent [39]. In the tumor testing space, debates about managing incidental discovery of PGVs in patients not consented for hereditary testing began to emerge.

Early editorials and reviews in the oncology literature addressing potential challenges in managing incidental PGVs found through tumor sequencing began appearing in 2013 [40, 41], with further input from the Clinical Sequencing Exploratory Research (CSER) Consortium in 2016 [42]. These authors emphasized that while sequencing of paired tumor-germline samples could clearly find PGVs, "tumor only" sequencing could also identify important PGV findings as the germline DNA is part of the tumor content. Recognizing that referral to genetic counseling was impractical for all patients having tumor testing, key points for oncologists to address prior to tumor sequencing were highlighted including:

- discussing the possibility of PGV findings
- developing a plan for disclosure of possible PGVs, including patient preferences for disclosure to family members if necessary
- identifying clinical genetics resources and colleagues for collaboration and referral as needed

An early study retrospectively reviewed history for patients who had tumor only sequencing results with potential germline significance to determine how many should have confirmatory germline testing, and what logistics would be needed to confirm possible germline events. Recognizing that full genetic coun-

seling for all patients having tumor testing was impractical and inefficient, they suggested that oncologists have a brief discussion of potential PGVs with *all* patients, then use family history assessment to target more detailed conversations or genetics referral to patients with highest risk for PGVs [43]. All of these publications also emphasized the need for research into informed consent approaches, decision algorithms to guide disclosure, and development of infrastructure to support disclosure at a scale not previously attempted in traditional clinical cancer genetics.

The American College of Medical Genetics and Genomics (ACMG) addressed PGV findings identified as part of paired tumor/germline sample sequencing in their 2013 recommendations for reporting of incidental findings in clinical exome and genome sequencing [44]. These recommendations defined a "minimum list" of disorders with clinical validity and utility for which preventive or treatment options were available, and suggested that PGVs in these genes should be disclosed to patients undergoing exome or genome sequencing regardless of the primary testing indication. This minimum list included 23 cancer risk genes, with a recent update bringing the total to 25 [45] (Table 8.1). At the time, this recommendation generated concerns about identifying PGVs outside the typical clinical genetics context without traditional pre-test genetic counseling [46]. Specific concerns cited by the oncology community in response to the ACMG recommendations included the potential burden placed on oncologists to discuss the possibility of PGVs, particu-

Table 8.1 American College of Medical Genetics and Genomics Minimum Gene List for Disclosure in Large Scale Sequencing

Gene	Syndrome
APC	Familial adenomatous polyposis
BMPR1A, *SMAD4*	Juvenile polyposis
BRCA1, BRCA2	Hereditary breast ovarian cancer syndrome
MEN1	Multiple endocrine neoplasia type 1
MLH1, MSH2, MSH6, PMS2	Lynch syndrome
MUTYH	MUTYH associated polyposis
NF2	Neurofibromatosis type 2
PTEN	PTEN hamartoma tumor syndrome
RB1	Retinoblastoma
RET	Multiple endocrine neoplasia type 2
SDHAF2, SDHB, SDHC, SDHD	Hereditary paraganglioma-pheochromocytoma syndrome
STK11	Peutz-Jeghers syndrome
TP53	Li-Fraumeni syndrome
TSC1, TSC2	Tuberous sclerosis
VHL	Von Hippel Lindau syndrome
WT1	Wilms tumor

larly in genes unrelated to cancer risk. Additional concerns included lack of data on potential clinical benefits, negative consequences, and preferences of patients and oncologists regarding reporting of PGVs in the setting of tumor sequencing [47]. With most studies of germline mutation prevalence occurring in high risk populations, there was little data available to understand how common PGV findings might be in broader patient populations, contributing to worries that oncologists and genetics providers would be unable to handle increased need for genetic services to manage PGV findings.

8.5 How Common Are Germline Findings in Patients with Cancer?

Since the initial ACMG guidelines were released, multiple studies have published results of paired tumor/germline testing in various populations, yielding better estimates of the prevalence of PGVs in patients with cancer that can inform practice and policy planning. An early study of 815 adults with 15 different cancer types reported nonsense, frame shift, and splice site mutations in 85 genes associated with cancer risk, and found 3% of patients carried a PGV [3]. A study of 1566 patients with 30 types of advanced cancers found PGVs in 12.6% using a larger list of 93 cancer risk genes, and including single nucleotide variants, which were the most common type of PGV identified in this study [48]. This study also noted that only 41% of patients with a PGV had a type of cancer concordant with their PGV finding, and that nearly every patient had a variant of uncertain significance (VUS) in one or more of the 187 Mendelian disease risk genes that were studied. Two additional studies focused attention on PGVs in cancer risk genes with the highest clinical impact. One study of 1000 patients found PGVs in 4.3% with a list of 36 genes associated with high risk cancer syndromes, noting that half of the PGVs were known prior to study participation and 63% were concordant with the cancer type diagnosed. This study found 40.5% of patients had a VUS, reflecting the fact that genes with more routine clinical testing are better characterized with fewer uncertain findings [49]. The second study examined 19 high penetrance cancer risk genes in 1000 patients, and found 4.3% carried a PGV with half of those known prior to tumor sequencing [50]. Overall in adult populations with no enrichment for high risk family history, PGVs in cancer risk genes were present in 3–12.6% [3, 48–52]. Differences across the studies including types of cancer diagnosed, number of genes investigated, and types of mutations included, account for the range in PGV findings. PGVs in the *BRCA1* and *BRCA2* genes were the most commonly identified finding across all adult studies. In pediatric populations, 8.5–10% of patients with cancer were found to have PGVs in cancer risk genes [53–55] with PGVs in the *TP53* gene causative of Li-Fraumeni syndrome being the most common.

8.6 Integrating Tumor and Germline Testing

Given the relatively high prevalence of PGVs potentially identifiable in tumor sequencing, and the potential benefits of PGV findings for medical management of patients and their family members, researchers have begun to study and address key barriers to successfully integrating tumor and germline testing. These issues include establishing patient preference for disclosure, ensuring key elements of pre-test informed consent are intact, and developing scalable genetic service delivery models, with the end goal of maximizing the potential benefits of PGV findings while minimizing potential negative outcomes for patients undergoing tumor sequencing. Many of these issues are not new to clinical genetic testing in general, but the "variety and uncertainty of potential results, broad implications of those results, and elevated expectations of personal benefit create some new or amplified challenges" [56].

8.7 Patient Preference, Right "Not to Know"

In the traditional model of germline genetic testing for hereditary cancer syndromes, patients have a dedicated discussion with a genetics provider to weigh the risks, benefits, and limitations of germline testing before deciding whether or not to proceed. Preferences for learning about PGVs are explicitly expressed, and are independent of any other indication for sequencing. With tumor sequencing, patients have testing for treatment planning as the specific indication, and the complexities of this discussion are appropriately prioritized over hypothetical germline findings. The original ACMG working group on incidental findings recommended mandatory reporting of PGVs on their minimum list, concluding that the duty of providers to prevent harm associated with high penetrance conditions where prevention or treatment could be offered outweighed concerns about violating patient autonomy and "right not to know" [44]. Debate among ACMG members and related stakeholders resulted in an updated recommendation for patient ability to *opt out* of germline findings not directly related to the testing indication [57]. Early evidence from three studies suggests that the majority of patients undergoing tumor sequencing choose to receive germline results when given the option including 97% of adults with refractory cancers [58], 90% of parents of pediatric patients with cancer [53], and 95% of patients with stage IV lung or colorectal cancer [59].

8.8 Informed Consent

Prior to wide use of NGS, traditional clinical genetics practice focused on testing for a single gene or small group of related genes (e.g. *BRCA1/2* or the Lynch Syndrome genes), allowing for detailed discussion of specific risks, benefits, and

limitations at the time of informed consent. Because single gene testing has been targeted at patients and families meeting risk criteria, these discussions have also typically occurred within a context of personal or family experience with the cancer in question. Use of multi-gene germline panels increases the likelihood of findings unrelated to the primary reason for testing, and has led to a necessary shift in the informed consent process to a more general discussion of *types* of potential findings, without the assumption that patients will have any experience with the condition identified. Hooker et al. [60] commented on how the technology shift is impacting genetic counseling practice where NGS is employed, noting that patients increasingly are asked to agree to some degree of uncertainty about findings, and to face findings that were not predictable based on personal or family history, with less expectation that their personal experience or feelings about a condition will impact their desire for or against receiving those results. These lessons from clinical germline exome and large panel testing appear to apply equally to patients undergoing tumor testing, where a general discussion about potential findings is relevant. The CSER Consortium has also studied how genetic counselors across consortium sites are adapting the informed consent process to meet patient needs. Their findings suggest that the main challenges include broad scope and uncertainty of results, and unrealistic expectations about the number and utility of results. Genetic counselors in the study focused consent discussion on addressing misperceptions and helping patients develop realistic expectations about types and implications of possible results [61]. One study of pediatric cancer patients found that 17% of parents declined to participate in a sequencing study, citing feeling "overwhelmed by child's diagnosis" as the most common reason, suggesting that the timing of consent discussion is important [62]. Opinions of assenting minors that differ from parents/guardians, seem particularly relevant for results that may have impact on minors when they reach adulthood [58].

8.9 Genetic Service Delivery

The NSGC anticipated an expanding need for genetic services across the clinical spectrum, and formed a Service Delivery Model Task Force as part of strategic initiatives in 2010–2011. The task force found that genetic counselors had already begun to expand the ways they provide services in order to improve access and efficiency, and that there were several variations of "best practice" models, with variables specific to clinic settings determining which model best suits a given situation [63]. In further follow up, they found that the goals centered on improving patient access, such as reducing travel distance and wait time for services, were driving genetic counselors to adapt service delivery [64]. Prior work in developing universal tumor screening programs for Lynch syndrome in all patients with colorectal and endometrial cancers offers useful lessons. These programs complete microsatellite instability testing and/or immunohistochemistry for mismatch repair proteins in tumors without formal pre-test genetic counseling, with follow up

genetics referral for patients with abnormal results suggesting underlying Lynch syndrome [65]. Some hospitals have also successfully implemented collaborative models for germline genetic testing with non-genetics providers providing pre-test consent for most patients, and referring on to genetics providers if there are any complicating factors or PGVs identified [66]. This suggests that brief pre-test communication discussion of potential for PGVs from a non-genetics provider, coupled with formal post-test genetic counseling for select patients may be an efficient model for integrated tumor/germline testing.

8.10 Case Examples

8.10.1 PGVs Can Have Relevance for Cancer Treatment

A 56 year old man with advanced prostate cancer consented to undergo paired tumor/germline sequencing (Fig. 8.1). Although his family history included a sister with breast cancer and a sister with ovarian cancer, neither he nor his family members had ever been referred for genetic counseling or cancer risk genetic testing prior to study participation. He opted to receive germline findings, and testing revealed a PGV in the *BRCA2* gene, along with copy loss of the second *BRCA2* allele in the tumor. This finding had primary relevance to treatment decision making, as the patient was eligible for use of PARP inhibitor therapy. Genetic testing for the *BRCA2* PGV could also be offered to his siblings, children, nieces, and nephews to give them options for risk reduction and prevention strategies.

Increasing evidence suggests that germline testing is indicated for all patients with specific cancer diagnoses including ovarian [29], metastatic prostate [67], and pancreatic [68] where prevalence of PGVs exceeds 10%, family history may not be suggestive, and confirmation of PGVs has potential treatment implications. Integrated sequencing of tumor and germline maximizes potential for medically informative results for these patients.

8.10.2 PGVs May Have Been Previously Missed

Case 2: A 68 year old woman with history of bilateral breast cancer and newly diagnosed primary peritoneal cancer consented to sequencing. She had prior clinical germline testing of the *BRCA1* and *BRCA2* genes, with no PGVs detected. Through integrated tumor/germline testing, she was found to carry a PGV in the *PALB2* (*P*artner *a*nd *L*ocalizer of *B*RCA*2*) gene. This gene has been associated with increased risk for breast cancer [69]. *PALB2* is part of currently available multi-gene panels for breast and ovarian cancer, but would not have been routinely offered clinically at the time the patient had her prior genetic testing. Studies of patients

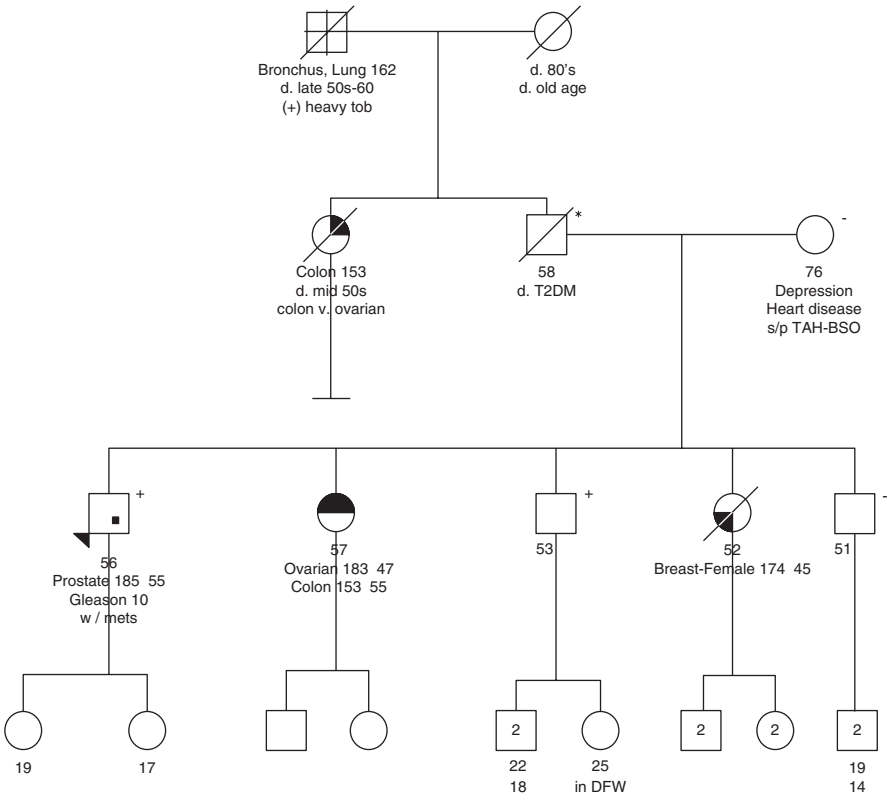

Fig. 8.1 Family history of Case 1. Proband is noted by the arrow. Family members subsequently tested positive for the *BRCA2* PGV are denoted with "+", while those tested negative for the *BRCA2* PGV are noted with "−"

with prior negative testing of *BRCA1/2* have found that 9–11% have PGVs in genes now available on multi-gene panels [70, 71]. While ethical arguments suggest health care providers should re-contact patients who may benefit from new genetic testing technologies, practical and logistical barriers prevent this from happening on a large scale in most practice settings [72]. Integrated tumor/germline sequencing can provide an opportunity to re-evaluate families where prior testing was uninformative.

Data on clinical validity and utility for *PALB2* is evolving, and there is less clarity around utility of PGVs in moderate penetrance genes even when identified through traditional clinical cancer genetic counseling and testing. Multiplex germline testing that includes moderate penetrance genes such as *PALB2* has gained wide acceptance in clinical practice [28] and has been recognized by the NCCN as potentially cost-effective [73]. However, few moderate penetrance genes have evidence based clinical recommendations. The NCCN now includes recommendations for breast cancer screening for *ATM*, *CDH1*, *CHEK2*, and *PALB2,* as well as consideration of risk-reducing salpingo-oophorectomy for *BRIP1*, *RAD51C*, and *RAD51D.*

While none of these moderate penetrance genes are currently included in ACMG recommended list, they are routinely interrogated on clinical panels that could be offered to patients with personal history of cancer, making it reasonable to offer disclosure to patients undergoing tumor sequencing as well.

Variants of uncertain significance represent another challenge in germline analysis. The ACMG working group acknowledged the challenges associated with interpretation of VUS, and subsequently made recommendations for disclosure of only variants classified as pathogenic or likely pathogenic [44]. Proponents of population based screening efforts for the *BRCA1/2* genes have also argued against reporting of VUS findings, given the challenges of interpretation even within high risk populations [74]. Issues of VUS have also expanded in traditional germline genetic testing for cancer predisposition, where multiplex panels are now the norm. Testing more genes at once increases the likelihood of finding a VUS. Few cancer predisposition genes have functional assays to investigate the specific impact of an amino acid change on the protein's function, and computer algorithms for predicting protein effects are not clinically validated. While some of these variants may eventually prove to have clinical relevance, current practice standards argue against reporting of germline VUS in patients undergoing tumor sequencing.

8.10.3 PGVs May Not Have Been Expected

A 54 year old woman with metastatic breast cancer initially diagnosed at age 49 consented to integrated tumor/germline sequencing (Fig. 8.2). The tumor showed a high number of somatic mutations (>1300), and the germline sequencing identified a PGV in the *MLH1* gene causative of Lynch syndrome. The high level of somatic mutations is consistent with mismatch repair deficiency in the tumor, a hallmark feature of Lynch syndrome cancers. Lynch syndrome is best known for causing increased risk for colon and endometrial cancers, neither of which were reported in this family. While the patient met criteria for breast cancer genetic testing, it is unlikely she would have been tested for Lynch syndrome unless a large pan-cancer multi-gene panel was offered. A 2013 literature review found that only one prospective study demonstrated increased breast cancer risk in patients with Lynch syndrome compared to the general population, but that MMR deficiency in breast tumors of patients with Lynch syndrome demonstrated that at least some breast cancer cases were likely caused by the underlying MMR PGV [75]. A subsequent 2015 study found that carriers of *MLH1* PGVs had a lifetime risk of breast cancer of 18.6%, moderately elevated over the general population [76]. Testing for Lynch syndrome, and relevant screening and prevention options, can now be offered to her siblings and other family members.

Traditional approaches to identifying patients with PGVs in cancer predisposition genes are likely to miss many families like this one who could benefit from genetic testing and related options for screening and prevention. Reliance on family history for referral also requires that accurate history be collected and

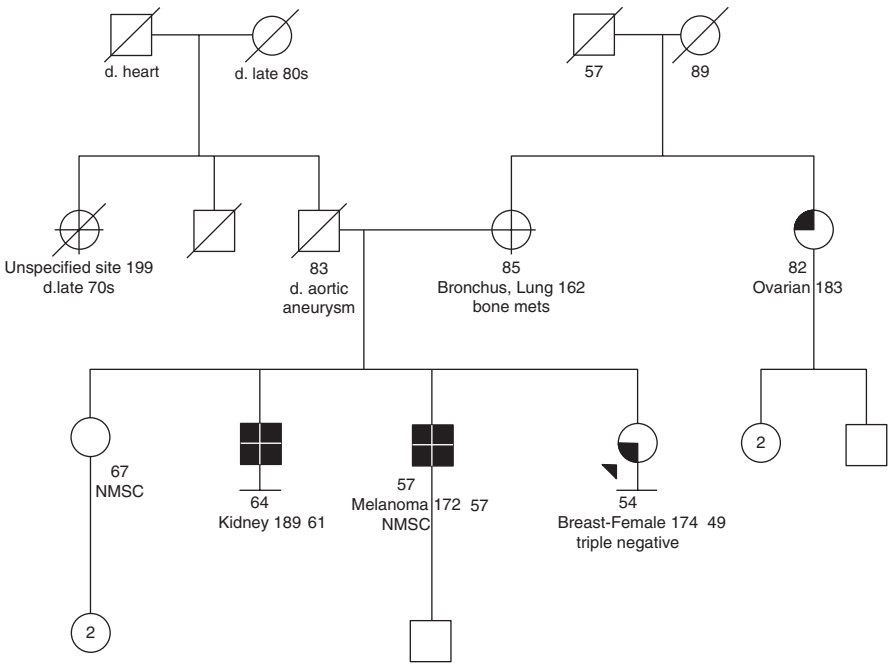

Fig. 8.2 Family history of patient with incidental *MLH1* PGV. The family history is negative for endometrial or colorectal, or any young onset cancer, the hallmarks of hereditary cancer syndromes

updated for patients, which may not always be the case. ASCO completed a Quality Oncology Practice Initiative (QOPI) for cancer family history and genetic counseling and testing referral that found 67–77% of patients with breast or colon cancer in 212 practices had incomplete family history recorded [77]. Even when documented, carriers of PGVs in *BRCA1/2* may not have suggestive family history due to small family size or few female relatives [78]. Patients with suggestive family history are also not always referred appropriately, with the ASCO QOPI study finding that only 43% meeting criteria were actually referred [77]. Integrated tumor/germline sequencing is well accepted by patients and provides an opportunity to identify families who could be missed through traditional genetic counseling and testing approaches.

8.11 Summary

Advances in sequencing technology have led to rapid changes in use of molecular testing in oncology, for both hereditary cancer risk (germline) and tumor (somatic) testing. Integrated tumor/germline testing offers improved ability to distinguish somatic variants that could be relevant to therapeutic options. Despite early

reservations about incidental discovery of germline PGVs through tumor sequencing, research suggests that most patients prefer to receive their germline findings when given an option. PGVs may have relevance for treatment of patients with cancer, as well as for screening and prevention options for their family members. PGVs identified in cancer risk genes during the process of tumor sequencing could be managed following multi-disciplinary models for genetic service delivery already implemented in other settings, with non-genetics providers managing a brief pre-test discussion of the potential for germline findings, followed by detailed post-test genetic counseling for patients found to have PGVs or other complicating factors. Continued research on implementation of integrated tumor/germline sequencing will be needed to ensure that benefits of this testing are fully realized, while minimizing potential risks for patients and families.

References

1. Van Allen EM, Wagle N, Levy MA (2013) Clinical analysis and interpretation of cancer genome data. J Clin Oncol 31(15):1825–1833. https://doi.org/10.1200/JCO.2013.48.7215
2. www.fda.gov [cited 2017 May]
3. Jones S, Anagnostou V, Lytle K, Parpart-Li S, Nesselbush M, Riley DR et al (2015) Personalized genomic analyses for cancer mutation discovery and interpretation. Sci Transl Med 7(283):283ra53. https://doi.org/10.1126/scitranslmed.aaa7161
4. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD et al (2015) PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 372(26):2509–2520. https://doi.org/10.1056/NEJMoa1500596
5. Kaufman B, Shapira-Frommer R, Schmutzler RK, Audeh MW, Friedlander M, Balmana J et al (2015) Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. J Clin Oncol 33(3):244–250. https://doi.org/10.1200/JCO.2014.56.2728
6. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips KA, Mooij TM, Roos-Blom MJ et al (2017) Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. JAMA 317(23):2402–2416. https://doi.org/10.1001/jama.2017.7112
7. Passaperuma K, Warner E, Causer PA, Hill KA, Messner S, Wong JW et al (2012) Long-term results of screening with magnetic resonance imaging in women with BRCA mutations. Br J Cancer 107(1):24–30. https://doi.org/10.1038/bjc.2012.204. PubMed PMID: 22588560; PubMed Central PMCID: PMCPMC3389408
8. Narod SA, Brunet JS, Ghadirian P, Robson M, Heimdal K, Neuhausen SL et al (2000) Tamoxifen and risk of contralateral breast cancer in BRCA1 and BRCA2 mutation carriers: a case-control study hereditary breast cancer clinical study group. Lancet 356(9245):1876–1881
9. Narod SA, Risch H, Moslehi R, Dorum A, Neuhausen S, Olsson H et al (1998) Oral contraceptives and the risk of hereditary ovarian cancer. Hereditary ovarian cancer clinical study group. N Engl J Med 339(7):424–428. https://doi.org/10.1056/NEJM199808133390702
10. Hartmann LC, Sellers TA, Schaid DJ, Frank TS, Soderberg CL, Sitta DL et al (2001) Efficacy of bilateral prophylactic mastectomy in BRCA1 and BRCA2 gene mutation carriers. J Natl Cancer Inst 93(21):1633–1637
11. Kauff ND, Domchek SM, Friebel TM, Robson ME, Lee J, Garber JE et al (2008) Risk-reducing salpingo-oophorectomy for the prevention of BRCA1- and BRCA2-associated breast and gynecologic cancer: a multicenter, prospective study. J Clin Oncol 26(8):1331–1337. https://doi.org/10.1200/JCO.2007.13.9626. PubMed PMID: 18268356; PubMed Central PMCID: PMCPMC3306809
12. Domchek SM, Friebel TM, Singer CF, Evans DG, Lynch HT, Isaacs C et al (2010) Association of risk-reducing surgery in BRCA1 or BRCA2 mutation carriers with cancer risk and mortality.

JAMA 304(9):967–975. https://doi.org/10.1001/jama.2010.1237. PubMed PMID: 20810374; PubMed Central PMCID: PMCPMC2948529
13. Ben David Y, Chetrit A, Hirsh-Yechezkel G, Friedman E, Beck BD, Beller U et al (2002) Effect of BRCA mutations on the length of survival in epithelial ovarian tumors. J Clin Oncol 20(2):463–466. https://doi.org/10.1200/JCO.2002.20.2.463
14. Golan T, Kanji ZS, Epelbaum R, Devaud N, Dagan E, Holter S et al (2014) Overall survival and clinical characteristics of pancreatic cancer in BRCA mutation carriers. Br J Cancer 111(6):1132–1138. https://doi.org/10.1038/bjc.2014.418. PubMed PMID: 25072261; PubMed Central PMCID: PMCPMC4453851
15. Pomerantz MM, Spisak S, Jia L, Cronin AM, Csabai I, Ledet E et al (2017) The association between germline BRCA2 variants and sensitivity to platinum-based chemotherapy among men with metastatic prostate cancer. Cancer 123:3532–3539. https://doi.org/10.1002/cncr.30808
16. Fong PC, Boss DS, Yap TA, Tutt A, Wu P, Mergui-Roelvink M et al (2009) Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med 361(2):123–134. https://doi.org/10.1056/NEJMoa0900212
17. Fong PC, Yap TA, Boss DS, Carden CP, Mergui-Roelvink M, Gourley C et al (2010) Poly(ADP)-ribose polymerase inhibition: frequent durable responses in BRCA carrier ovarian cancer correlating with platinum-free interval. J Clin Oncol 28(15):2512–2519. https://doi.org/10.1200/JCO.2009.26.9589
18. Jarvinen HJ, Aarnio M, Mustonen H, Aktan-Collan K, Aaltonen LA, Peltomaki P et al (2000) Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. Gastroenterology 118(5):829–834
19. Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM et al (2003) Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. N Engl J Med 349(3):247–257. https://doi.org/10.1056/NEJMoa022289. PubMed PMID: 12867608; PubMed Central PMCID: PMCPMC3584639
20. Riley BD, Culver JO, Skrzynia C, Senter LA, Peters JA, Costalas JW et al (2012) Essential elements of genetic cancer risk assessment, counseling, and testing: updated recommendations of the National Society of genetic counselors. J Genet Couns 21(2):151–161. https://doi.org/10.1007/s10897-011-9462-x
21. Ladabaum U, Wang G, Terdiman J, Blanco A, Kuppermann M, Boland CR et al (2011) Strategies to identify the Lynch syndrome among patients with colorectal cancer: a cost-effectiveness analysis. Ann Intern Med 155(2):69–79. https://doi.org/10.7326/0003-4819-155-2-201107190-00002. PubMed PMID: 21768580; PubMed Central PMCID: PMCPMC3793257
22. (1996) Statement of the American Society of Clinical Oncology: genetic testing for cancer susceptibility, Adopted on February 20, 1996. J Clin Oncol 14(5):1730–1736. discussion 7–40
23. Daly MB, Axilbund JE, Buys S, Crawford B, Farrell CD, Friedman S et al (2010) Genetic/familial high-risk assessment: breast and ovarian. J Natl Compr Cancer Netw 8(5):562–594
24. Robson ME, Bradbury AR, Arun B, Domchek SM, Ford JM, Hampel HL et al (2015) American Society of Clinical Oncology policy statement update: genetic and genomic testing for cancer susceptibility. J Clin Oncol 33(31):3660–3667. https://doi.org/10.1200/JCO.2015.63.0996
25. Trepanier A, Ahrens M, McKinnon W, Peters J, Stopfer J, Grumet SC et al (2004) Genetic cancer risk assessment and counseling: recommendations of the national society of genetic counselors. J Genet Couns 13(2):83–114. https://doi.org/10.1023/B:JOGC.0000018821.48330.77
26. Association for Molecular Pathology et al v. Myriad Genetics Inc et al (2013) Supreme Court of the United States
27. Cook-Deegan R, Niehaus A (2014) After myriad: genetic testing in the wake of recent supreme court decisions about gene patents. Curr Genet Med Rep 2(4):223–241. https://doi.org/10.1007/s40142-014-0055-5. PubMed PMID: 25401053; PubMed Central PMCID: PMCPMC4225052
28. Kurian AW, Hare EE, Mills MA, Kingham KE, McPherson L, Whittemore AS et al (2014) Clinical evaluation of a multiple-gene sequencing panel for hereditary cancer risk assessment. J Clin Oncol 32(19):2001–2009. https://doi.org/10.1200/JCO.2013.53.6607

29. Walsh T, Casadei S, Lee MK, Pennil CC, Nord AS, Thornton AM et al (2011) Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. Proc Natl Acad Sci U S A 108(44):18032–18037. https://doi.org/10.1073/pnas.1115052108
30. Yurgelun MB, Allen B, Kaldate RR, Bowles KR, Judkins T, Kaushik P et al (2015) Identification of a variety of mutations in cancer predisposition genes in patients with suspected Lynch syndrome. Gastroenterology 149(3):604–13 e20. https://doi.org/10.1053/j.gastro.2015.05.006
31. Pearlman R, Frankel WL, Swanson B, Zhao W, Yilmaz A, Miller K et al (2016) Prevalence and Spectrum of germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. JAMA Oncol 3:464. https://doi.org/10.1001/jamaoncol.2016.5194
32. Grant RC, Selander I, Connor AA, Selvarajah S, Borgida A, Briollais L et al (2015) Prevalence of germline mutations in cancer predisposition genes in patients with pancreatic cancer. Gastroenterology 148(3):556–564. https://doi.org/10.1053/j.gastro.2014.11.042
33. Roychowdhury S, Iyer MK, Robinson DR, Lonigro RJ, Wu YM, Cao X et al (2011) Personalized oncology through integrative high-throughput sequencing: a pilot study. Sci Transl Med 3(111):111ra21. https://doi.org/10.1126/scitranslmed.3003161
34. Garraway LA, Lander ES (2013) Lessons from the cancer genome. Cell 153(1):17–37. https://doi.org/10.1016/j.cell.2013.03.002
35. LaDuca H, Stuenkel AJ, Dolinsky JS, Keiles S, Tandy S, Pesaran T et al (2014) Utilization of multigene panels in hereditary cancer predisposition testing: analysis of more than 2,000 patients. Genet Med 16(11):830–837. https://doi.org/10.1038/gim.2014.40. PubMed PMID: 24763289; PubMed Central PMCID: PMCPMC4225457
36. Lowstuter KEC, Sturgeon D, Ricker C, Karam R, LaDuca H, Culver JO, Chao E, Sturgeon J, Speare V, Ma Y, Kingham K, Melas M, Idos GE, McDonnell KJ, Gruber SB (2017) Unexpected CDH1 mutations identified on multigene panels pose clinical management challenges. Precis Oncol 1:1–12. https://doi.org/10.1200/PO.16.00021
37. Yurgelun MB, Masciari S, Joshi VA, Mercado RC, Lindor NM, Gallinger S et al (2015) Germline TP53 mutations in patients with early-onset colorectal cancer in the colon cancer family registry. JAMA Oncol 1(2):214–221. https://doi.org/10.1001/jamaoncol.2015.0197. PubMed PMID: 26086041; PubMed Central PMCID: PMCPMC4465271
38. Whitworth J, Skytte AB, Sunde L, Lim DH, Arends MJ, Happerfield L et al (2016) Multilocus inherited neoplasia alleles syndrome: a case series and review. JAMA Oncol 2(3):373–379. https://doi.org/10.1001/jamaoncol.2015.4771
39. Domchek SM, Bradbury A, Garber JE, Offit K, Robson ME (2013) Multiplex genetic testing for cancer susceptibility: out on the high wire without a net? J Clin Oncol 31(10):1267–1270. https://doi.org/10.1200/JCO.2012.46.9403
40. Bombard Y, Robson M, Offit K (2013) Revealing the incidentalome when targeting the tumor genome. JAMA 310(8):795–796. https://doi.org/10.1001/jama.2013.276573. Epub 2013/08/29
41. Stadler ZK, Schrader KA, Vijai J, Robson ME, Offit K (2014) Cancer genomics and inherited risk. J Clin Oncol 32(7):687–698. https://doi.org/10.1200/JCO.2013.49.7271
42. Raymond VM, Gray SW, Roychowdhury S, Joffe S, Chinnaiyan AM, Parsons DW et al (2016) Germline findings in tumor-only sequencing: points to consider for clinicians and laboratories. J Natl Cancer Inst 108:4. https://doi.org/10.1093/jnci/djv351
43. Catenacci DV, Amico AL, Nielsen SM, Geynisman DM, Rambo B, Carey GB et al (2015) Tumor genome analysis includes germline genome: are we ready for surprises? Int J Cancer 136(7):1559–1567. https://doi.org/10.1002/ijc.29128
44. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL et al (2013) ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med 15(7):565–574.. Epub 2013/06/22. https://doi.org/10.1038/gim.2013.73
45. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP et al (2016) Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med 19:249–255. https://doi.org/10.1038/gim.2016.190

46. Burke W, Antommaria AH, Bennett R, Botkin J, Clayton EW, Henderson GE et al (2013) Recommendations for returning genomic incidental findings? We need to talk! Genet Med 15(11):854–859. https://doi.org/10.1038/gim.2013.113
47. Parsons DW, Roy A, Plon SE, Roychowdhury S, Chinnaiyan AM (2014) Clinical tumor sequencing: an incidental casualty of the American College of Medical Genetics and Genomics recommendations for reporting of incidental findings. J Clin Oncol 32(21):2203–2205. https://doi.org/10.1200/JCO.2013.54.8917
48. Schrader KA, Cheng DT, Joseph V, Prasad M, Walsh M, Zehir A et al (2016) Germline variants in targeted tumor sequencing using matched Normal DNA. JAMA Oncol 2(1):104–111. https://doi.org/10.1001/jamaoncol.2015.5208
49. Seifert BA, O'Daniel JM, Amin K, Marchuk DS, Patel NM, Parker JS et al (2016) Germline analysis from tumor-germline sequencing dyads to identify clinically actionable secondary findings. Clin Cancer Res 22(16):4087–4094. https://doi.org/10.1158/1078-0432.CCR-16-0015
50. Meric-Bernstam F, Brusco L, Daniels M, Wathoo C, Bailey AM, Strong L et al (2016) Incidental germline variants in 1000 advanced cancers on a prospective somatic genomic profiling protocol. Ann Oncol 27(5):795–800. https://doi.org/10.1093/annonc/mdw018
51. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM et al (2015) Integrative clinical genomics of advanced prostate cancer. Cell 161(5):1215–1228. https://doi.org/10.1016/j.cell.2015.05.001
52. Robinson DR, Wu Y, Lonigro RJ, Vat P, Cobain E, Everett J et al (2017) Integrative clinical genomics of metastatic cancer. Nature 548:297–303. https://doi.org/10.1038/nature23306
53. Mody RJ, Wu YM, Lonigro RJ, Cao X, Roychowdhury S, Vats P et al (2015) Integrative clinical sequencing in the Management of Refractory or relapsed cancer in youth. JAMA 314(9):913–925. https://doi.org/10.1001/jama.2015.10080
54. Parsons DW, Roy A, Yang Y, Wang T, Scollon S, Bergstrom K et al (2016) Diagnostic yield of clinical tumor and germline whole-exome sequencing for children with solid tumors. JAMA Oncol 2:616. https://doi.org/10.1001/jamaoncol.2015.5699
55. Zhang J, Walsh MF, Wu G, Edmonson MN, Gruber TA, Easton J et al (2015) Germline mutations in predisposition genes in pediatric cancer. N Engl J Med 373(24):2336–2346. https://doi.org/10.1056/NEJMoa1508054
56. Tomlinson AN, Skinner D, Perry DL, Scollon SR, Roche MI, Bernhardt BA (2016) "Not tied up neatly with a bow": professionals' challenging cases in informed consent for genomic sequencing. J Genet Couns 25(1):62–72. https://doi.org/10.1007/s10897-015-9842-8
57. ABo D (2015) ACMG policy statement: updated recommendations regarding analysis and reporting of secondary findings in clinical genome-scale sequencing. Genet Med 17(1):68–69. https://doi.org/10.1038/gim.2014.151
58. Everett JN, Gustafson SL, Raymond VM (2014) Traditional roles in a non-traditional setting: genetic counseling in precision oncology. J Genet Couns 23(4):655–660. https://doi.org/10.1007/s10897-014-9698-3
59. Gray SW, Park ER, Najita J, Martins Y, Traeger L, Bair E et al (2016) Oncologists' and cancer patients' views on whole-exome sequencing and incidental findings: results from the CanSeq study. Genet Med 18(10):1011–1019. https://doi.org/10.1038/gim.2015.207
60. Hooker GW, Ormond KE, Sweet K, Biesecker BB (2014) Teaching genomic counseling: preparing the genetic counseling workforce for the genomic era. J Genet Couns 23(4):445–451. https://doi.org/10.1007/s10897-014-9689-4
61. Bernhardt BA, Roche MI, Perry DL, Scollon SR, Tomlinson AN, Skinner D (2015.;167A) Experiences with obtaining informed consent for genomic sequencing. Am J Med Genet A 167(11):2635–2646. https://doi.org/10.1002/ajmg.a.37256
62. Scollon S, Bergstrom K, Kerstein RA, Wang T, Hilsenbeck SG, Ramamurthy U et al (2014) Obtaining informed consent for clinical tumor and germline exome sequencing of newly diagnosed childhood cancer patients. Genome Med 6(9):69. https://doi.org/10.1186/s13073-014-0069-3

63. Cohen SA, Marvin ML, Riley BD, Vig HS, Rousseau JA, Gustafson SL (2013) Identification of genetic counseling service delivery models in practice: a report from the NSGC Service delivery model task force. J Genet Couns 22(4):411–421. https://doi.org/10.1007/s10897-013-9588-0
64. Cohen SA, Huziak RC, Gustafson S, Grubs RE (2016) Analysis of advantages, limitations, and barriers of genetic counseling service delivery models. J Genet Couns 25(5):1010–1018. https://doi.org/10.1007/s10897-016-9932-2
65. Vig HS, Wang C (2012) The evolution of personalized cancer genetic counseling in the era of personalized medicine. Familial Cancer 11(3):539–544. https://doi.org/10.1007/s10689-012-9524-8. PubMed PMID: 22419176; PubMed Central PMCID: PMCPMC3905734
66. Cohen SA, McIlvried D, Schnieders J (2009) A collaborative approach to genetic testing: a community hospital's experience. J Genet Couns 18(6):530–533. https://doi.org/10.1007/s10897-009-9243-y
67. Pritchard CC, Mateo J, Walsh MF, De Sarkar N, Abida W, Beltran H et al (2016) Inherited DNA-repair gene mutations in men with metastatic prostate cancer. N Engl J Med 375(5):443–453. https://doi.org/10.1056/NEJMoa1603144
68. Hu C, Hart SN, Bamlet WR, Moore RM, Nandakumar K, Eckloff BW et al (2016) Prevalence of pathogenic mutations in cancer predisposition genes among pancreatic cancer patients. Cancer Epidemiol Biomark Prev 25(1):207–211. https://doi.org/10.1158/1055-9965.EPI-15-0455
69. Antoniou AC, Casadei S, Heikkinen T, Barrowdale D, Pylkas K, Roberts J et al (2014) Breast-cancer risk in families with mutations in PALB2. N Engl J Med 371(6):497–506. https://doi.org/10.1056/NEJMoa1400382
70. Moran O, Nikitina D, Royer R, Poll A, Metcalfe K, Narod SA et al (2017) Revisiting breast cancer patients who previously tested negative for BRCA mutations using a 12-gene panel. Breast Cancer Res Treat 161(1):135–142. https://doi.org/10.1007/s10549-016-4038-y
71. Yadav S, Reeves A, Campian S, Paine A, Zakalik D (2017) Outcomes of retesting BRCA negative patients using multigene panels. Familial Cancer 16(3):319–328. https://doi.org/10.1007/s10689-016-9956-7
72. Otten E, Plantinga M, Birnie E, Verkerk MA, Lucassen AM, Ranchor AV et al (2015) Is there a duty to recontact in light of new genetic technologies? A systematic review of the literature. Genet Med 17(8):668–678. https://doi.org/10.1038/gim.2014.173
73. Daly MB, Pilarski R, Axilbund JE, Berry M, Buys SS, Crawford B et al (2016) Genetic/familial high-risk assessment: breast and ovarian, version 2.2015. J Natl Compr Cancer Netw 14(2):153–162
74. King MC, Levy-Lahad E, Lahad A (2014) Population-based screening for BRCA1 and BRCA2: 2014 Lasker award. JAMA 312(11):1091–1092. https://doi.org/10.1001/jama.2014.12483
75. Win AK, Lindor NM, Jenkins MA (2013) Risk of breast cancer in Lynch syndrome: a systematic review. Breast Cancer Res 15(2):R27. https://doi.org/10.1186/bcr3405. PubMed PMID: 23510156; PubMed Central PMCID: PMCPMC3672741
76. Harkness EF, Barrow E, Newton K, Green K, Clancy T, Lalloo F et al (2015) Lynch syndrome caused by MLH1 mutations is associated with an increased risk of breast cancer: a cohort study. J Med Genet 52(8):553–556. https://doi.org/10.1136/jmedgenet-2015-103216. PubMed PMID: 26101330
77. Wood ME, Kadlubek P, Pham TH, Wollins DS, Lu KH, Weitzel JN et al (2014) Quality of cancer family history and referral for genetic counseling and testing among oncology practices: a pilot test of quality measures as part of the american society of clinical oncology quality oncology practice initiative. J Clin Oncol 32(8):824–829. Epub 2014/02/05. https://doi.org/10.1200/JCO.2013.51.4661
78. Weitzel JN, Lagos VI, Cullinane CA, Gambol PJ, Culver JO, Blazer KR et al (2007) Limited family structure and BRCA gene mutation status in single cases of breast cancer. JAMA 297(23):2587–2595. https://doi.org/10.1001/jama.297.23.2587. PubMed PMID: 17579227

Chapter 9
Ethical, Legal, and Social Implications of Precision Cancer Medicine

J. Scott Roberts and Michele R. Gornick

Keywords Ethics · Informed consent · Patient education · Health communication · Distributive justice · Resource allocation · Precision oncology · Return of results · Risk disclosure · Health disparities

9.1 Introduction

Like many other emerging biomedical technologies, precision cancer medicine (or precision oncology) poses a broad range of ethical, legal, and social implications (ELSI). These ELSI issues are present across multiple levels of health care, research, and social policy. For example, at the individual (patient) level, use of next-generation sequencing (NGS) presents challenges for securing truly informed consent and honoring patient preferences for the many different types of genetic information potentially yielded by NGS. At the institutional level, various stakeholders (e.g., professional organizations in genetics and oncology, test laboratories, clinical trials networks) have vested interests in deciding which types of sequencing results will be analyzed and returned to patients or research participants, and how the clinical implications of results will be managed. Finally, at a broader societal level, we must wrestle with questions of resource allocation and priority setting. Precision oncology represents a promising avenue for breakthroughs in cancer therapies, but what proportion of public funds should be earmarked for this approach versus other domains of cancer treatment, where significant disparities in health care access and outcomes still persist? In this chapter, we will address such ELSI issues, drawing upon both empirical research and ethical and policy commentary in the field.

J. S. Roberts (✉)
Department of Health Behavior & Health Education, University of Michigan School of Public Health, Ann Arbor, MI, USA
e-mail: jscottr@umich.edu

M. R. Gornick
Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, USA

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_9

9.2 Precision Cancer Medicine and Distributive Justice

9.2.1 Resource Allocation

Precision medicine has generated much excitement about its promise for advancing cancer care and improving patient outcomes. This enthusiasm was exemplified by President Obama in his 2015 State of the Union address announcing the launch of a national Precision Medicine Initiative (PMI), where the President commented, "Doctors have always tried to tailor their treatments as best they can to individuals. You can match a blood transfusion to a blood type—that was an important discovery. What if matching a cancer cure to our genetic code was just as easy, just as standard? What if figuring out the right dose of medicine was as simple as taking our temperature? [1]" At approximately the same time, prominent leaders at the National Institutes of Health (NIH), including NIH Director Francis Collins and National Cancer Institute (NCI) Director Harold Varmus, provided the rationale for a focus on precision oncology as a top PMI priority in a highly publicized commentary in the *New England Journal of Medicine* [2]. They noted that the PMI had been designed to address numerous potential barriers in oncology practice—unexplained drug resistance, genomic heterogeneity of tumors, insufficient means for monitoring treatment responses and tumor recurrence, and limited knowledge about the use of drug combinations—and that better tailoring of therapies based on genomic profiling would be a likely benefit of this large public research investment.

However, with its reliance on advanced, expensive technologies and highly specialized clinical and laboratory expertise, precision medicine—and precision oncology in particular—are not currently feasible options for many patient populations in the US. Precision medicine's focus on tertiary care instead of prevention and its relative lack of attention to broader social determinants of health (e.g., poverty, environmental stressors) have generated skepticism among some commentators about its likely return on investment, particularly when taking a population health perspective. For example, notable public health leaders Bayer and Galea (2015) have faulted the NIH's significant allocation of resources to the PMI as an example of misplaced priorities [3]. The authors concluded,

> "Enthusiasm [for the PMI]…derives from the assumption that precision medicine will contribute to clinical practice and thereby advance the health of the public. However, this enthusiasm is premature…if one views the world from the perspective of the broad pattern of morbidity and mortality, if one is concerned about why the US has sunk to the bottom of the list of comparable countries in terms of disease experience and life expectancy, or if one is troubled by the steep social gradient that characterizes who becomes sick and who dies. The burgeoning precision-medicine agenda is largely silent on these issues, focusing instead on detecting and curing disease at the individual level."

From this perspective, the significant allotment of public resources into the PMI—$215 million in its first year, with $70 million going to NCI [4]—could represent a missed opportunity to pursue investments focused on a public health (vs. biomedical) approach to addressing current and future cancer cases. Few precision

oncology interventions to date have demonstrated high levels of clinical utility, let alone cost-effectiveness. Of course, the field is still in its early stages, and thus such criticisms may be subject to revision as data emerges on the potential for precision oncology to improve both patient and population outcomes. Nevertheless, given other examples where leading proponents of genomic medicine have arguably over-promised on its benefits, it is likely that skeptics may continue to question the value of precision oncology in terms of distributive justice: that is, whether its benefits and burdens are likely to be spread equally across relevant patient populations [5].

9.2.2 Health Disparities

Another justice-related concern regarding precision medicine involves health disparities and inequities. Unequal access to precision oncology and its requisite clinical sequencing technologies may occur across a variety of patient and practice factors, including patient race, ethnicity, and socioeconomic status, and area-level resources (as represented by the concept of medically-underserved areas, or MUAs). Furthermore, type of medical practice—community vs. academic—is already known to be associated with differential access to advances in genomic medicine [6]. These barriers are certainly not unique to precision oncology but will need to be addressed if its envisioned future benefits are to be shared equitably across all segments of the cancer patient population.

Racial/ethnic disparities have been observed across multiple cancers [7]. Such disparities represent a social justice concern in that they are potentially avoidable, unnecessary, and unfair, and often attributable to modifiable factors beyond individual-level control. These factors include differential exposure to environmental risk factors, unequal insurance coverage and access to care, and implicit racial biases among health care providers [8]. It remains to be seen whether precision oncology will help address these disparities or unwittingly exacerbate them. One barrier to advancing knowledge in racial/ethnic minorities is their lack of adequate representation in large cancer genetics databases. For example, a recent analysis of The Cancer Genome Atlas (TCGA) found there were too few patients from various racial/ethnic minority groups (e.g., African Americans, Hispanics/Latinos) to detect even moderately common genomic alterations [9]. This lack of representation in research studies has potential implications for clinical practice, as racial/ethnic minority patients may be more likely to receive uninformative (or even erroneous) interpretation of their existing genomic variants, resulting in distressing uncertainty and/or suboptimal treatment. Oversampling of minority populations is therefore essential to develop a similar knowledge base as among non-Hispanic whites. To its credit, NIH has incentivized such sampling approaches within their Clinical Sequencing Exploratory Research (CSER) [10], a nationwide network of coordinated sites examining the clinical utility of NGS across a diverse range of patient populations and medical conditions. Funded sites are charged with recruiting patient samples consisting in some cases of at least 60% underrepresented minorities.

A second barrier to equitable expansion of precision medicine in racial/ethnic minority groups and patients of lower socioeconomic status is potential provider bias within patient-physician interactions. Physician recommendations regarding genomic sequencing have been recognized as a key step in the clinical process, but patient-provider communication about genomic sequencing and the possibility of racial and socioeconomic disparities in outcomes resulting from such communications (e.g., referrals for specialty services) have been understudied [11]. From multiple studies, we know that physician perceptions can be influenced, sometimes unconsciously, by patient race and socioeconomic status; these perceptions can include assumptions about patients' level of intelligence, personality traits, and likelihood of compliance with medical advice and complicated regimens [12–14]. Such assumptions may decrease the likelihood that physicians offer genomic sequencing technologies to their minority patients or patients with lower levels of educational attainment. Furthermore, physician concerns about minority and medically underserved patients' distrust of medical professionals or in medical research could also potentially make them less likely to offer sequencing to these patient populations. These assumptions, however, are not consistent with recent studies suggesting that minority group patients often do want to contribute their tumors for biospecimen studies [15] and to participate in clinical trials [12]. Education of physicians regarding the potential for implicit biases in their practices may help address the unintended consequences these biases can have on patient-provider interactions and clinical decision making.

Finally, the practice setting itself may help determine the level of available medical resources and access to precision medicine. Most clinics currently offering precision oncology treatments are located in major urban areas and/or large academic medical centers, posing potential barriers to rural populations who cannot afford the travel and lodging costs that might be incurred when trying to access such clinics. Even among physicians within academic practices, familiarity with (as well as attitudes toward) use of genomic technologies can vary widely [16]. In addition, genetic counselors—key providers of education and support for patients considering and undergoing genomic sequencing—are relatively few in number nationwide and uncommonly employed by community practices [6]. Moreover, a lack of established tumor boards and competing demands on oncologists' time decrease the likelihood that patients seen in community oncology practices will have equal access to advances in precision medicine, at least in the short term. Significant investment in the community health infrastructure may be necessary to fully take advantage of treatment advances in precision oncology.

9.3 Informed Consent and Patient Education

Obtaining informed consent is a standard requirement for patient participation in clinical trials. When individuals give their truly informed and well-considered consent, they should understand the procedure or study purpose, process, risks, and

benefits, as well as participation requirements and alternatives. Many of the ethical concerns regarding the integrity of the informed consent process in cancer care, such as challenges to patient understanding, knowledge, and ability to recall information about the study, are not unique to precision oncology [17]. Neither are the numerous communication issues associated with the readability, length, format, and language used in written consent documents [18–21].

9.3.1 Further Considerations in Precision Cancer Studies

However, there are several additional issues that require further consideration when consenting individuals to precision oncology research studies or clinical trials involving genomic information. Challenges include obtaining consent for studies involving a broad scope of risk information, as well as addressing the likely uncertainty of many sequencing results [22, 23]. For example, when sequencing an individual's genome, in essence a genome-wide disease screening is potentially being performed. This analysis opens the door not only to information related to the reason an individual is having his/her genome sequenced, but also information unrelated to the reason for testing (i.e., secondary findings). How to manage these secondary findings has become a matter of great debate within the bioethics and genome medicine communities.

Other concerns that merit further consideration when consenting individuals participating in precision oncology studies are the disclosure of individual level research results and addressing potentially unrealistic expectations of participants about the usefulness of genetic information [24, 25]. Patient misunderstandings in the research context often involve beliefs about the likelihood of direct study benefits (i.e. the "therapeutic misconception") [26–28]. The communication of individual research results in precision medicine studies can conflate research practice and clinical care, particularly if the patient's physician is also part of the research team. For example, participants might believe that genomic sequencing would not be offered if it did not carry the promise of clinical benefit, such as access to a new drug or treatment [28, 29]. Exploring participants' understanding of the purpose of research and their motivations for participation has been proposed as an avenue for measuring and preventing such misconceptions [30].

Concerns about whether our current models of informed consent [31, 32] in clinical research are adequate for genomics-related studies—particularly regarding safeguards for participants' confidentiality and personal autonomy—have been widely debated [33–36]. Issues raised regarding confidentiality include the potential for coded genomic data to be re-identified, and for breaches of private genetic information to enable employment or insurability discrimination—despite the Genetic Information Nondiscrimination Act of 2008 (GINA) [37], a federal law that protects individuals from genetic discrimination in health insurance and employment. Concerns about personal autonomy relate to participants' desires to have a choice about the types of genetic information they will receive, rights to their genetic infor-

mation, and the release of personal information to family members. Because of these concerns, alternative models for obtaining informed consent in the context of genomics research have been proposed. For example, the Informed Cohort Model allows participants to designate their preferences for which research results to receive, but has an added layer of ethical oversight by a governing body (e.g., Informed Cohort Oversight Board). Several precision medicine initiatives have implemented this model, including the Coriell Personalized Medicine Initiative and Boston Children's Hospital [38]. Other alternatives include a preference-setting model that allows participants to express their preferences for the return of individual genomic research results [39] via a "flexible-default" model where a "default" recommendation is provided based on clinical evidence and participants have the "flexibility" in certain cases to agree or disagree with the recommendation [32, 40]. A "tiered-binned" genetic counseling approach for informed consent and genetic counseling for multiplex testing has also been proposed. In this model, Tier 1 "indispensable" information is presented to all patients. In Tier 2, more specific information is provided to support the different informational needs among diverse patient populations. Clinically relevant information is "binned" into groups in order to minimize information overload, support informed decision making, and facilitate adaptive responses to testing [41].

Given the potential privacy risks noted above, several practical protections should be considered. These include (1) sharing information from identified or identifiable samples only with the participants and the persons designated by the participants (e.g., next of kin); (2) providing participants who participated in genomic studies as children (with consent from parents or legal caregivers) the opportunity to be informed about clinically significant genomic findings as adults when they reach the age of majority; (3) requirement of data use agreements and routine use of data access tracking systems to reduce the risk of re-identification through data sources like the electronic health record, and (4) greater attention to safeguards in studies involving vulnerable populations, including children, educationally disadvantaged groups, and groups where genomic findings could pose stigma issues [42].

9.3.2 Issues in Health Communication

The consent process in precision medicine also faces potential complications given the varying levels of genetic literacy in the general population and the inherent difficulties involved in conveying complex scientific concepts to laypersons [22, 43]. An example of a genetics concept that is often difficult for patients to fully understand is the concept of low or incomplete penetrance. Although penetrance is simply the proportion of individuals carrying a particular variant of a gene who also express a given trait, it quickly becomes confusing where a gene or genetic trait is expressed in only a subset of the population who carry the gene(s) in question. How an

individual can have an alteration in a known cancer related gene, but not develop cancer can be a challenging concept to convey given common public beliefs about genetic determinism. Penetrance is a challenging concept to communicate not only because the term itself is unfamiliar to most people, but also because it involves statistical probability. Indeed, participants have been found to misunderstand several aspects of statistics as related to genetic data, particularly if they are of low numeracy levels. For example, laypersons may encounter difficulties in grasping the probabilities involved in future disease risk, and they often overestimate their own level of personal risk, particularly if they have a positive family history for the disease in question. Patients are also prone to overestimation of the occurrence of heritable cancer syndromes due to germline variants, which may affect health care decisions and lead to inappropriate use of prevention and surveillance [44, 45]. Addressing this potential misconception may involve correcting patients' common misunderstandings about disease mechanisms. Finally, communication of risk statistics for cancer predisposition can be challenging even when only a single gene or variant is involved. Precision oncology further complicates matters by generating a wide variety of both germline and tumor findings.

Another health communication challenge for precision medicine is managing patient and participant expectations amidst the "hype" often surrounding genomic technologies. Precision medicine as described in the news media and through direct-to-consumer advertising is often viewed as the "wave of the future," with particularly great promise for improving patient outcomes. This portrayal may lead some patients and research participants to overestimate the benefits associated with sequencing, necessitating further clarification during the consent process about what a given study or procedure does and does not offer. Our own work in this area suggests that cancer patients taking part in clinical sequencing projects may indeed be prone to overrate the likelihood of direct study benefits. Roberts, JS, Gornick, MC, Le, LQ, Bartnik, NJ, Zikmund-Fisher, BJ, Chinnaiyan, AM; for the MI-ONCOSEQ Study team. Next-generation sequencing in precision oncology: Patient understanding and expectations. Cancer Med. 2019; 8: 227– 237. https://doi.org/10.1002/cam4.1947 For example, they may have unrealistically high expectations of learning clinically significant personal genomic results, and they may not fully appreciate the many potential barriers to becoming eligible for a clinical trial by virtue of their sequencing results.

Potential solutions to these challenges include assessing and addressing misperceptions during the consent process, helping participants set realistic expectations about the types of findings to be generated, and facilitating shared decision making amongst health care providers, participants and family members. Strategies to improve communication during the consent process include use of plain language, clear and concise presentation of study objectives, discussion with a study team member or health care provider, and disclosure of controversial information [46–50]. Empirical studies have suggested numerous proven strategies for improving risk communication and enhancing patient decision making; these include (1) presenting absolute risks instead of relative risks, (2) clarifying how treatment changes risks from preexisting baseline level, and (3) providing consent materials written in

a simplified manner so that even people of low literacy (eighth grade or lower level) can understand them [51].

9.3.3 Consenting Pediatric Participants

There are also unique ELSI concerns in the consent process when involving pediatric patients in precision oncology. These include the storage of biospecimens for future research, consideration of re-consenting participants when they reach the age of majority, and incorporating an adolescent's preferences about parental access to genomic test results [52, 53]. In precision oncology, one challenge is determining when and how information will be shared about a child's or adolescent's inherited predisposition to an adult-onset cancer syndrome, which may be uncovered as part of a genomic sequencing study.

9.4 Return of Sequencing Results

As alluded to earlier, a prominent ethical dilemma in precision medicine is deciding how much choice patients should have over how their whole-genome sequencing analyses are conducted and what types of results are returned to them following sequencing. In precision oncology, sequencing can yield not only tumor profiles, but also germline results of potential interest to both patients and their blood relatives [54]. With regard to the latter, a wide range of findings could be generated that would be of potential medical significance, including those related to conditions other than cancer (i.e., secondary findings). The American College of Medical Genetics and Genomics (ACMG) has issued recommendations for a minimum list of genes to be routinely reported as secondary findings anytime clinical exome or genome sequencing is conducted. This recently updated list [55] consists of 59 genes associated with high risks for medical conditions in which established interventions exist for disease prevention or management. These genes confer risks for a wide range of conditions including hereditary cancer syndromes of both child and adult onset, various cardiomyopathies, and even malignant hyperthermia. In best-case scenarios, genetic risks might be identified that allow for effective disease prevention (e.g., preventive surgeries for BRCA 1/2 carriers) for individuals and cascade screening for family members (e.g., first-degree relatives of patients with Lynch syndrome). However, even with this guidance, challenges can arise for providers, patients, and clinical laboratories in determining what genetic variants should be tested for and who should have access to what results. Disclosure of genomic information can bring with it unintended consequences, including misunderstanding of results and, in rare cases, genetic discrimination. For example, while federal laws such as GINA provide protections against discrimination by health insurers and employers (e.g., using genetic risk status to inform hiring and coverage

decisions), these protections do not extend to life, disability, and long-term care insurance. In the following sections, we consider various scenarios where ethical dilemmas may be particularly likely to arise.

9.4.1 Disclosing Risks for Adult-Onset Conditions in Childhood

Preliminary findings suggest that the clinical utility of precision oncology approaches may be promising in pediatric populations [56]. Clinical use of NGS in children would then necessitate deciding whether germline findings associated with adult-onset cancers (or other conditions) should be sought. The ACMG guidelines assert that findings suggestive of genetic risks for adult-onset conditions should be routinely returned, regardless of patient age. At the same time, other leading professional organizations in clinical genetics (e.g., American Society of Human Genetics, European Society of Human Genetics) have issued policy statements concluding that genetic testing should typically be deferred until adulthood, if results would not lead to direct medical benefits in childhood [57, 58]. One ethical justification for deferring testing until adulthood in these situations is that there would be no immediate clinical benefits to outweigh the potential harms (e.g., distress, stigma) of receiving genetic results in childhood. Another benefit could be the preservation of the child's autonomy, if learning genetic risk information is seen to infringe on the child's right to an "open future." From this perspective, it would be most appropriate to allow the child the opportunity to make his or her own genetic testing decisions once s/he has reached young adulthood. Yet one could also make the case for a more flexible policy for testing minors for adult-onset conditions, given that harms from genetic risk disclosure are relatively rare and parents often believe genetic information about their children should be accessible if desired [59].

9.4.2 Disclosing Results of Deceased Patients

Given that precision oncology approaches are often pursued to treat advanced, refractory cancers with poor prognoses, another ethical issue that can arise is deciding if and how to communicate germline findings to family members once a patient is deceased [60]. Oftentimes, patients' preferences regarding this matter are not formally ascertained and therefore clinicians may be forced to weigh potential privacy concerns of the patient (even in death) against blood relatives' interest in knowing more about their own risk status for various hereditary cancer syndromes. A recent survey [61] of participants in a pancreatic cancer biobank found that the vast majority (94%) would be agreeable to sharing their genetic results with blood relatives who wanted to know them, with nearly as many (88%) actually feeling

obligated to share them. However, a notable minority (7.5%) said they would not want their blood relatives to know their genetic results. In a related hereditary colorectal cancer exome sequencing study [62], participants were prospectively asked if they wanted to designate someone to receive their genetic results in the event of their own death. Overall, 92% of participants designated someone (most often a spouse), but 8% declined. Taken together, these findings suggest that the vast majority of individuals would support disclosure of their genetic results to family members after death, but that this preference cannot automatically be assumed for all patients.

The dilemma of when and how to return genomic results to family members occurs not only in clinical practice, but also in research contexts. To guide researchers on this topic, a national working group made up of leading experts in the field recently issued a set of recommendations [63]. Among these recommendations were the following: (1) a proactive approach for establishing participants' preferences for research results to be returned, if this would be a possibility in a given study; (2) a passive disclosure policy in responding to relatives' requests for a participant's research results, except in rare cases where highly actionable pathogenic findings have been identified and disclosure could result in the avoidance of imminent harms; and (3) return of results that are: (a) analytically valid (e.g., high level of certainty that the genetic result in question is accurate for that participant); (b) associated with a well-established and substantial risk of an important health problem with significant implications for the participant; (c) likely to benefit the participant by virtue of their medical actionability; (d) allowable under relevant state and federal laws such as the Clinical Laboratory Improvement Amendment (CLIA), Health Insurance Portability and Accountability Act (HIPAA), state genetic privacy laws, and the Common Rule; and (e) agreed to by the participant (i.e., s/he has affirmatively consented to disclosure).

9.4.3 Disclosing Risks for Diseases Other than Cancer

When conducting sequencing to inform precision oncology approaches, laboratories following the ACMG guidelines would be obliged to report secondary medically actionable findings (e.g., risks of hereditary cardiomyopathies). Such findings could inform future disease prevention/management decisions (e.g., earlier and/or more frequent screening) for both patients and their first-degree relatives. However, many oncologists ordering sequencing are likely to encounter challenges when discussing the meaning of germline findings related to conditions outside their area of expertise [64]. Ideally, patients could be referred for such discussions to medical genetics clinics with genetic counselors, but such specialists may not be readily available. Furthermore, the timing of such testing may also be challenging if the patient is experiencing significant medical and/or psychological burden from his or her (often advanced) cancer. In these cases, discussion of potential future risks for an unrelated condition (which may be decades away from a likely age of onset) is unlikely to be

a priority for either the patient or treating clinician. It may be useful, therefore, to obtain patient preferences regarding analysis and disclosure of secondary findings *before* sequencing is ordered, and to consider deferring discussion of significant secondary findings as appropriate. These clinical encounters should address which types of information the patient is interested in receiving, and a plan for notifying family members who might also be affected by learning this information.

9.4.4 Legal Issues for Clinicians and Researchers

The advent of NGS has raised not just ethical, but also legal questions about the responsibilities of clinicians and laboratories when sequencing is ordered. One concern is that health care professionals may face new legal liabilities should they fail to disclose secondary findings that offer an opportunity for interventions to improve health outcomes. By definition, all 59 genes on the current ACMG list would meet this criterion, but as discussed earlier, there might be good reasons why patients and clinicians would be reluctant to immediately pursue this information. To our knowledge, there have not yet been any cases where clinicians have faced legal action for failure to disclose secondary NGS findings. However, a review [65] of the medical imaging literature suggests some existing case law where clinicians have faced liability if, under the relevant standard of care, they were found to have either a) failed to appreciate the clinical significance of an actionable secondary finding, or b) failed to notify the patient of such a finding in a timely manner. In response to the ACMG guidelines, leading legal scholars have suggested ways in which health systems might structure the clinical use of sequencing so as to minimize tort liabilities for nondisclosure of significant secondary findings [66]. Responsibilities of researchers to disclose secondary findings are not likely to be as extensive as those in clinical practice, and many commentators have argued against a so-called "duty to hunt" for potentially significant sequencing results if they are not immediately relevant to a given study's research aims [67]. However, in certain cases failure to disclose actionable research results could still be viewed as a violation of the research ethics principle of respect for participants.

9.5 Conclusions

This chapter highlights some of the many ELSI issues associated with advances in precision oncology and use of genomic sequencing in medicine more generally. The rapid pace of progress in this area suggests a need for practitioners and researchers to actively monitor the latest developments in the field, with an eye toward how their ethical and legal responsibilities might be affected. There is also a need for legal analysis and empirical research to guide practice and policy moving forward. For example, clinical research in controlled settings might be helpful to establish greater

understanding of the likelihood and extent of the speculated benefits, harms, and limitations of precision oncology. Such research should not only address traditional patient outcomes related to cancer treatment response and morbidity/mortality, but also health communication challenges involved in interpretation and conveyance of results from laboratory to clinician to patient (and even to extended family). Integration of comprehensive sequencing in the care of cancer patients is a complex process. Sequencing information is voluminous, challenging to interpret, and often has ambiguous implications for cancer care. Those involved in sequencing analysis and clinical decision making based on its results may have vastly differing levels of experience with and understanding of genomics, and communications among these stakeholders are often not well coordinated. Research on the process and impact of cancer sequencing is therefore needed to identify priority areas and best practices for education, communication, and management of test results.

Acknowledgements The authors' work was supported by National Institutes of Health grant UM1 HG06508.

The authors would like to acknowledge Lan Le, Brent Doil, and Rachel Gringlas for their assistance with manuscript preparation.

References

1. The precision medicine initiative. https://www.whitehouse.gov/precision-medicine. Published 2015. Accessed 13 Dec 2016
2. Collins FS, Varmus H (2015) A new initiative on precision medicine. N Engl J Med 372(9):793–795. PMID:25635347
3. Bayer R, Galea S (2015) Public health in the precision-medicine era. N Engl J Med 373(6):499–501. PMID:26244305
4. Institute NC. NCI and the precision medicine initiative. (2016). https://www.cancer.gov/research/key-initiatives/precision-medicine. Published 3/10/2016. Accessed 09 Dec 2016
5. Tannock IF, Hickman JA (2016) Limits to personalized cancer medicine. N Engl J Med 375(13):1289–1294. PMID:27682039
6. American Society of Clinical Oncology (2016) The state of cancer care in America, 2016: a report by the american society of clinical oncology. J Oncol Pract 12(4):339–383. PMID:26979926
7. Siegel RL, Miller KD, Jemal A (2016) Cancer statistics, 2016. CA Cancer J Clin 66(1):7–30. PMID:26742998
8. Medicine Io (2003) Unequal treatment: confronting racial and ethnic disparities in health care. The National Academies Press, Washington, DC
9. Spratt DE, Chan T, Waldron L et al (2016) Racial/ethnic disparities in genomic sequencing. JAMA Oncol 2(8):1070–1074. PMID:27366979
10. Green RC, Goddard KA, Jarvik GP et al (2016) Clinical sequencing exploratory research consortium: accelerating evidence-based practice of genomic medicine. Am J Hum Genet 99(1):246. PMID:27392080
11. Gray SW, Martins Y, Feuerman LZ et al (2014) Social and behavioral research in genomic sequencing: approaches from the Clinical Sequencing Exploratory Research Consortium Outcomes and Measures Working Group. Genet Med 16(10):727–735. PMID:24625446

12. Bogart LM, Bird ST, Walt LC, Delahanty DL, Figler JL (2004) Association of stereotypes about physicians to health care satisfaction, help-seeking behavior, and adherence to treatment. Soc Sci Med 58(6):1049–1058
13. van Ryn M, Burke J (2000) The effect of patient race and socio-economic status on physicians' perceptions of patients. Soc Sci Med 50(6):813–828. PMID:10695979
14. Bogart LM, Catz SL, Kelly JA, Benotsch EG (2001) Factors influencing physicians' judgments of adherence and treatment decisions for patients with HIV disease. Med Decis Mak 21(1):28–36. PMID:11206944
15. Hagiwara N, Berry-Bobovski L, Francis C, Ramsey L, Chapman RA, Albrecht TL (2014) Unexpected findings in the exploration of African American underrepresentation in biospecimen collection and biobanks. J Cancer Educ 29(3):580–587. PMID:24243440
16. Gray SW, Hicks-Courant K, Cronin A, Rollins BJ, Weeks JC (2014) Physicians' attitudes about multiplex tumor genomic testing. J Clin Oncol 32(13):1317–1323. PMID:24663044
17. Emanuel EJ, Wendler D, Grady C (2000) What makes clinical research ethical? JAMA 283(20):2701–2711. PMID:10819955
18. Lorenzen B, Melby CE, Earles B (2008) Using principles of health literacy to enhance the informed consent process. AORN J 88(1):23–29. PMID:18603032
19. Stableford S, Mettger W (2007) Plain language: a strategic response to the health literacy challenge. J Public Health Policy 28(1):71–93. PMID:17363939
20. Holmes-Rovner M, Stableford S, Fagerlin A et al (2005) Evidence-based patient choice: a prostate cancer decision aid in plain language. BMC Med Inform Decis Mak 5:16. PMID:15963238
21. Davis TC, Williams MV, Marin E, Parker RM, Glass J (2002) Health literacy and cancer communication. CA Cancer J Clin 52(3):134–149. PMID:12018928
22. Bernhardt BA, Roche MI, Perry DL, Scollon SR, Tomlinson AN, Skinner D (2015) Experiences with obtaining informed consent for genomic sequencing. Am J Med Genet A 167A(11):2635–2646. PMID:26198374
23. Berkman BE, Hull SC, Eckstein L (2014) The unintended implications of blurring the line between research and clinical care in a genomic age. Per Med 11(3):285–295. PMID:25506378
24. Miller FA, Giacomini M, Ahern C, Robert JS, de Laat S (2008) When research seems like clinical care: a qualitative study of the communication of individual cancer genetic research results. BMC Med Ethics 9:4. PMID:18294373
25. Beskow LM, Friedman JY, Hardy NC, Lin L, Weinfurt KP (2010) Simplifying informed consent for biorepositories: stakeholder perspectives. Genet Med 12(9):567–572. PMID:20697289
26. Appelbaum PS, Roth LH, Lidz C (1982) The therapeutic misconception: informed consent in psychiatric research. Int J Law Psychiatry 5(3–4):319–329. PMID:6135666
27. Horng S, Grady C (2003) Misunderstanding in clinical research: distinguishing therapeutic misconception, therapeutic misestimation, and therapeutic optimism. IRB 25(1):11–16. PMID:12833900
28. Joffe S, Cook EF, Cleary PD, Clark JW, Weeks JC (2001) Quality of informed consent in cancer clinical trials: a cross-sectional survey. Lancet 358(9295):1772–1777. PMID:11734235
29. Appelbaum PS, Lidz CW, Grisso T (2004) Therapeutic misconception in clinical research: frequency and risk factors. IRB 26(2):1–8. PMID:15069970
30. Kim SY, Schrock L, Wilson RM et al (2009) An approach to evaluating the therapeutic misconception. IRB 31(5):7–14. PMID:19873836
31. Dressler LG (2009) Disclosure of research results from cancer genomic studies: state of the science. Clin Cancer Res 15(13):4270–4276. PMID:19549775
32. Halpern SD, Ubel PA, Asch DA (2007) Harnessing the power of default options to improve health care. N Engl J Med 357(13):1340–1344. PMID:17898105
33. Oliver JM, Slashinski MJ, Wang T, Kelly PA, Hilsenbeck SG, McGuire AL (2012) Balancing the risks and benefits of genomic data sharing: genome research participants' perspectives. Public Health Genomics 15(2):106–114. PMID:22213783
34. McGuire AL, Oliver JM, Slashinski MJ et al (2011) To share or not to share: a randomized trial of consent for data sharing in genome research. Genet Med 13(11):948–955. PMID:21785360

35. Edwards KL, Lemke AA, Trinidad SB et al (2012) Genetics researchers' and IRB professionals' attitudes toward genetic research review: a comparative analysis. Genet Med 14(2):236–242. PMID:22241102
36. Bernhardt BA, Tambor ES, Fraser G, Wissow LS, Geller G (2003) Parents' and children's attitudes toward the enrollment of minors in genetic susceptibility research: implications for informed consent. Am J Med Genet A 116A(4):315–323. PMID:12522783
37. The genetic information nondiscrimination act of 2008., 881, §1635.1-12
38. Kronenthal C, Delaney SK, Christman MF (2012) Broadening research consent in the era of genome-informed medicine. Genet Med 14(4):432–436. PMID:22422050
39. Bacon PL, Harris ED, Ziniel SI et al (2015) The development of a preference-setting model for the return of individual genomic research results. J Empir Res Hum Res Ethics 10(2):107–120. PMID:25742675
40. Gornick MC, Scherer AM, Sutton EJ et al (2016) Effect of public deliberation on attitudes toward return of secondary results in genomic sequencing. J Genet Couns 26(1):122–132. PMID:27307100
41. Bradbury AR, Patrick-Miller L, Long J et al (2015) Development of a tiered and binned genetic counseling model for informed consent in the era of multiplex testing for cancer susceptibility. Genet Med 17(6):485–492. PMID:25297947
42. Khan A, Capps BJ, Sum MY, Kuswanto CN, Sim K (2014) Informed consent for human genetic and genomic studies: a systematic review. Clin Genet 86(3):199–206
43. Tomlinson AN, Skinner D, Perry DL, Scollon SR, Roche MI, Bernhardt BA (2016) "Not tied up neatly with a bow": professionals' challenging cases in informed consent for genomic sequencing. J Genet Couns 25(1):62–72. PMID:25911622
44. Vries H, Mesters I, van de Steeg H, Honing C (2005) The general public's information needs and perceptions regarding hereditary cancer: an application of the Integrated Change Model. Patient Educ Couns 56(2):154–165. PMID:15653244
45. d'Agincourt-Canning L (2005) The effect of experiential knowledge on construction of risk perception in hereditary breast/ovarian cancer. J Genet Couns 14(1):55–69. PMID:15789156
46. Brown RF, Butow PN, Butt DG, Moore AR, Tattersall MH (2004) Developing ethical strategies to assist oncologists in seeking informed consent to cancer clinical trials. Soc Sci Med 58(2):379–390. PMID:14604623
47. Bergenmar M, Johansson H, Wilking N (2011) Levels of knowledge and perceived understanding among participants in cancer clinical trials – factors related to the informed consent procedure. Clin Trials 8(1):77–84. PMID:21109583
48. Jefford M, Moore R (2008) Improvement of informed consent and the quality of consent documents. Lancet Oncol 9(5):485–493. PMID:18452859
49. Grossman SA, Piantadosi S, Covahey C (1994) Are informed consent forms that describe clinical oncology research protocols readable by most patients and their families? J Clin Oncol 12(10):2211–2215. PMID:7931491
50. Epstein RM, Peters E (2009) Beyond information: exploring patients' preferences. JAMA 302(2):195–197. PMID:19584351
51. Irwin S, Kirsch AJ, Jenkins L, Kolstad A (1993) Adult literacy in America: a first look at the findings of the national adult literacy survey. National Center of Education Statistics; 8/30/1993
52. Brothers KB, Holm IA, Childerhose JE et al (2016) When participants in genomic research grow up: contact and consent at the age of majority. J Pediatr 168:226–231 e221. PMID:26477867
53. Clayton EW (2015) How much control do children and adolescents have over genomic testing, parental access to their results, and parental communication of those results to others? J Law Med Ethics 43(3):538–544. PMID:26479563
54. Raymond VM, Gray SW, Roychowdhury S et al (2016) Germline findings in tumor-only sequencing: points to consider for clinicians and laboratories. J Natl Cancer Inst 108:4. PMID:26590952

55. Kalia SS, Adelman K, Bale SJ et al (2016) Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. Genet Med 15(7):565–574
56. Mody RJ, Wu Y, Lonigro RJ et al (2015) Integrative clinical sequencing in the management of refractory or relapsed cancer in youth. JAMA 314(9):913–925
57. Botkin Jeffrey R, Belmont John W, Berg Jonathan S et al (2015) Points to consider: ethical, legal, and psychosocial implications of genetic testing in children and adolescents. Am J Hum Genet 97(1):6–21. PMID:PMC4570999
58. European Society of Human G (2009) Genetic testing in asymptomatic minors: recommendations of the european society of human genetics. Eur J Hum Genet 17(6):720–721. PMID:19277059
59. Wilfond B, Ross LF (2009) From genetics to genomics: ethics, policy, and parental decision-making. J Pediatr Psychol 34(6):639–647. PMID:18647793
60. Forrest LE, Delatycki MB, Skene L, Aitken M (2007) Communicating genetic information in families—a review of guidelines and position papers. Eur J Hum Genet 15(6):612–618. PMID:17392704
61. Breitkopf CR, Petersen GM, Wolf SM et al (2015) Preferences regarding return of genomic results to relatives of research participants, including after participant death: empirical results from a cancer biobank. J Law Med Ethics 43(3):464–475. PMID:26479556
62. Amendola LM, Horike-Pyne M, Trinidad SB et al (2015) Patients' choices for return of exome sequencing results to relatives in the event of their death. J Law Med Ethics 43(3):476–485. PMID:26479557
63. Wolf SM, Branum R, Koenig BA et al (2015) Returning a research participant's genomic results to relatives: analysis and recommendations. J Law Med Ethics 43(3):440–463. PMID:26479555
64. Gray SW, Park ER, Najita J et al (2016) Oncologists' and cancer patients' views on whole-exome sequencing and incidental findings: results from the CanSeq study. Genet Med 18(10):1011–1019. PMID:26866579
65. Clayton EW, Haga S, Kuszler P, Bane E, Shutske K, Burke W (2013) Managing incidental genomic findings: legal obligations of clinicians. Genet Med 15(8):624–629. PMID:23448723
66. Evans BJ (2013) Minimizing liability risks under the ACMG recommendations for reporting incidental findings in clinical exome and genome sequencing. Genet Med 15(12):915–920. PMID:24030435
67. Jarvik GP, Amendola LM, Berg JS et al (2014) Return of genomic results to research participants: the floor, the ceiling, and the choices in between. Am J Hum Genet 94(6):818–826. PMID:24814192

Chapter 10
Liquid Biopsy: Translating Minimally Invasive Disease Profiling from the Lab to the Clinic

Daniel Zainfeld, Umair Ghani, Irene Kang, and Amir Goldkorn

Keywords Liquid biopsy · Circulating tumor cells · Cell-free DNA · Cell-free RNA · Exosomes · Extracellular vesicles · Clinical trials

10.1 Introduction

Precision medicine aims to characterize the unique molecular profile of individual tumors in order to predict clinical course and inform therapeutic decisions. The success of this approach is dependent on an adequate characterization of the disease at presentation, as well as over the course of treatment, as adaptations to therapy drive new mutations and resistance mechanisms. Tissue obtained through surgical excision or biopsy remains the mainstay of initial diagnosis and molecular analysis. However, tumor location and size, patient safety, and costs often limit the feasibility of repeated biopsies throughout the disease course. In addition, significant molecular heterogeneity between tumor sites and even within individual tumors can compromise comprehensive assessment of disease state from biopsy alone [1–4]. Liquid biopsies offer an alternative means of collecting representative and highly relevant information in a safe, easily repeatable, low cost manner in the form of a simple peripheral blood draw, enabling examination of various analytes including circulating tumor cells (CTCs), cell free DNA (cfDNA), cell free RNA (cfRNA), and extracellular vesicles with a widening array of potential clinical applications. Due to their easy accessibility, liquid biopsies enable monitoring of disease course in real time to gauge response to treatments, elucidate temporal evolution of genetic or cellular adaptations in response to therapies, and guide subsequent treatments. Hence, this approach may ultimately serve a pivotal role in translating precision medicine principles to clinical practice.

D. Zainfeld · U. Ghani · I. Kang · A. Goldkorn (✉)
University of Southern California, Los Angeles, CA, USA
e-mail: agoldkor@usc.edu

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_10

10.2 Circulating Tumor Cells (CTCs)

The biology of metastasis is highly complex and requires specific characteristics of both the spreading cells (“seeds”) and the microenvironment that receives them (“soil”) [5]. At its core, metastatic spread requires tumor cell intravasation into the bloodstream, viable trafficking through the circulation, and eventual extravasation, implantation, and proliferation at distant sites [6–8]. These circulating tumor cells (CTCs), travelling singly or in clusters, are exceedingly scarce even in advanced disease, typically ranging between zero and a few dozen in a standard tube of blood, depending on the methods used to isolate them. Circulating tumor cells have been identified in the context of virtually all solid malignancies in advanced states and offer a unique window into active cancer biology and behavior [9]. CTCs can serve as prognostic markers, recapitulate primary or metastatic tumor biology, and potentially guide therapeutic decisions. CTC clusters, interestingly, have demonstrated 23–50 fold increase in metastatic potential, with capacity for travel even through capillary-sized vessels suggesting a significant role in metastatic spread [10, 11]. Advances in microfluidics, microscopy, and high content image analysis have significantly improved the identification, enumeration, and analysis of CTCs, opening new avenues for the study of cancer biology and clinical outcomes.

10.3 CTC Enrichment

Though described as early as the 1800s (Ashworth 1869), reliable methods for identification and recovery of CTCs have emerged only recently, nearly one and a half centuries later [12]. The sheer numbers of red blood cells (RBCs, 10^9) and white blood cells (WBCs, 10^6) in a standard blood sample pose the central technical challenge for enrichment and recovery of CTCs (0–10^2) [13, 14]. A number of strategies have been developed to identify and capture rare CTCs in the proverbial sea of red and white blood cells, each with its own benefits and limitations.

Cell surface marker based techniques are the most studied and clinically validated methods, wherein CTCs are enriched through immunomagnetic labeling, or alternatively, WBCs are depleted through immunomagnetic negative selection. The CellSearch® platform developed by Janssen Diagnostics (Raritan, NJ, USA) and recently acquired by Menarini Silicon Biosystems (Bologna, Italy) merits special emphasis, as this is currently the only FDA cleared device for CTC detection. CellSearch® enriches CTCs from blood by binding cell surface epithelial cell adhesion molecule (EpCAM) with antibody-coated ferrofluid nanoparticles, followed by staining for cytokeratins (CK), 4′,6-diamidino-2-phenylindole (DAPI) nuclear stain, and CD45, a cell surface leukocyte marker. Cells found to be EpCAM+, CD45-, DAPI+ and CK+ are deemed consistent with CTCs [13]. Limitations of immunomagnetic enrichment include the inability, at least in theory, to differentiate cancerous CTCs from benign epithelial cells in the circulation, as well as potential

failure to capture cells undergoing epithelial to mesenchymal transition (EMT), a key metastatic step characterized by loss of epithelial markers such as EpCAM [15–18]. In one example of this phenomenon, Ring et al. performed spike-in experiments with 10 breast cancer cell lines representing all major subtypes of breast cancer. Despite recovery of cancer cells from all subtypes, claudin-low cell lines, a subset marked by comparatively poor prognosis, had significantly lower capture rates compared to the other four groups [19]. Similarly, Yu et al. confirmed the importance of EMT in metastasis by characterizing both primary tumor and CTCs through mesenchymal or epithelial marker expression. High expression of mesenchymal markers was found among CTCs while primary tumor rarely expressed both epithelial and mesenchymal markers. Moreover, CTC monitoring demonstrated an association between mesenchymal CTCs and disease progression [18]. These observations, as well as the identification of cytokeratin-negative CTCs in a subset of patient samples, suggest that potentially relevant CTC subpopulations may not be fully captured by standard immunomagnetic techniques [16, 20–22].

Alternative, marker independent, techniques isolate CTCs based on unique biophysical characteristics such as size, deformability, or electric properties. Size-based systems using various filters and microfluidic chips enrich viable CTCs that are then specifically identified through staining prior to downstream analysis (eg. Parsortix, Angle, UK; ClearCell FX, Clearbridge Biomedics, Singapore). While these techniques are effective, CTC size variability has been demonstrated, raising the potential pitfalls in reliance on physical cell characteristics which remain poorly defined overall and vary significantly among CTCs [23]. Hybrid approaches utilizing both immunomagnetic and size-based properties have been explored as well. The CTC-iChip, for example, separates nucleated cells based on size using deterministic lateral displacement. At the same time, WBCs are tagged with specific antibodies enabling subsequent magnetopharesis to further enhance sample purity [24]. Still other systems employ differential dielectric properties to separate and recover CTCs. ApoStream (Apocell, Texas) exploits distinct dielectric properties of CTCs in comparison to PBMCs in a process termed dielectrophoresis field flow assist, whereby cells are selectively attracted to or repelled from a charged electrode allowing specific fractionation of cell types [25]. Another platform, the DepArray (Silicon Biosystems) isolates individual CTCs using a microfluidic cartridge with controllable electrodes to create dielectric cages around individual cells in a pre-enriched sample, thus allowing analysis and recovery of specific cells of interest [26].

In contrast to enrichment based instruments, some platforms implement a “no cell left behind” approach. Epic Sciences (San Diego, USA) employs high content scanning and algorithmic CTC identification based on immunofluorescent and morphologic features of whole blood smears. Similarly, Rarecyte (Seattle, USA) first enriches nucleated cells by density gradient centrifugation, then generates smears for automated multiplex scanning and CTC identification based on immunofluorescent and morphologic criteria. Automated digital microscopy and development of sophisticated scanning algorithms continue to create opportunities for rapid, high throughput imaging and analysis without the need for sample enrichment and attendant cell loss. However, scanning platforms, like all CTC methods,

Table 10.1 CTC Enrichment and isolation approaches

Immunomagnetic	**Principle: CTCs characterized by specific markers that allow differentiation from other PBMCs**	**Example**
Positive selection	Capture CTCs by labeling with immunomagnetic antibody specific marker (EpCAM) and isolating with magnet	CellSearch®
Negative selection	Enrichment for CTCs by exclusion of WBCs (CD45+), generally following RBC lysis	EasySep®
Physical property	**Principle: CTCs marked by unique physical characteristics that can be leveraged for separation**	
Filtration	Use of various microfilter technologies to capture larger, less deformable CTCs	Parsortix
Inertial sorting	CTC sorting dependent on advanced microfluidic principles	ClearCell®
Dielectrophoresis	Enrichment based on dielectric properties of CTCs	ApoStream®
Single cell capture	Isolation of individual CTCs in dielectric cages. Requires enriched sample	DEPArray™
Combination	Application of immunomagnetic and physical principles to enrichment (eg filtration followed by negative immunomagnetic selection)	CTC-iChip
Rapid scanning	Automated scanning algorithms to identify stained CTCs, often without enrichment	
Whole blood scanning	Enrichment free specimen analysis through high content algorithmic scanning	EPIC sciences™

CTC enrichment and identification techniques continue to expand. Work exploring clinical applications and optimal methods is ongoing

remain limited by lack of universally applicable CTC markers and still-evolving techniques for single cell recovery. Some of examples of the various CTC platforms are listed in Table 10.1 and have been recently reviewed in greater depth elsewhere [27] (Table 10.2).

10.4 Clinical Applications of CTCs: Enumeration

CTC counts have demonstrated prognostic utility in a number of solid malignancies and have been validated in multiple clinical trials [23, 28–32]. Enumeration has been dominated by the CellSearch immunomagnetic system, which received FDA clearance in metastatic breast, prostate, and colon cancer. Cristofanilli et al. completed one of the first multicenter prospective studies analyzing the value of CTC counts for predicting survival in the setting of metastatic breast cancer. Using the CellSearch platform, patients with greater than or equal to 5 CTCs/7.5 ml whole blood were found to have shorter progression-free and overall survival at multiple time points [33]. Similar results were later described by Lucci et al. in the setting of non-metastatic breast cancer, where the presence of one or more CTCs predicted early recurrence and decreased overall survival [34]. Likewise, de Bono et al.

Table 10.2 Select ctDNA detection studies in various tumor types

Study Reference	Analytic platform for ctDNA	Molecular alteration	Number of patients analyzed for ctDNA	Tumor Type	Stage	Sensitivity (patients with detected ctDNA/ positive tumor markers)	Summary
Dawson [83]	Digital PCR, tagged amplicon sequencing	SNV	30	Breast	IV	29/30 (97%)	ctDNA sensitivity for reflecting tumor burden and treatment response greater than CTC or CA 15-3 levels in metastatic breast cancer
Tie [84]	Digital PCR, massively parallel sequencing	SNV	53	Colorectal	IV	48/53 (92.3%)	Quantification of ctDNA in metastatic colorectal patients demonstrates prognostic value and predicts treatment response
Azad [85]	Digital PCR, tagged amplicon sequencing	SNV (AR exon 8 mutation)	62	Prostate	IV	11/62 (18%)	AR gene aberrations in cfDNA correlate with adverse outcomes and decreased median PFS in mCRPC

ctDNA circulating tumor DNA, *SNV* single nucleotide variant, *AR* androgen receptor, *PFS* progression free survival, *mCRPC* metastatic castrate resistant prostate cancer

performed CellSearch CTC counts in patients with castration-resistant prostate cancer (CRPC) prior to initiation of a new systemic therapy. Men with favorable CTC counts (<5 CTC/7.5 ml blood) were found to have greater overall survival (21.7 v 11.5 months; $P < 0.0001$) and change in CTC counts in response to therapy was predictive of survival as well [28]. Goldkorn et al. subsequently confirmed these findings in a prospective phase 3 trial, SWOG 0421, in men initiating first line chemotherapy with docetaxel and also showed that CTC telomerase activity was

prognostic in a subset of patients [32, 35]. In the setting of metastatic colorectal cancer, Cohen et al. completed a prospective multicenter study examining the prognostic value of CellSearch CTC enumeration in 430 patients and found that favorable CTC counts (<3 CTCs/7.5 ml) were associated with greater overall survival (18.5 v 9.4 months; P < 0.0001) [23]. Additional studies evaluating the prognostic utility of CTC enumeration have been completed in various stages of lung, breast, prostate, and colorectal cancers, as well as in pancreatic, melanoma, ovarian, bladder and gastric cancer using various technologies for enumeration with varied clinical endpoints [32, 36–43]. While limited sensitivity and specificity remain problematic in the accurate enumeration of CTCs, it is clear that even at current detection capacities, CTC counts often provide clinically relevant data regarding response to therapy, prognosis and potentially disease recurrence.

Studies are still ongoing to test whether CTC counts can successfully guide treatment decisions. One early attempt, SWOG S0500, did not demonstrate a clinical benefit for CTC enumeration as a predictive biomarker. In this multi-institutional randomized study, metastatic breast cancer patients with persistently elevated CTC counts after first line chemotherapy were randomized to either continue the same therapy or switch to an alternative chemotherapy. Prognostic value of CTC enumeration was again confirmed with significant differences among those patients without increased CTCs at baseline (Group A), those whose CTC counts were reduced at first follow-up after initiation of therapy (Group B), and those whose CTC counts had not been reduced at first follow-up after start of therapy (Group C) (Median OS for groups A, B, and C was 35, 23, and 13 months respectively) (Fig. 10.1). Group C with persistently elevated CTCs was further split into an arm with no treatment change vs. an arm with change to another treatment. However, no improvement in outcome was observed in patients in Group C who were switched to another chemotherapy based on CTC elevation (median OS 10.7 and 12.5 months in the two arms, respectively; p = 0.98) [44]. While these results were disappointing, they are perhaps not entirely surprising in the context of the underlying study: a cohort of patients with poor overall prognosis based on persistently high CTCs who were unlikely to respond to chemotherapy, and who were randomized to two relatively small treatment arms powered to detect a very large (70%) increase in survival. Based on these insights, new trials evaluating the predictive strength of CTC enumeration ideally would assign patients with elevated CTCs to large treatment arms powered to detect smaller predictive effects and utilizing treatment modalities thought most likely to exert a significant, differential impact on survival.

10.5 Beyond Enumeration: CTC Characterization and Analysis

Whereas enumeration constitutes a clinically useful surrogate for disease burden, molecular characterization of CTCs (protein, DNA, RNA) can serve as a true liquid biopsy that reflects the underlying biology of primary and metastatic tumor sites. As such, CTC-derived molecular phenotypes may elucidate patient-specific and

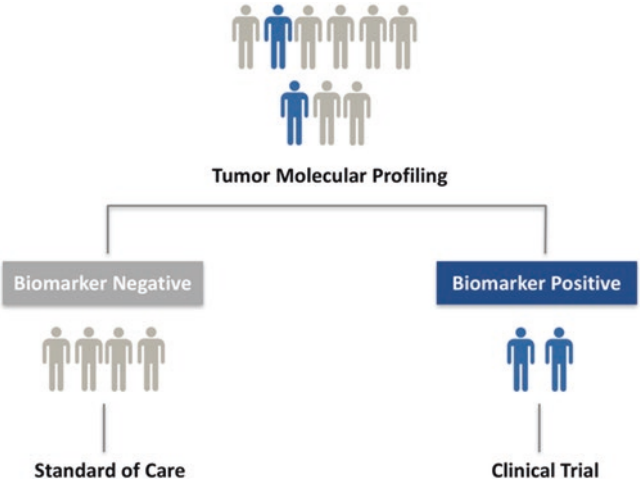

Fig. 10.1 AR-V7 Expression in CTC waterfall plots. (Figure 3 from Ref. [53])

tumor-specific drivers of progression and resistance and guide therapy decisions. Genetic alterations and expression patterns of numerous relevant cancer genes such as androgen receptor (AR), estrogen receptor (ER), progesterone receptor (PR), and epidermal growth factor receptor (EGFR) have clear implications for disease prognosis and potential treatment opportunities in various malignancies [45–47]. Characterization of CTCs affords an opportunity to evaluate these disease-relevant phenotypes and to assess changes in response to treatment in a manner that is safe, repeatable, and minimally invasive.

Multiple prostate CTC investigations have focused on androgen receptor (AR), a central transcriptional driver of progression. Darshan el al. used immunofluorescence AR staining in CellSearch captured CTCs to demonstrate a significant association between AR cytoplasmic localization and response to chemotherapy [48]. In another study, Miyamoto et al. used immunofluorescence staining to classify "AR-on" vs. "AR-off" CTCs captured on a microfluidic device from hormone-naive patients versus CRPC patients and found clear staining differences in their CTCs [49]. Crespo et al. examined AR expression in CTCs from CRPC patients and identified nuclear predominant AR localization in those who progressed on abiraterone, though significant AR heterogeneity was noted in this relatively small sample [50]. Punnoose et al. examined PTEN expression in CRPC. Using the Epic Sciences enrichment-free platform, CTCs were identified and PTEN status analyzed using fluorescence in situ hybridization (FISH). PTEN expression was then evaluated in fresh and archival tissue, revealing a concordance between CTCs and matched fresh tissue in 84% of patients as well as a prognostic capacity of PTEN loss for worse survival in CRPC patients [51]. Antonarakis and colleagues used an RNA-based CTC enrichment and detection test (AdnaTest) to analyze CTCs for the presence of AR-V7, a truncated ligand-independent constitutively active splice variant, and

demonstrated a clear association between expression of AR-V7 in CTCs and resistance to abiraterone or enzalutamide hormonal therapies [52]. These findings were confirmed and extended by Scher et al. who demonstrated that presence of CTC AR-V7 in CRPC patients predicts resistance to androgen receptor signaling inhibition and improved survival on alternative (taxane chemotherapeutic) therapies (Fig. 10.1) [53].

In breast cancer, Kalinsky et al. examined estrogen receptor (ER) and progesterone receptor (PR) status by immunohistochemistry in CTCs of patients with metastatic disease and found concordance with 68% of primary sites and 83% of metastatic sites, indicating potential for receptor status monitoring through liquid biopsy [54]. Another study of CTC ER expression in metastatic breast cancer revealed heterogeneity of ER expression in CTCs of patients with ER+ primary tumors, suggesting a potential escape mechanism from hormonal therapies [55]. Similarly, evaluation of HER-2 oncogene in breast cancer CTCs by FISH enabled quantification of HER-2 gene amplification, a potentially valuable CTC-based tool for selecting patients for HER-2 targeted therapies, in some cases despite originally having a HER-2-negative primary tumor [56]. In another study, characterization of CTCs with an eight gene profile successfully discriminated good versus poor response to aromatase inhibitor therapy among metastatic breast cancer patients [57]. These emerging CTC-based biomarker capabilities continue to be integrated into prospective trials for validation, and some may ultimately be incorporated into standard practice.

Molecular targets have been analyzed in CTCs from non-hormonally driven malignancies as well. In non-small cell lung cancer (NSCLC), Maheswaran et al. analyzed DNA from CTCs captured with an immunomagnetic chip for mutations in epidermal growth factor receptor (EGFR), a therapeutic target for antibody antagonists and small molecule receptor tyrosine kinase inhibitors [58]. EGFR mutations in CTCs were concordant with primary tumor mutations in 92% of cases, and their presence correlated with a reduced progression-free survival (7.7 months v 16.5 months, $P < 0.001$) [59]. In metastatic colorectal cancer, CTCs have been characterized for thymidylate synthase (TYMS) expression, which is linked to 5-fluorouracil resistance. Patients with CTCs staining positive for TYMS expression trended towards disease progression in comparison to those not expressing TYMS in CTCs, regardless of expression in primary tumor [60]. In ovarian cancer, Kuhlmann et al. found that ERCC1 expression on CTCs at primary diagnosis was predictive of platinum resistance ($p = 0.01$), whereas corresponding primary tumor tissue was not predictive [61].

Recent advances in DNA and RNA amplification and next generation sequencing have enabled a transition from specific targets (e.g. AR, ER) to broader high throughput discovery studies using minimal starting material isolated from few or even single cells. Ni et al. completed whole-genome amplification of single CTCs from patients with lung cancer, identifying known cancer-associated single nucleotide variations and insertions/deletions in CTC exomes that varied between individual cells. However, copy number variation (CNV) patterns were consistent and reproducible within individual patients and unique between adenocarcinoma

and small cell lung cancer [62]. Heitzer et al. completed genomic profiling of CTCs in colorectal cancer using array-comparative genomic hybridization. Through parallel sequencing of 68 colorectal cancer genes, mutations were found in APC, KRAD, and PIK3CA in both primary tumor and CTCs. Additional mutations which were initially missed in primary tumors and found only in CTCs were subsequently identified at the subclonal level in primary tumors and metastasis from the same patient on additional deep sequencing [63]. Subsequently, mutations in KRAS, BRAF, CD133 and Plastin3 were identified in a similar cohort of patients with colorectal cancer with significant discordance from primary tumor specimens, further highlighting the potential utility of liquid biopsy and CTC mutational analysis for temporal applications to personalized medicine [64]. Lohr et al. demonstrated feasibility of comprehensive genome analysis through liquid biopsy by capturing CTCs from men with metastatic prostate cancer and, for a select patient, completing whole exome sequencing of CTCs, lymph node metastasis, and multiple cores of primary tumor. They identified 10 early and 56 metastatic trunk mutations within primary tumor and metastatic sites. Analysis of CTC exomes from the same patient identified 90% of the early and 73% of the metastatic trunk mutations reflecting potential for accurate genomic profiling through evaluation of CTCs [65].

As with genomic analysis, sequencing of tumor transcriptomes has helped identify potential driving mechanisms of cancer progression. Despite the relative fragility of RNA and the low input material achievable from CTCs, analysis has been completed at even the single cell level. Yu et al., completed single molecule RNA sequencing on CTCs captured from an endogenous mouse pancreatic cancer model. The gene Wnt2 was found to be enriched in CTCs, suggesting its role as a potential therapeutic target in pancreatic cancer and demonstrating the value of transcriptome analysis [66]. Gorges et al. completed transcriptome analysis of CTCs from breast and prostate cancers, identifying distinct expression signatures as well as subsets of CTCs within patients marked by divergent expression of genes involved in EMT, cancer therapy resistance, and tumor progression with potential clinical significance [67]. Single cell RNA sequencing of CTCs from men with metastatic CRPC has also been completed revealing heterogeneity at the individual level including AR gene mutations and splice variants between cells. Analysis of patients with progressive disease revealed noncanonical Wnt signaling, a potential driving mechanism for drug resistance [68].

The realm of CTC analysis and characterization continues to rapidly expand, including new efforts aimed at CTC culture and drug susceptibility testing [69]. Significant challenges remain, however, including optimization of CTC isolation, discovery of reliable markers for CTC identification within various malignancies, and the need to reconcile implications of cellular phenotypic and genotypic heterogeneity as well as technical biases introduced by current amplification techniques. Ongoing work towards these goals in concert with prospective clinical trials and technologic advances to decrease cost and logistic constraints of CTC analysis will help to better define its role in clinical practice and ultimately to translate this promising facet of liquid biopsies into more effective patient care.

10.6 Cell Free DNA (cfDNA)

Cell free DNA (cfDNA) is composed of DNA fragments continuously shed into the bloodstream through cell lysis, necrosis, and apoptosis due to pathologic processes and normal cell turnover [70]. Circulating tumor DNA (ctDNA) is the portion of circulating DNA specifically derived from cancer cells by active secretion from macrophages that have phagocytized necrotic cells, or released directly from dying tumor cells or CTCs [71]. Compared with CTC enrichment, extraction of cfDNA from plasma is relatively straightforward. However, the actual cellular tissue source of DNA alterations is more difficult to ascertain, as collected cfDNA may derive from primary tumor, metastatic sites, or nonmalignant sources [72]. cfDNA levels vary widely depending on characteristics of tumor burden, stage, vascularity, cellular turnover, and response to therapy [73]. Hence, analysis of the small fractions of ctDNA within background levels of wild-type cfDNA requires highly sensitive techniques. These primarily target characteristic genetic or epigenetic modifications, such as mutations in tumor-suppressor genes, activated oncogenes, hypermethylation or chromosomal abnormalities, to confirm that cancer cells are indeed the source of the detected cfDNA [74]. Traditional methods including quantitative real-time polymerase chain reaction (PCR)-based, fluorescence-based, and spectrophotometric approaches have been utilized with some success [75, 76], and new methods like digital droplet PCR, microfluidic platforms for parallel PCR, and BEAMing (beads, emulsions, amplification and magnetics) are now able to detect point mutations in ctDNA at lower allele frequencies [77–79]. Next generation sequencing has facilitated identification of ctDNA alterations across wide genomic regions providing novel opportunities for comprehensive characterization of genomic profiles [80]. More recently, deep sequencing of cfDNA and examination of nucleosome footprints and promoter occupancies has been shown to reveal nuclear architecture, gene structure, expression, and even cell type of origin broadening the potential applicability of cfDNA evaluation [81, 82]. Recent studies in cfDNA and ctDNA have demonstrated clinical applications with potential therapeutic interventions in breast, colon, prostate, lung cancer and melanoma.

10.7 Clinical Applications of cfDNA

Dawson et al. studied the value of ctDNA as a marker of progression in metastatic breast cancer, compared to CTC counts and CA 15-3 levels [83]. Microfluidic digital PCR or tagged-amplicon deep sequencing was used to quantify specific somatic alterations in PIK3CA and TP53 in DNA isolated from plasma. ctDNA was detected in 95%, whereas 50% of patients had 5 or more CTCs per 7.5 ml and 42% had elevated levels of CA 15-3. At subsequent follow-up in the setting of progressive disease, increasing levels of ctDNA and CTC count were associated with inferior overall survival while CA 15-3 was not prognostic. ctDNA had superior sensitivity

for reflecting tumor burden compared with CA 15-3 and CTC levels, as well as earlier reflection of treatment response. Hence in this study ctDNA was an effective biomarker for serial temporal monitoring in metastatic breast cancer [83].

In the setting of metastatic colorectal cancer (mCRC), Tie et al. investigated the potential role of ctDNA quantification as an early predictor of treatment response [84]. In a prospective, multicenter study, patients receiving chemotherapy were evaluated for levels of ctDNA and CEA (a standard biomarker for colorectal cancer) at multiple time-points and correlated with baseline and follow-up imaging. Primary tumors were sequenced using a panel of 15 genes frequently mutated in mCRC in order to identify candidate genes for subsequent ctDNA analysis. Using Safe-SeqS, a massively parallel sequencing (MPS)-based assay which permits detection of low-frequency mutations, measurable mutations were identified in cfDNA of 92.3% of patients. Median ctDNA levels before cycle 2 were significantly lower than the median pretreatment level (0.54 versus 16.24) reflecting tumor response to treatment. No significant changes in CEA levels were seen across time points, whereas major reductions ($\geq$10-fold) in ctDNA prior to second chemotherapy cycle were associated with a trend towards increased progression-free survival (median 14.7 versus 8.1 months) [84]. Hence, based on this study in metastatic colorectal cancer, ctDNA may offer a useful adjunct to standard follow-up (imaging and biomarkers) for accurate disease monitoring.

As with CTCs, cfDNA characterization beyond quantification of ctDNA levels offers additional insights. Azad et al. examined AR gene aberrations in cfDNA to determine correlation to therapy resistance in mCRPC patients [85]. Patients stopping abiraterone acetate, enzalutamide, or other agents due to disease progression were examined. Array comparative genomic hybridization (aCGH) for chromosome copy number analysis and Roche 454 targeted next-generation sequencing of exon 8 in the AR were completed on collected cfDNA. AR amplification was more commonly seen in patients progressing on enzalutamide than those progressing on abiraterone or other agents (53% vs. 17% vs. 21%). In patients changing therapy to enzalutamide after cfDNA collection, AR gene aberration (copy number increase and/or an exon 8 mutation) in pretreatment cfDNA was associated with adverse outcomes including lower therapy response rates. Presence of AR gene aberrations also correlated with decreased median clinical or radiographic progression free survival (2.3 versus 7.0 months) [85]. Studies such as this have the potential to provide not only a useful clinical index for following disease progression, but also to elucidate specific mechanisms of response or resistance that may facilitate therapy selection.

10.8 Cell Free RNA and MicroRNA

Cancer-associated RNAs present in peripheral blood may offer prognostic or predictive value by reflecting tumor cell transcriptional profiles such as overexpression of particular driver genes or splice variants [86]. Whole blood RNA can be analyzed

from freshly collected specimens or alternatively using preservative tubes (e.g. PaxGene, Qiagen; RNA Streck, Streck Inc.) meant to minimize degradation and capture an RNA "snapshot" immediately after the blood is drawn. Olmos et al. used microarray-based expression profiling of whole blood samples from patients with advanced CRPC and patients undergoing active surveillance to identify expression patterns specific for aggressive disease. A nine-gene signature was developed predictive of worse overall survival (10.7 months vs 25.6 months) [87]. In another study by Ross et al., a panel of 168 inflammation and prostate cancer–related genes were assessed with optimized quantitative PCR, and a six-gene model (*ABL2*, *SEMA4D*, *ITGAL*, and *C1QA*, *TIMP1*, *CDKN1A*) discriminated patients with CRPC into a low-risk group with a median survival of more than 34.9 months and a high-risk group with a median survival of just 7.8 months [88]. Zander et al. analyzed expression profiles from Paxgene stabilized samples among patients with non-small cell lung cancer (NSCLC) identifying consistent differential gene expression among cancer patients in comparison to healthy controls. The expression profiles were then applied to diagnostic classification of patients and demonstrated high specificity for NSCLC diagnosis in comparison to healthy controls [89].

MicroRNAs (miRNAs) regulate gene expression and impact cell proliferation, differentiation, migration, and disease progression. These short (18–22 nucleotides) noncoding RNAs exhibit considerable stability even when exposed to high temperatures and extreme pH values and therefore may constitute an especially valuable analyte for biomarker purposes [90]. In prostate cancer, several miRNAs are upregulated and correlate with prognosis in patients with CRPC [91]. Therefore, Lin et al. assessed the value of circulating miRNA as an early biomarker for chemotherapy response in patients with CRPC. Taqman MicroRNA Array was used to measure the levels of 46 candidate miRNAs collected prior to and following chemotherapy treatment, and 12 of these were significantly associated with overall survival. Risk of death was 2.3–3.5 fold higher in patients found to have elevated pre-docetaxel levels of the miR-200 family and 2.6–3.3 times higher in patients with decreased or unchanged expression of miR-17 family after docetaxel treatment (Fig. 10.2) [92]. In NSCLC, Wang et al. conducted serum profiling of miRNAs involved in the TGF-β pathway, which plays a central role in cell proliferation, differentiation, apoptosis, and invasion. Seventeen miRNAs were significantly associated with survival and were used to generate an miRNA risk score correlated with a 2.5 fold increased risk of death and 7.8 month decreased median survival when elevated in the setting of advanced NSCLC [93] (Figs. 10.3 and 10.4).

10.9 The Rapidly Evolving Cell-Free Landscape

As illustrated by these studies, circulating DNA, RNA, and miRNA can reflect the evolution of disease-related genetic alterations and have significant potential as early non-invasive indicators of therapeutic response and survival. Recent efforts seek to apply cell-free analysis not only to profile existing cancers, but also to screen

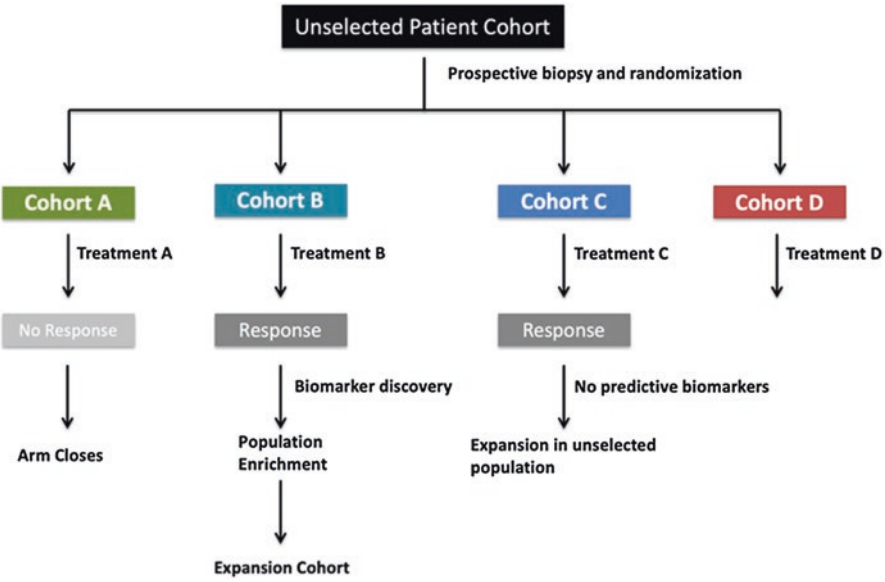

Fig. 10.2 miRNA expression Kaplan Meier curves. (Figure 3 from Ref. [92])

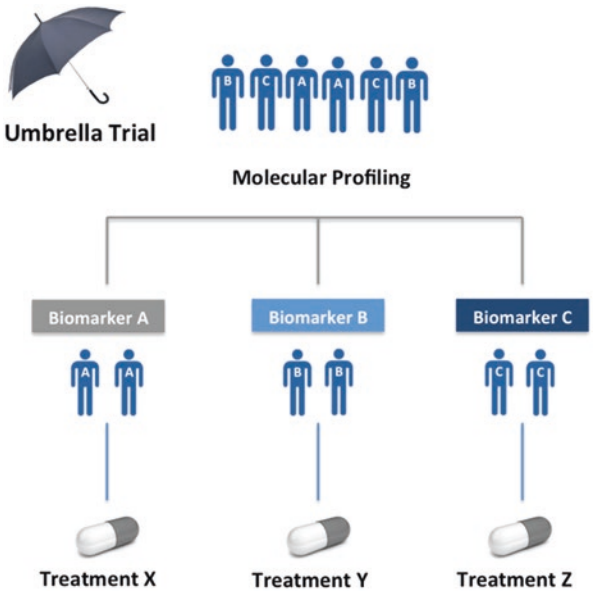

Fig. 10.3 Exosome AR-V7 expression waterfall plot. (Figure 3 from Ref. [122])

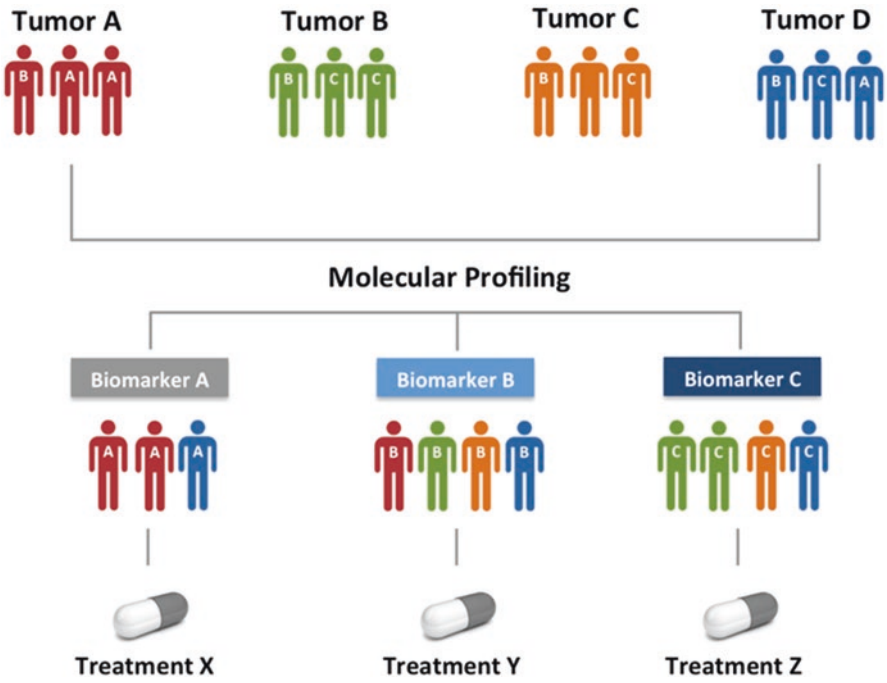

Fig. 10.4 Liquid biopsy analyte overview

and diagnose occult disease. For example, one ongoing prostate cancer trial is testing whether copy number instability in ctDNA of men with elevated PSA undergoing prostate biopsy correlates with prostate cancer diagnosis, potentially reshaping the landscape of prostate cancer screening (NCT02771769, clinicaltrials.gov). On a wider scale, GRAIL (Illumina) is pursuing the ambitious goal of early detection of all common cancers in asymptomatic individuals, hoping to identify malignancies at earlier, more curable stages [94]. Discovery, analytical and clinical validation, further refinement and confirmation of clinical utility of such screening panels, naturally, will require a highly orchestrated, extensive effort as cohorts numbering in the thousands are evaluated and followed prospectively over many years to determine clinical outcomes.

Given the rapid advances in ctDNA analysis, protocol standardization is essential because variations in collection and processing protocols can significantly impact DNA yield and outcomes [95]. Another challenge is posed by elevated levels of cfDNA from benign sources, which may be aggravated during inflammation and injury, further diluting ctDNA and interfering with meaningful detection and analysis [96]. DNA quantification methods (spectrophotometric, fluorescent dyes, quantitative PCR-based) often produce variable results as these measurements target only amplifiable DNA as opposed to total DNA [97]. One recent approach to this challenge has been evaluation of nucleosome footprints. Snyder et al. found

deep sequencing of cfDNA yields a dense, genome wide map of nucleosome occupancy that enables identification of cell-types of origin [81]. cfDNA nucleosome occupancy correlated well with nuclear architecture, gene structure, and gene expression observed in cells, suggesting strong potential to reflect cell type of origin and relative tissue contributions to cfDNA in the presence of malignancies. Similarly, Ulz et al. identified nucleosome promoter occupancy correlating with gene expression from hematopoietic cells in healthy donors and cancer driver genes in metastatic cancer donors. By sequencing DNA fragments after micrococcal nuclease digestion, they identified specific nucleosome patterns at promoter sites influencing gene regulation [82]. Although these techniques do not fully circumvent the caveats posed by cfDNA release mainly from dying cells, they do offer increasingly accurate information about the cell type(s) from which the cfDNA derives.

10.10 Extracellular Vesicles (EVs)

Extracellular vesicles comprise a group of structures ranging from 30 nanometers to a few micrometers enclosed in a lipid bilayer and released into the extracellular space by cells. Differentiated by size and mechanism of release, these vesicles may be broadly divided into exosomes (30–100 nm), ectosomes (0.05–1 μm), apoptotic bodies (1–4 μm), and large oncosomes (1–10 μm). These vesicles have been demonstrated to harbor distinct elements including DNA, RNA, proteins and lipids with direct involvement in numerous physiologic and malignant processes and are secreted by most cell types allowing identification in body fluids including blood and urine [98]. Exosomes are unique among EVs in that they are the product of intracellular multivesicular body fusion with the cell membrane as opposed to direct budding from the plasma membrane characteristic of other EVs. Though their significance was initially unclear, EVs are now known to comprise an integral aspect of intercellular communication mediated by miRNAs, mRNAs and proteins contained and transported within [99, 100]. The differential roles of EVs continues to be explored. For instance, in the setting of malignancies, exosome mediated cell to cell communication has been implicated in proliferation, invasion, motility, immune activation, drug resistance and metastasis, ultimately promoting tumor growth and disease progression [101–103].

In a pre-clinical model, Zomer et al. observed intercellular transfer of mRNAs via EVs, resulting in modulation of invasive and metastatic potential. Using a Cre-LoxP reporter system, labeled tumor cells as well as host non-tumor cells were shown to take up tumor-derived EVs, with transfer from benign cells occurring less frequently, consistent with prior reports that tumor cells generally secrete greater numbers of vesicles [99, 104]. Importantly, tumor cell lines deemed less malignant displayed increased migration and metastatic potential after uptake of EVs from more aggressive tumor cell lines. Analysis of mRNAs carried in EVs revealed enrichment for genes involved in migration and metastasis [105]. In another preclinical study, Costa-Silva et al. examined exosomes isolated from a murine pancre-

atic ductal adenocarcinoma cell line and subsequently injected into mice. Exosome uptake was observed preferentially in liver Kupffer cells, resulting in increased secretion of TGF-beta and fibronectin which served to attract macrophages and create a pre-metastatic niche in which liver metastasis ultimately formed [106]. EVs have a recognized role in immune activation and drug resistance through a number of mechanisms [107–109]. Dendritic cells that encounter antigens within exosomes induce an immune response including anti-tumoral activity, and mice vaccinated with tumor-derived exosomes subsequently develop anti-tumor effects mediated by CD8+ T cell activation. Though not yet fully understood, exosome-mediated drug resistance has been postulated to occur through neutralization of antibody-based drugs, delivery of drug resistant proteins, and export of drugs from cells [110]. Due to their accessibility in multiple body fluids, preferential release from tumor cells, and unique contents specific to cells of origin, EVs have the potential for myriad applications in the evaluation, care, monitoring and treatment of cancer [99].

10.11 EV Isolation

EVs have been isolated from a number of body fluids including plasma, serum, saliva and urine [111]. Techniques for isolation vary and must be modified according to desired vesicle for isolation but generally employ some form of ultracentrifugation, ultrafiltration, or immunoaffinity enrichment. Differential centrifugation is the most established method of isolating EVs, and its primary limitation is the requirement for ultracentrifuge instrumentation and prolonged spin times for sample processing. Alternatively, ultrafiltration methods result in excellent particle yields and efficiency, particularly when coupled with size exclusion chromatography, but often require greater quantity of initial material with variable sample purity [112]. Antibody dependent methods utilizing common surface markers such as CD9, CD63, CD81 and Rab5 have been used to isolate exosomes as well. While less time-consuming, these techniques frequently result in limited yields and are comparatively expensive. EV isolation methods continue to be refined with goals of improved yield, purity, and ease of collection. Due to the difficulties of isolation as well as noted variety of EVs with unique contents, confirmation of EV specimen type and quantification are necessary and may be performed through protein characterization, single vesicle characterization by transmission electron or atomic force microscopy, or functional assays [113].

10.12 Clinical Applications of EVs

The majority of EV clinical studies have focused on evaluating exosomes as potential biomarkers for disease status. Although they are known to contain protein and DNA from parent cells, most investigations have focused on the identification of

relevant RNA and miRNAs within exosomes. Hannafon et al. extracted miRNAs from exosomes of various breast cancer cell lines. Using high throughput sequencing, highly enriched miRNAs were identified and subsequently confirmed to be overexpressed in breast cancer patients in comparison to healthy controls [114]. Huang et al. identified miRNAs in circulating exosomes of patients with CRPC. Levels of two unique miRNAs, mi-1290 and mi-375, were significantly elevated and associated with poor overall survival [115]. Similarly, Matsumoto et al. examined exosome levels in patients with esophageal squamous cell cancer and found that expression of miR-375 was associated with poor prognosis [116]. In pancreatic ductal adenocarcinoma (PDAC), expression of macrophage migration inhibitory factor (MIF) within exosomes was associated with disease progression [106]. Evaluation of exosomal RNA has been undertaken in multiple disease states including melanoma, glioblastoma, lung, ovarian, and colorectal cancers. Consistently, patients with malignancies demonstrate elevated levels of specific exosomal miRNAs [117–121]. These may serve as biomarkers of disease progression, and their underlying mechanistic roles may offer new therapeutic targets in the management of advanced malignancies. Exosomal content beyond miRNAs also has been analyzed. Del Re et al. evaluated AR-V7 expression in exosome-derived RNA using digital droplet polymerase chain reaction (ddPCR). Similar to findings in CTCs among men with prostate cancer, exosomal derived RNA expression accurately predicted resistance to novel hormonal therapies. Among patients receiving abiraterone or enzalutamide for CRPC, AR-V7 variants were found in 39%, and their presence was associated with lower overall survival (median 8 months versus not reached). Response to hormonal therapy was just 7% in AR-V7+ patients versus 64% in AR-V7- patients [122].

Exosome and other EV analysis may ultimately develop into a uniquely valuable liquid biopsy analyte because of the variety of information which may be obtained (mRNA, miRNA, protein, DNA) combined with the comparatively robust nature of the source material. EV contents are sheltered and relatively well-preserved from degradation until they are evaluated. At the same time, methods of isolation must be refined and standardized to ensure purity, reproducibility and scalability for reliable clinical use. How the information gleaned from EVs will be integrated into clinical practice is as yet undefined but clearly holds great prognostic, diagnostic, and even therapeutic potential.

10.13 Conclusion

Precision medicine holds the promise of effective, individualized cancer therapy that evolves over time with a tumor's biology. This exciting new treatment paradigm cannot be adequately supported by traditional tumor biopsies that are constrained by patient safety, cost, and limited tissue sampling. Liquid biopsies may provide the solution: simple, noninvasive sampling of tumor cells or tumor-derived material circulating throughout the body, which can be analyzed repeatedly to guide therapy

as the disease progresses. Significant technical, biological and clinical challenges remain: In the technical realm, optimal methods for enrichment, recovery, amplification, and analysis of rare cells or cell-free material are still being refined and analytically validated. Biologically, the significance of detected alterations are still being elucidated: How do they compare with findings in tumor biopsies? How do they change over time as tumors evolve in response to treatment? What bioinformatics approaches are most suitable for addressing cellular heterogeneity and extrapolating from highly amplified single events to overall tumor driver phenotypes? Clinically, which of these phenotypes translate into presence of occult disease, better or worse outcomes for disease, or response to particular targeted therapies – questions that in the clinical realm can only be satisfactorily addressed through prospective trials. As reviewed in this chapter, significant progress has been made on all these fronts, and CTCs, cfDNA, cfRNA, and EVs already have been successfully captured and analyzed to yield important biological insights about evolving disease drivers, as well as preliminary prognostic and predictive clinical benefits. There is now little doubt that these analytes provide a wealth of valuable information, propelling liquid biopsy towards a central role in precision cancer care.

References

1. Marusyk A, Almendro V, Polyak K (2012) Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer 12:323–334
2. Yap TA, Gerlinger M, Futreal PA, Pusztai L, Swanton C (2012) Intratumor heterogeneity: seeing the wood for the trees. Sci Transl Med 4:127ps10
3. Bedard PL, Hansen AR, Ratain MJ, Siu LL (2013) Tumour heterogeneity in the clinic. Nature 501:355–364
4. Wang Y et al (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 512:155–160
5. Paget S (1989) The distribution of secondary growths in cancer of the breast. 1889. Cancer Metastasis Rev 8:98–101
6. Steeg PS (2006) Tumor metastasis: mechanistic insights and clinical challenges. Nat Med 12:895–904
7. Gupta GP, Massagué J (2006) Cancer metastasis: building a framework. Cell 127:679–695
8. Shen M, Cancer M (2015) The complex seeds of metastasis. Nature 520:298–299
9. Allard WJ et al (2004) Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res Off J Am Assoc Cancer Res 10:6897–6904
10. Aceto N et al (2014) Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. Cell 158:1110–1122
11. Au SH et al (2016) Clusters of circulating tumor cells traverse capillary-sized vessels. Proc Natl Acad Sci U S A 113:4947–4952
12. Ashworth T (1869) A case of cancer in which cells similar to those in the tumors were seen in the blood after death. Aust Med J 14:146–147
13. Ignatiadis M, Lee M, Jeffrey SS (2015) Circulating tumor cells and circulating tumor DNA: challenges and opportunities on the path to clinical utility. Clin Cancer Res Off J Am Assoc Cancer Res 21:4786–4800

14. Alix-Panabières C, Pantel K (2014) Challenges in circulating tumour cell research. Nat Rev Cancer 14:623–631
15. Thiery JP (2002) Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer 2:442–454
16. Mikolajczyk SD et al (2011) Detection of EpCAM-negative and cytokeratin-negative circulating tumor cells in peripheral blood. J Oncol 2011:252361
17. Serrano MJ et al (2014) EMT and EGFR in CTCs cytokeratin negative non-metastatic breast cancer. Oncotarget 5:7486–7497
18. Yu M et al (2013) Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. Science 339:580–584
19. Ring A et al (2015) EpCAM based capture detects and recovers circulating tumor cells from all subtypes of breast cancer except claudin-low. Oncotarget 6:44623–44634
20. Pecot CV et al (2011) A novel platform for detection of CK+ and CK- CTCs. Cancer Discov 1:580–586
21. Anantharaman A et al (2016) Programmed death-ligand 1 (PD-L1) characterization of circulating tumor cells (CTCs) in muscle invasive and metastatic bladder cancer patients. BMC Cancer 16:744
22. Alix-Panabières C, Pantel K (2014) Technologies for detection of circulating tumor cells: facts and vision. Lab Chip 14:57–62
23. Cohen SJ et al (2008) Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. J Clin Oncol Off J Am Soc Clin Oncol 26:3213–3221
24. Karabacak NM et al (2014) Microfluidic, marker-free isolation of circulating tumor cells from blood samples. Nat Protoc 9:694–710
25. Gupta V et al (2012) ApoStream(™), a new dielectrophoretic device for antibody independent isolation and recovery of viable cancer cells from blood. Biomicrofluidics 6:24133
26. The DEPArray system. Available at http://www.siliconbiosystems.com/deparray-system
27. Ferreira MM, Ramani VC, Jeffrey SS (2016) Circulating tumor cell technologies. Mol Oncol 10:374–394
28. de Bono JS et al (2008) Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. Clin Cancer Res Off J Am Assoc Cancer Res 14:6302–6309
29. Cristofanilli M et al (2005) Circulating tumor cells: a novel prognostic factor for newly diagnosed metastatic breast cancer. J Clin Oncol Off J Am Soc Clin Oncol 23:1420–1430
30. Smerage JB, Hayes DF (2006) The measurement and therapeutic implications of circulating tumour cells in breast cancer. Br J Cancer 94:8–12
31. Wang J et al (2013) Prognostic significance of circulating tumor cells in non-small-cell lung cancer patients: a meta-analysis. PLoS One 8:e78070
32. Goldkorn A et al (2014) Circulating tumor cell counts are prognostic of overall survival in SWOG S0421: a phase III trial of docetaxel with or without atrasentan for metastatic castration-resistant prostate cancer. J Clin Oncol Off J Am Soc Clin Oncol 32:1136–1142
33. Cristofanilli M et al (2004) Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med 351:781–791
34. Lucci A et al (2012) Circulating tumour cells in non-metastatic breast cancer: a prospective study. Lancet Oncol 13:688–695
35. Goldkorn A et al (2015) Circulating tumor cell telomerase activity as a prognostic marker for overall survival in SWOG 0421: a phase III metastatic castration resistant prostate cancer trial. Int J Cancer 136:1856–1862
36. Naito T et al (2012) Prognostic impact of circulating tumor cells in patients with small cell lung cancer. J Thorac Oncol Off Publ Int Assoc Study Lung Cancer 7:512–519
37. Wallwiener M et al (2014) Serial enumeration of circulating tumor cells predicts treatment response and prognosis in metastatic breast cancer: a prospective study in 393 patients. BMC Cancer 14:512

38. Tsai W-S et al (2016) Circulating tumor cell count correlates with colorectal neoplasm progression and is a prognostic marker for distant metastasis in non-metastatic patients. Sci Rep 6:24517
39. Bidard FC et al (2013) Circulating tumor cells in locally advanced pancreatic adenocarcinoma: the ancillary CirCe 07 study to the LAP 07 trial. Ann Oncol Off J Eur Soc Med Oncol ESMO 24:2057–2061
40. Koyanagi K et al (2010) Serial monitoring of circulating tumor cells predicts outcome of induction biochemotherapy plus maintenance biotherapy for metastatic melanoma. Clin Cancer Res Off J Am Assoc Cancer Res 16:2402–2408
41. Poveda A et al (2011) Circulating tumor cells predict progression free survival and overall survival in patients with relapsed/recurrent advanced ovarian cancer. Gynecol Oncol 122:567–572
42. Alva A et al (2015) Circulating tumor cells as potential biomarkers in bladder cancer. J Urol 194:790–798
43. Uenosono Y et al (2013) Clinical significance of circulating tumor cells in peripheral blood from patients with gastric cancer. Cancer 119:3984–3991
44. Smerage JB et al (2014) Circulating tumor cells and response to chemotherapy in metastatic breast cancer: SWOG S0500. J Clin Oncol Off J Am Soc Clin Oncol 32:3483–3489
45. Singh D et al (2002) Gene expression correlates of clinical prostate cancer behavior. Cancer Cell 1:203–209
46. Lopez NE, Weiss AC, Robles J, Fanta P, Ramamoorthy SL (2016) A systematic review of clinically available gene expression profiling assays for stage II colorectal cancer: initial steps toward genetic staging. Am J Surg 212:700–714. https://doi.org/10.1016/j.amjsurg.2016.06.019
47. Subramanian J, Simon R (2010) Gene expression-based prognostic signatures in lung cancer: ready for clinical use? J Natl Cancer Inst 102:464–474
48. Darshan MS et al (2011) Taxane-induced blockade to nuclear accumulation of the androgen receptor predicts clinical responses in metastatic prostate cancer. Cancer Res 71:6019–6029
49. Miyamoto DT et al (2012) Androgen receptor signaling in circulating tumor cells as a marker of hormonally responsive prostate cancer. Cancer Discov 2:995–1003
50. Crespo M et al (2015) Androgen receptor expression in circulating tumour cells from castration-resistant prostate cancer patients treated with novel endocrine agents. Br J Cancer 112:1166–1174
51. Punnoose EA et al (2015) PTEN loss in circulating tumour cells correlates with PTEN loss in fresh tumour tissue from castration-resistant prostate cancer patients. Br J Cancer 113:1225–1233
52. Antonarakis ES et al (2014) AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. N Engl J Med 371:1028–1038
53. Scher HI et al (2016) Association of AR-V7 on circulating tumor cells as a treatment-specific biomarker with outcomes and survival in castration-resistant prostate cancer. JAMA Oncol 2:1441–1449. https://doi.org/10.1001/jamaoncol.2016.1828
54. Kalinsky K et al (2015) Correlation of hormone receptor status between circulating tumor cells, primary tumor, and metastasis in breast cancer patients. Clin Transl Oncol Off Publ Fed Span Oncol Soc Natl Cancer Inst Mex 17:539–546
55. Babayan A et al (2013) Heterogeneity of estrogen receptor expression in circulating tumor cells from metastatic breast cancer patients. PLoS One 8:e75038
56. Meng S et al (2004) HER-2 gene amplification can be acquired as breast cancer progresses. Proc Natl Acad Sci U S A 101:9393–9398
57. Reijm EA et al (2016) An 8-gene mRNA expression profile in circulating tumor cells predicts response to aromatase inhibitors in metastatic breast cancer patients. BMC Cancer 16:123
58. Thomas A, Rajan A, Lopez-Chavez A, Wang Y, Giaccone G (2013) From targets to targeted therapies and molecular profiling in non-small cell lung carcinoma. Ann Oncol Off J Eur Soc Med Oncol 24:577–585

59. Maheswaran S et al (2008) Detection of mutations in EGFR in circulating lung-cancer cells. N Engl J Med 359:366–377
60. Abdallah EA et al (2015) Thymidylate synthase expression in circulating tumor cells: a new tool to predict 5-fluorouracil resistance in metastatic colorectal cancer patients. Int J Cancer 137:1397–1405
61. Kuhlmann JD et al (2014) ERCC1-positive circulating tumor cells in the blood of ovarian cancer patients as a predictive biomarker for platinum resistance. Clin Chem 60:1282–1289
62. Ni X et al (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc Natl Acad Sci U S A 110:21083–21088
63. Heitzer E et al (2013) Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. Cancer Res 73:2965–2975
64. Lyberopoulou A et al (2015) Mutational analysis of circulating tumor cells from colorectal cancer patients and correlation with primary tumor tissue. PLoS One 10:e0123902
65. Lohr JG et al (2014) Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nat Biotechnol 32:479–484
66. Yu M et al (2012) RNA sequencing of pancreatic circulating tumour cells implicates WNT signalling in metastasis. Nature 487:510–513
67. Gorges TM et al (2016) Accession of tumor heterogeneity by multiplex transcriptome profiling of single circulating tumor cells. Clin Chem 62:1504–1515
68. Miyamoto DT et al (2015) RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science 349:1351–1356
69. Hodgkinson CL et al (2014) Tumorigenicity and genetic profiling of circulating tumor cells in small-cell lung cancer. Nat Med 20:897–903
70. Kidess E, Jeffrey SS (2013) Circulating tumor cells versus tumor-derived cell-free DNA: rivals or partners in cancer care in the era of single-cell analysis? Genome Med 5:70
71. Alix-Panabières C, Pantel K (2016) Clinical applications of circulating tumor cells and circulating tumor DNA as liquid biopsy. Cancer Discov 6:479–491
72. Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P (2001) About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. Clin Chim Acta Int J Clin Chem 313:139–142
73. Heitzer E, Ulz P, Geigl JB (2015) Circulating tumor DNA as a liquid biopsy for cancer. Clin Chem 61:112–123
74. Swarup V, Rajeswari MR (2007) Circulating (cell-free) nucleic acids–a promising, non-invasive tool for early detection of several human diseases. FEBS Lett 581:795–799
75. Björkman L, Reich CF, Pisetsky DS (2003) The use of fluorometric assays to assess the immune response to DNA in murine systemic lupus erythematosus. Scand J Immunol 57:525–533
76. Chen Z et al (2009) Analysis of cancer mutation signatures in blood by a novel ultra-sensitive assay: monitoring of therapy or recurrence in non-metastatic breast cancer. PLoS One 4:e7220
77. Hindson BJ et al (2011) High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem 83:8604–8610
78. Forshew T et al (2012) Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 4:136ra68
79. Higgins MJ et al (2012) Detection of tumor PIK3CA status in metastatic breast cancer using peripheral blood. Clin Cancer Res Off J Am Assoc Cancer Res 18:3462–3469
80. Ignatiadis M, Dawson S-J (2014) Circulating tumor cells and circulating tumor DNA for precision medicine: dream or reality? Ann Oncol Off J Eur Soc Med Oncol 25:2304–2313
81. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J (2016) Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. Cell 164:57–68
82. Ulz P et al (2016) Inferring expressed genes by whole-genome sequencing of plasma DNA. Nat Genet 48:1273–1278

83. Dawson S-J et al (2013) Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 368:1199–1209
84. Tie J et al (2015) Circulating tumor DNA as an early marker of therapeutic response in patients with metastatic colorectal cancer. Ann Oncol Off J Eur Soc Med Oncol 26:1715–1722
85. Azad AA et al (2015) Androgen receptor gene aberrations in circulating cell-free DNA: biomarkers of therapeutic resistance in castration-resistant prostate cancer. Clin Cancer Res Off J Am Assoc Cancer Res 21:2315–2324
86. Rönnau CGH, Verhaegh GW, Luna-Velez MV, Schalken JA (2014) Noncoding RNAs as novel biomarkers in prostate cancer. Biomed Res Int 2014:591703
87. Olmos D et al (2012) Prognostic value of blood mRNA expression signatures in castration-resistant prostate cancer: a prospective, two-stage study. Lancet Oncol 13:1114–1124
88. Ross RW et al (2012) A whole-blood RNA transcript-based prognostic model in men with castration-resistant prostate cancer: a prospective study. Lancet Oncol 13:1105–1113
89. Zander T et al (2011) Blood-based gene expression signatures in non-small cell lung cancer. Clin Cancer Res Off J Am Assoc Cancer Res 17:3360–3367
90. Cortez MA et al (2011) MicroRNAs in body fluids–the mix of hormones and biomarkers. Nat Rev Clin Oncol 8:467–477
91. Zhang H-L et al (2011) Serum miRNA-21: elevated levels in patients with metastatic hormone-refractory prostate cancer and potential predictive factor for the efficacy of docetaxel-based chemotherapy. Prostate 71:326–331
92. Lin H-M et al (2014) Circulating microRNAs are associated with docetaxel chemotherapy outcome in castration-resistant prostate cancer. Br J Cancer 110:2462–2471
93. Wang Y et al (2013) Pathway-based serum microRNA profiling and survival in patients with advanced stage non-small cell lung cancer. Cancer Res 73:4801–4809
94. Press Release. Available at http://www.illumina.com/company/news-center/press-releases/press-release-details.html?newsid=2127903. Accessed 3 Jan 2017
95. Swinkels DW, Wiegerinck E, Steegers EAP, de Kok JB (2003) Effects of blood-processing protocols on cell-free DNA quantification in plasma. Clin Chem 49:525–526
96. Qin Z, Ljubimov VA, Zhou C, Tong Y, Liang J (2016) Cell-free circulating tumor DNA in cancer. Chin J Cancer 35:36
97. Devonshire AS et al (2014) Towards standardisation of cell-free DNA measurement in plasma: controls for extraction efficiency, fragment size bias and quantification. Anal Bioanal Chem 406:6499–6512
98. Ciardiello C et al (2016) Focus on extracellular vesicles: new frontiers of cell-to-cell communication in cancer. Int J Mol Sci 17:175
99. Balaj L et al (2011) Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. Nat Commun 2:180
100. D'Asti E, Chennakrishnaiah S, Lee TH, Rak J (2016) Extracellular vesicles in brain tumor progression. Cell Mol Neurobiol 36:383–407
101. Melo SA et al (2014) Cancer exosomes perform cell-independent microRNA biogenesis and promote tumorigenesis. Cancer Cell 26:707–721
102. O'Driscoll L (2015) Expanding on exosomes and ectosomes in cancer. N Engl J Med 372:2359–2362
103. Ma L (2016) MicroRNA and metastasis. Adv Cancer Res 132:165–207
104. Raposo G, Stoorvogel W (2013) Extracellular vesicles: exosomes, microvesicles, and friends. J Cell Biol 200:373–383
105. Zomer A et al (2015) In vivo imaging reveals extracellular vesicle-mediated phenocopying of metastatic behavior. Cell 161:1046–1057
106. Costa-Silva B et al (2015) Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. Nat Cell Biol 17:816–826
107. Clayton A, Mitchell JP, Court J, Mason MD, Tabi Z (2007) Human tumor-derived exosomes selectively impair lymphocyte responses to interleukin-2. Cancer Res 67:7458–7466

108. Robbins PD, Morelli AE (2014) Regulation of immune responses by extracellular vesicles. Nat Rev Immunol 14:195–208
109. Taylor DD, Gercel-Taylor C (2011) Exosomes/microvesicles: mediators of cancer-associated immunosuppressive microenvironments. Semin Immunopathol 33:441–454
110. Giallombardo M, Taverna S, Alessandro R, Hong D, Rolfo C (2016) Exosome-mediated drug resistance in cancer: the near future is here. Ther Adv Med Oncol 8:320–322
111. Rolfo C et al (2014) Liquid biopsies in lung cancer: the new ambrosia of researchers. Biochim Biophys Acta 1846:539–546
112. Lobb RJ et al (2015) Optimized exosome isolation protocol for cell culture supernatant and human plasma. J Extracellular Vesicles 4:27031
113. Lötvall J et al (2014) Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. J Extracellular Vesicles 3:26913
114. Hannafon BN et al (2016) Plasma exosome microRNAs are indicative of breast cancer. Breast Cancer Res BCR 18:90
115. Huang X et al (2015) Exosomal miR-1290 and miR-375 as prognostic markers in castration-resistant prostate cancer. Eur Urol 67:33–41
116. Matsumoto Y et al (2016) Quantification of plasma exosome is a potential prognostic marker for esophageal squamous cell carcinoma. Oncol Rep 36:2535–2543
117. Pfeffer SR et al (2015) Detection of exosomal miRNAs in the plasma of melanoma patients. J Clin Med 4:2012–2027
118. Skog J et al (2008) Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. Nat Cell Biol 10:1470–1476
119. Rabinowits G, Gerçel-Taylor C, Day JM, Taylor DD, Kloecker GH (2009) Exosomal microRNA: a diagnostic marker for lung cancer. Clin Lung Cancer 10:42–46
120. Taylor DD, Gercel-Taylor C (2008) MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer. Gynecol Oncol 110:13–21
121. Ogata-Kawata H et al (2014) Circulating exosomal microRNAs as biomarkers of colon cancer. PLoS One 9:e92921
122. Del Re M et al (2016) The detection of androgen receptor splice variant 7 in plasma-derived exosomal RNA strongly predicts resistance to hormonal therapy in metastatic prostate cancer patients. Eur Urol 71:680–687. https://doi.org/10.1016/j.eururo.2016.08.012

Chapter 11
Data Portals and Analysis

Angelica Ochoa and Nikolaus Schultz

Keywords Cancer genomics · Next-generation sequencing · Data visualization · Driver alterations · Variant annotation · Clinical actionability · Data repositories · Explorative data analysis · Databases · Data repositories · Integrative analysis

The development of powerful and scalable methods to analyze cancer genomic data has transformed the practice of medicine by bridging the gap between complex genomic data and their use in research and clinical translation. Of particular significance are the data portals and analysis platforms available that make genomic data more accessible to the research community, as they enable translational research by providing scientists with the information, tools, and frameworks necessary to perform integrative analyses on cancer genomic data.

This chapter will review publicly available data portals and translational research platforms for cancer genomics. Here we will focus on large-scale cancer genomics projects, research platforms and select, unique features for visualizing and analyzing data, as well as current limitations, challenges, and opportunities faced by the research community.

11.1 The Evolution of Cancer Genomics

Technological advancements over the last several decades have driven the evolution of cancer genomics and have significantly expanded our understanding of the molecular basis of cancer genetics. In particular, the development of high-throughput technologies and powerful computational methods have provided the means to investigate cancer genomes in a manner not previously possible. Before the advent of these technologies, traditional approaches to biological inquiries were limited by

A. Ochoa · N. Schultz (✉)
Marie-Josée & Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA
e-mail: schultz@cbio.mskcc.org

© Springer Nature Switzerland AG 2019

S. Roychowdhury, E. M. Van Allen (eds.), *Precision Cancer Medicine*,
https://doi.org/10.1007/978-3-030-23637-3_11

time-consuming molecular assays that can be difficult to scale up to interrogate a large set of genes of interest. Indeed, high-throughput technologies are not only more cost- and time-efficient than traditional approaches; they also offer a more comprehensive view of cancer genomes and have led to a series of breakthroughs in personalized medicine [1].

Unlike early sequencing technologies, high-throughput sequencing technologies are not limited to a single modality – the same set of data can be distilled into multiple molecular profiles that offer different, yet complementary, views of cancer genomes. For example, nucleotide substitutions, insertions and deletions (indels), copy-number alterations, and structural variants can all be derived from a single set of whole genome sequencing (WGS) or whole exome sequencing (WES) data [1, 2]. Integrating these highly informative datasets with clinical information creates the opportunity to ask and potentially answer deeper, more meaningful questions of the data in the context of clinical care. Resources that can capture this information to enable treatment decisions are essential but can be difficult to integrate into clinical settings without the adequate tools and infrastructure to manage and explore the torrent of omics data.

11.2 Navigating the Maze of Omics Data: Challenges and Opportunities

Efficiently deriving biological insights from cancer genomics data remains one of the major challenges in navigating the maze of omics data. A key step towards achieving this is differentiating between genomic alterations that contribute to cancer (drivers) and those that do not (passengers). Several methods to identify drivers rely on recurrence, i.e., identifying genomic alterations that occur at a higher rate than expected by chance across a set of tumors. These methods take many factors into consideration, such as mutation frequencies, gene lengths, types of mutations, and copy-number alterations [3]. To exclude genes from further downstream analyses, many methods also consider the functional impact of mutations and correlations with mRNA expression. However, given their nature even the most sophisticated recurrence-based methods cannot identify rare drivers.

Another step towards efficiently deriving biological insights from cancer genomics data is developing methods that address complications introduced by batch effects, uneven sequence coverage, differences in reference genome assembly used, tumor heterogeneity, and tumor purity [4]. Complications such as these have led to seriously flawed findings in the past when not removed from the data [5]. Methods that correct these complications will help ensure quality control going forward and improve the reproducibility of results in new studies. Additionally, such methods make it easier to harmonize data across datasets, thus facilitating comparisons of data from different sources which will be ideal going forward as more and more data becomes available.

11.3 Data Portals and Public Repositories

Cancer genomic data can be divided into the following levels of data depending on the extent of computational modifications and interpretations applied to the data [1]:

1. Raw
2. Processed or normalized
3. Interpreted
4. Summarized

While raw and individual-level sequencing data are often subjected to restricted access to help prevent compromising patient information, interpreted or distilled genomic data are usually publicly accessible. Data repositories have made it easier for investigators to access these publicly available data by creating a single-access point for users to download these data from. However, providing such resources can be logistically and technically difficult if adequate resources are not in place to store, manage, and maintain these data.

Making complex cancer genomics data easier to access remains one of the major challenges in downstream collection, integration, and analysis. Experience has shown that this can best be achieved by reducing the complexity of the data [3]. The resulting simplified genomic data can then be used to facilitate the identification of key pathways involved in the cancer initiation, maintenance, and progression, as well as the identification of potentially actionable targets for cancer treatments. Additionally, interpretations from these data can also be incorporated into clinical resources that enable clinicians to make genome-informed treatment decisions. However, to make cancer genomics data even more translational, there are many obstacles to overcome before cancer genomics data can reach their full potential.

11.3.1 Progress and Observations from Public Projects

Many large-scale coordinated efforts have been established that provide genomic data linked to high quality clinical information. A summary of available large-scale cancer genomic projects is available in Table 11.1.

As one of the first large-scale efforts of its kind, The Cancer Genome Atlas (TCGA), a collaboration between the National Cancer Institute (NCI) and National Human Genome Research Institute (NHGRI), was established in 2005 as a multi-institutional cross-collaborative project aimed at cataloguing genetic alterations in 33 cancer types [6]. Large-scale cancer genomics projects like TCGA have played a key role in expanding our knowledge and understanding of the cancer genome, and will be discussed in more detail in the following sections. Primary data sources for processed and/or pre-processed clinical and genomic data from these projects are accessible through their own project sites in addition to external data repositories such as the Genomic Data Commons (GDC) [13] Data Portal, the cBioPortal for Cancer Genomics [14], and Sage Synapse [15].

Table 11.1 Large-scale cancer genomic projects

Project	Summary
The Cancer Genome Atlas (TCGA) [6]	A multi-institutional cross-collaborative effort established by the National Cancer Institute (NCI) and National Human Genome Research Institute (NHGRI) to comprehensively catalogue genetic alterations in 33 cancer types
Therapeutically Applicable Research To Generate Effective Treatments (TARGET) [7]	An initiative jointly managed by the NCI and Cancer Therapy Evaluation Program (CTEP) comprised of disease-centered projects using comprehensive molecular characterization to determine genetic drivers in the development and progression of childhood cancers.
Cancer Genome Characterization Initiative (CGCI) [8]	An initiative supported by the National Institutes of Health (NIH) and NCI to molecularly characterize adult and pediatric cancers, including B-cell non-Hodgkin lymphoma, medulloblastoma, and HIV-associated cancers.
Cancer Cell Line Encyclopedia (CCLE) [9]	A collaboration between the Broad Institute and the Novartis Institutes for Biomedical Research and its Genomic Institute of the Novartis Research Foundation to conduct detailed genetic and pharmacologic characterization to develop and translate computational analyses that link genomic and pharmacologic data for ~1000 cancer cell lines.
International Caner Genome Consortium (ICGC) [10, 11]	A large coordinated effort to comprehensively elucidate genomic changes in tumors from 50 cancer types throughout the world. The ICGC for medicine (ICGCmed) was also announced in 2016, which aims to link at least 200,000 cancer genomes with clinical information by 2025.
Project GENIE (Genomics Evidence Neoplasia Information Exchange) [12]	A multi-phase, multi-year international data-sharing project supported by the American Association for Cancer Research (AACR) to develop a regulatory-grade registry that aggregates and links clinical-grade cancer genomic data with clinical outcomes from tens of thousands of cancer patients treated at multiple international institutions.

11.3.2 Public Repositories for Cancer Genomics Data

One of the key drivers in enabling precision cancer medicine is the development of data repositories for cancer genomics data. Indeed, large-scale efforts such as the aforementioned projects have shaped our understanding of cancer genetics and continue to expand our knowledge of the molecular basis of cancer. Many data repositories exist to facilitate access to data from these projects by providing a single access point where users can readily view and download cancer genomic datasets and clinical information (Table 11.2). By making these omics data available, data repositories have made collaborations across institutions and laboratories easier and they have been instrumental in the development of computational methods and research tools for cancer research.

In addition to hosting clinical and genomic data from a combination of the aforementioned projects, many public data repositories also host data from other cancer genomic projects and private datasets. The availability of various types of molecular

Table 11.2 Public repositories for cancer genomics data

Data Portal	Summary[a]
TARGET Data Portal https://ocg.cancer.gov/programs/target/data-matrix	Molecular profiling data and clinical information from TARGET projects, focusing on childhood cancers. It currently contains over 5000 samples from 8 projects and remains ongoing.
CCLE Data Portal https://portals.broadinstitute.org/ccle	Mutation, copy number, mRNA expression, and clinical information for ~1000 human cancer cell lines, plus drug response data for ~500 cell lines.
CGCL Data portal https://ocg.cancer/gov/programs/cgci/data-matrix	Molecular profiling data and clinical information for adult and pediatric cancers. It currently contains ~150 sample from the Pediatric Medulloblastoma Project (MB) and the Non-Hodgkin Lymphoma Project (NHL). The following CGCL projects remain ongoing: the Burkitt Lymphoma Characterization Project (BLGSP, expected completed in 2017) and the HIV+ Tumor Molecular Characterization Project (HTMPC)
ICGC Data Portal https://dcc.icgc.org	Molecular profiling data and clinical information from participating projects, including TCGA and TARGT. It currently contains data for more than 19,000 from 76 projects. In 2016, the ICGC also established the ICGC for Medicine (ICGCmed), which aims to link at least 200,000 cancer genomes with clinical information by 2025.
AACR Project GENIE http://www.cbioportal.org/genie/, http://synapse.org/genie	Molecular profiling data and clinical information from participating institutions. It currently contains data from ~19,000 samples contributed by eight major cancer centers.
GDC Data Portal https://gdc-portal.nci.nih.gov	Molecular profiling data and clinical information generated by TCGA and TARGET. It currently contains raw and processed data for more than 14,000 samples from 39 projects. The GDC also plans to make CCLE and CGCL data available in 2017, as well as data contributed by Foundation Medicine, Inc. and the Multiple Myeloma Research Foundation (MMRF). Currently, CCLE data and TCGA RPPA data are available through the GDC legacy archive.
dbGaP https://www.ncbi.nlm.nih.gov/gap	Raw sequencing data from ~780 studies including TCGA, TARGET, CGCI, and several other smaller sequencing studies.
EGA https://www.ebi.ac.uk/ega	Raw sequencing data from ~1430 studies and ~660 data providers, including ICGC and other smaller projects.
ArrayExpress http://www.ebi.ac.uk/arrayexpress	Raw and processed molecular profiling data from over 26,000 experiments using human samples. Used by several smaller cancer genomic studies.
GEO https://www.ncbi.nlm.nih.gov/geo	Raw and processed molecular profiling data for over 1 million human samples from ~1770 data sets. Used by several smaller cancer genomic studies.
Synapse http://www.synapse.org	Curated TCGA data sets, including from the pan-cancer analysis project.
GDAC Firehose http://gdac.broadinstitute.org	Aggregated and processed TCGA data sets, including automated standard analyses (recurrence, clustering, correlations, etc.). GDAC = Genome Data Analysis Center

[a]Number of projects and/or samples available as of November 2016

profiling data, depth of patient and sample clinical information, and overall coverage of cancer projects and smaller cancer genomic studies vary from repository to repository. A comparison of available public repositories and primary data sources is summarized in Table 11.3.

11.3.2.1 The Cancer Genome Atlas (TCGA) [6]

The Cancer Genome Atlas (TCGA) was launched in 2005 as a collaboration between the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) to explore the spectrum of genomic alterations in 33 cancer types from over 11,000 tumor samples. These data were made publicly available with the overarching goal of improving our ability to diagnose, treat, and prevent cancer. Amounting to over 2.5 petabytes of data, TCGA data have served as an immeasurable source of knowledge and have contributed to over 1000 studies by independent researchers and TCGA Research Network publications. The types of data collected for TCGA include gene expression, mutations, copy number alterations, methylation, protein expression, and clinical information. Although coming to a close in 2017, the success of TCGA has demonstrated the power of large-scale coordinated efforts in driving precision cancer medicine and will serve as a model for future NCI programs to follow. All of the TCGA data can be found at the Genomic Data Commons (GDC) Data Portal [13], a unified data repository developed by the NCI and other organizations containing data derived from various legacy and active NCI

Table 11.3 Comparison of public repositories and primary data sources

	Projects							Datatypes Supported							
Data Portal	TCGA	TARGET	CGCI	CCLE	ICGC	AACR Project GENIE	Other projects	Clinical	Mutation	Copy Number	Gene Expression	Methylation	Protein Expression	Primary Data	Processed Data
TARGET Data Portal		X						X	X	X	X	X		X	X
CGCI Data Portal			X					X	X	X	X			X	
CCLE Data Portal				X				X	X	X	X				X
ICGC Data Portal	X	X			X		X	X	X	X	X	X	X	X	X
cBioPortal	X	X		X	X	X	X	X	X	X	X	X	X		X
GDC Data Portal	X	X		X¶			X§	X	X	X	X	X	X¶	X	X
dbGaP	X	X	X				X	X	X	X	X	X	X	X	
EGA					X		X		X	X	X	X	X	X	X
ArrayExpress							X		X	X	X	X	X	X	X
GEO							X		X	X	X	X	X	X	
Synapse	X			X			X	X	X	X	X	X	X	X	X
GDAC Firehose	X							X	X	X	X	X	X		X

¶ Data is available through the GDC legacy archive. § Data from other supported projects will be available sometime in 2017

programs. Processed TCGA data are also hosted by the International Cancer Genome Consortium (ICGC) Data Portal [10], the database of Genotypes and Phenotypes (dbGaP) [16], Sage Synapse [15], Genome Data Analysis Center (GDAC) Firehose [17], and the cBioPortal for Cancer Genomics [14].

11.3.2.2 Therapeutically Applicable Research to Generate Effective Treatments (TARGET) [7]

The Therapeutically Applicable Research to Generate Effective Treatments (TARGET) initiative is a collaborative effort supported by a large, diverse consortium consisting of investigators from the NCI and extramural investigators. The TARGET initiative is an ongoing project managed by the NCI that was launched in 2006 with the overall goal of accelerating the development and application of new, more effective therapeutic strategies for treating pediatric cancers. Using comprehensive molecular characterization, TARGET aims to identify therapeutic targets and prognostic markers by determining the genetic alterations driving the initiation and progression of pediatric cancers. Currently, the TARGET initiative is characterizing subtypes of the following pediatric cancers: acute myeloid leukemia, osteosarcoma, acute lymphoblastic leukemia, neuroblastoma, and select kidney tumors.

Processed TARGET data are publicly available through their project site and their public ftp server to facilitate data access for investigators outside of the initiative and accelerate the discovery of novel therapeutic targets and enable rapid translation of these findings in the clinic. Publicly available data include clinical information, gene expression, copy number alterations, methylation, miRNA expression, and mutations. Additional molecular profile data available for select projects includes targeted resequencing data, ChIP-seq, TaqMan data, and kinome data. TARGET data can also be found at the ICGC Data Portal, GDC Data Portal, dbGaP, and the cBioPortal for Cancer Genomics. It is important to note that these repositories may not contain a complete set of available TARGET data due to ongoing efforts by these repositories to render them compatible with their respective platforms.

11.3.2.3 Cancer Genome Characterization Initiative (CGCI) [8]

The Cancer Genome Characterization Initiative (CGCI) is an ongoing project managed by the NCI that was launched in 2008 to support cutting-edge genomics research on adult and pediatric cancers. To date, the CGCI has successfully completed two projects focused on characterizing genetic alterations in medulloblastoma and B-cell non-Hodgkin lymphoma. Ongoing projects for CGCI are currently focused on characterizing HIV-associated cancers and Burkitt lymphoma. Data gathered for these projects includes clinical information, gene expression, copy number alterations, miRNA, and mutations. Data for the completed projects can be downloaded from the CGCI project site through their public ftp server as well as

through dbGaP. A limited amount of data is currently available for the ongoing projects through these sources as well.

11.3.2.4 Cancer Cell Line Encyclopedia (CCLE) [9]

The Cancer Cell Line Encyclopedia (CCLE) is an ongoing project that was established in 2011 as a collaboration between the Broad Institute, and the Novartis Institutes for Biomedical Research and its Genomics Institute of the Novartis Research Foundation, to develop robust preclinical model systems that reflect the genomic diversity of human cancers. Currently, the CCLE has conducted genetic characterizations for over 1000 cancer cell lines and pharmacological profiling for over 500 cancer cell lines. Available data from the CCLE project includes cell line annotations, sample information, copy number alterations, mutations, binary calls for copy number and mutation data, and pharmacological profiling. Data collection efforts for INDELs and binary calls for copy number and mutation data remain ongoing. CCLE data can be downloaded directly through the project site after setting up a free account. These data are also available through Sage Synapse, the GDC Legacy Archive (https://gdc-portal.nci.nih.gov/legacy-archive/), and the cBioPortal for Cancer Genomics [14].

11.3.2.5 International Cancer Genome Consortium (ICGC) [10]

In 2007, the International Cancer Genome Consortium (ICGC) was established to coordinate the generation of comprehensive catalogues of genomic alterations in tumors from 50 different cancer types and/or subtypes. To date, the ICGC has completed 41 projects characterizing genomic alterations in over 7000 tumor samples originating from 17 tumor sites. Potential datatypes available for each project are clinical information, mutations, copy number alterations, methylation, structural rearrangements, gene expression, protein expression, and miRNA. These data can be found in the ICGC Data Portal, the cBioPortal for Cancer Genomics, and EGA. The ICGC Data Portal provides tools for visualizing, querying, and downloading the data available through the portal. Additionally, users can also filter the data to download by primary site, cancer projects, datatype, tumor type, analysis software, specific mutations, and more.

11.3.2.6 AACR Project GENIE (Genomics Evidence Neoplasia Information Exchange) [12]

Project GENIE is a large-scale international data-sharing project launched in 2015 by the American Association for Cancer Research (AACR) to catalyze precision cancer medicine by providing the statistical power necessary to improve clinical decision-making, particularly for rare cancers and rare variants in common cancers.

Project GENIE is unique in that it aims to develop a regulatory-grade registry that aggregates and links clinical-grade cancer genomic data with clinical outcomes from over tens of thousands of cancer patients treated at multiple institutions. There are currently eight phase 1 project participants that are supporting ongoing genotyping efforts. Altogether the participating institutions have contributed clinical information and genomic data for nearly 19,000 samples and approximately 60 cancer types. The types of genomic data available for Project GENIE include mutations, copy number alterations, and structural rearrangements. These data were released to the public in early 2017 and will be available through the Project GENIE instance of the cBioPortal for Cancer Genomics (http://www.cbioportal.org/genie/) and Sage Synapse (https://www.synapse.org/#!Synapse:syn7222066).

11.3.2.7 Genomic Data Commons (GDC) Data Portal [13]

The GDC Data Portal is an open-source, open access centralized data repository developed by the NCI and other organizations containing data derived from various legacy and active NCI programs. The software for the GDC Data Portal is available through GitHub (https://github.com/NCIP/gdc-docs). Released in 2016, the GDC Data Portal provides a valuable resource that promotes the import and standardization of genomic and clinical data using a common set of bioinformatics pipelines to enable direct comparison of data across cancer research projects. The current implementation of the GDC Data Portal provides harmonized data from all 33 TCGA cancer projects, with the exception of protein expression data, and 6 out of 9 TARGET projects. There are ongoing efforts to harmonize TCGA protein expression data, CCLE data, and CGCI data to render them compatible with the current implementation of the data portal. In the meantime, these data can be found through the GDC Legacy Archive with the exception of CGCI data. Additionally, data contributed by Foundation Medicine, Inc. and the Multiple Myeloma Research Foundation (MMRF) are expected to become publicly available sometime in 2017.

Data can be viewed and retrieved through the GDC Data Portal web application or programmatically. Analysis for these data are linked out to a GDC instance of the cBioPortal (https://cbioportal.gdc.cancer.gov/cbioportal/) and the Broad Institute Firebrowse (http://firebrowse.org/). Currently supported datatypes include clinical information, mutations, gene expression, copy number alterations, structural rearrangements, protein expression, and methylation. Users can generate subsets of available data to view or retrieve by applying filters based on the cancer program, data category, primary tumor site, disease type, and experimental strategy, to name a few. Users may also submit data if they meet GDC data submission criteria and have registered the study successfully with dbGap. Once these conditions have been met, users can upload and validate their data within the GDC and submit it for processing. After processing is completed by the GDC, the user must release their data to the GDC within 6 months of submission to make it available through the GDC Data Portal as open or controlled data. For small projects, the data submission process can be done via the GDC Data Submission Portal, a web-based tool. For larger,

high volume projects, the GDC Data Transfer Tool (a client-based tool) can be used instead. Large organizations can also submit their data programmatically through GDC submission pipelines using the GDC Application Programming Interface (API).

11.3.2.8 Database of Genotypes and Phenotypes (dbGaP) [16]

The database of Genotypes and Phenotypes (dbGaP) is a repository developed by the National Center for Biotechnology Information (NCBI) and launched in 2006 to serve as a general repository for studies examining the association between phenotype and genotype. dbGaP provides a valuable resource to the research community by providing a centralized repository for accessing and browsing data in a uniform way. dbGaP assigns a unique and stable accession number for every study and every dataset within each study, which allows published studies to discuss or cite the primary data in a specific and uniform way. Many types of data are supported by dbGaP including, but not limited to, phenotype data, genome-wide association (GWAS) data, summary level analysis data, SRA (Short Read Archive) data, reference alignment (BAM) data, VCF (Variant Call Format) data, expression data, imputed genotype data, and image data. dbGap provides a combination of controlled and open access data for each study. Examples of publicly available data are study metadata, phenotype variable summary information, documentation related to the study, and association analyses that are in the public domain. Access to individual-level data requires that a user has either a NIH eRA Commons Account (for extramural researchers) or a NIH login (for intramural researchers) and that they be classified as a Principal Investigator (PI). Authorized access for non-PI's requires users to be approved for local access to view and download data files within the PI's lab. To date, dbGaP hosts data from over 800 studies, including data from large cancer research projects such as TCGA, TARGET, and CGCI.

11.3.2.9 European Genome-phenome Archive (EGA) [18]

The European Genome-phenome Archive (EGA) was launched in 2008 by the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) to support the voluntary archiving and the secure dissemination of data, including raw data as well as genotype calls provided by data contributors. Currently supported types of data include exome sequencing, whole genome sequencing, data from array-based technologies, transcriptomic data, and epigenomic data. To access datasets, users must be approved by a study's designated Data Access Committee (DAC). To date, EGA contains nearly 2800 datasets from over 1400 studies that were provided by approximately 650 data contributors.

11.3.2.10 ArrayExpress [19]

Established by the European Bioinformatics Institute (EMBL-EBI) in 2003, the ArrayExpress Archive of Functional Genomics Data is a repository recommended by most journals for data supporting peer-reviewed publications. ArrayExpress supports a wide range of raw and processed functional genomics data with the exception of raw data from high-throughput sequencing experiments, which are brokered to EGA, and are controlled-access datasets. The types of data currently supported include, but are not limited to, mutations, copy number alterations, methylation, gene expression, miRNA expression, and protein expression. Datasets can be browsed by organism, array used, assay technologies, and assayed molecule through the ArrayExpress web interface or programmatically through the ArrayExpress REST/JSON API. Currently, ArrayExpress contains data from over 26,000 experiments using human samples.

11.3.2.11 Gene Expression Omnibus (GEO) [20]

The Gene Expression Omnibus (GEO) project was launched in 2001 by the NCBI in response for the growing demand for a public repository for high-throughput gene expression data. GEO provides a user-friendly platform for submitting, storing, reviewing, and retrieving datasets from functional genomic experiments from array- and sequence-based technologies. GEO supports data from various projects and contributors in various formats including molecular profiling data for gene expression, non-coding RNA, methylation, mutations, and protein expression, among others. Currently, GEO contains data from approximately 1770 datasets generated using human samples.

11.3.2.12 Synapse [15]

Synapse is an open source software platform developed by Sage Bionetworks that enables the co-location of scientific content. Source code for Synapse is available through GitHub (https://github.com/Sage-Bionetworks/Synapse-Repository-Services). Some of the key features of Synapse include data versioning, provenance tracking, data annotation, query language, governance, group security, citation management, and open APIs for programmatic access. Beyond data-related functionalities, Synapse has proven to be a valuable resource for facilitating collaborative research due to its various other functionalities such as the management of analysis code and models, publication of these resources, and providing access to other tools that enable collaboration, such as group email and chat. At this time, Synapse contains over 900 public projects, including Project GENIE and various working groups for TCGA and ICGC.

11.3.2.13 GDAC Firehose (Genome Data Analysis Center) [17]

GDAC Firehose is a data repository established by the Broad Institute of MIT and Harvard in 2011 to provide aggregated and processed TCGA datasets. Since their launch in 2011, GDAC Firehose has been regularly running TCGA data through their analytical workflows, which are publicly available and can be found through the GDAC Firehose site. Some of the key features of GDAC Firehose include versioned and standardized datasets, analysis reports, and software packages and analysis code. Analysis reports are structured in an easy to read, user-friendly format and the reports can be downloaded directly along with the corresponding data and code used to generate the report. Users can download the data through the web user interface or programmatically by installing firehose_get, a command line tool developed by GDAC Firehose, or by using their RESTful API service through FireBrowse, a web tool developed by the Broad Institute for visualizing TCGA data, which will be discussed in further detail shortly.

11.3.2.14 The cBioPortal for Cancer Genomics [14]

The cBioPortal for Cancer Genomics is an open-access, open-source platform for interactively visualizing, analyzing, and downloading large-scale cancer genomics data sets. As of May 2017, the cBioPortal contains data for over 20,000 samples from 151 manually curated cancer studies, including data from TARGET, TCGA, GTEx, CCLE, and published data from literature. Data from the public instance of the portal (http://www.cbioportal.org/) are available through GitHub (https://github.com/cBioPortal/datahub) under the ODC Open Database License (ODbL). Data is also available for download through the portal site itself through the web interface and the cBioPortal API.

11.4 Data Analysis Platforms

As evident by the increasing availability of cancer genomic data and the importance of linking it to clinical information, the ability to perform comprehensive molecular profiling of tumor samples is ever more significant and necessary to further enable progress in precision cancer medicine. Translational research platforms have played a key role in enabling precision cancer medicine through the development of tools for visualizing and analyzing integrated omics data. Despite limitations, available resources and research platforms for studying cancer genomics have made a significant impact on realizing precision cancer medicine and are widely appreciated by the research community by helping biologists and clinicians interpret the information more effectively.

Some of the core functionalities provided by translational research platforms include:

- Integrating clinical and omics data from various sources
- Providing analysis frameworks and visualization tools
- Supplementing data with information cross-referenced from external databases (such as annotation mutation information with its clinical actionability)
- Enabling researchers and clinicians to explore data and generate hypotheses using various levels of analyses such as:
  - Cross-cancer study views
  - Gene-centric analyses
  - Cohort-level analyses
  - Patient-level analyses

The translational research platforms discussed in this section provide many of these functionalities, among others. A brief summary of publicly available research platforms can be found in Table 11.4. Although all of these research platforms provide many unique and complementary features and tools for cancer genomic data analysis, only a subset of these platforms will be described in depth to highlight key features. To find more information on each platform, please refer to the corresponding citations and platform sites.

11.4.1 Bridging the Gap for Translational Cancer Research

Users must consider many factors when selecting the right analysis platforms for their research. For example, if a user wishes to upload their own private dataset to visualize and analyze, will their data be secure? What datatypes are supported? Where is the data hosted on the research platform coming from and is it downloadable? Can users compare local data against platform databases and are local installations possible? Is the software for the platforms open-source? A comparison matrix of publicly available analysis platforms summarizes many of these points in Table 11.5.

11.4.1.1 cBioPortal for Cancer Genomics [14]

The cBioPortal for Cancer Genomics is an open-access, open-source platform for interactively visualizing, analyzing, and downloading large-scale cancer genomics data sets. Though originally developed at Memorial Sloan-Kettering Cancer Center (MSK), the cBioPortal software is now being developed and maintained by a multi-institutional team consisting of MSK, the Dana Farber Cancer Institute, Princess Margaret Cancer Centre in Toronto, Children's Hospital of Philadelphia, The Hyve in the Netherlands, and Bilkent University in Ankara, Turkey. Source code for the cBioPortal is available through GitHub (https://github.com/cBioPortal). Currently, the cBioPortal public site hosts data for over 20,000 samples from 151 cancer studies, including data from TCGA, CCLE, and TARGET. Data hosted on the public

Table 11.4 Publicly available research platforms for cancer genomics data

Analysis Platform	Summary
cBioPortal [14]	The cBioPoral for Cancer Genomics is an open-access, open-source platform for interactively visualizing, analyzing, and downloading large-scale cancer genomics data sets. Currently, the cBioPortal public sits hosts data for over 20,000 samples from 151 cancer studies, including data from large-scale cancer genomics projects such as TCGA, CCLE, and TARGET. The cBioPortal can be installed locally and the codebase and data for the cBioPortal are available through GitHub (https://github.com/cBioPortal/).
Broad Institute Firebrowse [21]	The Broad Institute TCGA GDAC Firebrowse is an analysis platform that sits above the TCGA GDAC Firehose, one of the deepest open cancer datasets available containing over 80,000 sample aliquots from over 11,000 cancer patients and 38 unique disease cohorts. FireBrowse provides a simple way to access, view, and download ~15,000 analysis reports created by Firehose, as well as publicly available data from TCGA.
COSMIC [22]	The catalogue of Somatic Mutations in Cancer (COSMIC) is a large and comprehensive knowledge base for exploring somatic mutations in human cancers. Data is manually curated by a curation team and users can be explored by cancer types, genes, specific mutations, and study names programmatically and through the web interface.
ICGC Data Portal [10]	The International Cancer Genome Consortium (ICGC) Data Portal is a web-based platform for visualizing, querying, and downloading data provided by ICGC member institutions and large-scale projects, such as TCGA and TARGET. As of May 2017, the ICGC Data Portal contains data from 70 projects and 21 primary cancer types. Local installation is available and the software is open-source through GitHub (https://github.com/icgc-dcc/dcc-portal).
UCSC Xena [23]	UCSC Xena is an interactive web-based data analysis and visualization platform for cancer genomics data. As of May 2017, UCSC Xena contains public data from 91 cohorts, including data from TCGA, ICGC, TARGET, and CCLE. The software for UCSC Xena is available through GitHub (https://github.com/ucscXena).
St. Jude PeCan Data Portal [24]	The St. Jude Pediatric Cancer (PeCan) Data Portal is a cancer genome data portal developed specifically for visualizing and exploring genomic alterations in pediatric cancers. As of May 2017, the PeCan data portal contains data for over 2000 pediatric tumors and 17 pediatric cancer types from the Washington University Pediatric Cancer Genome Project, as well as TARGET and smaller pediatric cancer projects.
Brown MAGI [25]	MAGI (Mutation Annotation & Genome Interpretation) is an open-source, open-access web application developed at Brown University for annotating, exploring, and analyzing private and public cancer genomics data. Local installation is available and the codebase is available through GitHub (https://github.com/raphael-group/magi).
OncoMine Research Edition [26]	OncoMine is a web-based cancer microarray database and data-mining platform for facilitating genome-wide expression analyses by integrating and unifying publicly available cancer profiling data across various cancer types and experiments.
IntOGen [27]	IntOGen is a web-based platform for integrating and mining data from cancer profiling experiments. As of May 2017, IntOGen hosts data for nearly 7000 samples across 26 cancer types and 48 projects. IntOGen can be installed locally and the software is open-source and is available here: https://bitbucket.org/intogen/intogen-pipeline/src.

(continued)

Table 11.4 (continued)

Analysis Platform	Summary
TumorPortal [28]	Tumorportal is a web-based tool developed at the Broad Institute that provides access to integrated genomic and clinical data from 22 cancer cohorts. Users can also explore private datasets through the TumorPortal beta version of the portal.
canEvolve [29]	canEvolve is a web portal developed to store data derived from functional genomics profiles from microarray and NGS platforms downloaded from GEO.
canSAR [30]	canSAR is an integrated knowledge base that provides multidisciplinary genomic and clinical annotations to support cancer translational research and enable drug discovery. Data sources include data from ArrayExpress [19], PDB [31], UniProt [32], COSMIC [22], Pathway Commons [33], ChEMBL [34], STRING [35], drug indications from the NCI (https://www.cancer.gov/), and clinical trial data from https://clinicaltrials.gov/.
G-DOC Plus [36]	The Georgetown Database of Cancer (G-DOC) is a platform developed to enable translational cancer research by facilitating access to integrated genomic and clinical cancer data, including medical and digital images for validation of analysis results. As of May 2017, G-DOC hosts data for over 10,000 patient from 55 private and public studies and resources, such as GEO and TCGA, among others.
OASIS Genomics [37]	OASIS Genomics is a web-based platform and central repository for multi-dimensional cancer genomics data developed by Pfizer Oncology Research Computational Biology in collaboration with Research Business Technology (RBT). OASIS Genomics hosts data from a variety of publicly available resources, including TCGA, CCLE, and GTEx, among others.
OncoScape [38]	OncoScape is a data visualization platform developed at the Fred Hutchinson Cancer Research Center to enable researchers to explore publicly available data from GDC, TCGA, UCSC Xena, and other resources in an intuitive and interactive manner. The code for OncoScape is open-source and is available through GitHub (https://github.com/FredHutch/Oncoscape) and data can be downloaded programmatically through their data API (http://resources.sttrcancer.org/api/data-explorer/).
Visual Omics Explorer (VEO) [39]	VEO is data visualization platform that provides a diverse set of data visualizations for dynamic and interactive data displays on desktops and mobile devices. Users are able to explore available data through the Google Genomics Cloud, as well as private datasets. Local installation is available and the code is open-source and available through GitHub (https://github.com/BCIL/VOE).
ISB Regulome Explorer [40]	The ISB Regulome Explorer is a web-based platform that provides tools for exploring genomic profiling data from TCGA, such as cross-cancer comparisons of genomic alterations ad associations among genomic and clinical features. The software for ISB Regulome Explorer is open-source and is available through GitHub. (https://github.com/cancerregulome/regulomeexplorer).
R2 Platform [41]	R2 is a web-based genomics analysis and visualization platform that was developed with biologist in mind. As of May 2017, R2 contains array-based expression data from over 500 experiments. Users can explore and download data hosted through R2, as well as explore their own private datasets.

Table 11.5 A comparison matrix of publicly available research platforms

	Basic Stats		Data Types										Data Views					Visualization				Analysis						Platform				
Platform	# Citations	# Studies	Non-synonymous mutations	Copy number alterations	mRNA expression	miRNA expression	DNA methylation	Protein /phosphoprotein expression	Structural rearrangements	Clinical attributes	Clinical timeline data	Gene-centric query / view	Cancer / cohort centric analysis	Cross-cancer view	Patient-centric view	Tumor tissues image view	Genomic & clinical visualization	Mutation summary diagrams	3D mutation visualization	Copy number segmentation	Mutual exclusivity and co-occurrence	Expression correlation analysis	Survival analysis	Enrichment analysis	Network or pathway analysis	Gene set / pathway significance	Correlation btw. genomic and clinical	Download of datasets	Web service interface	Open source license	Local installations	Local data
cBioPortal	2,166	151	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X	X	X	X	
Broad Institute Firebrowse		38	X	X	X	X	X	X		X		X	X				X	X		X		X		X	X	X	X	X			
COSMIC	880	210¶	X	X	X	X	X		X	X		X	X		X			X		X								X	X			
ICGC Data Portal	139	70	X	X	X	X	X	X	X	X		X	X	X	X		X	X					X	X	X	X		X		X	X	
UCSC Xena	49	91	X	X	X	X	X	X		X			X				X	X		X			X					X		X	X	X
St. Jude PeCan Data Portal	12	17	X		X				X			X	X	X				X		X												
Brown MAGI	3	31	X	X	X				X	X		X	X	X			X	X		X		X		X	X	X				X	X	X
GDC Data Portal	11	39	X	X	X	X	X	X	X	X					X													X	X			
OncoMine Research Edition	1,477	715*		X	X					X		X		X								X	X	X								
IntOGen	75	48	X									X		X				X												X	X	
TumorPortal	43	28§	X	X	X		X			X		X	X	X	X		X	X		X								X				X
CanEvolve	19	72	X	X	X	X		X				X	X	X									X	X	X	X		X				
canSAR	13	37	X	X	X	X						X																				
G-DOC Plus	2	55	X	X	X	X				X		X	X				X	X		X		X	X	X	X	X				X	X	
OASIS Genomics	4	30	X	X	X			X		X		X	X	X	X			X		X					X	X		X				
OncoScape	1	32	X	X						X	X		X				X	X		X			X	X	X	X			X	X		
Visual Omics Explorer (VEO)	1	0	X		X													X											X	X	X	
ISB Regulome Explorer		16	X	X	X	X	X	X		X		X	X				X					X			X		X			X	X	
R2 Platform		536*	X		X	X	X			X		X	X	X			X					X	X	X	X	X		X				X

¶ Studies with more than 100 genes only. ∗ Mostly expression data. § Only papers citing TumorPortal were counted

cBioPortal site can be downloaded programmatically using the cBioPortal API, through the web interface, or from the cBioPortal datahub GitHub repository (https://github.com/cBioPortal/datahub) under the ODC Open Database License (ODbL).

The cBioPortal was specifically designed to lower the barriers between researchers and access to complex multi-dimensional cancer genomics data by addressing data-integration issues and providing many unique tools for exploring these data. With the cBioPortal, users can visualize and analyze genetic alterations across samples, cancer types, data types, and genes. Additionally, what makes the cBioPortal such a powerful and unique tool is the ability to perform various levels of detailed analyses from cross-cancer cohort gene queries to detailed patient-level data views, such as clinical timeline and treatment information. Some of the many tools offered by the cBioPortal include OncoPrint diagrams, MEMo (Mutual Exclusivity Modules) analysis, customizable correlation plots, Kaplan-Meier plots, network analysis, and integrative genomics viewer (IGV) integration. The cBioPortal also integrates mutation annotations from a variety of resources to facilitate translational cancer research by helping researchers and clinicians identify potentially actionable mutations and mutational recurrence. These resources include OncoKB [42], CIViC [43], My Cancer Genome [44], and Cancer Hotspots [45]. Figure 11.1 shows example queries using the cBioPortal.

11.4.1.2 Broad Institute Firebrowse [21]

The Broad Institute TCGA GDAC Firebrowse is an analysis platform that sits above the TCGA GDAC Firehose, one of the deepest open cancer datasets available containing over 80,000 sample aliquots from over 11,000 cancer patients and 38 unique disease cohorts. FireBrowse provides a simple way to access, view, and download ~1500 analysis reports created by Firehose, as well as publicly available data from TCGA. Programmatic access is also available using the FireBrowse RESTful API service. Graphical tools are available through the FireBrowse web interface where scientists can explore and analyze cancer genomic data on a cohort-level, gene-centric level, and view correlations between clinical features and genomic data. Such tools available through FireBrowse include viewGene, a tool for exploring expression levels, and iCoMut, a tool for exploring comprehensive analysis profiles of each TCGA cohort (Fig. 11.2).

11.4.1.3 COSMIC [22]

The Catalogue of Somatic Mutations in Cancer (COSMIC) is a knowledge base designed to integrate cancer genomic data from various resources into a single platform, thus providing a “one stop shop” for exploring the impact of somatic mutations in human cancers. Data hosted by COSMIC is curated manually and semi-automatically from published datasets, pre-published data submitted by

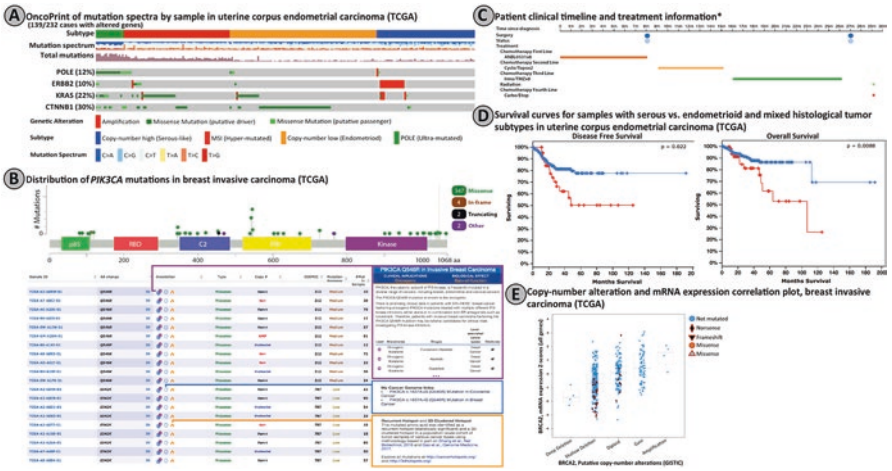

Fig. 11.1 Example queries using the cBioPortal. (a) OncoPrint of mutation spectra by sample in uterine corpus endometrial carcinoma (TCGA). **(b)** Distribution of PIK3CA mutations in breast invasive carcinoma (TCGA). Sample mutations are annotated with data from OncoKB (purple box), My Cancer Genome (blue box), and Cancer Hotspots (orange box). **(c)** Sample patient clinical timeline and treatment information. **(d)** Survival curves for samples with serous (red curve) vs. endometrioid and mixed histological tumor (black curve) subtypes in uterine corpus endometrial carcinoma (TCGA). **(e)** Copy-number alteration and mRNA expression correlation plot for breast invasive carcinoma (TCGA). Expression data is represented by Z-scores for all genes and copy-number data was computed with GISTIC. ∗Patient identifiers removed to protect patient information

authors to COSMIC for curation, and large cancer genome datasets, such as TCGA and ICGC, which account for approximately 50% of all available data in COSMIC. Higher level data, such as sample disease descriptions and classifications, are curated manually. As of May 2017, COSMIC contains curated data from nearly 25,000 published papers and over 1 million samples. Publicly available data from COSMIC can be accessed through the web interface, as well as programmatically through BioMart [46], and is integrated by a variety of platforms, such as the cBioPortal [14] and the ICGC Data Portal [10], as a resource for annotating genomic variants.

COSMIC provides an array of detailed gene-level information and relevant references for each gene in the literature. Such detailed information and visualizations for each gene overview includes variant information with genomic coordinates, the distribution of mutations across primary tissue types, the types of mutations (i.e., SNPs, INDELs, etc.), nucleic acid substitution types (i.e., A > G, A > C, etc.), drugs used to treat specific gene mutant tumors, and which genes develop drug resistance. Data displayed on gene reports can also be filtered by tissue type, somatic status, disease, mutation impact, and genomic alteration type, such as variant type, copy-number variation, gene expression, and methylation. Example visualizations using COSMIC are shown in Fig. 11.3.

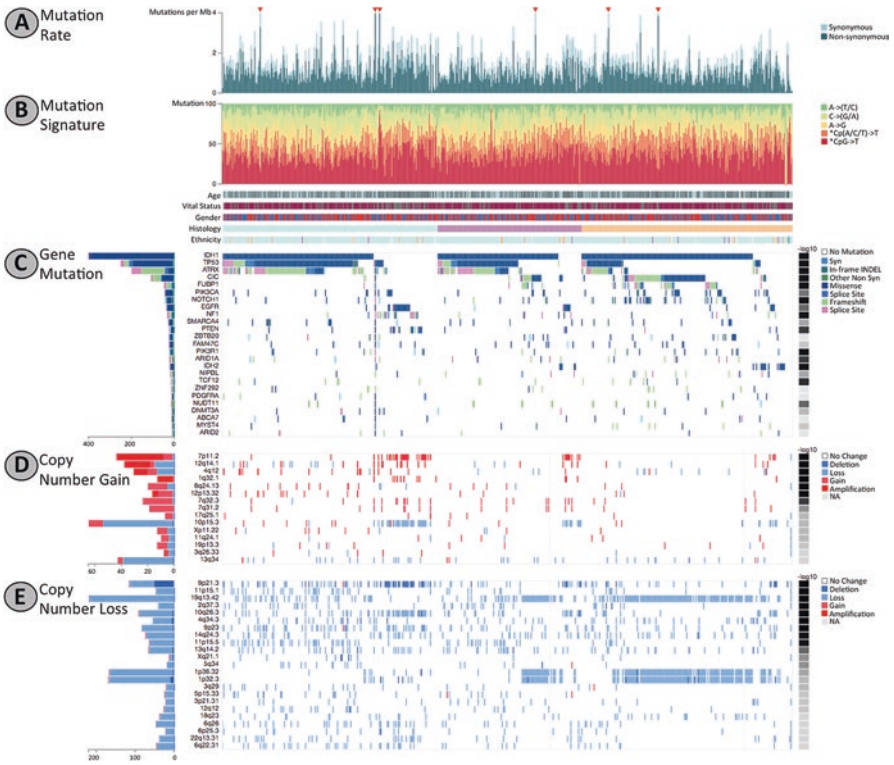

Fig. 11.2 Comprehensive analysis profile of TCGA Lower Grade Glioma cohort using iCoMut for FireBrowse. (a) Somatic mutation rate per patient stratified by synonymous and non-synonymous mutations. **(b)** Distribution of mutation substitutions per patient. **(c)** Significantly mutated genes identified by Mutsig2CV by q < = 0.1. **(d)** Somatic focal copy number gain events identified by GISTIC. **(e)** Somatic focal copy number loss events identified by GISTIC

11.4.1.4 ICGC Data Portal [10]

The ICGC Data Portal is an open-source, web-based tool for visualizing, querying, and downloading data provided by the ICGC member institutions, TCGA, TARGET, and other projects. Source code for the ICGC Data Portal is available through GitHub (https://github.com/icgc-dcc/dcc-portal). The ICGC Data Portal also provides a single access point for member institutions to manage and maintain their data locally while simultaneously sharing data with other member institutions and users. To ensure data uniformity and enable cross cohort comparisons, ICGC enforces set guidelines to ensure the same data models, controlled vocabularies, ontologies, and references are used by all ICGC member institutions. Users can download available data through the web interface and programmatically using the ICGC Data Portal API clients.

The ICGC Data Portal provides a variety of tools for exploring and downloading data by cohorts, patients, genes, mutations, and datatypes. Select visualizations and

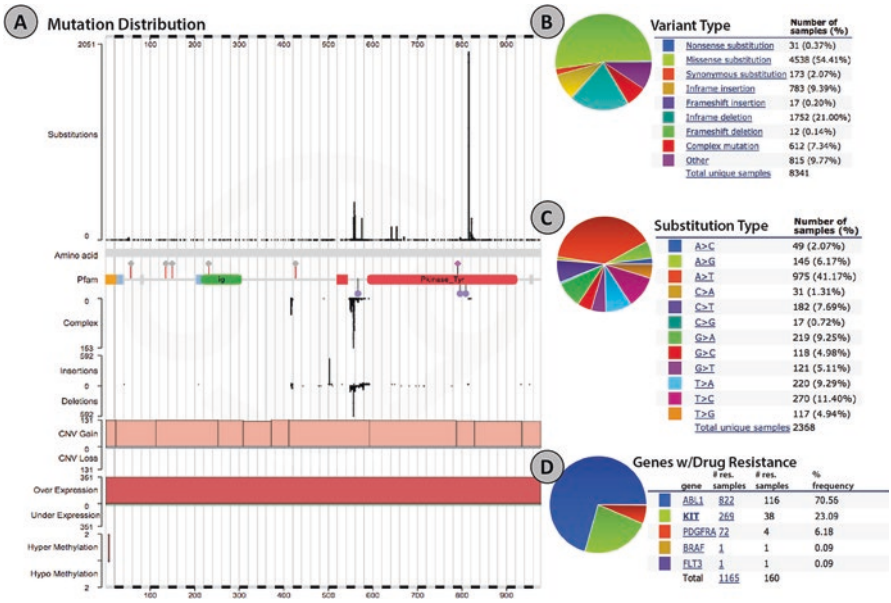

Fig. 11.3 COSMIC gene overview for *KIT*. (a) Full distribution of mutations for *KIT* across all tissues and cancer diseases. The top histogram displays the count of substitutions across the full length of the gene. The bottom peptide view shows the substitutions across the protein with color coded protein domains. **(b)** An overview of variant types that occur in *KIT*. **(c)** The breakdown of observed substitutions that occur in *KIT*. **(d)** Summary of genes that confer drug resistance to Imatinib, Nilotinib, and Sunitinib, which are drugs that are used to treat *KIT* mutant tumors

data summaries using the ICGC Data Portal are shown in Fig. 11.4. Users may also generate gene reports containing summaries of basic gene information, pathway annotations, mutations found in COSMIC, expression data, mutation frequencies, and cross references for genomic coordinates, pathways, and available publications, which are provided in a human-readable table format. Users can also perform enrichment analyses, cohort comparisons, and visualize genetic alterations through the data analysis tab to view the most commonly affected genes and pathways across cancer projects. The database can also be queried by specific datatypes, genes, samples, mutations, and other genomic alterations. Publicly available data can be downloaded from any of these reports for further analysis through the web interface.

11.4.1.5 UCSC Xena [23]

UCSC Xena is an open-source, web-based tool developed at the University of California Santa Cruz for visualizing and analyzing private and public cancer genomics datasets. Source code for UCSC Xena is available through GitHub (https://github.com/ucscXena). As of May 2017, UCSC Xena contains publicly

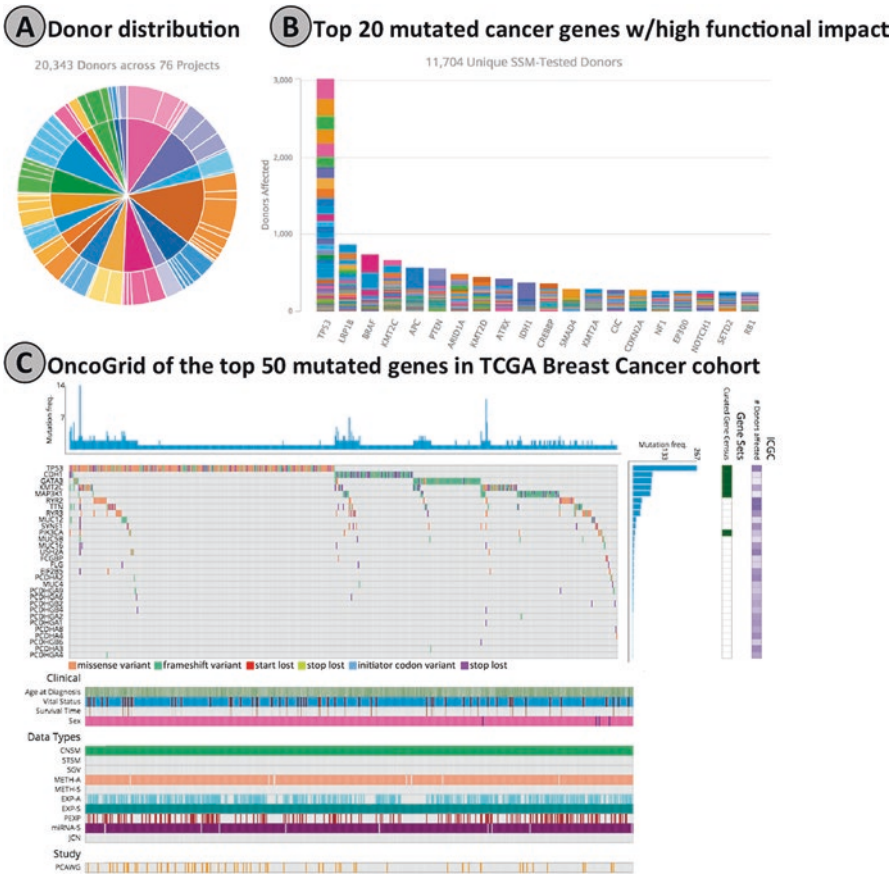

Fig. 11.4 ICGC data summaries and visualizations. (a) Donor distribution across 76 projects from ICGC members, TCGA, and TARGET. **(b)** The top 20 mutated cancer genes across all cohorts in the ICGC Data Portal. **(c)** OncoGrid view of the top 50 mutated genes in the TCGA Breast Cancer cohort. Mutation frequency and variant types are stratified by patients. Users may sort data by the provided clinical fields and available datatypes, as well as select regions to zoom into

available cancer genomics data from 91 projects and various cancer types, including data from TCGA, ICGC, TARGET, and CCLE. Users and/or institutions may also host their own installation of UCSC Xena with appropriate access control measures to help protect patient privacy and data security.

UCSC Xena provides a suite of tools for displaying and analyzing genome-wide experimental data across cohorts and datatypes. Visualizations include heatmaps, scatter plots, bar graphs, and 3D visualizations of mutations, as well as clinical data visualizations and survival analysis. Users can filter data by customized sample cohorts, clinical attributes, datatype, platform used, and more. Additionally, selecting specific variants when viewing SNPs and INDELs links out to the UCSC

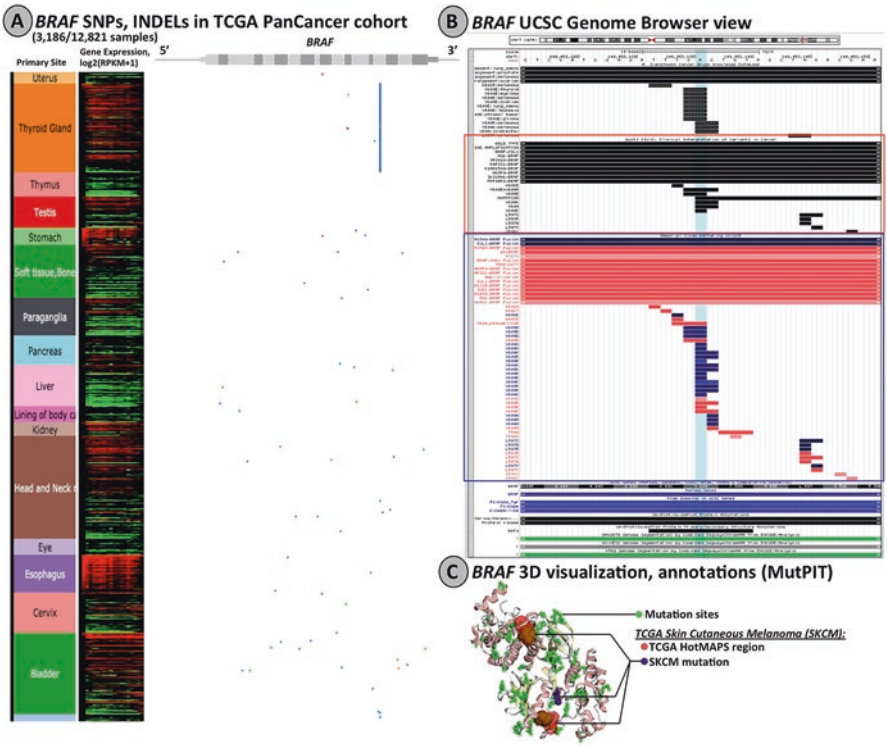

Fig. 11.5 TCGA PanCancer data visualizations with UCSC Xena. (a) Profile representation of SNPs and INDELs for BRAF in the TCGA PanCancer cohort using the UCSC Xena heatmap visualization tool. Data was filtered by samples with both exon expression data and somatic mutations in BRAF. High and low gene expression are represented by red and green, respectively. Somatic SNPs and INDELs are shown distributed along the BRAF genome. **(b)** BRAF UCSC Genome Browser with integrated CIViC (red box) and OncoKB (blue box) annotations. **(c)** 3D visualization of BRAF using MutPIT, a linked-out resource from the UCSC Xena heatmap tool. Mutations can be annotated by TCGA cancer type and filtered by sites

Genome Browser, where users will also see CIViC [43] and OncoKB [42] annotations integrated if available. Figure 11.5 shows example queries users can make with UCSC Xena.

11.4.1.6 St. Jude PeCan Data Portal [24]

The St. Jude Pediatric Cancer (PeCan) Data Portal is a cancer genome data portal developed specifically for visualizing and exploring genomic alterations in pediatric cancers using a tool called ProteinPaint. ProteinPaint was designed to simultaneously visualize SNPs, INDELs, fusion events, and gene expression data in parallel with annotations from COSMIC [22] and ClinVar [47]. An example of this is shown

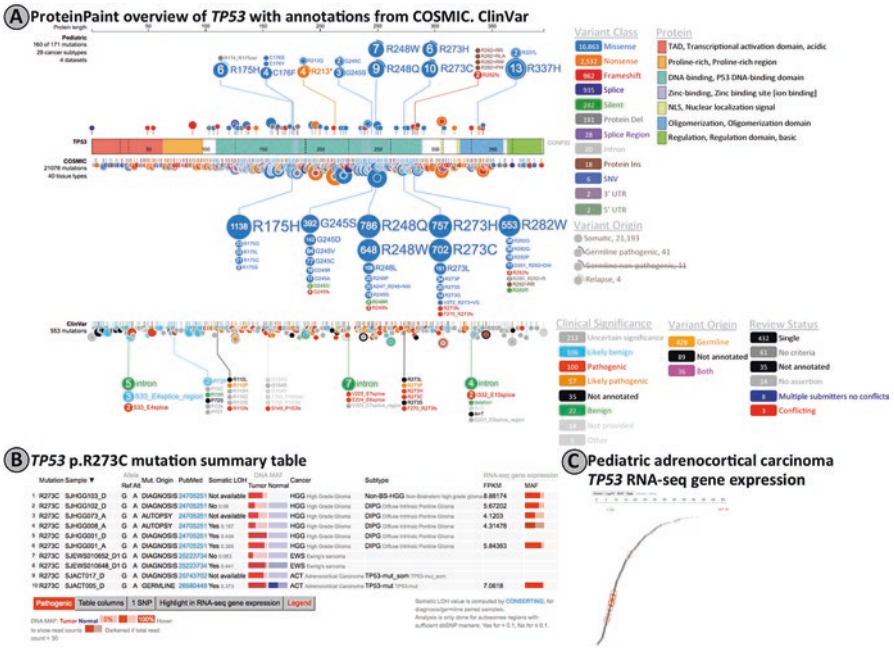

Fig. 11.6 St. Jude ProteinPaint overview of *TP53*. (a) Distribution of *TP53* mutations in pediatric cancers with annotations from COSMIC and ClinVar. **(b)** Example mutation summary table with read count heatmap for each sample variant. **(c)** *TP53* RNA-seq gene expression in pediatric cancers. Pediatric adrenocortical carcinoma *TP53* gene expression is indicated by red circles

in Fig. 11.6. Users can stratify data in a variety of ways such as by variants, project cohorts, cancer subtypes, and specimen types, for example. Additionally, users have the ability to add custom tracks to visualize their own data in bigWig, Stranded bigWigs, JSON-BED splice junction with read count, or VCF formats.

11.4.1.7 Brown MAGI [25]

MAGI (Mutation Annotation & Genome Interpretation) is an open-source, open-access web application developed at Brown University for annotating, exploring, and analyzing private and public cancer genomics data. Source code for installation is available through GitHub (https://github.com/raphael-group/magi). MAGI provides many interactive visualizations of cancer genomics data including real-time zooming, panning, and data filtering. Users can upload private datasets and compare or view their data in combination with TCGA Pan-Cancer data, collaboratively annotate genomic alterations and interactions, and view sample-level genomic aberrations. Private data upload is done through a simple web form without needing to install MAGI locally, and the types of data supported for upload include mutations, gene expression, methylation, and clinical attributes.

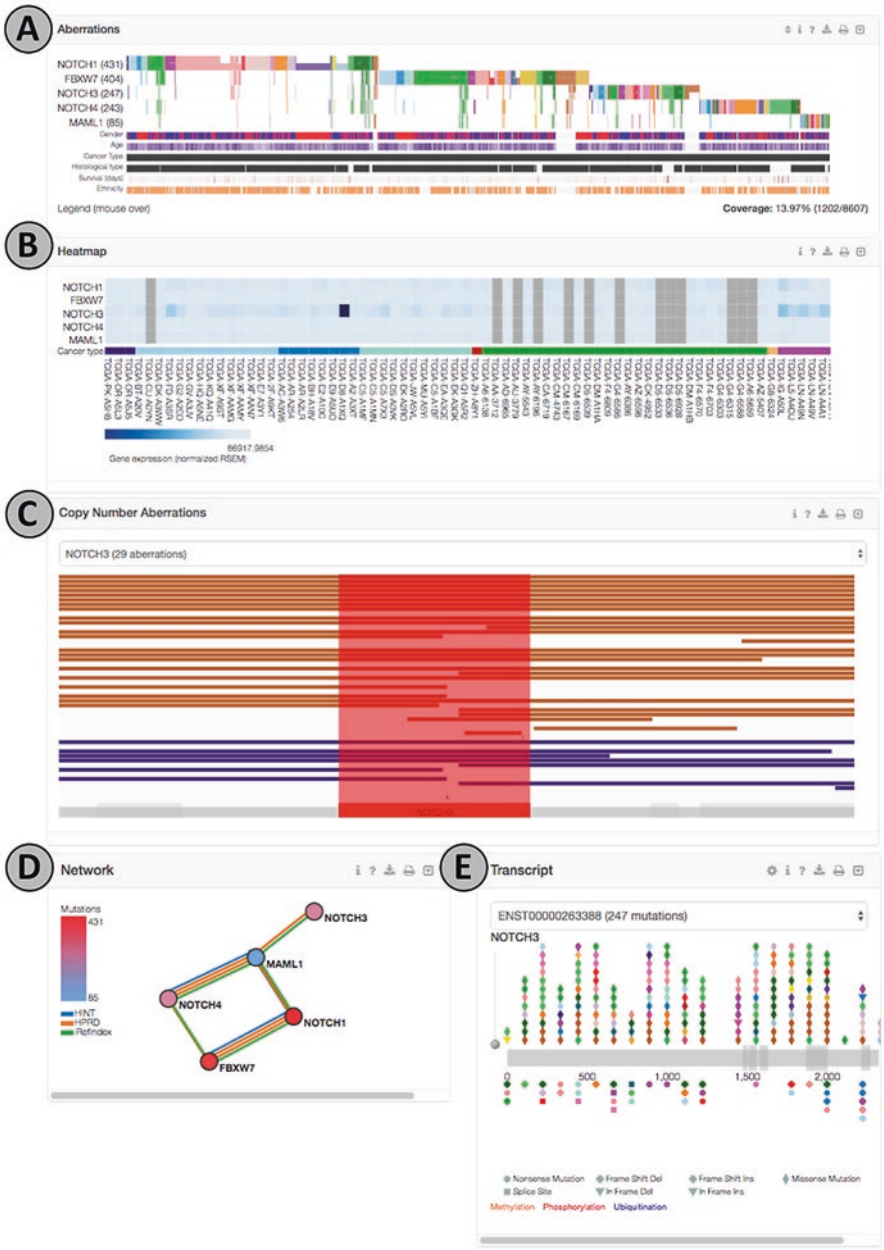

Fig. 11.7 An overview of the genomic alterations occurring in the Notch signaling pathway from the TCGA Pan-Cancer dataset. (a) Aberrations displays mutations in the queried genes across tumor samples, along with sample attribute data. **(b)** Heatmap displays gene expression data for the TCGA Pan-Cancer dataset. Users can also visualize private gene expression and methylation data. **(c)** Copy Number Alterations shows the copy number changes across tumor samples of a given gene. **(d)** Network shows gene interactions from multiple networks. Hovering over network interactions also displays a pop-up of annotations containing links to references in the literature. **(e)** Transcript displays the distribution of mutations and their types for a given gene

The types of visualizations include single-nucleotide variants, INDELs, copy-number changes, gene expression, and protein-protein interactions curated from literature and other sources. A screenshot of the MAGI web application displaying genomic alterations in the Notch signaling pathway from the TCGA Pan-Cancer dataset is shown in Fig. 11.7.

11.5 Clinical Actionability Resource Integration

As more data is collected from patient samples, the more powerful integrated knowledge systems become. Genomic alterations curated from sequencing data are studied to understand their impact on patient response and treatment, and this knowledge can be further integrated into knowledge bases and clinical reports. Figure 11.8 depicts the circular relationship of integrated knowledge systems.

Using this information, clinicians and researchers are better equipped to select the most beneficial treatment for their patients and can match their patients to clinical trials more easily based on their genomic profile. Furthermore, the aggregation of data from multiple resources, such as OncoKB [42] CIViC [43], My Cancer

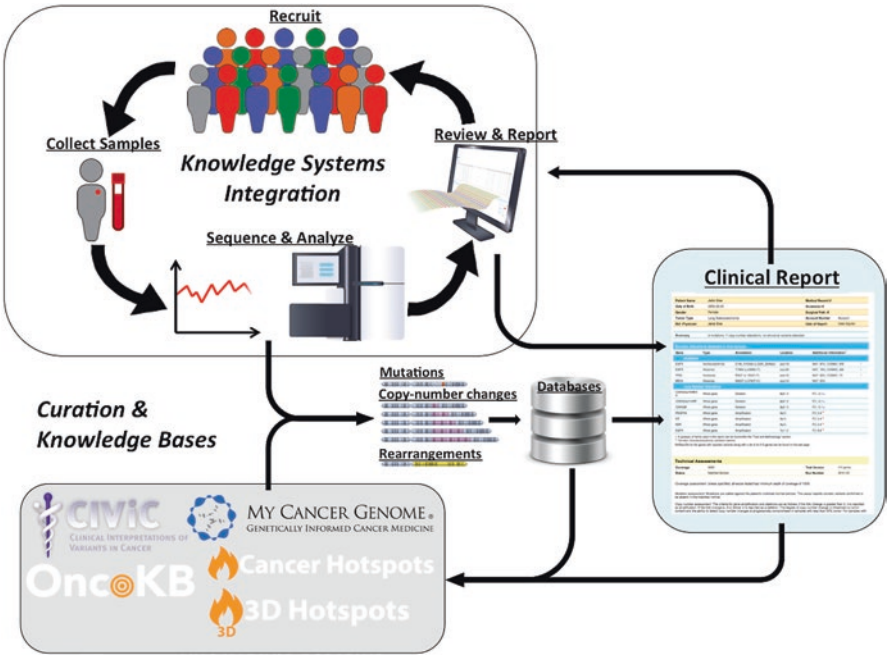

Fig. 11.8 Knowledge systems integration. A depiction of the circular relationship between the many systems involved with enabling precision cancer medicine. As more data is collected more patients, the more powerful and informative curated knowledge bases and clinical reports become

Genome [44], ClinVar [47], the Precision Medicine Knowledge Base [48, 49], the Jackson Laboratory Clinical Knowledgebase [50, 51], Cancer Genome Interpreter [52], Cancer Driver Log [53, 54], Tumor Portal [28, 55], Targeted Cancer Care [56], and Personalized Cancer Therapy [57, 58] provide researchers and clinicians with the tools to make more effective genome-driven medical decisions for their patients as these data continue to increase and expand their scope. Providing this information in a hierarchical matter allows researchers and clinicians to determine the clinical actionability of genomic alterations and better equips them to make decisions based on predicted prognosis and recurrence based on experimental evidence and the biological annotations of these alterations.

References

1. Chin L, Hahn WC, Getz G, Meyerson M (2011) Making sense of cancer genomic data. Genes Dev 25(6):534–555
2. Ng PC, Kirkness EF (2010) Whole genome sequencing. Methods Mol Biol 628:215–226
3. Gao J, Ciriello G, Sander C, Schultz N (2014) Collection, integration and analysis of cancer genomic profiles: from data to insight. Curr Opin Genet Dev 24:92–98
4. Hawkins RD, David Hawkins R, Hon GC, Ren B (2010) Next-generation genomics: an integrative approach. Nat Rev Genet 11:75–87
5. Petricoin EF et al (2002) Use of proteomic patterns in serum to identify ovarian cancer. Lancet 359(9306):572–577
6. The Cancer Genome Atlas.," *The Cancer Genome Atlas – National Cancer Institute*. [Online]. Available: https://cancergenome.nih.gov/. Accessed 28 Nov 2016
7. "Therapeutically Applicable Research to Generate Effective Treatments|Office of Cancer Genomics." [Online]. Available: https://ocg.cancer.gov/programs/target. Accessed 28 Nov 2016
8. Cancer Genome Characterization Initiative|Office of Cancer Genomics. [Online]. Available: https://ocg.cancer.gov/programs/cgci. Accessed 28 Nov 2016
9. Barretina J et al (2012) The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483(7391):603–607
10. Zhang J et al (2011) International Cancer Genome Consortium Data Portal--a one-stop shop for cancer genomics data. Database 2011:bar026
11. The International Cancer Genome Consortium for Medicine (ICGCmed). will link genomics data to clinical information, health and response to therapies|ICGC med. [Online]. Available: https://icgcmed.org/. Accessed 28 Nov 2016
12. AACR Project GENIE.. [Online]. Available: http://www.aacr.org/RESEARCH/RESEARCH/PAGES/AACR-PROJECT-GENIE.ASPX. Accessed 8 Dec 2016
13. Grossman RL et al (2016) Toward a shared vision for cancer genomic data. N Engl J Med 375(12):1109–1112
14. Cerami E et al (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2(5):401–404
15. Sage Synapse : Contribute to the Cure. [Online]. Available: https://www.synapse.org/. Accessed 7 Dec 2016
16. Mailman MD et al (Oct. 2007) The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 39(10):1181–1186
17. Broad Institute of MIT and Harvard., "Broad Institute TCGA Genome Data Analysis Center (2014): analysis overview for 15 July 2014."
18. Lappalainen I et al (Jul. 2015) The European Genome-phenome Archive of human data consented for biomedical research. Nat Genet 47(7):692–695

19. Kolesnikov N et al (2015) ArrayExpress update--simplifying data submissions. Nucleic Acids Res 43(Database issue):D1113–D1116
20. Edgar R, Domrachev M, Lash AE (Jan. 2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30(1):207–210
21. FireBrowse. [Online]. Available: http://firebrowse.org/. Accessed 12 Dec 2016
22. Forbes SA et al (2015) COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res 43(Database issue):D805–D811
23. Sanborn JZ et al (2011) The UCSC Cancer Genomics Browser: update 2011. Nucleic Acids Res 39(Database issue):D951–D959
24. Zhou X et al (2016) Exploring genomic alteration in pediatric cancer using ProteinPaint. Nat Genet 48(1):4–6
25. Leiserson MDM, Gramazio CC, Hu J, Wu H-T, Laidlaw DH, Raphael BJ (2015) MAGI: visualization and collaborative annotation of genomic aberrations. Nat Methods 12(6):483–484
26. Rhodes DR et al (2004) ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia 6(1):1–6
27. Gundem G et al (2010) IntOGen: integration and data mining of multidimensional oncogenomic data. Nat Methods 7(2):92–93
28. Lawrence MS et al (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505(7484):495–501
29. Samur MK et al (2013) canEvolve: a web portal for integrative oncogenomics. PLoS One 8(2):e56228
30. Tym JE et al (2016) canSAR: an updated cancer research and drug discovery knowledgebase. Nucleic Acids Res 44(D1):D938–D943
31. Berman HM et al (2000) The Protein Data Bank. Nucleic Acids Res 28(1):235–242
32. The UniProt Consortium (2017) UniProt: the universal protein knowledgebase. Nucleic Acids Res 45(D1):D158–D169
33. Cerami EG et al (2011) Pathway Commons, a web resource for biological pathway data. Nucleic Acids Res 39(Database issue):D685–D690
34. Gaulton A et al (2012) ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 40(Database issue):D1100–D1107
35. Snel B, Lehmann G, Bork P, Huynen MA (2000) STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. Nucleic Acids Res 28(18):3442–3444
36. Bhuvaneshwar K et al (2016) G-DOC Plus – an integrative bioinformatics platform for precision medicine. BMC Bioinform 17(1):193
37. Fernandez-Banet J et al (2016) OASIS: web-based platform for exploring cancer multi-omics data. Nat Methods 13(1):9–10
38. Schlicker A, Michaut M, Rahman R, Wessels LFA (2016) OncoScape: exploring the cancer aberration landscape by genomic data fusion. Sci Rep 6:28103
39. Kim B, Ali T, Hosmer S, Krampis K (2016) Visual Omics Explorer (VOE): a cross-platform portal for interactive data visualization. Bioinformatics 32(13):2050–2052
40. Cancer regulome tools. [Online]. Available: http://explorer.cancerregulome.org/. Accessed 12 Dec 2016
41. R2: genomics analysis and visualization platform. [Online]. Available: http://r2.amc.nl
42. Chakravarty D et al (2017) OncoKB: a precision oncology knowledge base. JCO Precis Oncol 2017(1):1–16
43. CIViC – Clinical Interpretations of Variants in Cancer. [Online]. Available: https://civic.genome.wustl.edu/. Accessed 2 Dec 2016
44. Micheel CM, Lovly CM, Levy MA (2014) My cancer genome. Cancer Genet 207(6):289
45. Chang MT et al (2016) Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. Nat Biotechnol 34(2):155–163
46. Shepherd R et al (2011) Data mining using the Catalogue of Somatic Mutations in Cancer BioMart. Database 2011:bar018
47. Landrum MJ et al (2016) ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 44(D1):D862–D868

48. Huang L et al (2017) The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. J Am Med Inform Assoc 24(3):513–519
49. Institute of Precision Medicine: Welcome to the Precision Medicine Knowledgebase. Precision Medicine Knowledgebase. [Online]. Available: https://pmkb.weill.cornell.edu/. Accessed 26 Nov 2017
50. Patterson SE, Liu R, Statz CM, Durkin D, Lakshminarayana A, Mockus SM (2016) The clinical trial landscape in oncology and connectivity of somatic mutational profiles to targeted therapies. Hum Genomics 10:4
51. Clinical Knowledgebase (CKB). The Jackson Laboratory. [Online]. Available: https://www.jax.org/clinical-genomics/clinical-offerings/ckb. Accessed 26 Nov 2017
52. Barcelona Biomedical Genomics Lab.: Cancer genome interpreter. Cancer genome interpreter. [Online]. Available: https://cancergenomeinterpreter.org/. Accessed 26 Nov 2017
53. Damodaran S et al (2015) Cancer Driver Log (CanDL): catalog of Potentially Actionable Cancer Mutations. J Mol Diagn 17(5):554–559
54. The Ohio State University: CanDL. Cancer Driver Log. [Online]. Available: http://candl.osu.edu/. Accessed 26 Nov 2017
55. Broad Institute: Welcome to TumorPortal: genes, cancers, DNA mutations and annotations. Tumor Portal. [Online]. Available: http://tumorportal.org/. Accessed 26 Nov 2017
56. Massachusetts General Hospital: targeted cancer care. Targeted cancer care & treatment options – Mass General Cancer Center. [Online]. Available: https://targetedcancercare.massgeneral.org/. Accessed 6 Sept 2017
57. Meric-Bernstam F et al (2015) A decision support framework for genomically informed investigational cancer therapy. J Natl Cancer Inst 107(7). https://doi.org/10.1093/jnci/djv168
58. The University of Texas MD Anderson Cancer Center: Personalized cancer therapy knowledge base for precision oncology. Knowledge Base for Precision Oncology. [Online]. Available: https://pct.mdanderson.org/. Accessed 26 Nov 2017