Translating Genomics for Precision Cancer Medicine

Sameek Roychowdhury¹ and Arul M. Chinnaiyan²

¹Department of Internal Medicine, Division of Medical Oncology, and Comprehensive Cancer Center, Ohio State University, Columbus, Ohio 43210; email: sameek.roychowdhury@osumc.edu

²Michigan Center for Translational Pathology, Department of Pathology, Comprehensive Cancer Center, Howard Hughes Medical Institute, Department of Urology, and Center for Computational Medicine and Biology, University of Michigan Medical School, Ann Arbor, Michigan 48109; email: arul@umich.edu

Keywords
cancer, precision oncology, resistance mechanisms

Abstract
The Human Genome Project not only provided the essential reference map for the human genome but also stimulated the development of technology and analytic tools to process massive quantities of genomic data. As a result of this project, new technologies for DNA sequencing have improved in efficiency and cost by more than a millionfold over the past decade, and these technologies can now be routinely applied at a cost of less than $5,000 per genome. Although the application of these technologies in cancer genomics research has continued to contribute to basic discoveries, opportunities for translating them for individual patients have also emerged. This is especially important in clinical cancer research, where genetic alterations in a patient’s tumor may be matched to molecularly targeted therapies. In this review, we discuss the integration of cancer genomics and clinical oncology and the opportunity to deliver precision cancer medicine.
INTRODUCTION

The Human Genome Project not only provided the essential reference map for the human genome but also stimulated the development of technology and analytic tools to process massive quantities of genomic data (45). As a result, new technologies have emerged that have improved the efficiency of DNA sequencing by more than a millionfold, and these technologies can now be routinely applied at a cost of less than $5,000 per genome (70). This is especially important in research on cancer, where genetic alterations may contribute to the initiation and progression of disease (48). The decreased cost and improved efficiency of next-generation sequencing technologies have enabled the characterization of a landscape of genomic alterations in cancer (42, 116). The integration of cancer genomics research and genomic technologies into clinical oncology promises the opportunity to deliver precision cancer medicine.

MOLECULAR CHARACTERIZATION OF CANCER

A Precision Taxonomy for Cancer

In 2011, the National Research Council published a report on building a framework for a new precision taxonomy of human disease based on the “explosion of molecular data” through the disciplines of genomics, proteomics, and metabolomics (76). In particular, the authors identified disease taxonomy as a part of this process and pointed toward an opportunity and need for application to human disease and clinical outcomes. They outlined the need for prospective clinical studies, the generation and collection of molecular and clinical data, and databases with restricted access for researchers.

As a disease, cancer has been a model for a new taxonomy through efforts such as The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC). Before the era of molecular biology and (later) genome-scale sequencing, cancer was classified as a disease by site of origin in the human body (Figure 1). This was influenced by the available technologies and tools and also by the clinical classification required by surgical management, which was the initial mainstay of clinical oncology. Subsequently, microscopy-based classification of disease further delineated cancer subsets based on histologic differences. For example, aggressiveness or risk of relapse was retrospectively linked to histologic grading as a prognostic biomarker, such as Gleason and Bloom–Richardson for prostate and breast cancer, respectively. The histologic classification was advanced with assessment of prototypic surface markers (immunohistochemistry), gross karyotypic changes, and selected DNA and RNA biomarkers. For instance, surface antigen markers for lymphoid subsets have heavily influenced the diagnostic classification of lymphoma. Characterization of chromosomal abnormalities in leukemia and sarcoma has aided in the diagnosis of disease subsets as well as prognostic assessment (7).

In the era of genomics, the molecular classification of cancer continues to expand and in turn has the potential to facilitate the development of novel biomarkers. In 2006, the National Cancer Institute and National Human Genome Research Institute funded TCGA as a national network of teams to combine research and technology expertise to catalog more than 40 common types of cancer. The development of an infrastructure of publicly accessible databases was essential to enable cancer research worldwide. Initial success garnered additional funding in 2009 for the ongoing characterization of additional tumor subtypes. Similarly, the international community organized the ICGC to coordinate and characterize more than 50 subtypes of cancer and make the data readily available to the research community. The privately funded Pediatric Cancer Genome Project collaborates with TCGA and the ICGC, focusing on pediatric cancer subtypes (27). Together, these and other research efforts have contributed to the discovery of novel cancer genes and
The evolution of molecular diagnostics in cancer. This time line for molecular pathology demonstrates the incorporation of new technologies for the classification of cancer. Microscopy was the initial tool and remains a mainstay for evaluating cancer specimens. Early genetic tools included chromosome karyotyping techniques that could be applied to readily available metaphase cells from leukemias and sarcomas. Immunology contributed specific monoclonal antibodies and enabled the identification of molecules through immunohistochemistry. Molecular biology approaches such as the polymerase chain reaction (PCR) and Sanger sequencing enabled the analysis of single-gene variants. High-throughput strategies using arrays facilitated the evaluation of thousands of genes. Finally, technology that enabled massively parallel sequencing has emerged since the Human Genome Project and has greatly advanced the molecular diagnosis of cancer. Additional abbreviation: FISH, fluorescence in situ hybridization. Karyotyping image taken from Reference 6, courtesy of the Department of Pathology and Laboratory Medicine of the Perelman School of Medicine at the University of Pennsylvania.

targets for therapy (Table 1). Integration of these data sets also facilitates analyses across different subtypes of cancer that would not be possible without these large collaborative projects (57, 119).

Even with more than 10,000 cancers profiled thus far, it is imperative to continue these projects because few genomic alterations are common in cancer; rather, alterations occur infrequently across different cancer subtypes as a “long tail” seen in pan-cancer analyses. With hundreds of different cancer subtypes, most cancers may in fact be rare or orphan cancers with regard to prevalence, and the complete characterization of cancer will require profiling many more cases.

**Paradigm for Targeted Therapy: Chronic Myeloid Leukemia**

Early discoveries on chromosomal aberrations in chronic myeloid leukemia (CML) revealed a pathognomonic chromosomal translocation involving \( BCR \) and \( ABL \) genes or the Philadelphia chromosome (77, 93). After 40 years of molecular characterization of \( BCR-ABL \) and subsequent drug development of an ABL kinase inhibitor (29), imatinib was clinically developed as an inhibitor of the constitutively active ABL kinase in CML (28). Imatinib and CML have served as the prototype for molecularly targeted therapeutics in oncology. Since the approval of imatinib, more than 30 targeted therapies have been approved for indications in oncology, including small-molecule kinase inhibitors, monoclonal antibodies, and antihormonal agents.
Table 1  Examples of how cancer genomics fuels translational therapeutics

<table>
<thead>
<tr>
<th>Genomic alteration</th>
<th>Pathway</th>
<th>Disease examples</th>
<th>Putative or proven drugs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKT3 fusion, mutation (4, 12)</td>
<td>Phosphoinositide 3-kinase</td>
<td>Breast cancer</td>
<td>AKT inhibitors</td>
</tr>
<tr>
<td>TSC1 mutation (13, 52)</td>
<td>mTOR</td>
<td>Bladder cancer, tuberous sclerosis</td>
<td>mTOR inhibitors</td>
</tr>
<tr>
<td>KRAS fusion (118)</td>
<td>RAS-MEK</td>
<td>Prostate cancer</td>
<td>RAF inhibitors, MEK inhibitors, phosphoinositide 3-kinase inhibitors</td>
</tr>
<tr>
<td>BRAF mutation (21, 112)</td>
<td>RAS-MEK</td>
<td>Hairy-cell leukemia</td>
<td>RAF inhibitors, MEK inhibitors</td>
</tr>
<tr>
<td>FGFR1–4 fusion (1, 120)</td>
<td>FGFR</td>
<td>Bladder, breast, ovarian, lung, and prostate cancers, cholangiocarcinoma</td>
<td>FGFR inhibitors</td>
</tr>
<tr>
<td>ERBB2 mutation (9, 30)</td>
<td>ERBB2</td>
<td>Breast, gastric, and lung cancers</td>
<td>ERBB2 inhibitors</td>
</tr>
<tr>
<td>ERBB3 mutation (54)</td>
<td>ERBB3</td>
<td>Gastric and colon cancers</td>
<td>ERBB3 inhibitors</td>
</tr>
<tr>
<td>JAK3 mutation (60)</td>
<td>JAK-STAT</td>
<td>Natural killer/T cell lymphoma</td>
<td>JAK or STAT inhibitors</td>
</tr>
<tr>
<td>IL7R mutation (121, 122)</td>
<td>JAK-STAT</td>
<td>Leukemia</td>
<td>JAK or STAT inhibitors</td>
</tr>
<tr>
<td>DDR2 mutation (47)</td>
<td>Receptor tyrosine kinase</td>
<td>Lung cancer</td>
<td>Some tyrosine kinase inhibitors</td>
</tr>
<tr>
<td>EPOR fusion (89)</td>
<td>JAK-STAT</td>
<td>Leukemia</td>
<td>JAK or STAT inhibitors</td>
</tr>
<tr>
<td>CDKN2A deletion (51)</td>
<td>Cyclin-dependent kinase</td>
<td>Melanoma</td>
<td>Cyclin-dependent kinase inhibitors</td>
</tr>
<tr>
<td>AR amplification, mutation (44, 55, 95)</td>
<td>Androgen</td>
<td>Prostate cancer</td>
<td>Androgen synthesis inhibitors, androgen receptor inhibitors</td>
</tr>
<tr>
<td>ESR1 mutation (91, 114)</td>
<td>Estrogen</td>
<td>Breast cancer</td>
<td>Estrogen synthesis inhibitors, estrogen receptor inhibitors</td>
</tr>
<tr>
<td>ALK fusion, mutation (16, 62, 66, 102)</td>
<td>ALK</td>
<td>Lung and colorectal cancers, neuroblastoma</td>
<td>ALK inhibitors</td>
</tr>
<tr>
<td>RET fusion, mutation (66, 108)</td>
<td>RET</td>
<td>Lung and thyroid cancers</td>
<td>RET inhibitors</td>
</tr>
<tr>
<td>ROS1 fusion (108)</td>
<td>ROS1</td>
<td>Lung cancer, cholangiocarcinoma</td>
<td>ROS1 inhibitors</td>
</tr>
<tr>
<td>NOTCH–2 fusion, mutation (34, 61, 87, 90)</td>
<td>NOTCH signaling</td>
<td>Leukemia, breast cancer</td>
<td>Notch signaling pathway inhibitors</td>
</tr>
</tbody>
</table>

Although the development of imatinib as a therapeutic for a genomic alteration heralded the era of molecularly targeted agents, several lessons would emerge over the ensuing decade. First, unlike CML, most cancers are not homogeneously propelled by a single genomic driver alteration; instead, they comprise rare disease subsets with a variety of genomic alterations. Second, single-agent therapies against a single genomic target have not been as successful in achieving cures or long-term survival as was imatinib in CML. Thus, CML has been the exception and not the rule, which highlights the importance of developing rational combination therapies and elucidating mechanisms of drug resistance to single agents.

Clinical Biomarkers

High-throughput technologies such as gene expression microarrays and genomic sequencing have the potential to further clinical biomarker development in cancer. Clinical biomarkers must first demonstrate analytic validity and then be evaluated for clinical utility in a clinical trial. Analytic
Figure 2
The development of genomics-driven biomarkers. The application of next-generation sequencing has the potential to support the discovery of multiple types of biomarkers. Diagnostic biomarkers help to classify cancers into types and subtypes. Predictive biomarkers guide the choice of therapy for a particular cancer. Prognostic biomarkers provide data on the risk of cancer relapse after initial therapy, such as surgery. Finally, pharmacogenomic biomarkers may influence the dosing or delivery route of a drug based on how the drug is absorbed, metabolized, and excreted. Positron-emission tomography (PET) scans taken from Reference 36.

validity implies that a biomarker test is reproducible and meets clinical-grade standards or certification (41); clinical utility implies that it provides clinicians with data that facilitate clinical decision making.

Biomarkers may have diagnostic, predictive, prognostic, or pharmacogenomic clinical utility (98) (Figure 2). Diagnostic biomarkers may indicate whether a disease is cancer and may provide information about the type or subtype of cancer. Predictive biomarkers provide information about which drug(s) may be appropriate to use for treatment. Prognostic biomarkers provide information about the risk of disease recurrence and whether a patient needs additional therapy or surveillance. Pharmacogenomic biomarkers provide information about drug dosing or risk of toxicity.
Molecular Biomarkers in Lung Cancer

During the past 30 years, there has been a molecular revolution in lung cancer that demonstrates how genomic technologies have changed the approach to characterizing cancer and developing novel therapies based on a new taxonomy (82). In the 1980s, the classification of lung cancer was based on histologic subtypes of adenocarcinoma, squamous-cell carcinoma, and small-cell carcinoma. These histologic subtypes provided oncologists with predictive value for therapies such as platinum- and radiation-based regimens. In the late 1980s, researchers identified that the KRAS oncogene was mutated and activated in up to 25% of lung cancers (74, 96). Additionally, inhibitors of epidermal growth factor receptor (EGFR) were developed, and patients who were sensitive to specific inhibitors of EGFR were more likely to have EGFR mutations that led to functional activation of the kinase without ligand (68, 80, 103). Subsequently, EGFR mutations were found in 10–15% of all lung cancers, leading to the development of a predictive molecular diagnostic test that could guide the selection of patients for whom EGFR inhibitors are most likely to be beneficial.

More recently, the molecular classification of lung cancer has expanded to include multiple putative targets for novel targeted therapies, including ALK fusions (102), RET fusions (66, 108), BRAF mutations (81), and FGFR alterations (fusion, mutation, or amplification) (31, 65, 120). A recent study demonstrated the broad-based implementation of genomic-based testing for lung cancer, its influence on molecular compared with histologic classification, and its survival benefits for patients with EGFR and ALK alterations (18).

Molecular Biomarkers in Prostate Cancer

Before 2005, chromosomal translocations were generally thought to be common in hematologic malignancies and leukemias but not in solid tumor malignancies. Through the use of gene expression microarrays and, later, genomic sequencing technologies, gene fusions or translocations were identified in up to 50% of prostate cancers. Generally, these involve an androgen-regulated gene (such as TMPRSS2) and genes encoding ETS family transcription factors (such as ERG or ETV1). The presence of these genomic alterations has 100% specificity for prostate cancer and therefore potential clinical utility as a diagnostic biomarker.

This is particularly important in prostate cancer, where screening is widely offered using the blood test for prostate-specific antigen (PSA), which is subject to false-positive results, leading to multiple invasive prostate biopsies. To improve the PSA screening test, Tomlins et al. (113) developed a multiplex assay, including blood PSA and urine evaluation for TMPRSS2-ERG fusion transcripts, that has improved specificity over PSA testing alone. Such a diagnostic tool has the potential to obviate the need for biopsy in patients with benign problems, such as benign prostate hypertrophy. Prospective study is needed to demonstrate whether this diagnostic tool selects for clinically significant or high-grade prostate cancer. Beyond a diagnostic biomarker, fusion genes in prostate cancer may also have therapeutic or predictive implications, indicating sensitivity to PARP inhibitors (indirect targeting) or to novel drugs that inhibit the function of ETS transcription factors (10, 59).

Molecular Biomarkers in Breast Cancer

Women with early-stage breast cancer (stage 1 or 2) have variable risk of relapse after curative intent surgery. They have the option of additional therapy with chemotherapy and antiestrogen drugs to prevent relapse. These additional or adjuvant therapies can add risk and morbidity to those women
who have a lower risk of relapse. Gene expression signatures generated from earlier microarray data may provide prognostic risk assessment to separate women with either higher or lower risk of relapse and thereby facilitate decision making about who needs additional chemotherapy treatment, potentially sparing some women unneeded treatment and potential toxicity (105).

Gene expression profiling has delineated subsets of breast cancer with prognostic and predictive implications (84, 104). Through the latest contributions from TCGA and the ICGC, this knowledge base has been extended with genomic sequencing data (33). Novel molecular targets are emerging as putative predictive biomarkers in breast cancer, including phosphoinositide 3-kinase signaling (3), FGFR amplifications or fusions (88, 120), Notch signaling (90), and cyclin-dependent kinase activation (106).

Rare Cancers

Although TCGA, the ICGC, and other efforts characterize the most common cancers, such as lung, breast, prostate, and colon cancers, there are hundreds of cancer subtypes that are considered rare and are not subject to these efforts. Furthermore, in contrast to common cancers, rare cancers such as sarcomas present a challenge to oncologists and pathologists because there are typically few retrospective studies available and little is known about molecular underpinnings.

Cancer epidemiologic workshops have defined rare cancers as those having an incidence of 15 per 100,000 per year, or less than 10,000 new cases (46). Based on these criteria, we estimate that rare cancers constitute up to 20–25% of the adult cancer patient population. Tumor sequencing studies that include such patients can provide clinically significant genomic sequencing data to patients and their doctors while at the same time expanding the molecular taxonomy of these poorly defined cancers. For example, Tiacci et al. (112) performed whole-exome sequencing on one index case of hairy-cell leukemia and identified a canonical activating mutation in BRAF. This finding was subsequently validated in 47 additional patients with hairy-cell leukemia, and this gene represents a novel therapeutic target for patients refractory to standard treatments for this disease. In another example, cancer genome sequencing strategies have identified novel recurrent gene fusions involving FGFR family genes in cholangiocarcinoma that result in constitutive kinase activity (1, 120). There are now clinical trials in development for FGFR and RAF kinase inhibitors for these rare subsets of cholangiocarcinoma and leukemia, respectively.

GENOMICS AND DRUG DEVELOPMENT

Imatinib and CML revolutionized drug development with a new focus on targeted therapies or “smart drugs” for cancer. This took place over a period of 40 years, from the discovery of the Philadelphia chromosome to the phase 1 trial of imatinib. The development of therapies targeting estrogen receptor and androgen receptor occurred over a similar time line, where the clinical effect was observed before researchers had an understanding of the molecular underpinnings. Two recent studies involving BRAF and ALK kinase inhibitors illustrate the potential for accelerating drug development.

Molecular Enrichment in BRAF-Mutant Melanoma and ALK-Rearranged Lung Cancer

The characterization of activating mutations and drug development for the BRAF oncogene highlights the opportunity for accelerated drug development enabled by the molecular selection of patients in clinical trials. In 2002, Davies et al. (21) performed a resequencing of RAS pathway
genes and identified recurrent mutations involving the kinase domain of \textit{BRAF}. Owing to the prevalence of \textit{BRAF}-activating mutations in melanoma and thyroid cancer, the phase 1 clinical trial of vemurafenib in 2010 (8) incorporated \textit{BRAF} molecular eligibility in the expansion cohort (36) and led to 80\% overall response rates in patients with \textit{BRAF} mutations. Soda et al. (102) identified gene fusions involving \textit{ALK} in 2007. The phase 1 trial in 2010 (62) involved screening more than 1,500 patients with lung cancer to identify 82 patients with \textit{ALK} gene fusions, with an overall response rate of 57\%, which led to eventual approval of the ALK inhibitor crizotinib for \textit{ALK}-rearranged lung cancer. Investigators are developing BRAF and ALK inhibitors in other cancer subtypes with the corresponding activating genomic alterations in \textit{BRAF} (lung cancer, hairy-cell leukemia) or \textit{ALK} (colon cancer) (11, 24, 66, 81, 112).

\textbf{Genomics-Driven Trials}

The ongoing challenge for genomics-driven trials is that the targetable genomic alterations in individual cancer subtypes are quite rare. Genomic alterations do not occur in the majority of a cancer subtype; rather, they occur in 1–20\% of a cancer [e.g., the \textit{BRAF} mutation occurs in 1\% of lung cancer (81)] or across multiple cancer subtypes (e.g., lung cancer, thyroid cancer, hairy-cell leukemia, chronic lymphocytic leukemia) (35). This has several implications for drug development (\textbf{Table 2}). First, because of the underlying prevalence of genomic targets across diseases, it will be challenging to complete large randomized clinical trials in a single disease. Second, because of the multitude of genes and types of genomic alterations for a given disease or across diseases, pathologists will need rapid, cost-effective approaches for cancer gene testing. Third, in some instances, a targeted therapy displays tremendous efficacy in a small subset of patients but may not have a well-defined predictive biomarker. Fourth, because of the smaller sample size of such clinical trials, it will be difficult to complete randomized trials, and alternative end points may be needed for such trials. Here, we outline several approaches to these problems.

\begin{table}[h]
\centering
\caption{Challenges and opportunities for genomics-driven trials}
\begin{tabular}{|l|l|l|}
\hline
\textbf{Challenge} & \textbf{Potential solutions} & \textbf{Examples} \\
\hline
Genomic targets are rare and occur across cancer subtypes & Multicenter trials through a disease consortium \\
& Basket trials for genomic targets across different cancer subtypes & Lung Cancer Mutation Consortium \\
& & BRAF inhibitor basket trial “A Study of Vemurafenib in Patients with BRAF V600 Mutation-Positive Cancers” (ClinicalTrials.gov ID NCT01524978) \\
& & N-of-1 registry for off-label therapy \\
\hline
There is no cost-effective method for pan-cancer and pan-genomic alteration testing & Multiplex assays utilizing genomic technologies & Academic cancer centers \\
& & Commercial testing labs \\
\hline
Effective targeted therapies are available but predictive biomarkers are unknown & Genomic evaluation of selected patients with exceptional response to targeted therapy & mTOR inhibitor (everolimus) in bladder cancer \\
\hline
A small sample size for trials places limitations on routine clinical end points & Trial end points that consider the magnitude of disease response to targeted therapy \\
& Assessment of signal and characterization of drug resistance & Studies of gastrointestinal stromal-cell tumors and imatinib \\
\hline
\end{tabular}
\end{table}
Disease-Based and Mutation-Based Basket Trials

Similar to the development of BRAF inhibitors for melanoma and ALK inhibitors for lung cancer, the majority of prospective trials with mutation eligibility are currently disease based (http://www.clinicaltrials.gov). For example, for BRAF-mutant cancers, there are multiple disease-based trials for melanoma, colon cancer, and thyroid cancer. These diseases have a high prevalence of the mutation (25–50%) or themselves are common enough (colon cancer) to support accrual. Lung cancer has a 1% incidence of BRAF-activating mutations and therefore presents a challenge for accruing patients. Because lung cancer is a common disease, accruing 1% is feasible only through a multicenter trial or consortium across 10 or more centers (e.g., the Lung Cancer Mutation Consortium) to facilitate screening and enrollment of patients (Figure 3b). In the majority of cases, however, either the genomic alterations or the diseases themselves are rare (e.g., BRAF-mutant leukemia or cholangiocarcinoma).

Some investigators are utilizing a pan-cancer (or basket) trial of multiple histologic subtypes with a mutational eligibility, such as BRAF-activating mutations (Figure 3c). A drawback to this approach is that BRAF mutations can behave differently in different diseases (e.g., melanoma compared with colorectal cancer) (15, 19, 85); however, such studies could facilitate understanding of primary resistance to therapy. Meanwhile, common diseases or alterations such as melanoma or colorectal cancer—for which basket trials would capture only rare disease or alterations—would continue to be studied in disease-based trials. Basket trials may also be practical for “signal finding” to determine which histologic subtypes of cancer are more likely to benefit from targeted therapy or to identify mechanisms of primary or acquired resistance.

Matching Patients, Mutations, and Drugs

Patients and oncologists need to consider several important challenges, including that not all mutations are significant, not all significant mutations respond to a matching therapy, and not all patients have ready access to a seemingly rational drug choice. Although recent work from TCGA has highlighted a spectrum of mutations that occur across cancer types, the majority of mutations may not be clinically or biologically significant, i.e., are passenger mutations or alterations (63). In contrast, driver mutations are thought to have conferred a selective advantage to cancer cells with respect to growth, survival, or drug resistance. Passenger mutations accumulate in cancer by co-occurring with driver mutations and do not necessarily provide a selective advantage. For example, melanoma and lung cancers have the highest rates of somatic point mutations that result from carcinogen exposures, but only a few mutations in each cancer are thought to be driver mutations (63). It is possible that we lack a complete understanding of the biology and therefore need to profile additional cancers to extend our knowledge of drivers or candidate drivers and their therapeutic implications (109).

The ability to match patients to therapies and the capacity to test treatment hypotheses for driver–drug matches are significant barriers, owing to limited access to investigational agents in disease-based trials and the costs of off-label therapy. Repurposing of approved targeted therapies has the potential to enable testing of novel treatment hypotheses (53). Collection of clinical outcome data for such N-of-1 scenarios is not systematic and thus could benefit from a centralized registry for tracking outcomes. Although case studies in a national-level registry do not have the statistical power of a prospective clinical trial, N-of-1 case studies can add value when selecting treatment hypotheses for further study in trials.
Traditional trial for unselected patients with a single cancer histology who are uniformly treated with a novel targeted therapy. (b) Genomics-enriched disease-based trial, in which patients with a single cancer histology are molecularly enriched for a genomic target and receive a matching novel targeted therapy. This design can prospectively test a putative predictive biomarker such as ALK fusion and ALK inhibitors. (c) Genomics-enriched basket trial in which patients with any cancer histology are molecularly enriched for a common genomic target and receive a matching novel targeted therapy. This design can prospectively test a putative predictive biomarker that occurs across diverse cancer histologies. (d) Exceptional-responders trial that evaluates patients with a single cancer histology from a traditional trial. The genomic characterization of the rare patients who have exceptional responses to a therapy can identify existing or novel predictive biomarkers.

Implementing Next-Generation-Sequencing-Based Molecular Diagnostics

Developing clinical-grade single-gene assays separately for mutation, amplification, and gene fusion testing is cost prohibitive and inefficient and delays drug development at academic cancer centers. As a result, the majority of cancer centers are investing in the development of pan-cancer testing strategies employing next-generation sequencing technology (71). This can help clinical
investigators identify patients with genomic alterations who are eligible for targeted therapies in development. Focused cancer gene panels (50–250 genes) are available in Clinical Laboratory Improvement Amendment–certified laboratories both commercially (38) and at individual academic cancer centers (5, 86, 101).

This approach underscores the role of molecular pathology as a discipline to facilitate drug development. Furthermore, it emphasizes the need to include genomic sequencing and bioinformatics training in pathology resident and fellowship education. In addition to focused cancer gene panels, there is a role for more unbiased sequencing strategies encompassing the whole exome (~22,000 genes), whole genome, and transcriptome (RNA sequencing) (72).

Genomic Evaluation of Exceptional Responders

For a traditional trial design in which all patients receive an investigational agent, retrospective evaluation of exceptional responders can facilitate predictive biomarker discovery (52) (Figure 3d). It is not uncommon for a phase 1 or 2 trial to have only one or two patients with an impressive disease response, with the remainder of patients being refractory to the agent. Oftentimes, this may be the end for a given investigational agent, but if a subset of patients can be identified through a predictive biomarker and further evaluated in another trial, investigational agents could be rescued for select patients.

Iyer et al. (52) reported on such a strategy for the treatment of patients with metastatic bladder cancer using the mTOR inhibitor everolimus. They evaluated a single patient with an exceptional response to everolimus (lasting longer than two years) through whole-genome sequencing of her tumor and identified a somatic mutation in TSC1, a known component of mTOR signaling. They expanded their analysis through targeted sequencing of 200 cancer genes in a larger cohort of bladder cancer samples and identified recurrent mutations involving TSC1. Thus, this approach has the potential to identify candidate predictive biomarkers, is relatively cost effective (compared with sequencing all patients in the trial), and is especially appealing to the National Cancer Institute program for trials (56).

Clinical End Points for Genomics-Driven Trials

Trials in which either the genomic target or cancer is rare may not be feasible with standard clinical end points, which depend on the randomization of patients. Some investigators have proposed alternative end points for trials of targeted therapies with molecular enrichment or selection of patients (99). One precedent comes from the phase 2 nonrandomized trial of two different doses of imatinib for patients with gastrointestinal stromal-cell tumors (GISTs) with activating mutations in CKIT (20). Imatinib was approved for GISTs based on its overall response rate (30–40%). For GISTs, the historical responses to other chemotherapies were dismal, and the magnitude of the duration of response was unprecedented. There was also a justified biological rationale for imatinib as a CKIT inhibitor in CKIT-activated GISTs.

We would make the case for additional end points for signal finding and understanding of acquired or secondary resistance. Signal-finding studies will enable larger multicenter trials to accrue the appropriate disease subsets that are likely to respond to the drug tested. Acquired or secondary resistance develops in patients who respond initially to treatment but whose disease later progresses. The intent of this end point is to elucidate mechanisms of acquired resistance in order to provide a basis for future rational drug combinations to thwart impending escape mechanisms (Table 3).
Table 3  Novel resistance mechanisms characterized through genomic sequencing strategies

<table>
<thead>
<tr>
<th>Genomic target</th>
<th>Cancer</th>
<th>Mechanisms of resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRAF mutation</td>
<td>Melanoma</td>
<td>BRAF amplification, MEK1 mutation (100, 117)</td>
</tr>
<tr>
<td>ALK fusion</td>
<td>Lung cancer</td>
<td>ALK mutations detected in resistant tumor (58)</td>
</tr>
<tr>
<td>ROS1 fusion</td>
<td>Lung cancer</td>
<td>ROS1 mutations detected in resistant tumor (2)</td>
</tr>
<tr>
<td>EGFR mutation</td>
<td>Lung cancer</td>
<td>KRAS mutations detected in plasma (23)</td>
</tr>
<tr>
<td>ESR1</td>
<td>Breast cancer</td>
<td>ESR1 mutations detected in resistant tumor (91, 114)</td>
</tr>
<tr>
<td>BTK</td>
<td>Chronic lymphocytic leukemia</td>
<td>BTK mutations detected in resistant tumor (14)</td>
</tr>
</tbody>
</table>

Benefits of Genomics in Clinical Trials

Although the incorporation of genomic sequencing and research biopsies adds costs to clinical trials, there are substantial cost savings for drug development. A significant component of development costs arises from expensive multicenter clinical trials involving hundreds or thousands of patients that result in limited clinical benefit or improvement (78). In contrast, genomics-based trials may facilitate the identification of the target population of patients and thereby decrease the size of the trials needed to observe a benefit or effect. Furthermore, understanding the mechanism of a drug’s effect or of resistance to a drug can contribute to subsequent trial design.

UNDERSTANDING MECHANISMS OF DRUG RESISTANCE

Combination Versus Single-Agent Therapy

Although the triumph of ABL inhibitors in enabling the long-term survival of patients with CML has been the model for targeted therapy in cancer, it is becoming clear that single-agent targeted therapy is unlikely to be curative in the majority of cancers. This is analogous to the paradigm of drug development in the 1960s for pediatric acute lymphoblastic leukemia, where empirical clinical trials over 30 years led to the advancement of a three-year combination chemotherapy regimen with five-year survival approaching 90% (39). This initial strategy was meant to be toxic to all leukemia cells before a subset have a chance to develop resistance to one agent. The strategy was appropriated from an approach in infectious diseases that used multiagent antibiotic therapy for tuberculosis, and it more recently emerged in a three-drug antiviral regimen for human immunodeficiency virus (HIV) when researchers observed the ability of the virus to evade one or two antiviral drugs (50, 83). Triple-drug regimens for HIV management have been successful in controlling viral replication over the long term (56). Thus, understanding drug resistance in cancer is fundamental to developing combination therapies with curative potential.

Research Tumor Biopsies

There are multiple mechanisms for cancer to develop resistance to targeted therapies, including preexisting genomic heterogeneity and evolution, stem cell compartments that are quiescent and resistant to therapy, and a microenvironment that can sustain tumors despite effective targeted therapy. Collection and profiling of tumor biospecimens in clinical trials are essential to understanding emerging drug resistance. Biospecimen collection in trials has historically been limited to leukemia studies, where the collection of blood and bone marrow specimens has been routine owing to ease of access and limited safety concerns. More recently, research tumor biopsies are
becoming more widely utilized in trials, with demonstrated safety for patients, the willingness of patients to participate in such trials, and an ethically sound rationale for understanding the basic mechanisms of how drugs succeed or fail (32, 79).

**Acquired Resistance to Targeted Therapies**

Resistance to ALK inhibition in ALK-rearranged lung cancer provides an example of acquired resistance to targeted therapies. Several groups have evaluated tumor specimens after patients developed resistance and identified secondary ALK mutations and amplifications, CKit amplification, and EGFR activation as mechanisms of resistance (17, 26, 58). However, the reason for the large proportion of acquired resistance in ALK-rearranged lung cancer is unknown (40). Most efforts to study this resistance have involved focused testing of known pathways in lung cancer. The advantage of high-throughput genomic sequencing strategies is their unbiased identification of candidate mechanisms of resistance, including point mutations, copy-number alterations, and translocations. Furthermore, combining these strategies with transcriptome sequencing could enable the evaluation of gene expression and alternative splicing.

For example, Wagle et al. (117) evaluated a patient with BRAF-mutant metastatic melanoma who initially displayed a profound response to the BRAF inhibitor vemurafenib but relapsed within two months. Pretreatment and postprogression biopsies were evaluated by DNA tumor sequencing of 137 oncogenes and tumor suppressors using next-generation sequencing. This revealed a mutation in MEK1, a component of RAF signaling downstream from BRAF. The authors demonstrated that this MEK1 mutation conferred constitutively active MEK activity, effectively bypassing BRAF inhibition. Additional mechanisms of resistance have been identified in BRAF-mutant melanoma (67), providing a rationale for a clinical trial combination of BRAF and MEK inhibition that resulted in improved clinical outcomes and reduced toxicity (37).

Two recent studies highlighted acquired mutations in women with metastatic estrogen-receptor-positive breast cancer through evaluation of postprogression tumor samples. Toy et al. (114) identified recurrent mutations in the ligand-binding domain of ESR1 in 14 of 180 patients from two clinical studies involving antiestrogen therapies. Robinson et al. (91) identified the same ESR1 mutations in a global clinical tumor sequencing study in women who developed resistance to standard-of-care antiestrogen therapies. The knowledge that persistent estrogen signaling occurs through ligand-binding domain mutations in ESR1 in metastatic breast cancer provides an immediate rationale for the development of novel antiestrogen therapies for women who have estrogen-receptor-positive breast cancer.

Acquired resistance to a closely related hormonal target, androgen receptor (encoded by the AR gene), had been observed in castrate-resistant prostate cancer nearly two decades earlier, associated with secondary point mutations or copy-number amplifications of AR (110, 115). This illustrates the importance of acquiring postprogression tumor samples from patients participating in clinical trials and receiving standard-of-care therapies, which can facilitate the study of mechanisms of acquired resistance and further potentiate drug development.

**Assessing Drug Resistance Beyond Tumor Biopsy**

In addition to assessment of postprogression tumor biopsy specimens, there are several strategies for studying emerging resistance. The establishment of patient-derived tumor xenografts (PDTXs) or patient avatars allows the characterization of acquired resistance and facilitates additional testing of novel drugs and combinations (69, 111). The research value of PDTX models is evidenced by the development of more than 50 PDTX models across leukemia and solid tumor malignancies...
and the commercial development of models for research. In addition to PDTX models, recent discoveries have facilitated the derivation of in vitro organoids or tumoroids from patient tumor specimens and created tremendous potential to maximize the evaluation of finite clinical specimens (97). Complementary to PDTX and organoids, genetically engineered mouse models (GEMMs) of cancer facilitate a more focused research question that can be specifically geared to study genetic modifiers or compensatory pathways for a given genomic alteration (75, 111). Together, PDTX models and GEMMs are facilitating coclinical trials that use parallel clinical trials to enable the testing of research hypotheses and finite tissue specimens.

Another strategy to augment the study of lethal cancer and drug resistance is rapid autopsy programs (94). Rapid autopsy programs depend on coordination between clinical oncologists and pathologists and early discussion of the study with patients. As an example, a rapid autopsy program for prostate cancer at the University of Michigan facilitated a study of castrate-resistant prostate cancer through exome sequencing (44). Furthermore, these programs can be easily integrated with PDTX programs. In addition to assessment of soft tissue through biopsy and rapid autopsy programs, liquid tumor biopsies utilizing whole blood or plasma have the potential to supplement or even replace invasive tumor biopsies. Several studies have demonstrated the ability to identify genomic alterations in free DNA in plasma or in circulating tumor cells (22, 64, 73). As the technology and methods for single-cell or small-input DNA sequencing improve with respect to analytic validity and precision, blood and plasma will become an easily accessible and safe means to evaluate disease response and drug resistance.

Tumor Heterogeneity and Evolution

Tumors exist with heterogeneous subclones in a microenvironment that includes selective pressures such as chemotherapy or targeted therapies. Genetic diversity in these tumor populations is expected based on DNA replication errors and defects in DNA repair found in some cancers. Gerlinger et al. (43) recently demonstrated this diversity and a pattern of branching evolution, through exome sequencing of multiple sites of disease in patients with metastatic kidney cancer, including evaluation of a primary kidney tumor and diverse metastatic deposits. They observed that well-known or predicted oncogenes and tumor suppressors that are likely to provide a selective advantage (driver mutation) are common to all primary and metastatic disease sites, such as the von Hippel–Lindau gene (*VHL*). In contrast, mutations that are less likely to provide a selective advantage (passenger mutations) were divergent between primary or metastatic sites.

Ding et al. (25) employed whole-genome sequencing of paired acute myeloid leukemia and relapsed samples to characterize clonal evolution after response to initial chemotherapy. Additional targeted capture facilitated deep sequencing of selected genes and revealed two patterns of tumor evolution: (a) mutations that were in the initial leukemia as subclones emerging as dominant mutations in relapse, and (b) new mutations unique to the relapsed leukemia and not observed in the initial leukemia (25). In another striking example, Romano et al. (92) performed whole-exome sequencing on two progression samples in a patient with *BRAF*-mutant melanoma and identified two separate mechanisms of resistance involving NRAS mutations and alternative *BRAF* splicing. These studies underscore the importance of research tumor biopsies in clinical trials.

Limitations of Genomics for Precision Cancer Medicine

Although the use of genomic sequencing in clinical oncology has the potential to pave a path toward precision cancer medicine, it is not without caveats. A focus on genomic alterations as targets of therapy does not consider small populations of cancer stem cells that may be challenging
Figure 4
Translational cancer genomics paradigm. This illustrates the developing paradigm where all patients with any type of cancer participate in studies that bank biospecimens, stockpile “-omics” data, collect clinical data, and integrate with clinical trials in oncology. This approach can facilitate the development of basic cancer research, the translation of that research into the clinic, and the advancement of novel targeted therapies in clinical trials.

Figure
Translational cancer genomics paradigm. This illustrates the developing paradigm where all patients with any type of cancer participate in studies that bank biospecimens, stockpile “-omics” data, collect clinical data, and integrate with clinical trials in oncology. This approach can facilitate the development of basic cancer research, the translation of that research into the clinic, and the advancement of novel targeted therapies in clinical trials.

to detect or characterize and may be resistant to therapy. The tumor microenvironment provides a soil for cancer growth and can contribute to drug resistance. For example, stromal and tumor coculture experiments have revealed a role for secreted ligands in resistance to targeted therapies (49, 107). Methodologies to separate tumor and microenvironment compartments will be needed to further study these interactions. In addition, there have been impressive advances in cancer immunotherapy, particularly with immune system checkpoint blockades and engineered T cell receptors. There is therefore an opportunity to consider combination strategies with targeted therapies and immunotherapies.

Future Directions and Opportunities
The widespread incorporation of genomic sequencing in clinical trials has the potential to hasten drug development, improve understanding of drug resistance, and provide rationales for future combination therapies (Figure 4). The ongoing molecular catalog of cancer is certainly far from complete. Furthermore, new classes of drugs that go beyond targeting kinases and instead specifically target elements of transcription factor function or chromatin remodeling are under development. Finally, the study of the noncoding genome has enormous potential for cancer research, as the biology of this part of the genome is largely unknown. As genomic sequencing and other high-throughput technologies improve in efficiency and cost, we will have a wealth of data to mine in working toward the goal of precision cancer medicine.

SUMMARY POINTS
1. Large-scale cancer genomics initiatives have identified and continue to identify novel targets for therapy across cancer subtypes.
2. There is a need for genomic characterization to identify driver genomic alterations in rare cancer subtypes that are currently not part of large-scale initiatives.

3. Molecular diagnostics utilizing genomic technologies can be used to prospectively identify candidate patients for a targeted therapy.

4. Genomic characterization of exceptional responders in clinical trials can be used to identify candidate predictive biomarkers.

**FUTURE ISSUES**

1. The characterization of 10,000 cancers thus far is only the tip of the iceberg. Continued investment is needed for the molecular characterization of human cancer.

2. To integrate genomics into clinical trials in order to understand drug resistance, we will need to support research tumor biopsies as a standard expectation for clinical trials.

3. Cooperative trial groups are efficient at carrying out multisite clinical trials, and we must provide the needed expertise and support for incorporating molecular diagnostics into these trials.

4. There is currently a limited amount of expertise being developed at genomics research centers. We need to develop clinically oriented genomics programs for training oncologists, pathologists, bioethicists, and bioinformaticians.

5. Although cancer genomics research has focused largely on the coding region of the genome, there is much to elucidate in the noncoding genome.

**DISCLOSURE STATEMENT**

A.M.C. serves as a consultant for Life Technologies, Hologic, MolecularMD, and Paradigm.

**ACKNOWLEDGMENTS**

We thank Karen Giles and Brianna Hunstad for administrative support. S.R. is supported by an American Cancer Society Mentored Scholar Research Grant, a Prostate Cancer Foundation Young Investigator Award, and Pelotonia at the Ohio State University. A.M.C. is supported by a Doris Duke Charitable Foundation Clinical Scientist Award and a Burroughs Wellcome Foundation Award in Clinical Translational Research. A.M.C. is a Howard Hughes Medical Institute Investigator, an American Cancer Society Research Professor, and an Alfred A. Taubman Scholar.

**LITERATURE CITED**


Contents

DNA and Other Strands: The Making of a Human Geneticist
Leon E. Rosenberg ................................................................. 1

An Opportune Life: 50 Years in Human Cytogenetics
Patricia A. Jacobs ..................................................................... 29

Determinants of Mutation Rate Variation in the Human Germline
Laure Séguéla, Minyoung J. Wyman, and Molly Przeworski ............ 47

No Gene in the Genome Makes Sense Except in the Light of Evolution
Wilfried Haerty and Chris P. Ponting ...................................... 71

Phenotypic Outcomes of Imprinted Gene Models in Mice: Elucidation of Pre- and Postnatal Functions of Imprinted Genes
Mary A.M. Cleaton, Carol A. Edwards, and Anne C. Ferguson-Smith .......................................................... 93

The Pivotal Regulatory Landscape of RNA Modifications
Sheng Li and Christopher E. Mason ......................................... 127

Age-Related Macular Degeneration: Genetics and Biology
Coming Together
Lars G. Fritsche, Robert N. Fariss, Dwight Stambolian, Gonçalo R. Abecasis, Christine A. Curcio, and Anand Swaroop .......................................................... 151

Disorders of Cholesterol Metabolism and Their Unanticipated Convergent Mechanisms of Disease
Frances M. Platt, Christopher Wassif, Alexandria Colaco, Andrea Dardis, Emyr Lloyd-Evans, Bruno Bembi, and Forbes D. Porter .......................................................... 173

The Diverse Genetic Landscape of Neurodevelopmental Disorders
Wen F. Hu, Maria H. Chabour, and Christopher A. Wabb ...................... 195

The Genetics of Microdeletion and Microduplication Syndromes:
An Update
Corey T. Watson, Tomas Marques-Bonet, Andrew J. Sharp, and Heather C. Mefford .......................................................... 215

The Genetics of Skin Fragility
Cristina Has and Leena Bruckner-Tuderman ........................................ 245
Mendelian Disorders of the Epigenetic Machinery: Tipping the Balance of Chromatin States

Jill A. Fahrner and Hans T. Bjornsson .................................................. 269

Deep Sequencing of HIV: Clinical and Research Applications

Shiven B. Chabria, Shaili Gupta, and Michael J. Kozal ................................... 295

Noninvasive Prenatal Screening by Next-Generation Sequencing

Anthony R. Gregg, Ignatia B. Van den Veyver, Susan J. Gross, Rajeevi Madankumar, Britton D. Rink, and Mary E. Norton ............................................. 327

Personalized Pharmacogenomics: Predicting Efficacy and Adverse Drug Reactions

Munir Pirmohamed .................................................................................. 349

Therapeutics Based on Stop Codon Readthrough

Kim M. Keeling, Xiaojiao Xue, Gwen Gunn, and David M. Bedwell ......................... 371

Translating Genomics for Precision Cancer Medicine

Sameek Roychowdury and Arul M. Chinnaiyan ........................................ 395

The Evolutionary Genomics of Cichlid Fishes: Explosive Speciation and Adaptation in the Postgenomic Era

Frederico Henning and Axel Meyer ......................................................... 417

The Evolutionary Origin of the Vertebrate Body Plan: The Problem of Head Segmentation

Takayuki Onai, Naoki Irie, and Shigeru Kuratani .......................................... 443

Emerging Issues in Public Health Genomics

J. Scott Roberts, Dana C. Dolinoy, and Betb A. Tarini ...................................... 461

The Ethical, Legal, and Social Implications Program of the National Human Genome Research Institute: Reflections on an Ongoing Experiment

Jean E. McEwen, Joy T. Boyer, Kathie Y. Sun, Karen H. Rothenberg, Nicole C. Lockhart, and Mark S. Guyer ............................................................... 481

Professional Medical Education and Genomics

Laurie A. Demmer and Darrel J. Waggoner ................................................. 507

Errata

An online log of corrections to Annual Review of Genomics and Human Genetics articles may be found at http://www.annualreviews.org/errata/genom
New From Annual Reviews:

**Annual Review of Statistics and Its Application**

Volume 1 • Online January 2014 • http://statistics.annualreviews.org

Editor: **Stephen E. Fienberg**, Carnegie Mellon University

Associate Editors: **Nancy Reid**, University of Toronto

**Stephen M. Stigler**, University of Chicago

The Annual Review of Statistics and Its Application aims to inform statisticians and quantitative methodologists, as well as all scientists and users of statistics about major methodological advances and the computational tools that allow for their implementation. It will include developments in the field of statistics, including theoretical statistical underpinnings of new methodology, as well as developments in specific application domains such as biostatistics and bioinformatics, economics, machine learning, psychology, sociology, and aspects of the physical sciences.

Complimentary online access to the first volume will be available until January 2015.

**TABLE OF CONTENTS:**

- A Systematic Statistical Approach to Evaluating Evidence from Observational Studies, David Madigan, Paul E. Stang, Jesse A. Berlin, Martijn Schuemie, J. Marc Overhage, Marc A. Suchard, Bill Dumouchel, Abraham G. Hartzema, Patrick B. Ryan
- The Role of Statistics in the Discovery of a Higgs Boson, David A. van Dyk
- Brain Imaging Analysis, F. DuBois Bowman
- Statistics and Climate, Peter Guttorp
- Climate Simulators and Climate Projections, Jonathan Rougier, Michael Goldstein
- Probabilistic Forecasting, Tilmann Gneiting, Matthias Katzfuss
- Bayesian Computational Tools, Christian P. Robert
- Bayesian Computation Via Markov Chain Monte Carlo, Radu V. Craiu, Jeffrey S. Rosenthal
- Build, Compute, Critique, Repeat: Data Analysis with Latent Variable Models, David M. Blei
- Structured Regularizers for High-Dimensional Problems: Statistical and Computational Issues, Martin J. Wainwright
- High-Dimensional Statistics with a View Toward Applications in Biology, Peter Bühlmann, Markus Kalisch, Lukas Meier
- Next-Generation Statistical Genetics: Modeling, Penalization, and Optimization in High-Dimensional Data, Kenneth Lange, Jeannette C. Papp, Janet S. Sinsheimer, Eric M. Sobel
- Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond, Elena A. Erosheva, Ross L. Matsueda, Donatello Telesca
- Event History Analysis, Niels Keiding
- Statistical Evaluation of Forensic DNA Profile Evidence, Christopher D. Steele, David J. Balding
- Using League Table Rankings in Public Policy Formation: Statistical Issues, Harvey Goldstein
- Statistical Ecology, Ruth King
- Estimating the Number of Species in Microbial Diversity Studies, John Bunge, Amy Willis, Fiona Walsh
- Dynamic Treatment Regimes, Bibhas Chakraborty, Susan A. Murphy
- Statistics and Related Topics in Single-Molecule Biophysics, Hong Qian, S.C. Kou
- Statistics and Quantitative Risk Management for Banking and Insurance, Paul Embrechts, Marius Hofert

Access this and all other Annual Reviews journals via your institution at www.annualreviews.org.