ANNUAL REVIEWS

Annual Review of Pathology: Mechanisms of Disease

Enabling Precision Oncology Through Precision Diagnostics

Noah A. Brown[1] and Kojo S.J. Elenitoba-Johnson[2]

[1]Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA; email: noahbr@med.umich.edu

[2]Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; email: Kojo.Elenitoba-Johnson@pennmedicine.upenn.edu

Annu. Rev. Pathol. Mech. Dis. 2020. 15:97–121

The *Annual Review of Pathology: Mechanisms of Disease* is online at pathol.annualreviews.org

https://doi.org/10.1146/annurev-pathmechdis-012418-012735

Keywords

precision medicine, precision oncology, personalized medicine, next-generation sequencing, cancer, molecular pathology

Abstract

Genomic testing enables clinical management to be tailored to individual cancer patients based on the molecular alterations present within cancer cells. Genomic sequencing results can be applied to detect and classify cancer, predict prognosis, and target therapies. Next-generation sequencing has revolutionized the field of cancer genomics by enabling rapid and cost-effective sequencing of large portions of the genome. With this technology, precision oncology is quickly becoming a realized paradigm for managing the treatment of cancer patients. However, many challenges must be overcome to efficiently implement the transition of next-generation sequencing from research applications to routine clinical practice, including using specimens commonly available in the clinical setting; determining how to process, store, and manage large amounts of sequencing data; determining how to interpret and prioritize molecular findings; and coordinating health professionals from multiple disciplines.

INTRODUCTION

Precision medicine is a health-care model in which patient management is customized based on patient-specific analytics that include genomic data (1).

Precision medicine has been defined as "an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environments and lifestyle for each person" (2, p. 443). The prospects of achieving this goal, particularly the approach that takes individual variability into account, have been significantly enhanced by the recent advent of high-dimensional multiparametric approaches that facilitate interrogation of, for example, entire genomes (genomics), transcriptomes (transcriptomics), proteomes (proteomics), and metabolomes (metabolomics). The contemporaneous development of computational algorithms that facilitate analysis of the large data sets generated from these approaches and the ubiquitous availability of digital and mobile technologies further enhance the ability to electronically archive and disseminate the data.

In particular, nucleic acid–based high-throughput technologies, especially next-generation sequencing (NGS), offer great promise for radically impacting the routine implementation of precision medicine. Among different spheres, oncology is an immediately apparent area for application, given the high aggregate frequency of cancers, the fact that cancer remains a leading cause of mortality, the increasing incidence linked to population aging, and the availability of targeted treatments that aim to treat cancers by inactivating the driver alterations that are specific, critical vulnerabilities for the cancer cells.

In cancer, the genomic testing of cancer tissue can be employed to render a diagnosis, inform prognosis, predict the efficacy of targeted therapies, and monitor disease burden longitudinally. Genetic testing of germline DNA can also be used in oncology to determine an individual's risk for various cancer types. The increasing pace of drug discovery and clinical trials in specific cancers or orphan diseases is contributing to the implementation of high-throughput approaches for identifying robust biomarkers for cancer diagnostics. Accordingly, precision diagnostics driven by NGS is poised to radically alter the routine practice of cancer diagnostics in the clinical laboratory. In this review, we discuss the infrastructural, technical, and operational considerations for developing, implementing, and deploying genomic precision diagnostics within a clinical laboratory environment to support state-of-the-art clinical oncological care and research trials.

ASSAY CONFIGURATIONS FOR PRECISION ONCOLOGY

Single-Gene Assays Versus Next-Generation Sequencing

Single-gene assays have been the mainstay of molecular diagnostics for decades. Various methods are employed depending on the type of molecular alteration, the type of specimen to be tested, the burden of disease in tested specimens, and how the result is to be used. Common methods include the amplification of DNA or RNA using polymerase chain reaction (PCR) followed by sequencing (e.g., Sanger, pyrosequencing), fragment analysis using electrophoresis, restriction enzyme digestion, or hybridization with sequence-specific probes. Fluorescence in situ hybridization is also frequently employed in molecular oncology.

These single-gene tests often have significant advantages over large-scale genomic sequencing, including faster turnaround time; lower cost; and lower complexity in developing, running, and interpreting the test. Since each gene is evaluated separately, these assays can also be customized to meet the needs of the clinical context in which they are employed. For instance, single-gene assays can be designed to be extremely sensitive and highly quantitative, and once they are standardized across laboratories, can be used to monitor the response of oncology patients; an example is testing for *BCR-ABL1* in patients with in chronic myeloid leukemia. Single-gene assays can also

be carefully designed to overcome challenges associated with problematic regions of the genome (e.g., the high G:C content associated with *CEBPA*) and variant types that may be challenging to detect by NGS (e.g., the large internal tandem duplications within *FLT3*) (3, 4).

The use of single-gene tests continues to be a useful approach for clinical contexts in which information about a limited number of genes is clinically relevant. However, as more and more genes become relevant for a given clinical context, NGS becomes an increasingly attractive approach. The resources associated with designing, optimizing, validating, and running several single-gene tests become prohibitive with more than a handful of genes. Molecular testing of advanced-stage non–small cell lung cancer (NSCLC) provides a good example of the quickly increasing need for the molecular profiling of several genes. Initially, testing for *EGFR* exon 19 deletions and the L858R mutation was the only molecular information relevant for targeted tyrosine kinase inhibitor (TKI) therapy (5–7). However, within a few years, effective targeted therapies approved by the US Food and Drug Administration (FDA) arose and are effective in treating lung cancers with other *EGFR* mutations (8, 9), *ALK* and *ROS1* rearrangements (10–17), and *BRAF* mutations (18, 19). Still other molecular alterations have associated targeted therapies with promising early clinical data—including *MET* exon 14 skipping mutations (20, 21), high-level *MET* amplification (22), *RET* rearrangements (23–25), and *ERBB2* (*HER2*) mutations (26)—leading to new treatment recommendations from the National Comprehensive Cancer Network (NCCN) (27).

NGS also offers some technical advantages over single-gene methods. For instance, NGS is more sensitive than many single-gene methodologies, such as Sanger sequencing. Other single-gene methods, such as allele-specific PCR, offer better sensitivity, but can be employed only to detect a limited number of specific mutations. The simultaneous sequencing of many regions in parallel with NGS also enables sequencing of many genes even when DNA or RNA is limited.

Gene Panels Versus Unbiased Genomic and Transcriptomic Analyses

NGS approaches may include whole-genome sequencing (WGS), whole-exome sequencing (WES), whole-transcriptome sequencing (RNA-seq), and target gene panels. The first human genome to be sequenced was a massive multicenter project that cost billions of dollars and took decades (28). WGS can now be accomplished by a single laboratory in a matter of days for as little as $1,000. The potential advantage of WGS is its ability to detect essentially any genomic alteration, including novel or rare alterations. However, this approach is expensive and cannot practically be done at sufficient depth to reliably detect low-level alterations in samples containing a mixture of tumor cells and normal cells. In addition, this approach identifies a huge number of sequence variants. In NSCLC, for instance, whole-genome studies have demonstrated a median of 888 and 15,659 mutations in NSCLC samples from, respectively, nonsmokers and smokers (29). The vast majority of these variants are not clinically actionable, and making this distinction when faced with this much data can be overwhelming. In the absence of comparisons between tumor and normal DNA, distinguishing somatic mutations from germline polymorphisms adds even further complexity.

WES is an unbiased approach that has also found utility in some laboratories as a platform for the interrogation of cancers. WES limits sequencing to the ~1.5% of the genome that lies in the exons of genes. Nevertheless, this approach also generates a large number of potential variants, the vast majority of which currently do not have annotated clinical implications. Exome sequencing will also fail to detect pathogenic variants, such as structural rearrangements with intronic break points. While input and DNA quality requirements are lower than those for WGS, the drawbacks of this approach include the fact that the depth of sequencing obtained with WES is much lower than that obtained by targeted panels. Thus, in cases with a low tumor burden,

the allele frequencies obtained with WES may result in missed calls for clinically actionable mutations. Finally, WES is more expensive than targeted assays.

RNA-seq involves sequencing RNA transcripts, which may include mRNA or small RNAs. RNA-seq data provide a snapshot of gene expression patterns and can be an alternative to data obtained from microarray-based gene expression profiling, which has limited dynamic range and less specificity due to cross-hybridization (30). Also, RNA-seq enables the detection of gene fusions, including those that are novel. Finally, RNA-seq can provide information about noncoding RNAs (ncRNAs), including long noncoding RNAs, and microRNAs. Abnormal ncRNA function has been implicated in the pathogenesis of cancer and non-neoplastic diseases (31). While the role of ncRNA in cancer remains far from fully characterized, some studies have suggested that profiling ncRNA may have diagnostic or prognostic applications (32–34). Like WGS and WES, RNA-seq generates a large amount of complex data and is not routinely performed by most clinical laboratories.

Most laboratories performing NGS for clinical purposes use targeted sequencing panels. In addition to having lower sequencing costs and faster turnaround times, this approach provides much higher coverage of genomic regions of interest, thereby enabling the detection of lower-level molecular alterations: an essential capability for somatic testing of clinical samples that contain a mixture of neoplastic and non-neoplastic cells. The dramatically decreased complexity of the data generated by gene panels markedly reduces the time and effort associated with interpretation. By focusing on those molecular alterations that are clinically actionable, this approach mitigates the risk of actionable data being overshadowed by a deluge of data without clinical application. Finally, these types of panels can be designed with capabilities to meet various clinical needs. For example, solid tumor NGS panels can be designed to be run with very small amounts of nucleic acid, thus enabling the testing of small biopsy and aspirate samples (35). Targeted panels can also be designed with sufficient sensitivity to detect molecular alterations that are present at extremely low levels, for example, in cell-free DNA (cfDNA) (36). The principal drawback of targeted gene panels is the need for a priori knowledge of the molecular alterations to be detected because novel alterations may be missed. In most cases, novel variants are not expected to have clear clinical implications, but there are exceptions, such as *ALK* gene fusions in NSCLC even in cases in which the fusion partner is novel (27). Targeted panels may also require updating as new, clinically actionable molecular analytes emerge.

APPLICATIONS OF MOLECULAR ONCOLOGY TESTING

Diagnosis

For a number of neoplastic diagnoses, molecular findings either are a requirement for rendering the diagnosis (e.g., *BCR-ABL1* in chronic myeloid leukemia) or molecular data can assist oncologists or pathologists, or both, in rendering a diagnosis in cases in which the clinical presentation and pathologic features are equivocal (e.g., the *KIAA1549-BRAF* fusion in pilocytic astrocytoma). However, in many instances, molecular data must be used cautiously and correlated with other clinicopathologic data. For example, while many mutations associated with myelodysplastic syndromes are readily detectable in specimens from peripheral blood or bone marrow, these mutations are also found in otherwise healthy individuals with clonal hematopoiesis of indeterminate potential (37–39).

Prognosis

Molecular alterations found in neoplastic cells frequently correlate with various prognostic indices. For instance, *IDH1* mutations are associated with more favorable outcomes in glioma

patients (40, 41). Not only can these prognostic associations be applied to counseling patients, but the mutations may also carry diagnostic or therapeutic implications, or both. For instance, *IDH1* and *IDH2* mutation status is now incorporated into the World Health Organization's classification of central nervous system tumors (42). The prognostic implications of mutations in *NPM1*, *FLT3*, *CEBPA*, *KIT*, *TP53*, *RUNX1*, and *ASXL1* are used to stratify patients with acute myeloid leukemia (AML) who are under the age of 60 into risk categories that, in turn, determine the appropriate postremission therapy according to the NCCN's guidelines (43).

Therapy

Molecular profiling of neoplastic cells is frequently used to guide therapeutic decisions. This profiling often identifies an oncogenic alteration that can be specifically targeted by small molecule inhibitors or monoclonal antibodies. For example, BRAF inhibitors and MEK inhibitors (acting downstream of BRAF) are efficacious in treating melanomas with a *BRAF* V600E mutation (44, 45), and several BRAF and MEK inhibitors are currently approved by the FDA for this application. Other molecular alterations may predict resistance to a particular therapy. In colorectal cancer, *KRAS* and *NRAS* mutations are associated with primary resistance to anti-EGFR antibody therapy (46, 47).

Molecular findings also frequently guide the use of FDA-approved therapies for applications outside of the approved indications (off-label indications). For example, imatinib is a TKI that is FDA approved for the treatment of several neoplasms with TKI-sensitive mutations in *KIT*, including gastrointestinal stromal tumor, chronic myeloid leukemia, and systemic mastocytosis. While imatinib has not been FDA approved for treating melanoma, several studies have shown that patients with melanoma harboring specific mutations in *KIT* respond to imatinib (48–50). This off-label therapy is recommended in the current NCCN guidelines for metastatic melanoma (51).

Targeted therapies may also be employed in an iterative fashion based on both those alterations initially identified as well as those acquired during treatment. In NSCLC, an *EGFR* exon 19 deletion or L858R mutation is associated with sensitivity to first- and second-generation EGFR inhibitors, for example, erlotinib, gefitinib, and afatinib (5, 6). However, the acquisition of a secondary *EGFR* T790M mutation is a common mechanism of acquired resistance to these therapies. The emergence of third-generation EGFR inhibitors, such as osimertinib, has provided oncologists with an effective therapy to treat NSCLC with acquired resistance to EGFR inhibitors (52).

Another class of treatment for which molecular results may be informative is immune-enhancing therapies. These include checkpoint inhibitors that block inhibitory immune signaling, such as the interaction between programmed cell death 1 protein (PD-1) and its ligand PD-L1 [e.g., pembrolizumab (Merck), nivolumab (Bristol-Myers Squibb)] and cytotoxic T lymphocyte–associated protein 4 [CTLA-4; e.g., ipilimumab (Bristol-Myers Squibb)]. Several studies suggest that tumors with microsatellite instability (53) or a high tumor mutation burden (54–57), or both, are strong predictors of response to these agents, and both microsatellite instability status and tumor mutation burden can be identified using NGS data sets.

The response to therapies that inhibit poly (ADP-ribose) polymerase (PARP) is also associated with molecular findings. Several studies have shown that some of these drugs (e.g., olaparib, rucaparib) are more efficacious in cancers with homologous recombination deficiency due to mutations in genes such as *BRCA1* and *BCRA2* (58–60). Other drugs (e.g., niraparib) have proven efficacious regardless of homologous recombination deficiency or *BRCA1* or *BRCA2* status (61).

Finally, molecular results are increasingly being used to screen patients for molecular alterations that may qualify them for clinical trials. These may include clinical trials for particular cancer types with specific molecular alterations or trials that enroll patients with advanced

cancer who have particular targetable molecular alterations, regardless of cancer type. These include the Molecular Analysis for Therapy Choice (MATCH; **http://www.cancer.gov/about-cancer/treatment/clinical-trials/nci-supported/nci-match**) trial conducted by the National Cancer Institute (NCI) and the Targeted Agent and Profiling Utilization Registry (TAPUR) trial conducted by the American Society of Clinical Oncology (ASCO).

Precision Diagnostics Enabling Clinical Trials in Oncology

The implementation of NGS-based genomic diagnostics that identify tumor-associated variants and structural alterations activating specific pathways in cancers, along with the rapid development of targeted novel therapies, have given rise to paradigm shifts in clinical trial design in oncology. Observations from numerous studies in clinical and research environments show that tumors apparently arising from the same organ can be heterogeneous in their pathogenesis and genetic signatures. Indeed, within the same tumor there may be intratumoral genetic heterogeneity. Also, tumors with different organ and tissue origins may demonstrate identical genetic alterations and molecular drivers.

In the category of clinical trials known as umbrella trials, patients with a particular tumor type are subjected to genomic sequencing and assigned to multiple therapies that target the specific driver abnormalities identified in each individual tumor. This paradigm maximizes the opportunity for a patient to receive and benefit from precision therapies based on the identification by NGS of oncogenic drivers. Examples of umbrella trials include the Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis 2 (I-SPY2) (62) and Lung Cancer Master Protocol (Lung-MAP) studies.

Another approach in which therapies are driven by precision diagnostics is the so-called basket trial. In this approach, multiple tumors or tumors with different histopathologic classifications are treated with one specific drug targeting a common group of oncogenic driver alterations. This approach may be complementary to the umbrella trial approach and seeks to determine whether a targeted therapy offers a benefit across diverse tumor types that share a targetable molecular aberration. This tissue-agnostic paradigm is exemplified by the NCI's MATCH trial, the Pediatric MATCH trial (63), and a multicenter basket trial of the TRK inhibitor larotrectinib (64). Precision diagnostics is also playing an increasingly critical role in adaptive trials wherein data accumulating during the trial are utilized to modify the course of the clinical trial (62, 65).

Disease Monitoring

Molecular testing is also frequently used for monitoring cancer patients after treatment begins. The most established example is the monitoring of patients with chronic myeloid leukemia with quantitative *BCR-ABL1* testing after initiating TKI therapy (66, 67). Patients are evaluated every 3 months (67), and a major molecular response (defined as a 3-log reduction in *BCR-ABL1* values) is strongly associated with a more favorable outcome (68). TKI resistance is suspected if there is an incomplete initial response, a 1-log increase in *BCR-ABL1* while on treatment, or hematologic evidence of relapse (69). For these patients, *ABL1* kinase domain sequencing is used to identify resistance mutations and guide TKI therapy decisions (70, 71).

Several studies have also highlighted the prognostic significance of measurable residual disease (MRD; previously termed minimal residual disease) in AML and acute lymphoblastic leukemia (ALL). MRD can be evaluated using flow cytometry or molecular methods. In *BCR-ABL1*-negative ALL, molecular MRD monitoring has traditionally been performed using allele-specific oligonucleotide PCR of rearranged immunoglobulin and T cell receptor genes; however, this approach is time consuming and labor intensive (72–74). The use of NGS for these genes has

emerged as a more practical alternative (75–80). In the past, molecular MRD detection in AML has been performed using reverse transcription PCR assays for oncogenic gene fusions (i.e., *PML-RARA*, *RUNX1-RUNX1T1*, and *CBFB-MYH11*) (81–84). Studies have also demonstrated the utility of monitoring *NPM1* mutation status in *NPM1*-mutated AML (85–88). Recently, studies have investigated the utility of multigene NGS panels for detecting MRD based on the presence or absence of mutations in multiple different genes (89, 90). However, this approach requires defining those alterations that are specific to leukemic cells (rather than in nonleukemic hematopoietic clones) and present in all leukemic cells (rather than in subclonal populations) (89, 90).

Peripheral blood from patients with solid tumor cancers is also increasingly being used as a noninvasive sample for monitoring disease burden. The evaluation of neoplastic DNA can be accomplished through the collection of circulating tumor cells or cfDNA in plasma. Patients on treatment (particularly targeted therapies) can be monitored both for recurrence (91–93) and for acquired resistance mutations (94, 95), which may enable therapy to be adjusted (e.g., osimertinib may be used for patients with NSCLC who acquire an *EGFR* T790M mutation). Not only does the use of cfDNA offer a noninvasive alternative for detecting resistance mutations but also this method may mitigate sampling challenges associated with tumor heterogeneity (96).

Cancer Predisposition

Some patients diagnosed with cancer have a genetic predisposition. Identifying these patients can have important implications for treatment, follow-up, screening for other cancers, and genetic counseling. Lynch Syndrome (also known as hereditary nonpolyposis colorectal cancer) is an example of a cancer predisposition syndrome for which cancer patients are frequently screened. This syndrome results from a deleterious germline mutation within a gene (*MLH1*, *MSH2*, *MSH6*, or *PMS2*) encoding a mismatch repair protein, and it is associated with an increased risk of several cancers, particularly colorectal and endometrial carcinomas (97). Screening for Lynch syndrome is now recommended for all patients diagnosed with colorectal cancer (98, 99) or endometrial cancer (100). It is recommended that patients diagnosed with Lynch syndrome receive more frequent follow-up, more rigorous screening for other cancers, and genetic counseling (99). In addition, cancers in these patients display microsatellite instability, which is predictive of the response to immune checkpoint inhibitors (53).

Another common example of screening for cancer predisposition is testing for mutations in *BRCA1* and *BRCA2*. Mutations in these genes are associated with a much higher risk of several cancers, including breast, ovarian, and prostate. Again, detecting these mutations has implications for follow-up, screening for other cancers, and genetic counseling. As stated above, both germline and somatic mutations in these genes are also associated with the response to PARP inhibitors (58–60).

NEXT-GENERATION SEQUENCING TECHNOLOGIES

While the performance of NGS varies depending on the particular assay, the basic steps include isolation or enrichment of neoplastic cells, extraction of nucleic acid, preparation of a sequencing library, sequencing, and data processing.

Isolation or Enrichment of Neoplastic Nucleic Acid

The detection of molecular alterations in neoplastic cells is contingent on having a sufficient proportion of nucleic acid from neoplastic cells within the extracted specimen to meet or exceed the limit of detection of the test being performed. Because molecular testing is frequently employed

for patients with advanced-stage cancer for whom resection is not indicated, many specimens consists only of small biopsies and aspirates. These limited specimens are often challenging substrates that have only a limited percentage of neoplastic cells (thus, requiring enrichment) or a limited absolute number of neoplastic cells, resulting in low nucleic acid yield, or both.

Various methods can be employed to enrich for neoplastic cells. For formalin-fixed, paraffin-embedded (FFPE) tissues, manual macrodissection is frequently performed and involves circling the relevant area for extraction on a hematoxylin and eosin–stained slide. The corresponding region from subsequent stained or unstained slides is then scraped into a container for extraction (101). This method has low precision and is generally limited to specimens in which a fairly large region contains sufficient neoplastic cells. Microdissection with the aid of a stereomicroscope provides additional precision, enabling the use of specimens in which small foci must be extracted (e.g., a microscopic metastasis within a lymph node) (102). Digitally guided and laser-capture microdissection provide even greater precision, enabling the extraction of even single cells (103, 104).

Traditionally, the use of blood for molecular testing has been limited to hematologic neoplasms in which neoplastic cells circulate in the blood in relatively high numbers. However, blood is increasingly being used for testing in cases of solid tumor cancers as it offers a noninvasive method of obtaining and characterizing tumor DNA and for monitoring disease burden. The collection of circulating tumor cells relies on capturing these cells using specific cell surface markers. In addition to providing a source of neoplastic nucleic acid, the enumeration of these cells can be employed for disease monitoring (105–108). Tumor DNA can also be found within the plasma of cancer patients as cfDNA. While the use of cfDNA is often complicated by the exceedingly low proportion of tumor DNA, this limitation can be overcome with ultrasensitive molecular assays. These assays require an extremely high sequencing depth in order to achieve the needed sensitivity and may employ error correction using molecular barcodes to avoid false-positive results from PCR errors (109). These assays also require prompt removal of white blood cells to prevent the dilution of cfDNA with DNA released from lysed white blood cells. These assays may be useful when surgical pathology or cytology specimens contain insufficient neoplastic cells for molecular analysis (e.g., FFPE specimens that are exhausted). Several studies of various cancer types, including breast (110–113), colorectal (91, 110, 114) and NSCLC(109, 110, 115), have demonstrated high concordance between paired tumor and cfDNA specimens.

Both DNA and RNA are frequently used in molecular testing. DNA is more stable than RNA, which can easily be degraded by ribonucleases; as a result, DNA is most frequently used to detect single nucleotide variants (SNVs), insertion/deletion variants (indels), and copy number variants. RNA is a useful substrate for measuring gene expression as well as gene fusions. While DNA may also be used to detect gene fusions, this can often be done more efficiently with RNA due to the removal of introns by splicing.

Library Preparation

Following extraction, sequencing libraries must be prepared. In order to perform targeted sequencing, the DNA or RNA regions of interest must first be enriched using either hybrid capture or amplification. Hybrid capture starts by fragmenting nucleic acid to create a library of fragments of a size that is optimal for sequencing. Fragmentation can be accomplished by various means: physical (e.g., acoustic shearing and sonication), enzymatic (e.g., using a cocktail of nonspecific endonucleases or transposase tagmentation reactions), or chemical (116). The fragments of interest are then captured by hybridizing to complementary oligonucleotides, either in solution or on an array. Solution-based hybridization is more commonly employed in molecular oncology due to its lower cost and lower requirements for DNA and RNA input.

PCR-based library preparation involves a pool of primers designed to specifically amplify the regions of interest. In both capture- and amplification-based methods, adapters and barcodes are ligated to library fragments in order to, respectively, perform subsequent sequencing and enable sample identification of each fragment after pooling.

Each library preparation method has its pros and cons. Amplification enables the evaluation of samples with extremely small quantities of nucleic acid. For instance, the Oncomine Focus Assay (ThermoFisher Scientific) can be run with 10 ng of DNA or RNA, or less (117). Amplification-based library preparations are also generally less expensive and less time consuming. However, variants within a primer annealing site or large deletions that extend into a primer annealing site may cause false-negative results (118). The detection of gene fusions using amplification has the advantage of using far fewer sequence reads since reads are generated only for the specific fusion(s) present in sample, if any. However, novel fusion partners or fusion transcripts will not be detected with traditional reverse transcription PCR since primers must be designed for each gene fusion. Anchored multiplex PCR is a library preparation method that is compatible with low nucleic acid input from FFPE and enables the detection of gene rearrangements without prior knowledge of the fusion partner (119). Capture-based library preparation is also capable of detecting novel gene fusions since detection relies on capture probes annealing to one of the two genes in fragments that contain a sequence corresponding to both genes.

Sequencing

The two principal sequencing chemistries underlying the function of NGS platforms in clinical use today are sequencing by synthesis (Illumina platform) and pH-sensing technologies (Ion Torrent, ThermoFisher Scientific). Both methods require the generation of many immobilized copies of each template molecule to be sequenced using solid-phase amplification. The Illumina platform relies on immobilizing sequencing templates in a flow cell followed by solid-phase bridge amplification to create clusters of identical copies. Ion Torrent employs emulsion PCR to generate identical copies of each template molecule immobilized on separate beads. Illumina's sequencing by synthesis is accomplished using fluorescent-labeled, reversible terminator deoxynucleoside triphosphates, while Ion Torrent relies on detecting voltage changes as each base is added. Both technologies provide accurate sequencing results for targeted gene panels and large-scale genomic or transcriptomic analyses, although Ion Torrent tends to have a higher error rate for calling insertions and deletions in regions containing homopolymers (120).

Bioinformatics

Translating raw data from sequencing instruments into informative results that can be applied to patient care requires several processing steps (**Figure 1**). The large amounts of data generated by NGS as well as the complexity of base calling, alignment and assembly, variant calling, quality control, and annotation pose significant bioinformatics challenges. These challenges have been addressed through the creation of a number of commercial, open source, and laboratory-developed computational tools that are used to assemble a bioinformatics pipeline for routine handling of clinical NGS data (**Figure 2**).

Base calling (the conversion of raw signal files into FASTQ text files) is often performed using proprietary, vendor-specific data processing programs that are integrated into the sequencing pipeline. However, all sequencing chemistries are prone to low-level errors. Ion Torrent sequencing is especially prone to indel errors within homopolymers (120, 121); although this problem has been significantly mitigated in recent years with refined base calling algorithms and sequencing chemistries. In the Illumina platform, substitution errors are more frequent than indel errors, and

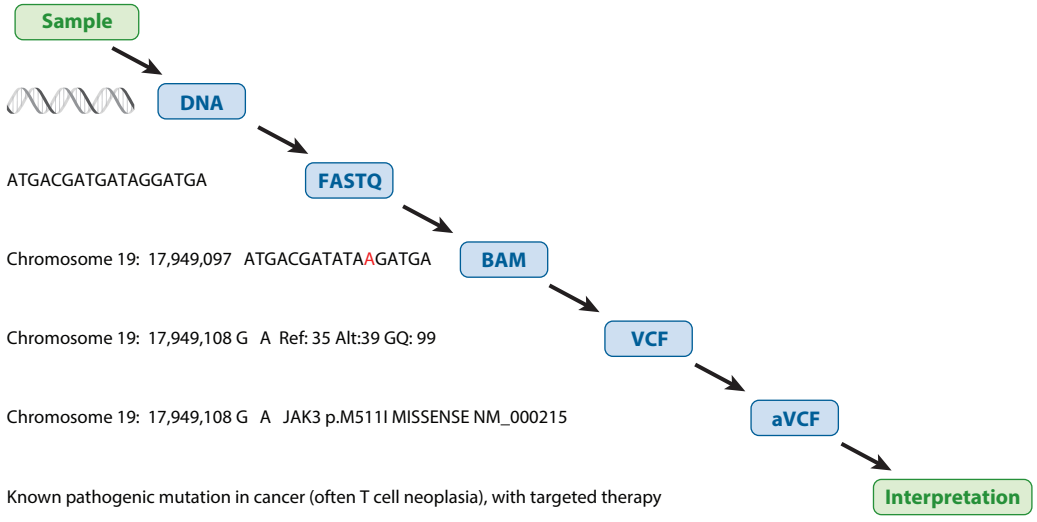

Figure 1

Workflow for analyzing genome sequencing data, from sample processing to result interpretation. Interpretation may include comment about targeted treatment opportunities. Genotype–phenotype correlations are known for several conditions. Increasingly, these data are available in public databases. FASTQ is a text-based format that includes biological sequence information. Abbreviations: aVCF, annotated variant call format; BAM, binary alignment map; VCF, variant call format.

particular substitutions (A>C) are more common (122). Error rates are greater near the ends of reads due to the accumulation of synthesis asynchrony. Base calling algorithms also generate error probabilities, which are often converted to Phred quality scores (Q scores) (123). Sequence artifacts can also be related to suboptimal template DNA extracted from poor quality specimens. In FFPE tissue, sequence artifacts are common, particularly fragmentation and deamination artifacts (C>T or G>A on the opposite strand) (124).

Base calling is followed by the mapping of reads from a FASTQ file to the human genome, resulting in a binary alignment map (or BAM) file (**Figure 3**). Incorrect read mapping or the inability to map sequence reads can also lead to the identification of artifactual variants or the failure to detect true variants. The detection of large indel variants, translocations, and variants in regions homologous to other regions in the genome, can all be affected by mapping issues. Nevertheless, large structural and copy number variants can be accurately detected and visualized using a number of approaches, including Circos plots (125) (**Figure 4*a***). Mapping transcriptome data is associated with additional challenges related to reads that span splice junctions or fusion junctions (**Figure 4*b***). Several algorithms have been created specifically to correctly map splice site–crossing reads and to correctly identify gene fusions (126–130).

Variant calling is the next bioinformatics step. SNVs are perhaps most readily detected by current variant calling algorithms (**Figure 4*c***). However, other errors, such as amplification errors and deamination, can create false-positive SNV calls. While these artifacts generally produce a low variant frequency, they may be more prevalent in poor quality specimens, presumably due to a lack of stoichiometric competition from good quality template molecules. Minimum variant requirements (e.g., variant allele frequency, quality score, reads, strand bias) as well as minimum specimen requirements (e.g., sample depth, depth per region, read length, percentage of base calls greater than Q20) can minimize false-positive calls. Optimally, each variant call should be curated

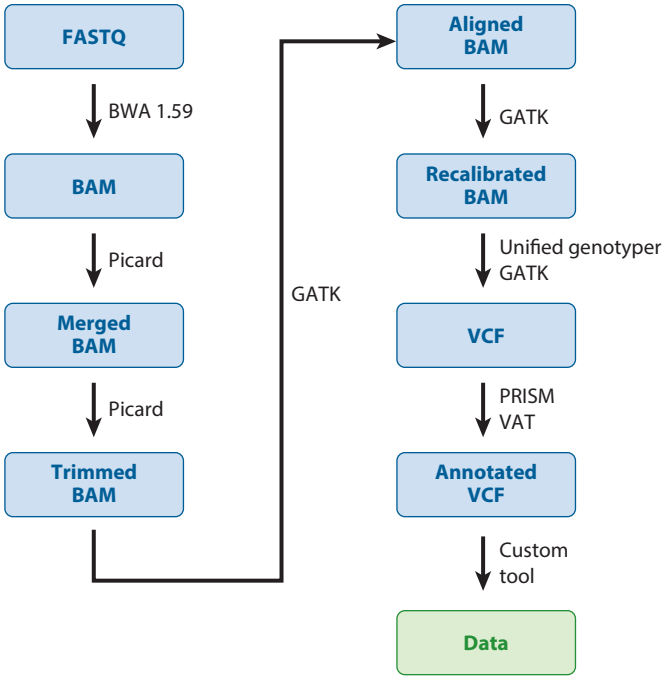

Figure 2

Flowchart of a bioinformatics pipeline, from raw next-generation sequencing data to annotated variants. FASTQ is a text-based format that includes biological sequence information (often nucleotide sequences) and quality scores in ASCII format. The BWA (Burrows–Wheeler Aligner) is a program that aligns relatively short nucleotide sequences against a long reference sequence, such as the human genome. The BWA utilizes two algorithms that together align contiguous and gapped sequences. BAM (binary alignment map) is the raw genomic sequencing data and consists of the compressed binary representation of SAM (sequence alignment map). SAM is a flexible, text-based format for storing large nucleotide sequence alignments generated by different alignment programs. Picard is a tool that is composed of Java-based command-line utilities and a Java application programming interface that facilitates the creation of new programs that can read and write SAM files. The GATK (Genome Analysis Toolkit, Broad Institute) is used for variant identification. It is one of many tools used for identifying single nucleotide variants (SNVs) and insertion/deletion variants (indels) in genomic and RNA sequencing data. VCF (variant call format) is a tab-delimited format that represents several types of variants including SNVs, small indels, and structural variants. Each VCF data line contains several fields that include significant information, such as the chromosome identity, the position and identity of the variant, the reference and alternative alleles, and variant and reference quality information. VAT (variant annotation tool) is a framework that generally operates in a cloud environment and functionally annotates the variants as, for example, synonymous, nonsynonymous, insertion, or deletion.

by an experienced professional based not only on technical benchmarks but also on each variant's functional effect, reported frequency in the tumor type being analyzed, variant allele frequency in comparison with the estimated percentage of neoplastic cells, and clinicopathologic context, among other considerations.

Indel mutations within NGS data can be significantly more challenging than SNV mutations to identify and to accurately annotate (131, 132). Various variant calling algorithms are commonly employed to detect indels including Pindel, the Genome Analysis Toolkit (GATK) UnifiedGenotyper, and the GATK HaplotypeCaller (133, 134). Studies have suggested that some

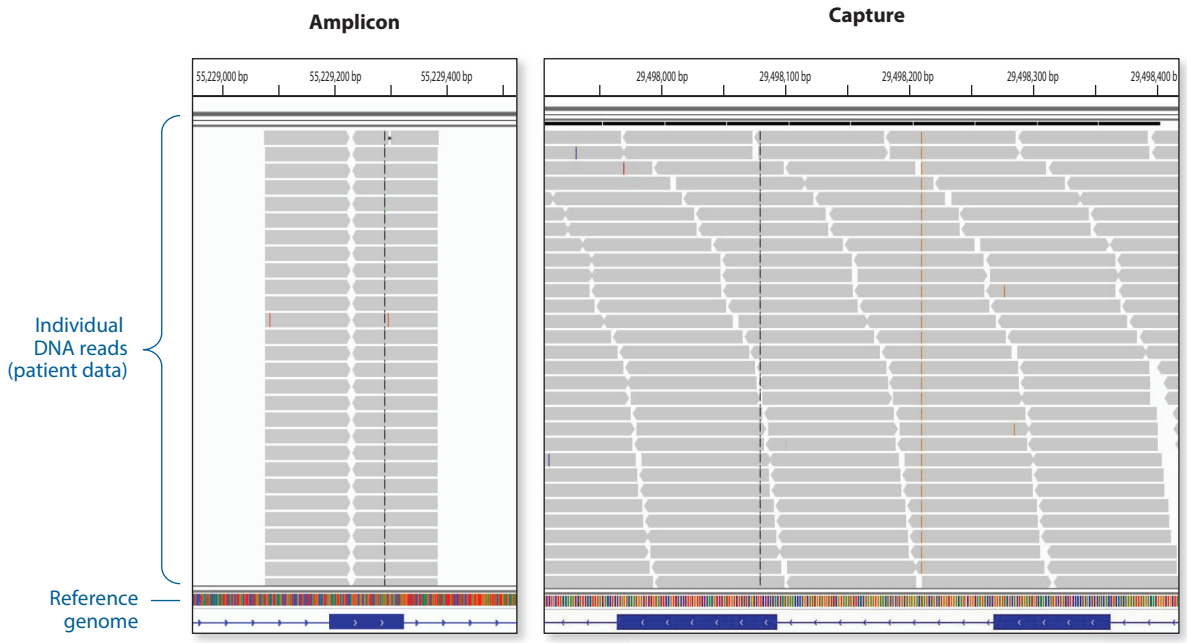

Figure 3

Sequence mapping profiles comparing amplicon-based and capture-based base next-generation sequencing strategies. Amplicon-based sequencing produces sequence information delineated by the 5′ and 3′ ends of the primers. Thus, redundancy across regions does not occur. By comparison, capture-based sequencing provides overlapping sequence coverage with multiple redundancies contributed by other probes utilized in the capture of the genomic bait.

indel variant callers have greater accuracy for smaller indels while others perform better with large indels (134, 135). Therefore, multiple algorithms may be used to obtain maximum sensitivity (134). However, this approach can lead to greater false-positive indel calls. Again, technical requirements for each variant call as well as manual curation by an experienced professional can be invaluable for ensuring the accuracy of indel calls.

Most clinical laboratories also employ one or more approaches for visualizing sequencing data. Visual inspection of read support data using tools such as the Integrative Genomics Viewer can facilitate the distinction of true variants from artifacts, or even the detection of indels missed by variant callers. Manual review is also often required to designate variants according to accepted nomenclature and naming conventions, such as that of the Human Genome Variation Society (136). Circos plots are another common mode of visualizing sequencing data (125) (**Figure 4*a***). Global analysis of variants detected by NGS within a sample is also increasingly being applied to precision oncology. The assessment of microsatellite instability has traditionally been performed by comparing PCR analyses from several well-established microsatellites in tumor and normal tissue (137) or by evaluating tumor tissue for the loss of mismatch repair protein expression (138). However, several studies have reported high concordance of these conventional techniques with algorithmic evaluation of NGS data (139–142). This can typically be accomplished by normalizing the data to a population of controls, obviating the need to evaluate normal tissue from each patient. This information can be used for cancer predisposition screening—for example, for Lynch syndrome (143)—as well as for predicting response to immune checkpoint inhibitor therapies such as the PD-1 antibody pembrolizumab (53). The tumor mutation burden, defined as the incidence

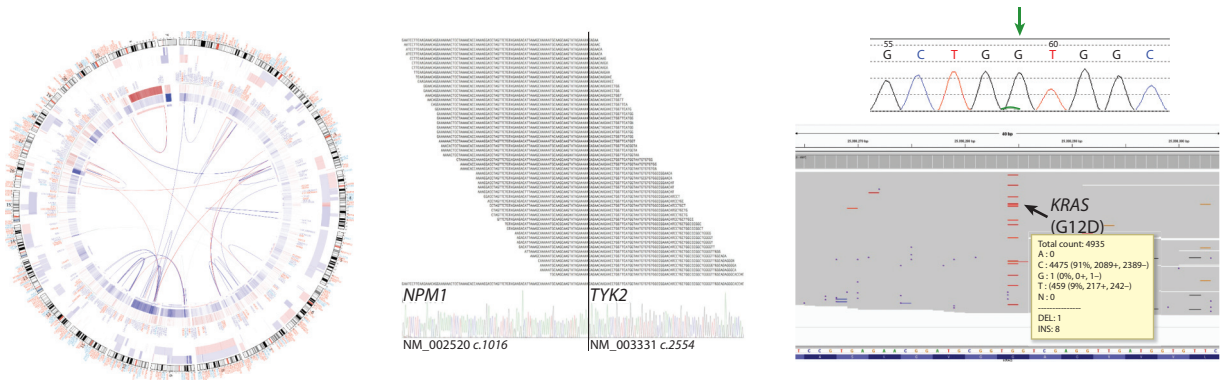

Figure 4

Display outputs for genome sequencing. (*a*) Whole-genome sequencing. Circos diagram representing data from whole-genome sequencing of a T cell neoplasm (125). In these plots, the 22 autosomes are disposed end to end, and the sex chromosomes complete the circle. The plots demonstrate the genomic complexity associated with the case and depict the structural alterations detected. The arcs within the circle represent translocations connecting chromosomal regions. The thickness of the arcs represents the read support for the chromosomal events. The outer data track represents the relative sequencing coverage of chromosomal regions normalized to publicly available genome sequencing data from healthy individuals. Light purple indicates regions of euploidy; dark purple indicates regions of chromosomal gain; and red indicates regions of chromosomal loss. The inner data tracks represent large structural alterations between spatially distinct genomic regions affecting the coding regions of one or more genes. Light purple lines indicate a single gene involved in structural alteration. Some of the genes implicated in the structural alterations are identified in this display. (*b*) RNA sequencing. The panel shows the identification of a gene fusion between *NPM1* and *TYK2* in a T cell lymphoma by whole-transcriptome sequencing. Read support from RNA sequencing illustrates the forward fragment reads spanning the break point (*vertical line*) between the *NPM1* and the *TYK2* genes. Sanger sequencing confirmation is shown at the bottom. (*c*) Targeted sequencing. Targeted panel sequencing showing a mutation in *KRAS* (G12D) visualized on the Integrative Genomics Viewer. Read counts are shown in the yellow box with the calculated variant allele fraction. Next-generation sequencing more readily shows the variant allele, which is barely visible (*green peak*) on the Sanger sequence electropherogram (*green arrow*).

of SNVs per megabase, has also been employed as a predictor of response to immune checkpoint inhibitors (54–57). While WES is the gold standard for assessing the tumor mutation burden, analyses from large, targeted panels show good correlation (144, 145).

As with every other aspect of molecular testing, each component of a bioinformatics pipeline requires optimization and thorough validation (146).

Technical Innovation and Validation Standards for Patient Safety

The majority of NGS assays in the United States consist of laboratory-developed tests. As a result, assays in different laboratories are quite variable in terms of content, methodology, bioinformatics analysis, quality control practices, interpretative procedures, and reporting. This variability largely reflects the tremendous pace at which sequencing technology and its clinical implications are evolving. Many laboratories have not only developed their own NGS assay but also have gone through several iterations of assays in order to have an assay that takes advantage of the latest technological advances and includes all of the clinically relevant content. While the process of designing, optimizing, and validating an NGS assay is time consuming, labor intensive, and expensive, the innovation involved in this process has helped to advance the technology as laboratories and oncologists share their experience and learn from other institutions. However, this variability in laboratory practices raises concerns about quality and patient safety in some laboratories.

Organizations such as the Association for Molecular Pathology, the ASCO, and the College of American Pathologists (CAP) have taken up the call to police NGS practices through education, the publication of consensus guidelines, and in the case of CAP, the regulation and inspection of laboratories (146, 147). This structure enables medical innovation while maintaining quality control and patient safety.

INTEGRATING GENOMIC RESULTS INTO REPORTS AND CLINICAL DECISION SUPPORT

NGS is a powerful tool that enables the simultaneous interrogation of many regions of the human genome. However, as the volume of data from NGS testing grows, so does the challenge of distinguishing the findings that are clinically meaningful and prioritizing their clinical utility. Given the large number of genetic variants that occur in cancer genomes and the increasingly large tail of low-frequency or nonrecurrent mutations that are detected using NGS, it is becoming quite clear that a systematic approach to prioritizing variants is necessary to effectively implement NGS-based precision diagnostics in routine clinical contexts. Molecular pathologists in collaboration with their oncology colleagues, are tasked with evaluating this abundance of data, distilling it to what is clinically relevant, and communicating this information in the most cogent and manageable manner possible. Several components are required to properly integrate genomic results into clinical reports. These include proper annotation and prioritization of the genomic variants, development and maintenance of a well-curated and current knowledge base in real time, and generation of succinct clinical reports with pertinent information offering direction toward appropriate therapies and clinical trials.

Filtering Variants Without Clinical Utility

A great deal of information generated from NGS data is not clinically useful and should not be reported because doing so would only make extracting clinically useful information that much more difficult for all practitioners. These findings that should not be reported include artifacts, synonymous variants, most intronic variants (other than splice site mutations and functional gene rearrangements), and benign germline polymorphisms.

The distinction between germline and somatic variants is not always clear (unless non-neoplastic tissue is also evaluated), and some germline variants can be clinically relevant, including those associated with cancer predisposition syndromes. Without a comparison between tumor and normal tissue, distinguishing somatic from germline variants relies largely on constitutional databases (e.g., the 1000 Genomes Project, the Exome Aggregation Consortium's ExAC browser, the Single Nucleotide Polymorphism Database dbSNP, and ClinVar) and cancer-specific databases [e.g., the Catalogue of Somatic Mutations in Cancer (COSMIC) and cBioPortal]. However, information from these databases must be interpreted with caution as some somatic variants (e.g., *JAK2* V617F) are included in germline databases, and both deleterious and benign germline variants may be included in cancer-specific databases. Even the acquired resistance mutation *EGFR* p.T790M can be seen as a rare a germline event, and this must be considered when it is identified at cancer diagnosis (148). Ultimately, tumor-only testing should not be used to conclusively infer the presence of a germline variant, but it can be used as an indication for definitive germline testing.

Somatic variants are frequently inferred to be driver or passenger mutations (149). A driver mutation is one that confers a selective growth advantage for the mutated cell, while a passenger

Table 1 Recommended framework for tiered reporting of genetic variants in the clinical laboratory setting

Tier[a]	Classification (therapeutic, prognostic, diagnostic)	Evidence level	Description of level of evidence
I	Variants of strong clinical significance	A	US Food and Drug Administration-approved therapies, professional guidelines
		B	Well-powered studies with consensus from experts in the field
II	Variants of potential clinical significance	C	Multiple small studies with some consensus, investigational therapies, and clinical trials
		D	Preclinical studies, small studies, or a few case reports without consensus
III	Variants of unknown clinical significance		Not observed at a significant allele frequency in general population or specific subpopulation databases or in pan-cancer or tumor-specific variant databases No convincing evidence of cancer association
IV	Benign or likely benign variants		Observed at significant allele frequency in general population or specific subpopulation databases No evidence of cancer association

Adapted with permission from Reference 147.

[a]Tiers I and II represent variants that are consistently reported, and those in Tier III are sometimes listed as variants of unknown clinical significance. Tier IV variants are annotated in databases but reported only infrequently.

mutation has no growth advantage and is observed within a tumor due to its co-occurrence with a driver. Because passenger mutations are random events that are not selected for during oncogenesis, they generally are not recurrent, while driver mutations are. Therefore, the frequency of a particular variant within databases of somatic mutations (such as COSMIC and cBioPortal) and the large cancer-specific studies (e.g., in The Cancer Genome Atlas) are eminently useful in evaluating the potential oncogenesis of a mutation. Functional studies provide the most definitive evidence of oncogenesis; however, these studies may not be available for unusual alterations. While driver mutations are generally the most clinically relevant, passenger mutations may also be informative. For example, passenger mutations are relevant for calculating the tumor mutation burden, which can be predictive of the response to immune checkpoint inhibitor therapies (54, 56, 150).

Classifying Variants

The classification of variants according to perceived clinical significance is based on the availability of relevant therapies and national guidelines developed on the basis of clinical and preclinical studies. Several classification schemas are employed by various institutions performing clinical sequencing. The MD Anderson Cancer Center's Knowledge Base for Precision Oncology classifies variants based on the level of clinical or preclinical evidence supporting the efficacy of a therapy for a particular tumor type and a particular variant (151). Other groups incorporate clinical trial eligibility criteria (152). As shown in **Table 1**, guidelines for interpreting and reporting sequence variants in cancer were recently published in a joint consensus recommendation from the Association for Molecular Pathology, ASCO, and CAP (147). These guidelines recommend classifying variants based on clinical and experimental evidence into four tiers, each with differing therapeutic, diagnostic, or prognostic significance.

Information Management Platform for Incorporating Molecular Genomic Data into Electronic Health Records

Most genomic sequencing test results are currently delivered as a text document and stored in portable document format (that is, as PDFs). For optimal benefit to be obtained from genomic sequencing data, this information needs to be incorporated into electronic health records (EHRs) as discrete data elements that facilitate flexible computational interrogation, transformation, and analysis. The delivery of results in an electronic format facilitates real-time interaction, retrieval, and data interrogation by pathologists and clinicians. Information management systems need to be configured to permit seamless interoperability between data housed within laboratory information systems and EHRs (e.g., Epic). Several advantages accrue from implementing interoperable platforms. These include the ability to perform monitoring activities to investigate the processes of nonrandom variations, to run queries for statistically valid trends in mutations and outcomes, and to identify promptly patients with specific mutations who are eligible for trials and can benefit from novel therapies. Regardless of the specific platform implemented, it is recommended that the discrete data elements comprising the laboratory result inputs be formatted according to the standards of the Systematized Nomenclature of Medicine–Clinical Terms (known as SNOMED CT) and the Human Genome Variation Society nomenclature for archiving data in EHRs.

Multidisciplinary Integration

Written reports are not always the best way to communicate complex information derived from molecular testing or to integrate these findings with other important patient information. Many institutions that apply molecular data to clinical care are increasingly initiating multidisciplinary conferences, often termed molecular tumor boards. These face-to-face, real-time interactions can greatly facilitate the communication of complex molecular results. These conferences also take advantage of the expertise and perspectives of multiple disciplines, including oncology, molecular pathology, surgical pathology, pharmacology, medical genetics, and social work (**Figure 5**).

REQUIREMENTS FOR ADDITIONAL DATA SOURCES AND DATA SHARING

The availability of molecular data enables oncologists to tailor treatments to individual patients. However, this practice increasingly leads to patients receiving therapy outside of the current standard of care, including, for example, the use of drugs off label or for compassionate care. Because these treatments are often given outside of the context of a clinical trial, the clinical outcomes from these *N*-of-1 trials are often unpublished or published as individual, heterogeneous case reports that cannot be easily collated and that suffer from publication bias and a lack of experimental controls. However, many potentially targetable alterations are so rare that a formal clinical trial is not possible. An interinstitutional database that enables health-care professionals to collect and share these clinical experiences is needed. A number of clinical trials have been created that attempt to accumulate sufficient numbers of patients harboring these rare but targetable alterations by enrolling patients from across the United States, regardless of cancer type. Examples include the MATCH trial conducted by the NCI and the TAPUR trial conducted by ASCO. Nonetheless, the conduct of these trials depends on widespread molecular profiling of cancers in order to identify potential participants.

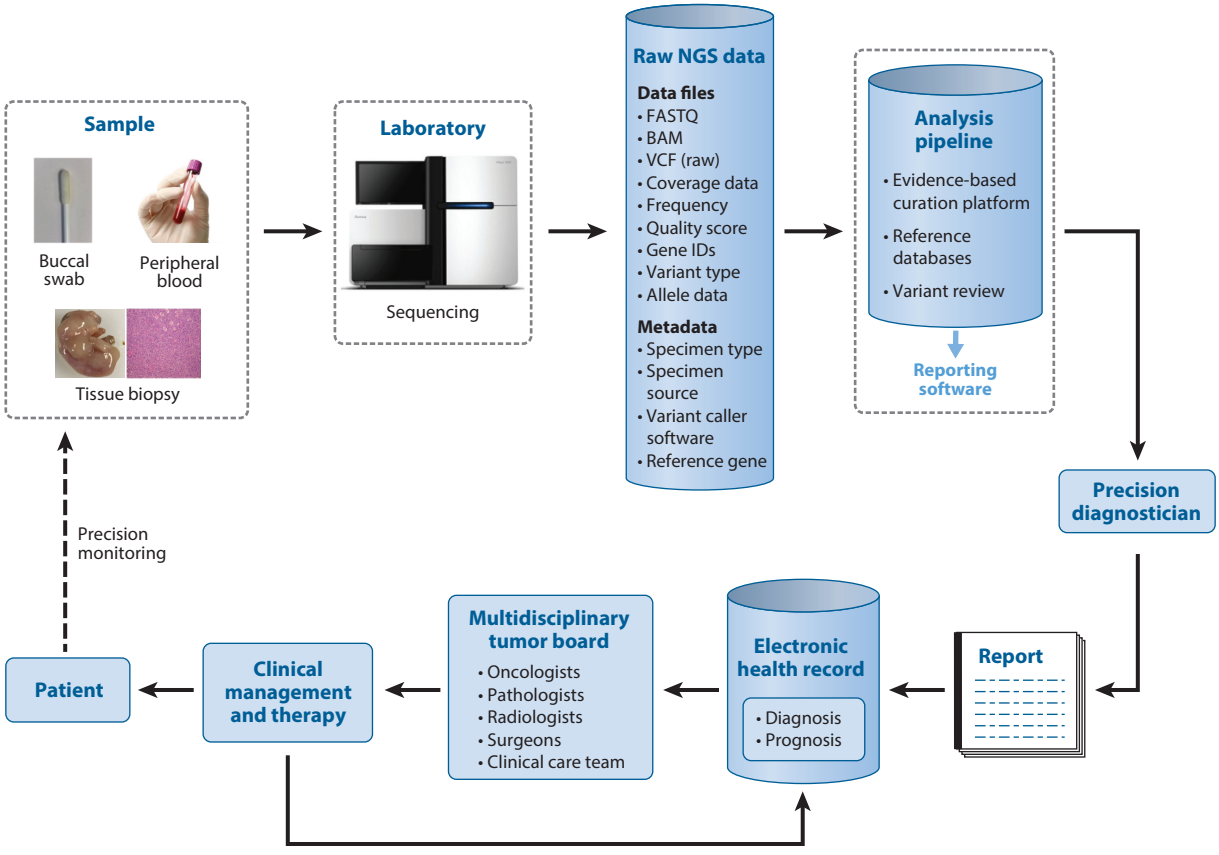

Figure 5

Overall process for clinical implementation of precision oncology using next-generation sequencing (NGS). A patient's sample undergoes sequencing followed by bioinformatics analysis using an analysis pipeline. The data are then interpreted by a precision diagnostician, and the diagnostic, prognostic, and therapeutic implications of the findings are communicated to clinicians in other disciplines through a report and multidisciplinary conferences. This information is then used to create a patient management plan that is customized to the patient's molecular findings.

CONCLUSIONS

Revolutionary advances in genomic testing have made possible the routine use of genomic data to guide patient management. Since the completion of the first human genome sequence in 2004 (28), the advent of massively parallel DNA sequencing—commonly referred to as NGS—has drastically reduced the speed and costs associated with molecular testing. However, the large quantities of data produced by this technology require the implementation of data management and infrastructure that facilitate the efficient utilization of data obtained from clinically relevant settings, and the data also present significant challenges that must be overcome to ensure that the most salient genomic information is applied to patient care.

The acquisition, interpretation, and application of genomic data have emerged as critical components in the care of cancer patients. This process involves several steps and improves the quality of care by enabling clinical management to be tailored to individual cancer patients (**Figure 5**). Enormous innovations in molecular oncology have propelled precision medicine forward while also engendering new challenges, including the need to use specimens commonly available in the

clinical setting; determining how to process, store, and manage large amounts of sequencing data; determining how to interpret and prioritize molecular findings; and coordinating health professionals from multiple disciplines. Precision oncology remains a dynamic field as both molecular technology and its applications continue to evolve rapidly.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

K.S.J.E.-J. is partially supported by a grant from the US National Cancer Institute (R01CA231021).

LITERATURE CITED

1. Lu YF, Goldstein DB, Angrist M, Cavalleri G. 2014. Personalized medicine and human genetic diversity. *Cold Spring Harb. Perspect. Med.* 4:a008581
2. Garrido P, Aldaz A, Vera R, Calleja MA, de Álava E, et al. 2018. Proposal for the creation of a national strategy for precision medicine in cancer: a position statement of SEOM, SEAP, and SEFH. *Clin. Transl. Oncol.* 20:443–47
3. Yan B, Hu Y, Ng C, Ban KH, Tan TW, et al. 2016. Coverage analysis in a targeted amplicon-based next-generation sequencing panel for myeloid neoplasms. *J. Clin. Pathol.* 69:801–4
4. Spencer DH, Abel HJ, Lockwood CM, Payton JE, Szankasi P, et al. 2013. Detection of *FLT3* internal tandem duplication in targeted, short-read-length, next-generation sequencing data. *J. Mol. Diagn.* 15:81–93
5. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, et al. 2004. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 350:2129–39
6. Linardou H, Dahabreh IJ, Bafaloukos D, Kosmidis P, Murray S. 2009. Somatic EGFR mutations and efficacy of tyrosine kinase inhibitors in NSCLC. *Nat. Rev. Clin. Oncol.* 6:352–66
7. Popat S, Mok T, Yang JC, Wu YL, Lungershausen J, et al. 2014. Afatinib in the treatment of EGFR mutation-positive NSCLC—a network meta-analysis. *Lung Cancer* 85:230–38
8. Saxon JA, Sholl LM, Janne PA. 2017. *EGFR* L858M/L861Q cis mutations confer selective sensitivity to afatinib. *J. Thorac. Oncol.* 12:884–89
9. Yang JC, Sequist LV, Geater SL, Tsai CM, Mok TS, et al. 2015. Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon *EGFR* mutations: a combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. *Lancet Oncol.* 16:830–38
10. Shaw AT, Kim DW, Nakagawa K, Seto T, Crino L, et al. 2013. Crizotinib versus chemotherapy in advanced *ALK*-positive lung cancer. *N. Engl. J. Med.* 368:2385–94
11. Solomon BJ, Mok T, Kim DW, Wu YL, Nakagawa K, et al. 2014. First-line crizotinib versus chemotherapy in *ALK*-positive lung cancer. *N. Engl. J. Med.* 371:2167–77
12. Ou SH, Ahn JS, De Petris L, Govindan R, Yang JC, et al. 2016. Alectinib in crizotinib-refractory *ALK*-rearranged non-small-cell lung cancer: a phase II global study. *J. Clin. Oncol.* 34:661–68
13. Soria JC, Tan DSW, Chiari R, Wu YL, Paz-Ares L, et al. 2017. First-line ceritinib versus platinum-based chemotherapy in advanced *ALK*-rearranged non-small-cell lung cancer (ASCEND-4): a randomised, open-label, phase 3 study. *Lancet* 389:917–29
14. Shaw AT, Kim DW, Mehra R, Tan DS, Felip E, et al. 2014. Ceritinib in *ALK*-rearranged non-small-cell lung cancer. *N. Engl. J. Med.* 370:1189–97

15. Kim DW, Tiseo M, Ahn MJ, Reckamp KL, Hansen KH, et al. 2017. Brigatinib in patients with crizotinib-refractory anaplastic lymphoma kinase–positive non-small-cell lung cancer: a randomized, multicenter phase II trial. *J. Clin. Oncol.* 35:2490–98
16. Shaw AT, Ou SH, Bang YJ, Camidge DR, Solomon BJ, et al. 2014. Crizotinib in *ROS1*-rearranged non-small-cell lung cancer. *N. Engl. J. Med.* 371:1963–71
17. Mazieres J, Zalcman G, Crino L, Biondani P, Barlesi F, et al. 2015. Crizotinib therapy for advanced lung adenocarcinoma and a *ROS1* rearrangement: results from the EUROS1 cohort. *J. Clin. Oncol.* 33:992–99
18. Planchard D, Besse B, Groen HJM, Souquet PJ, Quoix E, et al. 2016. Dabrafenib plus trametinib in patients with previously treated $BRAF^{V600E}$-mutant metastatic non–small cell lung cancer: an open-label, multicentre phase 2 trial. *Lancet Oncol*. 17:984–93
19. Planchard D, Smit EF, Groen HJM, Mazieres J, Besse B, et al. 2017. Dabrafenib plus trametinib in patients with previously untreated $BRAF^{V600E}$-mutant metastatic non-small-cell lung cancer: an open-label, phase 2 trial. *Lancet Oncol*. 18:1307–16
20. Frampton GM, Ali SM, Rosenzweig M, Chmielecki J, Lu X, et al. 2015. Activation of *MET* via diverse exon 14 splicing alterations occurs in multiple tumor types and confers clinical sensitivity to MET inhibitors. *Cancer Discov*. 5:850–59
21. Paik PK, Drilon A, Fan PD, Yu H, Rekhtman N, et al. 2015. Response to MET inhibitors in patients with stage IV lung adenocarcinomas harboring *MET* mutations causing exon 14 skipping. *Cancer Discov*. 5:842–49
22. Suryavanshi M, Shah A, Kumar D, Panigrahi MK, Metha A, Batra U. 2017. *MET* amplification and response to MET inhibitors in stage IV lung adenocarcinoma. *Oncol. Res. Treat.* 40:198–202
23. Drilon A, Subbiah V, Oxnard GR, Bauer TM, Velcheti V, et al. 2018. A phase 1 study of LOXO-292, a potent and highly selective RET inhibitor, in patients with *RET*-altered cancers. *J. Clin. Oncol.* 36(Suppl.):102
24. Subbiah V, Gainor JF, Rahal R, Brubaker JD, Kim JL, et al. 2018. Precision targeted therapy with BLU-667 for *RET*-driven cancers. *Cancer Discov*. 8:836–49
25. Drilon AE, Liu S, Doebele R, Rodriguez C, Fakih M, et al. 2017. A phase 1b study of RXDX-105, a VEGFR-sparing potent RET inhibitor, in RETi-naïve patients with *RET* fusion–positive NSCLC. *Ann. Oncol.* 28(Suppl. 5):v605–49
26. Li BT, Shen R, Buonocore D, Olah ZT, Ni A, et al. 2018. Ado-trastuzumab emtansine for patients with *HER2*-mutant lung cancers: results from a phase II basket trial. *J. Clin. Oncol.* 36:2532–37
27. Ettinger DS, Wood DE, Airsner DL, Wallace A, Bauman J, et al. 2019. NCCN Clinical Practice Guidelines in Oncology: non–small cell lung cancer. Version 5.2019. *Natl. Compr. Cancer Netw.* **https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf**
28. Int. Hum. Genome Consort. 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431:931–45
29. Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, et al. 2012. Genomic landscape of non–small cell lung cancer in smokers and never-smokers. *Cell* 150:1121–34
30. Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10:57–63
31. Esteller M. 2011. Non-coding RNAs in human disease. *Nat. Rev. Genet.* 12:861–74
32. Chen X, Ba Y, Ma L, Cai X, Yin Y, et al. 2008. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res.* 18:997–1006
33. Volinia S, Galasso M, Sana ME, Wise TF, Palatini J, et al. 2012. Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. *PNAS* 109:3024–29
34. Martens-Uzunova ES, Jalava SE, Dits NF, van Leenders GJ, Moller S, et al. 2012. Diagnostic and prognostic signatures from the small non-coding RNA transcriptome in prostate cancer. *Oncogene* 31:978–91
35. de Leng WW, Gadellaa-van Hooijdonk CG, Barendregt-Smouter FA, Koudijs MJ, Nijman I, et al. 2016. Targeted next generation sequencing as a reliable diagnostic assay for the detection of somatic mutations in tumours using minimal DNA amounts from formalin fixed paraffin embedded material. *PLOS ONE* 11:e0149405

36. Aarthy R, Mani S, Velusami S, Sundarsingh S, Rajkumar T. 2015. Role of circulating cell-free DNA in cancers. *Mol. Diagn. Ther.* 19:339–50
37. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, et al. 2015. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* 126:9–16
38. Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, et al. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* 371:2477–87
39. Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, et al. 2014. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* 20:1472–78
40. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, et al. 2009. *IDH1* and *IDH2* mutations in gliomas. *N. Engl. J. Med.* 360:765–73
41. Sanson M, Marie Y, Paris S, Idbaih A, Laffaire J, et al. 2009. Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. *J. Clin. Oncol.* 27:4150–54
42. Louis DN, Ohgaki H, Wiestler OD, Cavenee WK, eds. 2016. *WHO Classification of Tumours of the Central Nervous System*. Lyon, France: IARC (Int. Agency Res. Cancer). Revis. 4th ed.
43. Tallman MS, Wang ES, Altman JK, Appelbaum FR, Bhatt VR, et al. 2019. NCCN Clinical Practice Guidelines in Oncology: acute myeloid leukemia. Version 3.2019. *Natl. Compr. Cancer Netw.* **https://www.nccn.org/professionals/physician_gls/pdf/aml.pdf**
44. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, et al. 2011. Improved survival with vemurafenib in melanoma with *BRAF* V600E mutation. *N. Engl. J. Med.* 364:2507–16
45. Flaherty KT, Infante JR, Daud A, Gonzalez R, Kefford RF, et al. 2012. Combined BRAF and MEK inhibition in melanoma with *BRAF* V600 mutations. *N. Engl. J. Med.* 367:1694–703
46. Douillard JY, Oliner KS, Siena S, Tabernero J, Burkes R, et al. 2013. Panitumumab-FOLFOX4 treatment and *RAS* mutations in colorectal cancer. *N. Engl. J. Med.* 369:1023–34
47. Peeters M, Price TJ, Cervantes A, Sobrero AF, Ducreux M, et al. 2010. Randomized phase III study of panitumumab with fluorouracil, leucovorin, and irinotecan (FOLFIRI) compared with FOLFIRI alone as second-line treatment in patients with metastatic colorectal cancer. *J. Clin. Oncol.* 28:4706–13
48. Guo J, Si L, Kong Y, Flaherty KT, Xu X, et al. 2011. Phase II, open-label, single-arm trial of imatinib mesylate in patients with metastatic melanoma harboring *c-Kit* mutation or amplification. *J. Clin. Oncol.* 29:2904–9
49. Carvajal RD, Antonescu CR, Wolchok JD, Chapman PB, Roman RA, et al. 2011. *KIT* as a therapeutic target in metastatic melanoma. *JAMA* 305:2327–34
50. Hodi FS, Corless CL, Giobbie-Hurder A, Fletcher JA, Zhu M, et al. 2013. Imatinib for melanomas harboring mutationally activated or amplified *KIT* arising on mucosal, acral, and chronically sun-damaged skin. *J. Clin. Oncol.* 31:3182–90
51. Coit DG, Thompson JA, Albertini MR, Algazi A, Andtbacka R, et al. 2019. NCCN Clinical Practice Guidelines in Oncology: cutaneous melanoma. Version 2.2019. *Natl. Compr. Cancer Netw.* **https://www.nccn.org/professionals/physician_gls/pdf/cutaneous_melanoma.pdf**
52. Mok TS, Wu YL, Ahn MJ, Garassino MC, Kim HR, et al. 2017. Osimertinib or platinum-pemetrexed in *EGFR* T790M–positive lung cancer. *N. Engl. J. Med.* 376:629–40
53. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, et al. 2017. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357:409–13
54. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, et al. 2015. Cancer immunology: mutational landscape determines sensitivity to PD-1 blockade in non–small cell lung cancer. *Science* 348:124–28
55. Bouffet E, Larouche V, Campbell BB, Merico D, de Borja R, et al. 2016. Immune checkpoint inhibition for hypermutant glioblastoma multiforme resulting from germline biallelic mismatch repair deficiency. *J. Clin. Oncol.* 34:2206–11
56. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, et al. 2014. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* 371:2189–99
57. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, et al. 2015. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350:207–11

58. Kaufman B, Shapira-Frommer R, Schmutzler RK, Audeh MW, Friedlander M, et al. 2015. Olaparib monotherapy in patients with advanced cancer and a germline *BRCA1/2* mutation. *J. Clin. Oncol.* 33:244–50
59. Mateo J, Carreira S, Sandhu S, Miranda S, Mossop H, et al. 2015. DNA-repair defects and olaparib in metastatic prostate cancer. *N. Engl. J. Med.* 373:1697–708
60. Swisher EM, Lin KK, Oza AM, Scott CL, Giordano H, et al. 2017. Rucaparib in relapsed, platinum-sensitive high-grade ovarian carcinoma (ARIEL2 Part 1): an international, multicentre, open-label, phase 2 trial. *Lancet Oncol.* 18:75–87
61. Mirza MR, Monk BJ, Herrstedt J, Oza AM, Mahner S, et al. 2016. Niraparib maintenance therapy in platinum-sensitive, recurrent ovarian cancer. *N. Engl. J. Med.* 375:2154–64
62. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. 2009. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clin. Pharmacol. Ther.* 86:97–100
63. Redig AJ, Janne PA. 2015. Basket trials and the evolution of clinical trial design in an era of genomic medicine. *J. Clin. Oncol.* 33:975–77
64. Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, et al. 2018. Efficacy of larotrectinib in *TRK* fusion–positive cancers in adults and children. *N. Engl. J. Med.* 378:731–39
65. Younes A, Berry DA. 2012. From drug discovery to biomarker-driven clinical trials in lymphoma. *Nat. Rev. Clin. Oncol.* 9:643–53
66. Hughes T, Deininger M, Hochhaus A, Branford S, Radich J, et al. 2006. Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting *BCR-ABL* transcripts and kinase domain mutations and for expressing results. *Blood* 108:28–37
67. Baccarani M, Pileri S, Steegmann JL, Muller M, Soverini S, Dreyling M. 2012. Chronic myeloid leukemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 23(Suppl. 7):vii72–77
68. Hughes TP, Kaeda J, Branford S, Rudzki Z, Hochhaus A, et al. 2003. Frequency of major molecular responses to imatinib or interferon alfa plus cytarabine in newly diagnosed chronic myeloid leukemia. *N. Engl. J. Med.* 349:1423–32
69. Saglio G, Fava C. 2012. Practical monitoring of chronic myelogenous leukemia: when to change treatment. *J. Natl. Compr. Cancer Netw.* 10:121–29
70. Milojkovic D, Apperley J. 2009. Mechanisms of resistance to imatinib and second-generation tyrosine inhibitors in chronic myeloid leukemia. *Clin. Cancer Res.* 15:7519–27
71. Soverini S, Hochhaus A, Nicolini FE, Gruber F, Lange T, et al. 2011. *BCR-ABL* kinase domain mutation analysis in chronic myeloid leukemia patients treated with tyrosine kinase inhibitors: recommendations from an expert panel on behalf of European LeukemiaNet. *Blood* 118:1208–15
72. Pongers-Willemse MJ, Verhagen OJ, Tibbe GJ, Wijkhuijs AJ, de Haas V, et al. 1998. Real-time quantitative PCR for the detection of minimal residual disease in acute lymphoblastic leukemia using junctional region specific TaqMan probes. *Leukemia* 12:2006–14
73. Bruggemann M, Droese J, Bolz I, Luth P, Pott C, et al. 2000. Improved assessment of minimal residual disease in B cell malignancies using fluorogenic consensus probes for real-time quantitative PCR. *Leukemia* 14:1419–25
74. Verhagen OJ, Willemse MJ, Breunis WB, Wijkhuijs AJ, Jacobs DC, et al. 2000. Application of germline IGH probes in real-time quantitative PCR for the detection of minimal residual disease in acute lymphoblastic leukemia. *Leukemia* 14:1426–35
75. Wood B, Wu D, Crossley B, Dai Y, Williamson D, et al. 2018. Measurable residual disease detection by high-throughput sequencing improves risk stratification for pediatric B-ALL. *Blood* 131:1350–59
76. Logan AC, Vashi N, Faham M, Carlton V, Kong K, et al. 2014. Immunoglobulin and T cell receptor gene high-throughput sequencing quantifies minimal residual disease in acute lymphoblastic leukemia and predicts post-transplantation relapse and survival. *Biol. Blood Marrow Transplant.* 20:1307–13
77. Pulsipher MA, Carlson C, Langholz B, Wall DA, Schultz KR, et al. 2015. IgH-V(D)J NGS-MRD measurement pre- and early post-allotransplant defines very low- and very high-risk ALL patients. *Blood* 125:3501–8

78. Wu D, Emerson RO, Sherwood A, Loh ML, Angiolillo A, et al. 2014. Detection of minimal residual disease in B lymphoblastic leukemia by high-throughput sequencing of IGH. *Clin. Cancer Res.* 20:4540–48
79. Wu D, Sherwood A, Fromm JR, Winter SS, Dunsmore KP, et al. 2012. High-throughput sequencing detects minimal residual disease in acute T lymphoblastic leukemia. *Sci. Transl. Med.* 4:134ra63
80. Kotrova M, Muzikova K, Mejstrikova E, Novakova M, Bakardjieva-Mihaylova V, et al. 2015. The predictive strength of next-generation sequencing MRD detection for relapse compared with current methods in childhood ALL. *Blood* 126:1045–47
81. Grimwade D, Jovanovic JV, Hills RK, Nugent EA, Patel Y, et al. 2009. Prospective minimal residual disease monitoring to predict relapse of acute promyelocytic leukemia and to direct pre-emptive arsenic trioxide therapy. *J. Clin. Oncol.* 27:3650–58
82. Platzbecker U, Avvisati G, Cicconi L, Thiede C, Paoloni F, et al. 2017. Improved outcomes with retinoic acid and arsenic trioxide compared with retinoic acid and chemotherapy in non-high-risk acute promyelocytic leukemia: final results of the randomized Italian-German APL0406 trial. *J. Clin. Oncol.* 35:605–12
83. Yin JA, O'Brien MA, Hills RK, Daly SB, Wheatley K, Burnett AK. 2012. Minimal residual disease monitoring by quantitative RT-PCR in core binding factor AML allows risk stratification and predicts relapse: results of the United Kingdom MRC AML-15 trial. *Blood* 120:2826–35
84. Willekens C, Blanchet O, Renneville A, Cornillet-Lefebvre P, Pautas C, et al. 2016. Prospective long-term minimal residual disease monitoring using RQ-PCR in *RUNX1-RUNX1T1*-positive acute myeloid leukemia: results of the French CBF-2006 trial. *Haematologica* 101:328–35
85. Ivey A, Hills RK, Simpson MA, Jovanovic JV, Gilkes A, et al. 2016. Assessment of minimal residual disease in standard-risk AML. *N. Engl. J. Med.* 374:422–33
86. Shayegi N, Kramer M, Bornhauser M, Schaich M, Schetelig J, et al. 2013. The level of residual disease based on mutant *NPM1* is an independent prognostic factor for relapse and survival in AML. *Blood* 122:83–92
87. Krönke J, Schlenk RF, Jensen KO, Tschürtz F, Corbacioglu A, et al. 2011. Monitoring of minimal residual disease in *NPM1*-mutated acute myeloid leukemia: a study from the German–Austrian acute myeloid leukemia study group. *J. Clin. Oncol.* 29:2709–16
88. Balsat M, Renneville A, Thomas X, de Botton S, Caillot D, et al. 2017. Postinduction minimal residual disease predicts outcome and benefit from allogeneic stem cell transplantation in acute myeloid leukemia with *NPM1* mutation: a study by the Acute Leukemia French Association Group. *J. Clin. Oncol.* 35:185–93
89. Jongen-Lavrencic M, Grob T, Hanekamp D, Kavelaars FG, Al Hinai A, et al. 2018. Molecular minimal residual disease in acute myeloid leukemia. *N. Engl. J. Med.* 378:1189–99
90. Morita K, Kantarjian HM, Wang F, Yan Y, Bueso-Ramos C, et al. 2018. Clearance of somatic mutations at remission and the risk of relapse in acute myeloid leukemia. *J. Clin. Oncol.* 36:1788–97
91. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, et al. 2008. Circulating mutant DNA to assess tumor dynamics. *Nat. Med.* 14:985–90
92. Tie J, Wang Y, Tomasetti C, Li L, Springer S, et al. 2016. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci. Transl. Med.* 8:346ra92
93. Chaudhuri AA, Chabon JJ, Lovejoy AF, Newman AM, Stehr H, et al. 2017. Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling. *Cancer Discov.* 7:1394–403
94. Diaz LA Jr., Williams RT, Wu J, Kinde I, Hecht JR, et al. 2012. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 486:537–40
95. Oxnard GR, Paweletz CP, Kuang Y, Mach SL, O'Connell A, et al. 2014. Noninvasive detection of response and resistance in *EGFR*-mutant lung cancer using quantitative next-generation genotyping of cell-free plasma DNA. *Clin. Cancer Res.* 20:1698–705
96. De Mattos-Arruda L, Weigelt B, Cortes J, Won HH, Ng CK, et al. 2014. Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: a proof-of-principle. *Ann. Oncol.* 25:1729–35

97. Lynch HT, Smyrk T, Lynch J, Fitzgibbons R Jr., Lanspa S, McGinn T. 1995. Update on the differential diagnosis, surveillance and management of hereditary non-polyposis colorectal cancer. *Eur. J. Cancer* 31:1039–46
98. Sepulveda AR, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, et al. 2017. Molecular Biomarkers for the Evaluation of Colorectal Cancer: Guideline From the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and American Society of Clinical Oncology. *Arch. Pathol. Lab. Med.* 141:625–57
99. Giardiello FM, Allen JI, Axilbund JE, Boland CR, Burke CA, et al. 2014. Guidelines on Genetic Evaluation and Management of Lynch Syndrome: a consensus statement by the US Multi-Society Task Force on Colorectal Cancer. *Am. J. Gastroenterol.* 109:1159–79
100. Abu-Rustum NR, Yashar CM, Bean S, Bradley K, Campos SM, et al. 2019. NCCN Clinical Practice Guidelines in Oncology: uterine neoplasms. Version 3.2019. *Natl. Compr. Cancer Netw.* **https://www.nccn.org/professionals/physician_gls/pdf/uterine.pdf**
101. Morikawa T, Shima K, Kuchiba A, Yamauchi M, Tanaka N, et al. 2012. No evidence for interference of H&E staining in DNA testing: usefulness of DNA extraction from H&E-stained archival tissue sections. *Am. J. Clin. Pathol.* 138:122–29
102. Rabien A, Kristiansen G. 2016. Tissue microdissection. *Methods Mol. Biol.* 1381:39–52
103. Datta S, Malhotra L, Dickerson R, Chaffee S, Sen CK, Roy S. 2015. Laser capture microdissection: big data from small samples. *Histol. Histopathol.* 30:1255–69
104. Geiersbach K, Adey N, Welker N, Elsberry D, Malmberg E, et al. 2016. Digitally guided microdissection aids somatic mutation detection in difficult to dissect tumors. *Cancer Genet.* 209:42–49
105. Hardingham JE, Grover P, Winter M, Hewett PJ, Price TJ, Thierry B. 2015. Detection and clinical significance of circulating tumor cells in colorectal cancer—20 years of progress. *Mol. Med.* 21(Suppl. 1):S25–31
106. Alva A, Friedlander T, Clark M, Huebner T, Daignault S, et al. 2015. Circulating tumor cells as potential biomarkers in bladder cancer. *J. Urol.* 194:790–98
107. Yu N, Zhou J, Cui F, Tang X. 2015. Circulating tumor cells in lung cancer: detection methods and clinical applications. *Lung* 193:157–71
108. Lianidou ES, Strati A, Markou A. 2014. Circulating tumor cells as promising novel biomarkers in solid cancers. *Crit. Rev. Clin. Lab. Sci.* 51:160–71
109. Narayan A, Carriero NJ, Gettinger SN, Kluytenaar J, Kozak KR, et al. 2012. Ultrasensitive measurement of hotspot mutations in tumor DNA in blood using error-suppressed multiplexed deep sequencing. *Cancer Res.* 72:3492–98
110. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, et al. 2014. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* 6:224ra24
111. Higgins MJ, Jelovac D, Barnathan E, Blair B, Slater S, et al. 2012. Detection of tumor *PIK3CA* status in metastatic breast cancer using peripheral blood. *Clin. Cancer Res.* 18:3462–69
112. Rothe F, Laes JF, Lambrechts D, Smeets D, Vincent D, et al. 2014. Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. *Ann. Oncol.* 25:1959–65
113. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, et al. 2013. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N. Engl. J. Med.* 368:1199–209
114. Thierry AR, Mouliere F, El Messaoudi S, Mollevi C, Lopez-Crapez E, et al. 2014. Clinical validation of the detection of *KRAS* and *BRAF* mutations from circulating tumor DNA. *Nat. Med.* 20:430–35
115. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, et al. 2014. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* 20:548–54
116. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, et al. 2014. Library construction for next-generation sequencing: overviews and challenges. *BioTechniques* 56:61–64
117. Williams HL, Walsh K, Diamond A, Oniscu A, Deans ZC. 2018. Validation of the Oncomine™ focus panel for next-generation sequencing of clinical tumour samples. *Virchows Arch.* 473:489–503
118. Ikegawa S, Mabuchi A, Ogawa M, Ikeda T. 2002. Allele-specific PCR amplification due to sequence identity between a PCR primer and an amplicon: Is direct sequencing so reliable? *Hum. Genet.* 110:606–8

119. Zheng Z, Liebers M, Zhelyazkova B, Cao Y, Panditi D, et al. 2014. Anchored multiplex PCR for targeted next-generation sequencing. *Nat. Med.* 20:1479–84
120. Wommack KE, Bhavsar J, Ravel J. 2008. Metagenomics: Read length matters. *Appl. Environ. Microbiol.* 74:1453–63
121. Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* 8:R143
122. Minoche AE, Dohm JC, Himmelbauer H. 2011. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems. *Genome Biol.* 12:R112
123. Ewing B, Green P. 1998. Base-calling of automated sequencer traces using *Phred*. II. Error probabilities. *Genome Res.* 8:186–94
124. Do H, Dobrovic A. 2015. Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin. Chem.* 61:64–71
125. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, et al. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res.* 19:1639–45
126. Grant GR, Farkas MH, Pizarro AD, Lahens NF, Schug J, et al. 2011. Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq unified mapper (RUM). *Bioinformatics* 27:2518–28
127. Sboner A, Habegger L, Pflueger D, Terry S, Chen DZ, et al. 2010. FusionSeq: a modular framework for finding gene fusions by analyzing paired-end RNA-sequencing data. *Genome Biol.* 11:R104
128. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, et al. 2011. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLOS Comput. Biol.* 7:e1001138
129. Kinsella M, Harismendy O, Nakano M, Frazer KA, Bafna V. 2011. Sensitive gene fusion detection using ambiguously mapping RNA-Seq read pairs. *Bioinformatics* 27:1068–75
130. Kim D, Salzberg SL. 2011. TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome Biol.* 12:R72
131. Mullaney JM, Mills RE, Pittard WS, Devine SE. 2010. Small insertions and deletions (INDELs) in human genomes. *Hum. Mol. Genet.* 19:R131–36
132. Albers CA, Lunter G, MacArthur DG, McVean G, Ouwehand WH, Durbin R. 2011. Dindel: accurate indel calls from short-read data. *Genome Res.* 21:961–73
133. Kay Ye, Schulz MH, Long Q, Apweiler R, Ning Z. 2009. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* 25:2865–71
134. Ghoneim DH, Myers JR, Tuttle E, Paciorkowski AR. 2014. Comparison of insertion/deletion calling algorithms on human next-generation sequencing data. *BMC Res. Notes* 7:864
135. Li S, Li R, Li H, Lu J, Li Y, et al. 2013. SOAPindel: efficient identification of indels from short paired reads. *Genome Res.* 23:195–200
136. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, et al. 2016. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum. Mutat.* 37:564–69
137. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, et al. 1998. A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 58:5248–57
138. Pinol V, Castells A, Andreu M, Castellvi-Bel S, Alenda C, et al. 2005. Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. *JAMA* 293:1986–94
139. Salipante SJ, Scroggins SM, Hampel HL, Turner EH, Pritchard CC. 2014. Microsatellite instability detection by next generation sequencing. *Clin. Chem.* 60:1192–99
140. Hempelmann JA, Scroggins SM, Pritchard CC, Salipante SJ. 2015. MSIplus for integrated colorectal cancer molecular testing by next-generation sequencing. *J. Mol. Diagn.* 17:705–14
141. Kautto EA, Bonneville R, Miya J, Yu L, Krook MA, et al. 2017. Performance evaluation for rapid detection of pan-cancer microsatellite instability with MANTIS. *Oncotarget* 8:7452–63
142. Huang MN, McPherson JR, Cutcutache I, Teh BT, Tan P, Rozen SG. 2015. MSIseq: software for assessing microsatellite instability from catalogs of somatic mutations. *Sci. Rep.* 5:13321

143. Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, et al. 1993. Clues to the pathogenesis of familial colorectal cancer. *Science* 260:812–16
144. Garofalo A, Sholl L, Reardon B, Taylor-Weiner A, Amin-Mansour A, et al. 2016. The impact of tumor profiling approaches and genomic data strategies for cancer precision medicine. *Genome Med*. 8:79
145. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, et al. 2017. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med*. 9:34
146. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, et al. 2017. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J. Mol. Diagn*. 19:341–65
147. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, et al. 2017. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: a Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J. Mol. Diagn.* 19:4–23
148. Oxnard GR, Miller VA, Robson ME, Azzoli CG, Pao W, et al. 2012. Screening for germline *EGFR* T790M mutations through lung cancer genotyping. *J. Thorac. Oncol.* 7:1049–52
149. Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, et al. 2010. Accumulation of driver and passenger mutations during tumor progression. *PNAS* 107:18545–50
150. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, et al. 2015. PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* 372:2509–20
151. MD Anderson Cancer Cent. Personalized cancer therapy: knowledge base for precision oncology. *MD Anderson Cancer Center*. **https://pct.mdanderson.org/**
152. Ghazani AA, Oliver NM, Pierre JP St, Garofalo A, Rainville IR, et al. 2017. Assigning clinical meaning to somatic and germ-line whole-exome sequencing data in a prospective cancer precision medicine study. *Genet. Med.* 19:787–95