The next steps in next-gen sequencing of cancer genomes

D. Neil Hayes1,2 and William Y. Kim1,2,3,4

1Lineberger Comprehensive Cancer Center, 2Department of Medicine, 3Department of Genetics, and 4Department of Urology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.

The Journal of Clinical Investigation

The necessary infrastructure to carry out genomics-driven oncology is now widely available and has resulted in the exponential increase in characterized cancer genomes. While a subset of genomic alterations is clinically actionable, the majority of somatic events remain classified as variants of unknown significance and will require functional characterization. A careful cataloging of the genomic alterations and their response to therapeutic intervention should allow the compilation of an “actionability atlas” and the creation of a genomic taxonomy stratified by tumor type and oncogenic pathway activation. The next phase of genomic medicine will therefore require talented bioinformaticians, genomic navigators, and multidisciplinary approaches to decode complex cancer genomes and guide potential therapy. Equally important will be the ethical and interpretable return of results to practicing oncologists. Finally, the integration of genomics into clinical trials is likely to speed the development of predictive biomarkers of response to targeted therapy as well as define pathways to acquired resistance.

Introduction

The dawn of the 21st century brought about a critical turning point in our understanding of the molecular basis of cancer and the potential to translate cancer genetics discoveries into therapeutic advances. The dramatic efficacy of imatinib in the treatment of Philadelphia chromosome–positive chronic myelogenous leukemia (CML) (1, 2), as well as the recognition that activating mutations in EGFR predict the therapeutic response of patients with non–small-cell lung cancer (NSCLC) to EGFR inhibitors (3, 4), were thought to be harbingers of a time when targeted therapy could be paired with genomic information to identify responsive patients and, effectively, personalize medicine. While additional success stories exist, cancer is a complex disease, and the envisioned tsunami of the genomic revolution has taken time to reach shore.

Cancer researchers have benefited tremendously from the advent of high-throughput sequencing. The genomes of thousands of tumors have now been sequenced both by individual groups and through collective efforts such as the International Cancer Genome Consortium (ICGC) and The Cancer Genome Atlas (TCGA) (1, 2, 5). At the same time, there has been an explosion in the development and clinical testing of targeted agents (6, 7). Therefore, two of the fundamental building blocks of an effective genomics-driven oncology program are now available. While conceptually the integration of genomics-driven oncology into clinical practice would seem straightforward, its implementation has proven to have real-world complexities such as the unintended discovery of germline variants of unknown significance, the sparse amounts of archival tumor tissue, the difficulty of establishing the biologic significance (or “actionability”) of somatic events, and the translation of a cohesive and accurate depiction of the genomic portrait of a tumor to treating physicians and their patients. As we strive to integrate next-generation (next-gen) sequencing technology into clinical medicine, forethought about desired outcomes, technical and ethical considerations, and a structured multidisciplinary plan will be needed to extract the most beneficial information from the large amounts of data and genomic results produced.

Actionable events

High-throughput sequencing techniques have now evolved to the point where cancer genomes can be sequenced quickly, sensitively, and accurately. Data from ongoing research initiatives such as the ICGC and the TCGA have demonstrated that the mutation rates vary greatly between tumor types (from 0.28 to 8.15 mutations per megabase in acute myeloid leukemia [AML] and lung squamous carcinoma, respectively), as does the mutational spectrum (8). Some cancers also exhibit significant intratumoral heterogeneity (9). Viewed broadly, these findings show that cancer genomes are convoluted and harbor both driver and passenger mutations.

At this juncture, the cost of high-throughput sequencing is easily within reach of academic medical centers. Furthermore, standard operating protocols (SOPs) have been developed to ensure reproducibility of results with validation in clinical laboratory improvement amendment–certified (CLIA-certified) settings. Therefore, the current bottleneck in delivering genomics-based cancer medicine is rarely data generation, but rather results interpretation, validation, and determination of actionability. If the overarching hypothesis of genomics-driven oncology is that the genomic landscape of tumors can be converted into actionability maps or an “actionability atlas,” as we have chosen to call it, then the careful cataloging of genomic variants and their resultant ability to predict prognosis or treatment response is imperative. Much like the Catalogue of Somatic Mutations in Cancer (COSMIC) contains a comprehensive catalog of over 136,000 somatic coding mutations in over 500,000 tumor samples, we predict

Conflict of interest: The authors have declared that no conflict of interest exists.

that the effectiveness of genomics-based oncology will rely on an operative database linking somatic mutations to therapeutic successes or failures (10).

So what attributes make a genomic event clinically actionable? While clearly investigator dependent, we propose some common basic tenets. First, the alteration in a gene or pathway should be clinically informative, either at the level of diagnosis (e.g., mutation of the von Hippel–Lindau tumor-suppressor gene is specific for the clear-cell histologic subtype of renal cell carcinoma; ref. 11), prognosis, or for the prediction of treatment response (e.g., EGFR inhibition in EGFR-mutant NSCLC; refs. 3, 4, 12) or resistance (e.g., KRAS mutations confer resistance to EGFR inhibition in colorectal cancer [CRC]; refs. 13-16). Second, there should be precedent (preferably in the specific tumor type) that the somatic event is a driver mutation. Finally, the genetic aberration should optimally be druggable in a specific manner. Based on these strict criteria, the number of truly actionable mutations is surprisingly low (tier 1 lists those currently defined as actionable at the University of North Carolina at Chapel Hill [UNC] using our next-gen sequencing platform UNCseq) (Tables 1 and 2). Nonetheless, a careful, prospective cataloging of genomic results and a more pathway-driven (as opposed to organ site–driven) approach to therapy will likely lead to a more expansive, yet refined, list of clinically actionable events.

To date, the vast majority of targeted therapy successes are small-molecule kinase inhibitors. While these stories tell us that such inhibitors can be used successfully to treat cancers driven by specific molecular events, whether such precision targeting can be achieved outside the realm of kinases remains to be seen. Exciting examples of inhibitors of other enzymatic reactions, such as inhibition of the H3K79 methyltransferase DOT1L in mixed-lineage leukemia (MLL), rearranged acute leukemias, or the isocitrate dehydrogenase 1 and 2 (IDH1/2) enzymes in IDH-mutant malignancies (i.e., glioblastoma and AML), have entered clinical trials and shown initial indications of effectiveness in both early-stage clinical trials and preclinical models (17–20). In addition, the enhanced sensitivity of BRCA1/2-deficient cancers to poly(ADP-ribose) polymerase (PARP) inhibition suggests that targeted therapies might be used to target tumor-suppressor gene loss as well (21).

Communication of results to clinicians: molecular tumor boards

A key step in the infrastructure of genomics-based oncology is the communication of genomic alterations to treating physicians and their patients.

At UNC, we have chosen to stratify genomic alterations based on their likelihood of clinical actionability. Molecular alterations that are considered standard of care (SOC), for example, EGFR or KRAS testing in NSCLC and CRC, respectively (22, 23), are readily determined to be both actionable and reportable. The majority of oncologists also have a good understanding of the clinical implications of the results. These alterations are considered tier 1 (Tables 1 and 2). A second, more challenging group of genomic alterations, tier 2, involves genetic alterations of potential clinical significance; for example, a previously reported mutation that is detected in a class of tumors in which it has not been reported (e.g., EGFR mutations in bladder cancer). This list can be further stratified by whether the targeted drug(s) is commercially available (tier 2A), in clinical trials (tier 2B), or not yet in clinical trials but the target is still considered druggable (tier 2C). Finally, genetic alterations that are considered to have prognostic value are reported as well (tier 2D).

The assignment of genetic alterations to the above categories is relatively straightforward for previously reported variants or copy number alterations that have been compiled into online databases such as COSMIC. More challenging is the attribution of clinical actionability to variants of unknown significance (i.e., those that have not been previously been reported) in known oncogenes or tumor-suppressor genes or to variants clearly altering the function of genes that have not been implicated previously in cancer. We have found it useful to discuss these cases and generate a framework for decision making at a so-called “molecular tumor board” (MTB) composed of clinical and translational oncologists, molecular pathologists, and bioethicists. Within these tumor boards, decisions about the actionability of genomic events and

<table>
<thead>
<tr>
<th>Table 1. Gene lists for reportable tiers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tier 1</strong></td>
</tr>
<tr>
<td>ABL</td>
</tr>
<tr>
<td>ALK</td>
</tr>
<tr>
<td>BRAF</td>
</tr>
<tr>
<td>CSF3R</td>
</tr>
<tr>
<td>EGFR</td>
</tr>
<tr>
<td>FLT3</td>
</tr>
<tr>
<td>ERBB2</td>
</tr>
<tr>
<td>JAK2</td>
</tr>
<tr>
<td>KRAS</td>
</tr>
<tr>
<td>KIT</td>
</tr>
<tr>
<td>ROS1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
respectively, serve as proof of principal that biomarker profiling can occur in real time and that adaptive trial design can validate prespecified hypotheses (32, 33). Nonetheless, additional randomized, controlled studies are needed to further confirm the true predictive value of the biomarkers studied in those trials. Finally, it will likely take some work to define the optimal next-gen biomarker cutoff points for results other than canonical mutations. In the BATTLE and I-SPY2 trials, many of the prespecified biomarker cutoffs were based on a wealth of past clinical experience (e.g., ER and HER2 immunohistochemistry) accumulated the necessity of CLIA confirmation are made by consensus. Nonetheless, in a reasonable fraction of detected variants of unknown significance, the decision to validate and report remains somewhat arbitrary (Figure 1).

Integration of next-gen sequencing into clinical trials

If discrete genetic alterations can truly predict therapeutic vulnerabilities, then next-gen sequencing of tumors should allow either the prospective identification and enrichment for potential responders or the retrospective evaluation to identify genetic features associated with response or resistance to therapy. The ability to perform continual mining of data after their initial use for treatment choice is of particular import, as novel therapies are approved for use and as we gain additional knowledge about the utility of genomic biomarkers. Additionally, an unintended consequence of next-gen sequencing of cancer genomes is likely to be a shift from clinical trials based on organ site to those that are based on alterations in oncogenic pathways (Figure 2). An example of this is the treatment of BRAF-mutant cancers with MEK or BRAF inhibitors in multiple tumor types such as melanoma, CRC, thyroid, and urothelial carcinoma (24–27). Nonetheless, it is important to proceed with caution, as genomic alterations may be exquisitely context specific. For example, recent clinical trials have shown that BRAF-mutant melanomas are highly sensitive to BRAF kinase inhibitors (vemurafenib) (28), as are BRAF-mutant papillary thyroid cancers. Thus, many oncologists were surprised when BRAF-mutant CRCs were found to be less responsive to vemurafenib (29). Careful laboratory studies have since revealed that this relative resistance to BRAF inhibition in CRC is secondary to compensatory upregulation of EGFR signaling, which in preclinical models can be cotargeted with PI3K/AKT pathway inhibitors (30, 31). These findings underscore the danger in generalizing therapeutic vulnerabilities across all tumors and demonstrate the importance of validating targets in the context of individual tumor types (Figure 2). Therefore, while it is tempting to imagine a brave new oncologic world in which tumors are categorized by driver pathway alterations rather than organ site, cancer is complex, and this simplistic view will need refinement.

While not based on a next-gen sequencing platform, several large, adaptively randomized trials using molecular biomarkers, such as BATTLE (biomarker-integrated approaches of targeted therapy for lung cancer elimination) and I-SPY2 TRIAL (investigation of serial studies to predict your therapeutic response with imaging and molecular analysis 2) for NSCLC and breast cancer, respectively, serve as proof of principal that biomarker profiling can occur in real time and that adaptive trial design can validate prespecified hypotheses (32, 33). Nonetheless, additional randomized, controlled studies are needed to further confirm the true predictive value of the biomarkers studied in those trials. Finally, it will likely take some work to define the optimal next-gen biomarker cutoff points for results other than canonical mutations. In the BATTLE and I-SPY2 trials, many of the prespecified biomarker cutoffs were based on a wealth of past clinical experience (e.g., ER and HER2 immunohistochemistry) accumulated

<table>
<thead>
<tr>
<th>Tier</th>
<th>Reported</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>Variation targeted by commercially available drug that is approved to treat this genetic variation</td>
</tr>
<tr>
<td>2A</td>
<td>Yes</td>
<td>Variation potentially treatable by commercially available targeted drug, but drug not indicated for this use</td>
</tr>
<tr>
<td>2B</td>
<td>Yes</td>
<td>Variation potentially treatable by targeted drug that is in clinical trials</td>
</tr>
<tr>
<td>2C</td>
<td>No</td>
<td>Variation potentially treatable by targeted drug, but no targeted drug has reached the clinic</td>
</tr>
<tr>
<td>2D</td>
<td>No</td>
<td>Variation of prognostic significance</td>
</tr>
</tbody>
</table>

Figure 1. Genomics-based oncology workflow and integration of an actionability atlas. The sequencing of cancer genomes requires coordinated efforts and constant modification. Patients are consented to access archival tumor tissue or fresh tumor tissue obtained from a new biopsy. Genomic profiles are generated by next-gen sequencing of tumor and normal (germ-line) tissues and are discussed at MTBs. The decision to report variants back to the patient and physician is made by consensus at the MTB in part based on a categorization of genes predetermined by an independent panel of clinical experts (the Clinical Committee for Genomic Research [CCGR]). Potential outcomes of the reporting of an actionable variant include: enrollment in a clinical trial, off-label use of a current FDA-approved agent, or continued SOC treatment. Systematic cataloging of outcomes into a proposed actionability atlas should aid future decision making.
over decades (32, 33). Nevertheless, the integration of next-gen sequencing into prospective clinical trials is a reality and, when performed rigorously, will likely yield immense value.

Unresolved issues

**Tumor heterogeneity and the development of resistance.** High-resolution tumor-sequencing studies have now defined the startlingly broad scope of intratumoral genomic heterogeneity, which likely contributes to treatment failure as well as the development of drug resistance (34). In retrospect, intratumoral heterogeneity should not have been unexpected, since pathologists have for decades remarked on the variant histologic differentiation that can coexist within what appears to be a solitary tumor. Genomic variability has clear implications for genomics-driven oncology, as genomic characterization of tumors (at least for solid tumors) is typically made from a limited tumor sampling and therefore likely represents only a small percentage of the genomic diversity within a tumor. Furthermore, genomic heterogeneity is not limited to primary tumors, but is clearly an attribute of metastases (35). Therefore, it seems likely that genomics-based oncology will contend with intratumoral heterogeneity and the limited sampling of tumor tissue as intrinsic considerations. While shared mutations should be easily annotated, the detection of private mutations will be dependent on the lesion sampled.

Clinical responses to targeted therapy can often be dramatic (e.g., EGFR-mutant lung cancers and response to erlotinib; refs. 3, 4), yet treatment with targeted therapy invariably results in the development of drug resistance, presumably through positive selection for rare subclonal cell populations within the primary tumor harboring secondary mutations, activation of downstream effectors, or activation of a bypass oncogene (36). In the current paradigm of genomics-based oncology, a single genomic snapshot of a patient’s tumor is taken, with all future actionability decisions made on that single sample; however, recent studies demonstrate the value of performing rebiopsies of tumors at the time of clinical progression on targeted therapy (37). Such studies have defined novel mutations that reduce the efficacy of kinase inhibitors (i.e., T790M mutations in NSCLC treated with EGFR kinase inhibitors) but have also defined dramatic changes in histologic subtypes (i.e., a therapy-driven switch between NSCLC and small-cell lung cancer) (37). Therefore, for genomics-based oncology to fully achieve its promise, a genomic snapshot of a patient’s tumor requires repeated updates.

**Circulating tumor DNA.** While the next-gen sequencing of cancer genomes has traditionally been performed on banked tissue (either fresh-frozen or formalin-fixed, paraffin-embedded [FFPE]), it has recently been applied to detect circulating tumor DNA (ctDNA) in the peripheral blood of cancer patients. Beyond the obvious ease of sample acquisition (pneumotomy rather than obtaining archival tumor tissue or a new biopsy), the next-gen sequencing of ctDNA is likely to serve a number of broader applications. For example, recent work suggests that it can be used to noninvasively monitor tumor burden (38–40) and, perhaps most excitingly, that it can be used to detect mechanisms of secondary resistance to targeted therapy (e.g., activation of the RAS/MEK/ERK pathway in EGFR inhibitor–treated metastatic CRC patients) (40, 41). While not prohibitive, a current limitation to most of the reported ctDNA techniques is the need for relatively deep and broad sequencing of putative ctDNA.

**Tumor microenvironment.** Harnessing the immune system for cancer therapy has undergone a recent renaissance. Impressive therapeutic outcomes have been observed across several immunotherapy platforms, including adoptive cellular therapy (e.g., chimeric antigen receptors [CARs] in hematologic malignancies) (42–46), immune checkpoint blockade (e.g., anti–PD1/PD-L1 antibodies in numerous solid tumors) (47, 48), and various tumor vaccination strategies (49–51). This has led to an intense search for predictive biomarkers of immune responsiveness. While intratumoral expression of PD-L1 has been proposed to have a high predictive value for the efficacy of anti–PD-L1 therapy (48), additional biomarkers are under active investigation.

Unfortunately, next-gen sequencing of bulk cancers does not capture important information on the immune infiltrate, such as the level of immune cell infiltration or the cellular milieu of immune cells present, which may be important biomarkers of not only response to immunotherapy but also to cytotoxic chemotherapy (52, 53). Nonetheless, recent studies confirm the utility of exome sequencing for discovery of tumor neoantigens, although it should be noted that the majority of neoantigens identified are in genes that are absent in current targeted exon–sequencing panels (54, 55). Overall, it seems likely that the integration of other genomic platforms such as RNA sequencing and platforms such as proteomics and polychromatic flow cytometry will allow for a more comprehensive “immune portrait” of tumors, including the molecular assessment of lymphocyte clonality and diversity (56) as well as the assessment of other key microenvironmental components such as tumor-associated fibroblasts and endothelial cells, which may be key predictors of response to both immune and antiangiogenic therapies (refs. 57, 58, and Figure 3).
perceived level of genomic confidence, with approximately 22% of surveyed physicians reporting that they were “not very confident” or “not confident at all” in their knowledge of genomics. These findings reiterate that vehicles for consensus such as multidisciplinary MTBs and evidence-based guidelines are key components to a successful genomics-driven oncology program. It will be of interest to assess similar perceptions in community-based oncology practices where, at least to our knowledge, genomics tumor boards are not part of routine practice.

The next steps
At many academic medical centers, the building blocks for genomics-driven oncology are now in place. In reality, this has been relatively easy, as capital expenditure to purchase next-gen sequencing machines and reagents is merely a line item on a cancer center’s annual budget. It seems that the coming years will be the most difficult, for to truly benefit, patients will depend on the use of human capital for data processing, the timely interpretation and reporting of actionable results, and, perhaps most importantly, the effective communication about the implications of these results to patients. Indeed, we can sequence as many cancer genomes as we like, but without talented bioinformaticians to decipher the data and so-called “genomic navigators” to guide us, the clinical actionability of unannotated variants remains questionable.

Genomics- or “omics”-driven oncology? The value of bioinformaticians is likely to increase exponentially, as genomics-driven oncology expands to encompass platforms outside of DNA-based mutations and gene copy number alterations. The future of genomics, or perhaps more appropriately, “omics”-
The Journal of Clinical Investigation

REVIEW

driven oncology, likely holds a multiphase approach that will allow comprehensive characterization of a tumor on multiple levels; the resulting genomic profile would then be integrated with other “omic” data to predict the functional consequences of genomic alterations on oncogenic pathways. Thus, our current one- or, at best, two-dimensional map of a cancer genome will soon transform into a true genomic landscape replete with multidimensional annotation.

Genomic confidence and genomic navigators. The survey by Gray and colleagues suggests that even in tertiary care centers, many physicians do not feel confident about their knowledge of genomics (61). While not directly surveyed, this low genomic confidence would suggest a relative reticence toward both ordering next-gen sequencing of cancer genomes as well as discussing their implications directly with patients. Much like medical geneticists, who have specialized in the evaluation, testing, and counseling of patients with a suspected genetic disorder, it would seem that the field of genomics-driven oncology might also usher in a new cadre of genomic navigators. The role of these genomic navigators might be to shepherd patients through the process of testing as well as the potential complications of testing (e.g., detection of a suspected germline variant). Ultimately, however, discussion of the relative actionability of a genomic variant and whether to forgo SOC treatment for a personalized therapy will require significant input from the treating oncologist.

Development of an actionability atlas. While the future of genomics-driven oncology holds much promise, the human genome is vast, and the majority of cancer genomes are likely to harbor somatic variants of unknown significance. To continue to move the field forward rationally, it is imperative that we develop a central repository such as the above-proposed actionability atlas to catalog novel variants, their purported actionability, and, if attempted, their response to therapy. This should be easy in the context of clinical trials, but will take a concerted effort as more case reports or so-called “N-of-1” studies are undertaken. Despite the catchy name, these N-of-1 studies are to some extent case reports and remain anecdotal. Moreover, while many N-of-1 studies are supported by elegant in vitro work backing the genomic hypotheses, most clearly require prospective validation or validation in large, retrospective datasets before widespread adoption in clinical practice. Nonetheless, only through the careful cartography of cancer genomic landscapes and their responses to therapy can we develop an actionability atlas and genetic taxonomy of cancer based on both tumor type and genetic makeup.

Acknowledgments
We acknowledge Jon Serody and Ned Sharpless for useful discussions and the entire UNCSeq team (in particular, Michele Hayward and Ashley Salazar) for their dedication and work. This work was supported by NIH grants R01 CA142794 (to W.Y. Kim), U24 CA143848 (to D.N. Hayes), U24 CA143848-02S1 (to D.N. Hayes), and U10 CA181009-01 (to D.N. Hayes); the University Cancer Research Fund (to W.Y. Kim and D.N. Hayes); AACR Kure It (to W.Y. Kim); and the Bladder Cancer Advocacy Network (BCAN) Innovation Award (to W.Y. Kim). W.Y. Kim is a Damon Runyon Merck Clinical Investigator.

Address correspondence to: D. Neil Hayes or William Y. Kim, Lineberger Comprehensive Cancer Center, University of North Carolina, CB# 7295, Chapel Hill, North Carolina 27599-7295, USA. Phone: 919.966.3786; E-mail: hayes@med.unc.edu (D.N. Hayes). Phone: 919.966.4765; E-mail: wykim@med.unc.edu (W.Y. Kim).


