Clinical Implementation of Comprehensive Strategies to Characterize Cancer Genomes: Opportunities and Challenges

Laura E. MacConaill, Paul Van Hummelen, Matthew Meyerson, and William C. Hahn

An increasing number of anticancer therapeutic agents target specific mutant proteins that are expressed by many different tumor types. Recent evidence suggests that the selection of patients whose tumors harbor specific genetic alterations identifies the subset of patients who are most likely to benefit from the use of such agents. As the number of genetic alterations that provide diagnostic and/or therapeutic information increases, the comprehensive characterization of cancer genomes will be necessary to understand the spectrum of distinct genomic alterations in cancer, to identify patients who are likely to respond to particular therapies, and to facilitate the selection of treatment modalities. Rapid developments in new technologies for genomic analysis now provide the means to perform comprehensive analyses of cancer genomes. In this article, we review the current state of cancer genome analysis and discuss the challenges and opportunities necessary to implement these technologies in a clinical setting.

Significance: Rapid advances in sequencing technologies now make it possible to contemplate the use of genome scale interrogation in clinical samples, which is likely to accelerate efforts to match treatments to patients. However, major challenges in technology, clinical trial design, legal and social implications, healthcare information technology, and insurance and reimbursement remain. Identifying and addressing these challenges will facilitate the implementation of personalized cancer medicine.

THE CASE FOR INDIVIDUALIZED CANCER MEDICINE

Work from many laboratories has identified genetic alterations that occur at an appreciable frequency in specific types of cancers. For example, a reciprocal translocation between chromosome 9 and 22, known as the Philadelphia chromosome and resulting in the BCR-ABL fusion gene, occurs in ~95% of chronic myelogenous leukemias (CML; refs. 1, 2); oncogenic KIT mutations are present in ~85% of gastrointestinal stromal tumors (GISTs; refs. 3, 4); mutations in the serine-threonine kinase BRAF are present in >50% of cutaneous melanoma (5); activating mutations in the epidermal growth factor receptor (EGFR) have been identified in ~15% of non–small cell lung cancers (NSCLC; refs. 6–8), and ERBB2 is amplified in 15–20% of breast cancers (9–12). Biochemical studies confirmed that these genetic alterations result in constitutively active molecules, and tumors that harbor such mutations depend on the activity of these proteins for survival.

On the basis of these observations, efforts to target these molecules with either small-molecule inhibitors or antibodies have led to several agents that induce significant clinical responses. For example, the tyrosine kinase inhibitor (TKI) imatinib induces clinical responses in Philadelphia chromosome–positive CML (13) and GISTs that harbor KIT mutations (14, 15); PLX4032 has been shown to induce responses in cutaneous melanomas with BRAF mutations (16, 17); the EGFR TKIs erlotinib and gefitinib show activity in NSCLCs that harbor activating mutations and small insertions/deletions in EGFR (6–8); and the ERBB2 inhibitors trastuzumab and lapatinib show clinical responses in breast cancers with amplification/overexpression of ERBB2 (18). Moreover, the presence of mutations in proteins other than the intended therapeutic target can affect the response to a particular therapeutic regimen. As an example, lung and colorectal cancers that harbor mutations in EGFR as well as KRAS or BRAF fail to respond to treatment with anti-EGFR–directed agents (19–22).

ABSTRACT

An increasing number of anticancer therapeutic agents target specific mutant proteins that are expressed by many different tumor types. Recent evidence suggests that the selection of patients whose tumors harbor specific genetic alterations identifies the subset of patients who are most likely to benefit from the use of such agents. As the number of genetic alterations that provide diagnostic and/or therapeutic information increases, the comprehensive characterization of cancer genomes will be necessary to understand the spectrum of distinct genomic alterations in cancer, to identify patients who are likely to respond to particular therapies, and to facilitate the selection of treatment modalities. Rapid developments in new technologies for genomic analysis now provide the means to perform comprehensive analyses of cancer genomes. In this article, we review the current state of cancer genome analysis and discuss the challenges and opportunities necessary to implement these technologies in a clinical setting.

Significance: Rapid advances in sequencing technologies now make it possible to contemplate the use of genome scale interrogation in clinical samples, which is likely to accelerate efforts to match treatments to patients. However, major challenges in technology, clinical trial design, legal and social implications, healthcare information technology, and insurance and reimbursement remain. Identifying and addressing these challenges will facilitate the implementation of personalized cancer medicine.
patients on the basis of factors such as primary tumor diagnosis or site, nodal status, histologic subtype, or hormone receptor status. In hematologic malignancies, blood count and microscopy are standard diagnostic tools used to categorize lymphomas and leukemias by cell lineage; a number of these diseases can also be classified by cytogenetics [acute myeloid leukemia (AML) and CML] or immunophenotyping of malignant cells (lymphoma, myeloma, and chronic lymphocytic leukemia).

Currently, technologies to profile samples for the clinical selection of patients for targeted therapies assess the mutational status of one or a few genes (capillary sequencing and pyrosequencing) or interrogate a specific histologic or pathologic phenotype [immunohistochemistry and fluorescence in-situ hybridization (FISH)]. Figure 2 highlights current and emerging clinical technologies to detect various cancer DNA alterations. For example, FISH is used to detect the translocation and resultant fusion of BCR and ABL in CML in clinical settings (25, 26). The BCR-ABL translocation is also found in acute lymphoblastic leukemia (ALL) at a lower frequency (25–30%); thus molecular characterization of BCR-ABL is of critical diagnostic importance in ALL, for which the presence or absence of this alteration mutation will dictate therapy. In a similar manner, amplification of ERBB2 in breast cancers (27) and fusions involving the anaplastic lymphoma kinase (ALK) gene (28) are detected by FISH and identify patients who are likely to respond to anti-ERBB2 agents or small-molecule inhibitors of ALK, respectively.

Several additional tests have been developed for other oncogenes. Detection of nucleotide substitution mutations, insertions, or deletions within the kinase domain of EGFR for gefitinib treatment is currently determined by capillary gel sequencing. In the clinical setting, KIT mutations in...
The technological revolution in the field of genomics began over 30 years ago with the development of the Sanger method for DNA sequencing (37), followed 20 years later by the development of microarrays (38). Mass-spectrometric genotyping emerged as a useful technology shortly afterward (39), and initial next-generation sequencing methodologies were reported in 2005 (refs. 40, 41; see Fig. 3 for a timeline of technological advances and concomitant cancer genome landmark discoveries). Concurrent advances in the field of sequencing resulted in the emergence of several massively parallel sequencing technologies, which yield vastly greater amounts of information. These sequencing technologies have advanced DNA sequencing capacity at an unprecedented rate (refs. 40–43; for a comprehensive review on second-generation sequencing, see Shendure and Ji, ref. 44) and lowered the cost per base of sequencing data by 10,000-fold in the past decade (45).

The completion of a highly refined human genome sequence (46, 47) was a landmark achievement in establishing a baseline reference genome to which other sequences could be compared. The availability of such reference genomes has enabled the concomitant characterization of genomic alterations from many different diseases, including cancers (48). Technological advances in experimental and informatics methodologies over the past 10 years have made possible the
International efforts to sequence normal genomes [the 1000 Genomes Project (www.1000genomes.org)] as well as cancer genomes [International Cancer Genome Consortium (ICGC)] promise to provide a growing number of reference whole-genome sequences. In the past 4 years, for example, whole cancer genomes derived from AML (60, 61), lung (62–64), breast (65–67), melanoma (68), and multiple myeloma (69) have been reported in detail. Large-scale cancer genome studies, such as The Cancer Genome Atlas (TCGA) and the ICGC, are applying next-generation sequencing technologies to tumors from 50 different cancer types to generate >25,000 genomes at the genomic, transcriptomic, and epigenomic level, and will provide the foundation for a complete catalog of oncogenic mutations (70).

Table 1. Overview of cancer therapeutics available or in development for exemplary cancer genes in each category of genomic alterations

<table>
<thead>
<tr>
<th>Category of genomic alteration</th>
<th>Exemplary cancer gene</th>
<th>Cancer</th>
<th>Targeted therapeutic agent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translocation/fusion</td>
<td>BCR-ABL</td>
<td>CML</td>
<td>Imatinib, dasatinib, nilotinib</td>
</tr>
<tr>
<td></td>
<td>PML-RAR</td>
<td>Acute promyelocytic leukemia</td>
<td>All-trans retinoic acid (ATRA)</td>
</tr>
<tr>
<td></td>
<td>EML4-ALK</td>
<td>Breast, colorectal, lung</td>
<td>Crizotinib (phase III), foretinib (phase II)</td>
</tr>
<tr>
<td></td>
<td>FIP1L1-PDGFR</td>
<td>Chronic eosinophilic leukemia</td>
<td>Imatinib</td>
</tr>
<tr>
<td>Amplification</td>
<td>EGFR</td>
<td>Lung, colorectal, glioblastoma, pancreatic</td>
<td>Cetuximab, gefitinib, erlotinib, panitumumab, lapatinib</td>
</tr>
<tr>
<td></td>
<td>ErbB2</td>
<td>Breast, ovarian</td>
<td>Trastuzumab, lapatinib</td>
</tr>
<tr>
<td></td>
<td>KIT</td>
<td>GISTs, glioma, HCC, RCC, CML</td>
<td>Imatinib, nilotinib, sunitinib, sorafenib</td>
</tr>
<tr>
<td></td>
<td>SRC</td>
<td>Sarcoma, CML, ALL</td>
<td>Dasatinib</td>
</tr>
<tr>
<td></td>
<td>PIK3CA</td>
<td>Breast, ovarian, colorectal, endometrial...</td>
<td>PI3-kinase inhibitors, none approved; experimental: LY294002</td>
</tr>
<tr>
<td>Point mutation</td>
<td>EGFR</td>
<td>Lung, glioblastoma</td>
<td>Cetuximab, gefitinib, erlotinib, panitumumab, lapatinib</td>
</tr>
<tr>
<td></td>
<td>KIT</td>
<td>GISTs, glioma, HCC, RCC, CML</td>
<td>Imatinib, nilotinib, sunitinib, sorafenib</td>
</tr>
<tr>
<td></td>
<td>PDGFR</td>
<td>GISTs, glioma, HCC, RCC, CML</td>
<td>Imatinib, nilotinib, sunitinib, sorafenib</td>
</tr>
<tr>
<td></td>
<td>BRAF</td>
<td>Melanoma, pediatric astrocytoma</td>
<td>PLX4032 (phase III)</td>
</tr>
<tr>
<td></td>
<td>MET</td>
<td>Lung</td>
<td>Cetuximab, gefitinib, erlotinib, panitumumab, lapatinib</td>
</tr>
<tr>
<td></td>
<td>KRAS</td>
<td>Colorectal, pancreatic, GI tract, lung...</td>
<td>Resistance to erlotinib, cetuximab (colorectal)</td>
</tr>
<tr>
<td></td>
<td>RAS/RAF</td>
<td>CTCL</td>
<td>Selumetinib (phase II)</td>
</tr>
<tr>
<td></td>
<td>PTEN (mTOR)</td>
<td>Endometrial, prostate, NSCLC, renal</td>
<td>Ridaforolimus, temsirenilimus, everolimus</td>
</tr>
<tr>
<td></td>
<td>P38 (Akt, mTOR)</td>
<td>Endometrial, prostate, NSCLC, renal</td>
<td>Ridaforolimus, temsirenilimus, everolimus</td>
</tr>
<tr>
<td></td>
<td>PTCH1, SMO (Hedgehog)</td>
<td>Basal cell carcinoma</td>
<td>GDC-0449 (vismodegib) (phase II)</td>
</tr>
<tr>
<td>Genotype</td>
<td>VEGF-2578</td>
<td>Breast</td>
<td>Bevacizumab</td>
</tr>
<tr>
<td></td>
<td>VEGF-1154</td>
<td>Breast</td>
<td>Bevacizumab</td>
</tr>
</tbody>
</table>

Abbreviations: ALL, acute lymphoblastic leukemia; CML, chronic myelogenous leukemia; CTCL, cutaneous T-cell lymphomas; GIST, gastrointestinal stromal tumor; HCC, hepatocellular carcinoma; NSCLC, non–small cell lung cancer; RCC, renal cell carcinoma.

characterization of cancer genomes. In initial studies, a gene-focused approach used first-generation (Sanger) sequencing, which resulted in the identification of many driver events in cancers, such as mutations of BRAF in melanoma (5), EGFR in NSCLC (6–8), JAK2 in myeloproliferative disorders (49, 50), FGFR2 in endometrial carcinoma (51), ALK in neuroblastoma (52, 53), and PIK3CA in several cancers (54). However, with the development of second-generation sequencing technologies, it is now feasible to sequence exomes (known exons in the genome), transcriptomes (expressed genes in the genome), or whole genomes of cancer samples. Massively parallel DNA sequencing platforms have become widely available, and the number of both normal and cancer genomes that have been completed is now in the thousands (42, 55–59). International efforts to sequence normal genomes [the 1000 Genomes Project (www.1000genomes.org)] as well as cancer genomes [International Cancer Genome Consortium (ICGC)] promise to provide a growing number of reference whole-genome sequences. In the past 4 years, for example, whole cancer genomes derived from AML (60, 61), lung (62–64), breast (65–67), melanoma (68), and multiple myeloma (69) have been reported in detail. Large-scale cancer genome studies, such as The Cancer Genome Atlas (TCGA) and the ICGC, are applying next-generation sequencing technologies to tumors from 50 different cancer types to generate >25,000 genomes at the genomic, transcriptomic, and epigenomic level, and will provide the foundation for a complete catalog of oncogenic mutations (70).
The near future are the subject of this review. Three pivotal facing scientists, researchers, clinicians, and oncologists in based on tumor genomics. The challenges and opportunities need to be overcome to realize the goal of tailored treatments side effects. Significant challenges currently exist that will maximize efficacy of treatment while minimizing undesirable.

The result of this explosion in the molecular characterization of cancer is an emerging vision for personalized cancer medicine. In this paradigm, the specific genomic alterations driving a patient’s tumor will be analyzed, and a targeted therapy or therapies will be recommended in accordance with the genomic characterization of each tumor. This process will maximize efficacy of treatment while minimizing undesirable side effects. Significant challenges currently exist that will need to be overcome to realize the goal of tailored treatments based on tumor genomics. The challenges and opportunities facing scientists, researchers, clinicians, and oncologists in the near future are the subject of this review. Three pivotal components will be required to enable a paradigm shift to personalized cancer medicine:

- Every patient is profiled to identify genetic alterations present in a specific cancer.
- Caregivers have access to this information in a form that provides relevant contextual information.
- This information can be obtained within a time frame that permits its incorporation into the decision-making process.

An increased understanding of the biological driver events for some cancers, coupled with advances in technologies used to detect somatic cancer alterations, has led to the establishment of some initial personalized cancer medicine programs at several cancer centers, including the Dana-Farber/Brigham and Women’s Cancer Center (71, 72), Massachusetts General Hospital (73, 74), Memorial Sloan-Kettering Cancer Center (75–77), MD Anderson Cancer Center (78), Oregon Health & Science University (79), and Vanderbilt Cancer Center (73). Each of these programs uses a genotyping platform to profile patient samples for alterations in a panel of potentially “actionable” or “druggable” gene mutations that may inform a therapeutic paradigm for patients. Currently, these initiatives are limited by the number of genes interrogated, the type of somatic alteration investigated (mostly mutations and small insertions/deletions), throughput and cost of technology, and types of cancers being screened. Additional efforts are under way to translate the recent advances in genome technologies (such as massively parallel, whole-genome sequencing) into the clinical setting to guide patient treatment options.

Although advances in technology and the biological understanding of cancer pathogenesis have brought the goal of personalized cancer medicine closer, important challenges
remain, pertaining to each step described above, before this vision of individualized cancer care is actualized. These challenges are discussed in detail below.

**CHALLENGES IN IMPLEMENTATION OF PERSONALIZED CANCER MEDICINE**

**Technological Challenges in Sample Preparation**

Diagnostic interventions that successfully introduce tumor mutation profiling to clinical practice must circumvent several technical and logistical difficulties. Important limitations in the deployment of genomic technologies in the clinic are the minute amounts of tumor material available for genomic profiling and the availability of tissue in forms amenable to available technologies.

A barrier to large-scale genomic interrogation is the specimen itself. Many of the research-oriented, large-scale cancer-sequencing efforts thus far focus on surgically resected samples that yield a large amount of tissue for analysis. In the absence of a surgical resection, however, biopsy specimens taken for diagnostic purposes are often small and are usually composed of both normal and malignant cells in a variable ratio. Moreover, for some common cancers, such as those of breast and prostate, primary tumors are often identified at an early stage when they are small. Incorporation of genomic profiling into clinical decision making will therefore require reliable genomic profiling of small specimens, diffuse biopsy samples, and fine-needle aspirates.

In addition to the availability of limiting quantities of material for genomic profiling, the quality of nucleic acids can also vary greatly from sample to sample. The standard practice of banking formalin-fixed, paraffin-embedded (FFPE) clinical specimens (whether from surgical resections or biopsies) is advantageous for a clinical laboratory that performs histologic diagnoses. The majority of specimens are fixed for purposes other than genetic characterization, and this process often results in degraded and/or chemically modified nucleic acids (80–82) that are unlikely to yield the quality or quantity of DNA necessary for many high-throughput genomic technologies (83). Thus, any clinical test must be able to accept as input nucleic acid material derived from FFPE and/or archival tumor specimens.

In addition, samples derived from tumors are rarely homogeneous. For example, many tumors contain large areas of necrotic tissue that reduce the overall amount of usable DNA. Furthermore, specimens contain a mixture of normal cells and cancer cells, and the identification method must therefore have sufficient sensitivity to detect mutations in samples in which the tumor often represents less than half of the material. Many cancers are also populations of heterogeneous cells (84), with events that occur throughout the time line of tumor evolution (85). Finally, the emergence of resistance mutations is a clinically relevant genotype that may have impact on a therapeutic regimen and is, in some cases, driven by a mutation that may be present initially (whether in the target of therapy or an entirely different gene) in a small subclone of cancer cells (86–89). As these mutations may be present in only a fraction of tumor cells examined (90), even greater sensitivity is required to detect low-level events (91).

**The Selection of Genomic Approach and Platform**

Until recently, a major limiting factor inherent in the detection of cancer genomic alterations has been the availability of technologies to assay samples for all types of alterations; however, recent major technological innovations have all but eliminated this as a bottleneck in the translation of individualized cancer profiling to the clinical setting.

Capillary sequencing, used to detect point mutations, insertions, deletions, and substitutions at the DNA level, is still considered the gold standard in clinical laboratories and is used to detect *EGFR* mutations in lung cancer and *KRAS* mutations in colorectal cancer, among others. More sensitive sequencing technologies, such as pyrosequencing, are also routinely used in the clinical laboratory for detection of small base-pair changes in genes. Comparative genomic hybridization (92), FISH, and array comparative genomic hybridization (93) allow the detection of genomic imbalances, including copy-number alterations and deletions at a low resolution. Mass-spectrometric analysis of DNA enables the detection of base-pair changes, as well as small (<50 bp) insertions and deletions in genes (71, 72, 94, 95).

However, the implementation of these technologies in a clinical setting will be challenging for several reasons. Some limitations can be attributed to the inherent technical aspects of PCR and DNA manipulations in general. Moreover, the data output from hybridization-based tests is subject to error, and both comparators (such as wild-type or normal) and replicates are necessary to inform a conclusive test result. In the case of mass-spectrometric tests and sequencing, the signal generated is plotted as a composite intensity peak and (similar to hybridization-based methodologies) is subject to saturation limits and necessitates a subjective interpretation.

Second-generation sequencing technologies (such as Illumina and SOLiD) are better candidates to incorporate into the clinical setting. They permit the detection of the full range of genomic alterations, including mutations, structural rearrangements, and copy-number changes, using a single approach (refs. 60–64, 66, 68, 96; for a comprehensive review of genomic sequencing strategies and technologies, see ref. 96 and references therein) to detect epigenetic modifications with chromatin immunoprecipitation (ChIPSeq; ref. 97) and alterations in RNA (98), such as transcript expression, allele-specific expression, and alternative splicing. The theoretical sensitivity of second-generation technology platforms can be increased by sequencing to a higher coverage. In addition, protocols for analyzing samples derived from FFPE are available (83).

Current difficulties with these sequencing platforms include the lack of validation efforts, high running costs, and slow turnaround time. The reliability and reproducibility of second-generation sequencing platforms have yet to be established. International comparative studies may therefore need to be organized, similar to the MicroArray Quality Control project (MAQC) for the assessment of quality and reproducibility between microarray platforms and microarray data-generating laboratories (99). The MAQC effort was instrumental in the scientific community’s acceptance of and confidence in microarray data.
Costs and throughput in sequencing the full genome will need to decrease by an order of magnitude before these technologies can be used routinely in clinical laboratories. Many research and diagnostic goals, however, may be achieved much sooner by sequencing a specific subset of the genome in large numbers of individuals or at a greater sequencing depth. Several advances, including barcoding DNA for multiplex sequencing (100, 101) and hybridization-based sequence capture methods, reduce the complexity of the total genome (102–105) and decrease the cost, increase the number of applications of next-generation sequencing, and retain sensitivity sufficient to detect low abundant somatic mutations (106). Transcriptome sequencing is a sensitive application for detecting intragenic fusions (including in-frame fusion events that lead to oncogene activation; refs. 107, 108) and for generating gene expression profiles (109).

Second-generation sequencing technologies currently have a slow turnaround time (1–2 weeks) and are still quite costly. To accommodate a clinical setting, technological advances are required to achieve the performance necessary for clinical implementation, including short turnaround time (on the order of days), high accuracy and sensitivity, inter- and intra-laboratory reproducibility, and cost-effectiveness. Additional modifications in chemistries and detection methodologies, as well as a decrease in the experimental surface areas, may enable such advances. Recently, smaller, faster versions of available technologies have been released. These “personal sequencers” (produced by Intelligent Biosystems and Illumina, among others) have outputs sufficient for analysis of 1 exome that can be generated within a day. Going forward, faster (third-generation) sequencing technologies may have the ability to interrogate cancer genome alterations on a single-molecule or per-cell basis, and possibly enable noninvasive methods for cancer diagnosis and monitoring.

**Analytical and Informatics Issues in Cancer Genome Diagnostics**

A challenge deriving from the interrogation of low-quality tumor samples subject to the biological confounders detailed above is the analysis of data from cancer genome characterization efforts. Because of issues such as tumor heterogeneity, ploidy, and stromal contamination, a heterozygous somatic mutation in a tumor specimen with 60% admixed normal cells will be detected at an overall allele frequency of 20%. Thus mutation-calling algorithms can be confounded by the presence of additional (normal) signal, and disambiguating a somatic mutation signal from background can be difficult. Similar issues apply to the detection of other types of alterations, such as copy-number changes and chromosomal rearrangements. With regard to current technologies, massively parallel sequencing can overcome or mitigate some of these problems. The data output from these technologies is essentially digital—that is, one can count the number of reads, alleles, or nucleotides at any position on the aligned sequence. The advantage of this digital readout of data is that counting the number of reads that identify mutant alleles can help to overcome obstacles of low tumor quantity, variable ploidy, tumor heterogeneity, and normal cell admixture. Challenges in the detection of somatic alterations in tumor specimens can thus be addressed using statistically rigorous algorithms and bioinformatics approaches (for a comprehensive review of considerations in the analysis of cancer genomes, see ref. 96). Complete cancer genome sequencing requires that both the tumor and the matched normal sample be significantly oversampled. Specifically, the nucleotide sequence of the same fragment of DNA is sequenced repeatedly to gain a confident, high-quality readout sequence (42), which allows identification of somatic variants (as opposed to germline alterations, which may be private or common) that can be selected as candidates for further interrogation. From a diagnostic perspective, it may be prudent to have a high-quality cutoff for clinically relevant somatic mutations, or candidate events can be validated using an orthogonal technology, such as genotyping, pyrosequencing, or capillary sequencing, preferably in Clinical Laboratory Improvement Amendments (CLIA)–certified laboratories.

The vast amount of information produced from whole-genome sequencing studies presents a challenge of its own. Indeed, new algorithms are at present being developed to translate raw sequence data into meaningful reads that can be used to analyze the variety of genomic endpoints. A number of groups and approaches are currently in development, including mrFAST for copy-number variation by Alkan et al. (110); the genome analysis toolkit (GATK) developed by DePristo and coworkers for single-nucleotide polymorphism (111) and genotyping discovery; and several ChipSeq analysis tools reviewed by Kim et al. (112).

Besides new algorithm development, the computational cost for storage and processing of data is another challenge and may approach that of sequencing itself. These factors suggest that it may be more parsimonious to keep a record of only the deviations of a patient’s tumor (and normal) from a reference genome. Another important consideration for clinical sequencing is a robust sample tracking, management, and laboratory information management system, as maintaining the identity of a sample is critical. The abundance of sequence information that will be collected using a sequencing technology in the clinical setting is advantageous: it permits identification of somatic tumor alterations, as well as matching of the tumor with the corresponding normal (or germline) events that will be seen in both matched samples, thus ensuring that sample identity can be maintained.

**Interpretation of Biological Data**

Whole-genome studies have shown that the number and type of alterations in cancer genomes are often diverse and complex. Signals of driver alterations (those that dysregulate growth or promote tumorigenesis) are often difficult to distinguish from passenger events (alterations that are inconsequential and do not contribute to tumorigenesis). The significance of a somatic event must therefore be assessed in the context of the background of additional somatic events that may not have physiological consequence. Integrating datasets and genomic endpoints may help in identifying driver mutations or pathways. The initial findings of TCGA, for example, demonstrated the utility of this approach by using integrative analyses of DNA sequence, copy number, gene expression, and DNA methylation in glioblastomas to provide a network view of altered pathways (113). However, this approach requires sufficient biological knowledge of
the relevant pathways, relationships, and interactions to be able to interpret and understand signals. Nevertheless, for diagnostic purposes, distinguishing driver from passenger mutations may not be that relevant, but for prognosis, it is important to understand the master regulating genes or pathways to predict disease progression and response to targeted personalized therapy. Determining driver events or molecular targets is an emerging area recognized by the National Cancer Institute in efforts such as the Cancer Target Discovery and Development Network to prospectively pursue this path and to test strategies for selecting therapeutic targets (114).

A related question concerning the use of genomics to guide therapeutic decisions for cancer is whether mutations predictive of pharmacologic sensitivity in one tumor type can be extended to predict sensitivity in other tumor types. Many oncogenic KIT mutations that are predictive of imatinib response in GISTs have been identified. The subsequent discovery of identical KIT mutations in some acral and mucosal melanomas (115) allowed assessment of this hypothesis. Several reports indicate that inhibition of KIT using imatinib or nilotinib can elicit clinical responses in KIT-mutant melanoma (116, 117). These observations suggest that known driver genomic events may have the potential to denote therapeutic vulnerability regardless of the tissue type in which they occur, although clinical studies will be necessary to confirm this idea in many cases. Other reports, however, have contradictory findings and indicate that cellular context does play a role: for example, ovarian cancer patients with PIK3CA and either KRAS or BRAF mutations respond to treatment with a phosphoinositide 3-kinase (PI3K) pathway inhibitor, whereas colorectal cancer patients with PIK3CA and KRAS mutations do not respond (118).

A critical step toward defining correct personalized anticancer therapy is the identification of the additional genes and pathways altered in the tumor, and the elucidation of their particular oncogenic role. Genetic interactions and compensatory mechanisms may therefore be challenging for determining cancer treatment (see ref. 119 for a review), and the treatment paradigm for a patient thus may not be decided on the mutational status of a single gene, but instead on the context in which that mutation is found. As detailed above, mutations in genes other than the predicted therapeutic target can dramatically affect the response of a patient to a specific therapy, as in the case of RAS and RAF mutations indicating no response to EGFR inhibitor therapy in colorectal cancer and NSCLC (19–22). Treating metastatic melanoma patients with a BRAF-selective inhibitor, in the presence of RAS mutations, may unintentionally activate the mitogen-activated protein/extracellular signal regulated kinase (ERK) kinase (MEK)/ERK signaling pathway, instead of inhibiting tumor formation (120). The presence of additional mutations in intracellular signaling pathways activated by receptor tyrosine kinases can exhibit patterns of cross-talk. To this point, it has been reported that inhibition of certain pathways, such as the mTOR route, may lead to increased mitogen-activated protein kinase (MAPK) activity (121). Therefore, ablation of signals through both routes may be required for efficient antitumor activity. Moreover, combined inhibition of PI3K and MAPK routes has shown superior antitumor effect compared with individual targeting of either pathway (122). Another example is the sensitivity of BRCA1/2-deficient tumor cells to PARP inhibitors. PARP has been shown to be critical to cell survival in a compensatory mechanism for BRCA1 or BRCA2 loss-of-function mutations (123).

Accurate representation of the significance of a variant is critical—translating this information to physicians, clinicians, and oncologists requires a thoughtful analysis of the utility of available data. As the current approach to clinical characteristics that drive analyses today transitions to a more nuanced and realistic representation of the complexity of individual human physiology, robust decision support systems will be vital as an adjunct to the clinician’s evaluation. A system based on levels of evidence, for example, might indicate to a provider how much information has been collated for a specific mutation, tissue type, and therapeutic target. This information would aid in treatment decisions; as more exome-based and genome-based sequencing approaches translate to the clinical arena, a database linking to approved indications and guidelines may assist in collating the mass of information into a more interpretable framework. Regardless of the sophistication of informatics support, thoughtfully designed and carefully monitored clinical trials will be necessary to safely and effectively disambiguate the complex network of molecular alterations, signaling pathways, cellular context, and response to a particular therapeutic regimen.

**Disease Monitoring**

The molecular signature of a tumor during progression can be dynamic, one consequence of which is the emergence of resistance during therapy. Clinical responses to targeted anticancer therapeutics are frequently confounded by de novo or acquired resistance (15, 124, 125). Logistically, then, it would be ideal to sample a tumor at diagnosis and then track the progression of disease by recurrent sampling. This practice would aid in the management of disease and might indicate the emergence of resistance mutations. The clinical promise of selective RAF inhibitors, for example, has widespread ramifications for patient treatment, yet single-agent targeted therapy is almost invariably followed by relapse caused by acquired drug resistance. Identification of resistance mechanisms in a manner that elucidates alternative druggable targets may inform effective long-term treatment strategies (126).

Recent studies have shown that mutant BRAF-specific inhibitors can activate CRAF through the formation of dimeric RAF complexes, and this process is enhanced by the presence of an oncogenic RAS mutation. Thus these inhibitors should be used for treating cancers driven by BRAF, but should be avoided in cancers caused by RAS mutations (127, 128). Similarly, resistance to RAF inhibition can be achieved by multiple MAP3K-dependent mechanisms of MEK/ERK reactivation (129) but might be intercepted through combined therapeutic modalities for MAPK pathway inhibition. Thus, identification of the correct combination of therapeutic agents targeting several key steps in signaling pathways will be a new paradigm for cancer management. However, whether acquired resistance is due to compensatory mechanisms/mutations or due to the selection of existing mutations within a heterogenic tumor is not yet clear.
Several recent studies suggest that cancer genomes evolve as the cancer progresses. Deep sequencing of two primary breast cancers and subsequent metastases showed either novel mutations or enrichment for low-frequency mutations in the metastases, indicating that analysis of such tissues in some cancers might identify additional mutations that may be important drug targets (65, 66).

In many cases, it may be convenient to diagnose or detect the molecular signature of a tumor using noninvasive procedures. To this end, much effort is focused on the development of technologies to detect genomic alterations in circulating tumor cells (CTCs; refs. 130–132) or plasma DNA (133–135). The efficacy of using tumor-derived CTCs in the monitoring of disease and the early detection of resistance mutations during treatment have been reported for EGFR in lung cancer (130) as well as other solid tumor metastases (132); prospective clinical trials are under way in metastatic breast cancer. An FDA-approved technology for detection has also shown promise in characterizing the molecular profile (specifically HER2 expression) of CTCs in breast cancer patients (136). Next-generation sequencing could be the ideal technology to screen CTCs now that enough DNA can be extracted from flow-sorted single nuclei (137).

**Clinical Trial Design**

As our knowledge of the myriad alterations that drive cancer biology increases, our ability to accurately interpret these somatic alterations (and transfer patients to the appropriate targeted therapy) will become more refined. Thus, personalized diagnostic technologies may aid in the stratification and enrollment of patients with specific molecular alterations for a clinical trial with a targeted therapeutic regimen. One might therefore design clinical trials based on a targeted genotype, rather than the cancer type, and to select patients based on that genotype. It is clear that selecting patients with different cancers by mutation will be challenging owing to the heterogeneity of the natural history of the cancers, as well as the question of whether different responses will be found in tumors derived from different tissues even if they harbor the same genetic alteration.

To date, success with molecularly characterized targeted therapeutic trials has occurred with cancer types in which the hallmark mutation is present in a large subset of specimens as defined by tissue type or histologic classification. For example, *KIT* mutations occur in ~85% of GISTs (14, 138), and targeted therapies such as imatinib mesylate are effective only in patients with alterations in the intended targets (mutated *KIT* or *PDGFRα*; refs. 139, 140). Evidence is increasing, however, that these potentially actionable events occur at lower frequencies in other cancer types, and thus patients harboring these alterations are potentially suitable candidates for a targeted therapy; this concept is the basis for the paradigm of personalized or individualized cancer medicine. NSCLC patients, for instance, can harbor activating mutations in *EGFR*, but also *KRAS*, *BRAF*, *PIK3CA*, *ERBB2*, or translocations involving *ALK* (28, 141, 142). Similarly, *SLC45A3-BRAF* and *ESR1-RAF1* fusions were observed in frequencies below 2% in patients with advanced prostate cancer, gastric cancer, and melanoma, and may therefore be sensitive to RAF or MEK inhibitors (143).

Improvements in our understanding of chemical biology and drug targets should also aid in the interpretation of clinical trial data and enhance our understanding and interpretation of the biological response to specific treatment paradigms. The first *BRAF*-targeting drug to enter the clinic, for example, was sorafenib, a type II multikinase inhibitor (144–146). In clinical trials, sorafenib was ineffective in *BRAF*-mutated melanoma (147), but effective in renal cell (148) and hepatocellular carcinoma (149). Thus the interpretation of clinical trial data can be confounded by the relative specificity and potency of inhibitors in specific cancer types as well as off-target effects. The efficacy of a specific *BRAF*V600E inhibitor in clinical trials demonstrates the power of using targeted, mechanistic inhibitors (30, 150); the necessity of adding combination therapies to offset the cellular resistance to such drugs (151) indicates that our ability to interpret these data is still limited. Carefully designed and controlled clinical trials, in combination with robust genomic profiling technologies, will be needed to define the effective mutation-drug combinations.

It is imperative that new tests, drugs, and procedures in patients be subjected to rigorous appraisal, and that new signatures of relapse risk, such as that described for colon cancer, should also be tested in prospective clinical trials (152). The efficacy of such clinical trials has been shown in stratification of NSCLC patients based on the mutational status of *EGFR* (153, 154). Additional large-scale, prospective randomized clinical trials are under way for breast cancer gene expression tests.

With the cost of genome sequencing decreasing dramatically, it may be feasible in the not-too-distant future to augment clinical trials of cancer drugs with complete cancer genome or transcriptome sequencing in order to identify these determinants. It has yet to be determined, however, if clinical trials will be required to show the value of comprehensive cancer genome profiling, and what form those trials should take. We refer the reader to recent reviews that discuss new approaches to clinical trial design and development of new therapeutics (155–158).

**Ethical, Legal, and Social Implications and Intellectual Property Challenges**

Patients and subjects must give consent properly, as well as be educated with regard to the import and impact of genomic testing. Privacy issues may also be a concern for the patient. The United States FDA recently updated its requirements for informed consent documents used in clinical trials of drugs, biologics, and medical devices, to include a statement informing participants that information from trials will be entered into a database. For somatic mutations and current clinical tests assaying a small number of DNA changes, this information can be linked to the patient’s medical record when contained within in a secure database with restricted access. Looking ahead to sequencing-based approaches, we suspect that the need to protect this patient’s data and confidentiality will arise. In the United States, the Genetic Information Nondiscrimination Act was passed in 2008 to protect against discrimination (health insurance and employment) based on genetic information, and was seen as a landmark achievement in enabling patients to take advantage of personalized medicine without fear of discrimination.

As more genome-wide association studies are completed across populations, individual germline differences between
patients will also be factored into the equation—lower penetrance alleles in many coding and noncoding regions can help inform the progression of disease, response to therapy, and metabolism of specific compounds. Testing for disease-causing mutations in the BRCA1 and BRCA2 genes (implicated in familial breast and ovarian cancer syndromes) is an early example of the contribution of germline features to disease prognosis. Discovery of a disease-causing mutation in a family can inform at-risk persons about whether they are at higher risk for cancer and may prompt individualized prophylactic therapy. Another consequence of using sequencing-based technologies is the generation of ancillary or unintentional data. This imight come in the form of sequence alterations that were not being screened or tested for but that may impact the health or treatment of the patient (and possibly related individuals) in other ways.

Public education on the potential benefits of personalized cancer medicine and individualized treatments will be an important facet of its widespread acceptance. Furthermore, the implementation of these technologies will require several levels of additional oversight, such as obtaining Institutional Review Board–approved consent for appropriate testing, with a view to ensuring ethical research conduct and adequate subject protection (for a review on electronic health records and healthcare information technology considerations in personalized medicine, see ref. 159). Considerations such as privacy concerns and the unknown clinical significance of genome sequence data must be weighed against the ethical principles of respect for autonomy and the right of every patient to receive relevant personal medical information. As next-generation sequencing technologies and genomic science filter into clinical medicine and standard of care, guidance from programs such as the Department of Energy- and NIH-funded Ethical, Legal and Social Implications Research Program will be needed in order to adequately protect subjects and patients.

A number of additional hurdles currently confound the process of implementing personalized cancer profiling in the clinical setting, including a pre-existing intellectual property (IP) environment that gives a company the sole right to genetic tests of specific genes and mutations. The arena of gene-based IP is currently in flux with regard to clinical diagnostics and patented genes. Addressing these changes in the architecture of cancer care may require a substantial shift in the cultural understanding of individualized cancer medicine and how it affects each party in the arena of cancer care.

Oversight and Regulatory Challenges

In the United States, several large stakeholders currently use the output from clinical research to make decisions and set policy. Three of the largest are the U.S. FDA, the Agency for Healthcare Research and Quality, and the Centers for Medicare & Medicaid Services (CMS). Whereas the FDA regulates diagnostic devices sold as kits (Medical Device Amendments of 1976), the CMS regulates diagnostic tests that are developed and performed in clinical laboratories under CLIA certification. Generally, FDA clearance of in vitro diagnostic devices includes evaluation of the performance claims of the assay, whereas CLIA laboratory inspections focus on reference laboratory quality standards. In addition, many state health departments require their own certifications for tests performed on patients from their state, and these inspections review assay validation reports for in-house–developed tests in detail. The Clinical and Laboratory Standards Institute publishes standards on assay validation and performance (160). Furthermore, large reference laboratories generally have adopted assay validation and acceptance policies to comply with pharmaceutical industry guidelines and expectations, including International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, Good Clinical Practice and Good Laboratory Practice, and International Organization for Standardization accreditation.

The proliferation of a consensus interpretative framework providing recommendations to clinicians and oncologists would be helpful for personalized cancer medicine. One possible structure would be a form similar to that of the NCCN or ASCO, in which a centralized body (or bodies) collates information on clinical trials and issues guidelines on clinical policy in oncology based on a system akin to “levels of evidence.” Guidelines on personalized cancer medicine could be issued in the form of an annual guide, for example, and made available in an affirmed database. These guidelines would be based on the strength of evidence available from preclinical and clinical trials. The result of this interface between cancer centers would be a gradual shift from empirical cancer treatment options to more individualized (and effective) therapies.

An unintentional but nonetheless important consequence of creating a cross-disciplinary, integrated framework is the empowerment of any participating cancer center with clinical trials of sufficient volume. A linked database could then act as a venue to interpret the findings of clinical trials and to determine the therapeutic significance of an individual mutation or a spectrum of alterations. This framework could encourage the emergence of professional and patient-centric networks and support groups, and function as an interface between individual cancer centers and a more global network of participants.

Overall, realizing the goal of personalized cancer medicine will require the systematic engagement of all interested parties: patients, oncologists, cancer centers, large cooperative clinical trials groups (such as the National Surgical Adjuvant Breast and Bowel Project), pharmaceutical companies, insurance providers, and government and regulatory agencies. Some tentative movement has been made in this direction, with the emergence of the drug-diagnostic codevelopment paradigm (161), whereby drug development groups team with diagnostic entities to develop a companion test for their therapeutic. The proliferation of targeted agents in development and clinical practice necessitates concomitant implementation of companion diagnostic approaches that enrich for subpopulations most likely to respond to a drug.

**CONCLUDING REMARKS**

Ongoing global genome characterization efforts are revolutionizing both tumor biology and the optimal paradigm for cancer treatment to an unprecedented degree. The pace of
advance is empowered in large part through technological innovations that render complete cancer genome characterization feasible on a large scale. The technologies available in the research setting today will become the clinical tests of tomorrow. In the near future, traditional diagnostic and prognostic factors will be supplemented by tumor genomic signatures, well-characterized molecular therapeutic targets, chemotherapy response assays, and the identification of deranged molecular pathways. Despite the inherent complexity of cancer genomics, incorporating this knowledge of the molecular basis of cancer into clinical decision making will speed the advent of more effective anticancer therapies. The challenges (in many different fields) relate to the implementation of comprehensive screening of a patient’s cancer genome for diagnostic and/or therapeutic use. These challenges lie in the areas of technologies, clinical trial design, legal and social implications, healthcare information technology, and insurance and reimbursement.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed by L.E. MacConaill and P. Van Hummelen. M. Meyerson is a consultant for and received research support from Novartis; the founding advisor of, consultant for, and equity holder in Foundation Medicine; a coinventor and patent holder on the use of EGFR mutations for lung cancer diagnosis, licensed to Genzyme Genetics/Labcorp. W.C. Hahn is a consultant for Novartis and Thermo-Fisher and has received research support from Novartis.

Acknowledgments

We thank our many colleagues whose work could not be cited owing to space limitations.

Received May 12, 2011; accepted August 4, 2011; published online September 15, 2011.

REFERENCES

Clinical Implementation of Cancer Genomes


92. Consortium M, Shi L, Reid LH, Jones WD, Shipp R, Warrington JA, et al. The MicroArray Quality Control (MAQC) project shows...


