Whole-Genome Sequencing in Cancer

Eric Y. Zhao, Martin Jones, and Steven J.M. Jones

Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, British Columbia V5Z 4S6, Canada

Correspondence: mjones@bcgsc.ca

Genome sequencing of cancer has fundamentally advanced our understanding of the underlying biology of this disease, and more recently has provided approaches to characterize and monitor tumors in the clinic, guiding and evaluating treatment. Although cancer research is relying more on whole-genome characterization, the clinical application of genomics is largely limited to targeted sequencing approaches, tailored to capture specific clinically relevant biomarkers. However, as sequencing costs reduce, and the tools to effectively analyze complex and large-scale data improve, the ability to effectively characterize whole genomes at scale in a clinically relevant time frame is now being piloted. This ability effectively blurs the line between clinical cancer research and the clinical management of the disease. This leads to a new paradigm in cancer management in which real-time analysis of an individual's disease can have a rapid and lasting impact on our understanding of how clinical practices need to change to exploit novel therapeutic rationales. In this article, we will discuss how whole-genome sequencing (WGS), often combined with transcriptome analysis, has been used to understand cancer and how this approach is uniquely positioned to provide a comprehensive view of an evolving disease in response to therapy.

EARLY APPLICATION OF WHOLE-GENOME SEQUENCING IN ONCOLOGY

The application of whole-genome sequencing (WGS) to derive a more complete understanding of cancer has been a central goal of cancer researchers since before the human genome was first decoded in 2003 (Lander et al. 2001). It would take a further 5 years and a sea change in genome-sequencing technology before the first application of next-generation WGS to a cancer sample was described. Ley and colleagues reported the analysis of a cytogenetically normal acute myeloid leukemia (AML) in 2008 (The Cancer Genome Atlas Network 2013), only 6 months after the first human whole-genome sequence by next-generation technologies was published (Wheeler et al. 2008). At that time, the bioinformatics tools and genomic resources to facilitate the in-depth analysis of whole-genome data could be considered in their infancy compared with today's standards. Even so, the insights gained into both the approach taken to sequencing the tumor and the biology of the tumor itself were profound when compared with the targeted sequencing approaches commonly applied to cancer research at the time.

Today, the resources required for WGS analysis have decreased substantially. Alongside a

steady reduction in the cost of WGS, there have been improvements in the technologies for generating and processing quality raw data as well as the tools and companion datasets that contextualize findings for biological and clinical interpretation. However, a majority of cancer genomics efforts remain focused around targeted deep sequencing and whole-exome sequencing (WES) (Morris et al. 2017; Raphael et al. 2017).

LARGE-SCALE EFFORTS TO CHARACTERIZE GENOMIC EVENTS IN CANCER

Large-scale efforts using genome sequencing to characterize a wide variety of adult and pediatric cancers began in earnest as early as 2005. This included projects such as The Cancer Genome Atlas (TCGA; see cancergenome.nih.gov), the International Cancer Genome Consortium (ICGC; see icgc.org), Catalog of Somatic Mutations in Cancer (COSMIC; see cancer.sanger.ac.uk), and Therapeutically Applicable Research to Generate Effective Treatments (TARGET; see ocg.cancer.gov/programs/target), to name but a few. Not only have such efforts progressed our understanding of cancer as a genomic disease, they also provide the data needed for developing tools and resources that facilitate the rapid detection and analysis of potentially relevant genomic events (Cerami et al. 2012; Gao et al. 2013; Gonzalez-Perez et al. 2013; Rubio-Perez et al. 2015). However, because the bulk of the data produced is focused on the coding region of the genome, the available data are underpowered to inform how untranslated, intronic, and intergenic regions might impact the molecular pathogenesis of disease (Nik-Zainal et al. 2016). In many cases, the data also lack comprehensive clinical annotation, which is required for linking genomic events to specific cancer types, prognoses, and treatment responses (Robinson et al. 2017). Furthermore, the majority of samples in these cohorts are from primary untreated disease and do not offer insight into how tumors respond to often complex and disparate treatment regimens (Robinson et al. 2017). Additional cancer cohorts of samples from multiple time points and biopsy sites that include rich clinical information are therefore still required to better define tumor biology and the relationship to treatment history and response.

WHOLE-GENOME CHARACTERIZATION OF CANCERS

The number of tumor whole-genome sequences that have been published and made publicly available has steadily increased over the past 10 years. These analyses have led to surprising insights into cancer biology, particularly from the analysis of structural variants (SVs) in tumor genomes (Chong et al. 2017). They range in scope from characterization of cancer cell lines (Pleasance et al. 2010a,b) and N-of-One case reports with rich clinical detail (Ellis et al. 2012), to ultradeep sequencing of a single tumor to uncover clonal heterogeneity (Griffith et al. 2015). Larger scale efforts are also emerging, focused both at the characterization of the somatic mutation, including noncoding and SVs (Banerji et al. 2012; Alexandrov et al. 2013b; Wang et al. 2014; Nik-Zainal et al. 2016) and germline-specific analysis for the discovery of predisposing and pharmacogenomic information (Foley et al. 2015). For example, Nik-Zainal et al. (2016) sequenced the whole genomes of 560 breast cancers, 260 supplemented with transcriptome sequencing. They showed how such approaches fill gaps in our understanding of the genome between the exons and expand the known repertoire of biological mechanisms underlying tumorigenesis with potential clinical use (Alexandrov et al. 2013b; Alexandrov and Stratton 2014; Davies et al. 2017; Zolkind and Uppaluri 2017).

WHOLE-GENOME CHARACTERIZATION OF CANCERS IN THE CLINIC

The scope of sequencing and its applications has also broadened into the clinic. For specific cancer types, the use of targeted genomic panels for both germline susceptibility and known "actionable" somatic mutations is becoming routine in many cancer centers (De Leeneer et al. 2011; Bosdet et al. 2013). The development and application of larger-scale gene panels is also seeing routine use on a large number of patient samples

(Zehir et al. 2017). These high-throughput approaches drive discovery in the clinical context and quantify the frequency and prevalence of both well-characterized and novel variants in cancer-related genes. However, they capture a tiny fraction of the genomic complexity that can exist in an individual tumor.

The first published description of an attempt to characterize a whole genome for a clinical application was in 2010 (Jones et al. 2010). In this analysis, Jones et al. sequenced an adenocarcinoma of the tongue, identifying genomic amplification and concurrent abundant expression of the RET oncogene as a potential driver of the disease. This analysis led to a personalized treatment approach for the patient using kinase inhibitors targeting the RET protein. Subsequent analysis of a posttreatment sample after disease progression provided comprehensive insight into how the tumor evolved to circumvent the treatment regimen in a way that a targeted approach could not have achieved. The success of this study led to a pilot for the Personalized OncoGenomics clinical trial, now in its fifth year, which aims to leverage whole-genome analysis with the intent-to-treat based on the genomic information. Sequencing of the initial 100 patients on this trial required development of pipelines and comprehensive interpretation tools (Laskin et al. 2015) and provided the framework for a more unbiased approach to cancer type selection. One benefit of this approach is in facilitating the sequencing of rare tumor types that might not otherwise be profiled so extensively (Bose et al. 2015; Wrzeszczynski et al. 2017). More recently, other whole-genome clinical projects have been announced or are underway (Wrzeszczynski et al. 2017) with perhaps the most ambitious project to date being the 100,000 Genomes Project (see genomicsengland .co.uk), which is sequencing whole genomes for the purpose of better understanding rare and infectious diseases, and includes a common cancer component.

SUMMARY

With an ongoing debate about the cost-effectiveness and analytical complexity of WGS for characterizing cancer, the unique capabilities of such an approach require consideration. What benefits could these capabilities provide if introduced into standard clinical practice? This is particularly timely given the insights over the past few years into the potential clinical actionability of mutation signatures that can elicit immune responses in the presence of immune checkpoint inhibitors (Rizvi et al. 2015). Here, we will highlight the unique opportunities (Fig. 1) for discovery from WGS, as they relate to the understanding of the underlying biology of a tumor, the potential impact on drug discovery and success of clinical trials and how those discoveries can impact the treatment of an individual's disease. We also consider the resources needed to build the infrastructure for WGS in routine clinical practice.

UNIQUE CAPABILITIES OF WGS

Toward a Digital Karyotype: High-Resolution Structural Variation

Few images are more deeply enscribed in the history of human genetics than the karyotype, which emphasizes the functional importance of genomic organization. Over the years, a plethora of technologies have enabled inspection of SVs and copy number variants (CNVs) in the genome. Some, such as fluorescence in situ hybridization, are precisely targeted, whereas others, like array-comparative genomic hybridization, are comprehensive at varying resolutions. WGS promises to deliver precise and comprehensive characterization of both SVs and CNVs. Although many challenges stand in the way of achieving a complete "digital karyotype" of cancer, WGS has made dramatic advances toward this goal.

The cancer genome frequently features complex and interlocking patterns of somatic SVs, which expands the realm of possible cancer-driving alterations. Since the discovery of the Philadelphia chromosome, the characterization of oncogenic fusions has been central to cancer diagnosis and treatment. Aside from fusions, SVs also modulate gene regulation by rearranging the noncoding genome. Variants that
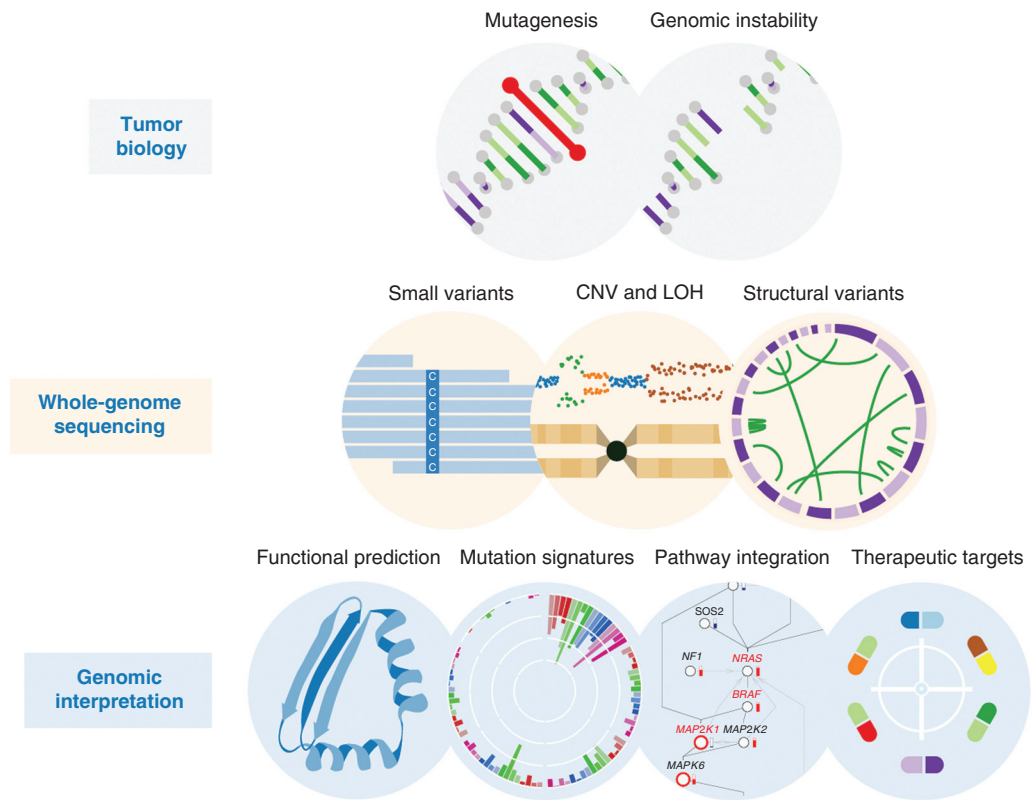
Figure 1. Whole-genome sequencing (WGS) data reveal diverse forms of genomic alteration. Tumor genomes show frequent mutation and genomic instability, which drive the hallmarks of cancer. WGS can catalog various forms of genomic alteration, enabling integrative analyses of tumor biology.

impact the copy number or relative positioning of promoters and other regulatory elements can alter gene expression (Alaei-Mahabadi et al. 2017). Last, the more recent discovery of complex SVs in cancer suggests that a broad repertoire of variation exists, which has previously gone unappreciated.

Paired-end WGS has become the standard for comprehensively and precisely cataloguing SVs and CNVs. Although targeted arrays and WES can provide comparative read counts for CNV analysis, they lack the resolution to detect microamplifications and microdeletions and suffer from sequencing depth bias. These challenges, along with a need for computational methods to address them, limit the accuracy of CNV calls and result in high false discovery rates (Zare et al. 2017). Although WES can detect gene fusions (Chmielecki et al. 2013), it misses fusions

affecting splice sites, promoters, and other functionally critical loci.

SV analysis methods fall broadly into four categories: read density, split reads, paired-end reads, and de novo assembly (Liu et al. 2015; Tattini et al. 2015). Recent methods combine these strategies to capitalize on the unique strengths of each. For example, DELLY improves sensitivity by considering paired end reads while enabling base-pair breakpoint calling precision by examining split reads (Rausch et al. 2012). A promising strategy, which has gained substantial recent interest, is the use of genome-wide local assembly at candidate breakpoints (Chong et al. 2017; Wala et al. 2017) to leverage some benefits of de novo assembly at substantially lower computational cost.

A major challenge in digital karyotype construction is poor concordance between SVs and

CNV breakpoints (Alaei-Mahabadi et al. 2017), which suggests mismatched accuracies and thresholds between modalities. This mismatch causes further difficulty in reconstructing complex rearrangements such as chromothripsis (Stephens et al. 2011) and chromoplexy (Baca et al. 2013). Without refined methods or costly assembly-based approaches, clinically relevant complex mutations can be missed by conventional next-generation sequencing (NGS) methods (Robertson et al. 2017). Graph-based methods are an emerging approach, which may help to align SV and CNV calls to better reconstruct complex variants and even resolve the temporal ordering of rearrangement events (Greenman et al. 2012; Maciejowski and Imielinski 2016).

SVs and CNVs play important roles in the somatic alteration of cancer genes and are best comprehensively characterized by WGS. Evolving methodologies promise to move the field toward increasingly accurate estimates of the digital karyotype of cancer.

Genomic Instability

Genome instability and mutagenesis are hallmark characteristics of cancers (Hanahan and Weinberg 2011). Endogenous and exogenous exposures coupled with failures of in-built DNA repair mechanisms give rise to rampant mutation and rearrangement. Total mutational burden (TMB) is a differentiating quality between and within tumor types. Analyses of somatic variant frequencies across tumor-normal pairs reveal somatic mutation rates from 0.001 to 1000 mutations per megabase (Alexandrov et al. 2013a; Lawrence et al. 2013), which vary by cancer type and individual tumor features. SVs also show heterogeneous patterns of occurrence, with substantial differences in mutation burden and distributions even between tumor subtypes (Waddell et al. 2015; Nik-Zainal et al. 2016).

The clinical implications of genomic instability have received increased attention in recent years. The advent of immunotherapy has bolstered interest in TMB as a potential predictive biomarker, especially as it relates to the emergence of candidate immunogenic neoantigens. Programmed death ligand 1 (PD-L1) inhibitor immunotherapy response is also associated with mutagenesis arising from mismatch repair deficiency (MMRD), which has conventionally been detected by immunohistochemistry staining of MMRD genes and sequencing to detect microsatellite instability (MSI) sites (Le et al. 2015; Iyer et al. 2017). Homologous recombination deficiency (HRD) is another source of mutation with therapeutic promise. It involves loss of *BRCA1*, *BRCA2*, *RAD51*, and other genes linked to poly(ADP ribose) polymerase (PARP) inhibitor sensitivity (Robson et al. 2017).

The recent interest in genomic instability as a therapeutic target reflects substantial recent advancement in the use of NGS to precisely catalog mutation. Beyond actionable drivers, WGS can detect somatic variants arising from specific mutagenic mechanisms. Because each mutagenic source produces a characteristic mutational pattern, unsupervised learning techniques can decipher signatures associated with distinct etiologies. Mutation signatures have also been referred to as "genomic scars," a term suggestive of lasting and detectable DNA damage. This approach considers the somatic cancer genome as a functional readout of its mutagenic exposures and the integrity of its DNA maintenance and repair pathways (Helleday et al. 2014).

The discovery of mutation signatures of somatic single nucleotide variants (SNVs) has revealed the interacting roles of diverse carcinogenic etiologies. To date, 30 "consensus" signatures have been cataloged (see cancer.sanger.ac.uk/cosmic/signatures) thanks to large-scale analyses of publicly available cancer datasets (Alexandrov et al. 2013a). Although SNV signatures can be deciphered from targeted gene sequencing (Zehir et al. 2017), broad coverage across the whole genome provides the most precise characterization (Alexandrov et al. 2013b). Signatures of SV have also been identified, which require sensitive and specific SV capture and calling across the whole genome (Nik-Zainal et al. 2016). Meanwhile, the demand for genome-guided cancer therapy has motivated the development of N-of-One approaches to mutation signature characterization (Rosenthal et al. 2016).

Mutation signatures layer etiology-specific information atop TMB. The presence of signa-

tures specific to certain cancer types (i.e., mutations associated with cigarette smoking, ultraviolet radiation) may therefore help inform the classification of cancers of uncertain primary. Mutation signatures have also begun to spur research into clinical applications. In the following subsections, we will describe two specific examples of DNA repair defects associated with putative therapeutic implications: MMRD and HRD.

Mismatch Repair Deficiency

Mismatch repair (MMR) repairs erroneous base pairing and its deficiency produces cancer risk (Lynch syndrome) and distinct signatures of mutation (Signature 6 and MSI). MMR repairs DNA lesions caused by the mispairing of bases, which can result from errors in replication or mutagenic DNA damage processes. Germline mutations in MMR genes *MSH2*, *MLH1*, *MSH6*, and *PMS2* give rise to Lynch syndrome. Genetic panel sequencing and immunohistochemistry are commonly used to assess the genotype and expression of these genes. MMRD is also associated with MSI, the genome-wide shrinkage or expansion of short repetitive sequences known as microsatellites. MSI is conventionally detected by polymerase chain reaction (PCR) amplification and electrophoresis at five standard microsatellite regions to map the variability in microsatellite length.

MMRD tumors are remarkable for their immunogenicity, possibly by expressing neoepitopes resulting from base modifications. This feature suggests a therapeutic link to immune checkpoint inhibitors. Many recent studies have shown improved checkpoint inhibitor response in patients with MMRD tumors across diverse cancer types (Le et al. 2017). The ability to assess both the genotypes and molecular phenotypes associated with MMRD is becoming a critical element of personalized tumor analysis.

NGS is informative both for genotyping MMR genes and quantifying their impacts on genomic maintenance. Reads captured by NGS can be used to analyze microsatellite length distributions. Some MSI detection strategies have used targeted deep sequencing of clinically validated MSI loci (Gan et al. 2015), whereas others screen reads for MSI sites incidentally captured during sequencing (Niu et al. 2014; Salipante et al. 2014). Both approaches yield results similar to gold standard clinical PCR methods. Mutation signatures also provide insight into MMR. Two out of 21 mutation signatures deciphered from publicly available cancer genomes and exomes have been putatively linked to MMRD (Alexandrov et al. 2013a).

The analysis of MMR by NGS technologies is becoming increasingly prevalent in precision oncology initiatives. The recent MSK-IMPACT analysis of 10,000 cancers by NGS found 102 individuals with MMRD, concordantly classified by mutation signatures and MSI detection. This work supported two recent conference abstracts reporting associations between MSI detected by NGS and response to immune checkpoint inhibitors in urothelial and esophagogastric carcinoma (Iyer et al. 2017; Ku et al. 2017).

Homologous Recombination Deficiency

Homologous recombination (HR) refers to the exchange of similar or identical nucleotide sequences and has long been used as a laboratory technique for site-directed mutagenesis. However, in cancer, HR facilitates error-free repair of DNA double-strand breaks, as well as inter- and intrastrand cross-links. The well-known cancer genes *BRCA1* and *BRCA2* centrally coordinate HR, and mutations in these genes confer an up to 85% lifetime risk of breast and ovarian cancer (see cancer.ca/en/cancer-information/cancer-101/what-is-a-risk-factor/genetic-risk/brca-gene-mutations and cancer.gov/cancertopics/factsheet/Risk/BRCA). Other HR genes have also been associated with cancer, including *ATM*, *RAD50*, *RAD51*, *PALB2*, *XRCC2/3*, and the FANC gene family (Cerbinskaite et al. 2012; Roy et al. 2012).

HRD has been identified as a promising target for the administration of PARP inhibitors (Gelmon et al. 2011; Kaufman et al. 2013), as well as DNA-damaging agents such as platinum-based chemotherapy and anthracyclines (Kennedy et al. 2004; Farmer et al. 2005; Yang et al. 2011). This link is motivated by substantial

recent evidence linking germline BRCA1 and BRCA2 variants with sensitivity to PARP inhibitors and platinum-based chemotherapy in breast and ovarian cancers (Byrski et al. 2010; Arun et al. 2011; Von Minckwitz et al. 2014; Sikov et al. 2015; Tutt et al. 2015).

The first mutation signatures of HRD discovered by NGS were characteristic large-scale CNV patterns detected by hybrid-capture panels with relatively evenly spaced probes. This approach yielded three quantifiable scores (HRD-LOH, HRD-TAI, and HRD-LST), which correlated with *BRCA1/2* mutation status (Timms et al. 2014) and response to platinum-containing neoadjuvant chemotherapy in primary breast cancers (Telli et al. 2016). However, in the advanced breast cancer setting, HRD scores were not associated with platinum response based on the phase 3 TNT trial (Tutt et al. 2015).

More recently, the advent of SNV and SV mutation signature analyses has uncovered additional HRD-associated signatures (Alexandrov et al. 2013b; Nik-Zainal et al. 2016). Overlapping microhomology at deletions is another feature observed in cancers with HRD (Stephens et al. 2012; Davies et al. 2017). By combining SNV signatures, SV signatures, the HRD score, and overlapping microhomology, Davies et al. (2017) formulated a model capable of sensitive and specific prediction of BRCA1/2 mutation status called HRDetect. This analysis improves our understanding of the downstream mutational impacts of HRD and shows a clear link with its etiology. Moreover, a recent WGS study in 93 advanced-stage breast cancers found HRDetect to be associated with response to platinum-based chemotherapies (Zhao et al. 2017).

The detection of MMRD and HRD both represent important advances in the understanding of genomic instability and the applicability of NGS techniques. Although targeted sequencing and exome analysis have proven sufficient for MMRD analysis, the latest HRD signatures demand genome-wide capture of SNVs, SVs, indels, and CNVs. WGS is well suited for each of these tasks and can provide a wholesale solution obviating the need for additional assays.

Variation in Noncoding Regions

The Human Genome Project estimated that protein-coding exons comprise 1.2% of the human genome (Human Genome Sequencing Consortium 2004). Interpreting the remaining noncoding DNA remains a challenge, but advances in epigenetics have revealed extensive roles of DNA beyond protein coding. Findings from The Encyclopedia of DNA Elements (ENCODE) Project suggest that the majority of the genome is involved in biochemical interactions, such as protein binding or chromatin remodeling (ENCODE Project Consortium 2012).

An example of the significance of intergenic DNA in cancer are variants in the promoter region of telomerase reverse transcriptase (*TERT*), a gene central to the immortalization of cancer cells (Horn et al. 2013; Huang et al. 2013). These highly recurrent mutations are associated with two- to fourfold increases of *TERT* transcriptional activity and are found in multiple cancer types (Vinagre et al. 2013). Analyses from TCGA and others have identified additional recurrent regulatory element mutations in additional genes across multiple cancer types (Weinhold et al. 2014; Rheinbay et al. 2017), indicating that the number of relevant intergenic events is likely to increase.

The ability to detect noncoding mutations, especially in regulatory regions, is a strength of WGS. Large-scale efforts have produced a functional map of noncoding variation, which has facilitated the discovery of driver mutations in regulatory regions of the genome. However, further research is necessary to facilitate the interpretation of novel noncoding variants found by WGS in a personalized cancer medicine setting.

Complementarity with the Transcriptome

Alongside advancements in WGS, RNA sequencing technologies including whole-transcriptome sequencing (WTS) can provide complementary insights in a personalized medicine setting. WTS enables genome-wide quantification of gene expression, which substantially expands the capture of potentially actionable

molecular aberrations. Substantial efforts exist aiming to use tumor expression profiles to refine cancer diagnoses and subtyping. Translation of these efforts to interpretable and clinically actionable parameters such as the widely used PAM50 gene set in breast cancer (Chia et al. 2012) can further advance these aims. The integration of gene expression data into functional clusters or pathways using statistical methods or visualizations can help identify dysregulated cancer pathways. This can provide rationale for guiding targeted cancer treatment, including the use of experimental therapeutics (Tomasetti et al. 2017). When treatments fail, gene expression analysis can also elucidate resistance mechanisms and potentially suggest follow-up targets (Jones et al. 2010).

The existence and functional impacts of many events observed in the genome can be further analyzed by WTS. Amplified and deleted genes can be assessed for differential expression. The presence of oncogenic or deleterious mutations on the expressed transcript can be confirmed. Exon skipping and intron retention can be identified and potentially linked to splice site variants. Transcriptome assembly can facilitate detection of potentially oncogenic alternative transcripts. The presence of oncogenic gene fusions can be confirmed, and their expression verified. The effects of promoter and enhancer mutations on gene expression can be quantified. Tumor suppressors such as *BRCA1* can be assessed for potential epigenetic silencing.

However, WTS brings its own limitations. Foremost among them is the uncertainty that transcribed genes and variants will be faithfully translated in proportional quantities. Comparisons between RNA-level and protein-level expression frequently reveals weak, nonlinear correlations. WTS also cannot capture posttranslational processing and modifications, which have significant functional impacts. Correlation of WGS with proteomics technologies may enable further advances in personalized cancer care, but various technological, scientific, and logistic obstacles must first be overcome to catalyze clinical translation (Duarte and Spencer 2016; Kwon et al. 2016; Nice 2016).

Cancer Immunology

Genome sequencing has been at the heart of recent advances aiming to personalize cancer immunotherapy. In particular, the joint analysis of human leukocyte antigen (HLA) sequence and neoantigen repertoire holds promise for predictive patient stratification. Computational approaches for HLA genotyping can be performed using WGS, WES, or RNA-seq data (Szolek et al. 2014). Neoantigens can be cataloged based on peptide sequence modifications profiled by WGS or WES, and their binding affinity to major histocompatibility complex (MHC) proteins can be predicted using machine learning–based methods (Jurtz et al. 2017). Although most such analyses are limited to exonic regions, whole-genome analysis can be valuable for the discovery of novel features relevant to tumor immunology. For example, WGS can expand the profiling of neoantigens by incorporating fusion antigens (Zhang et al. 2016).

Emerging Capabilities of WGS

The past decade's explosion of short-read NGS technology has spurred more recent developments in long-read sequencing. Long-reads enable more reliable SV detection, as well as resolution of repeat-rich regions. Alternatively, linked-read technologies extend paired-end short-read sequencing by grouping multiple distant short reads by molecule, which enables haplotype phasing and improvements to SV detection (Spies et al. 2017) and de novo assembly (Jackman et al. 2017). Both long-read and linked-read sequencing could substantially improve the resolution of complex genomic events, moving toward precise and comprehensive cancer digital karyotypes.

The quantification of inter- and intratumor heterogeneity is also advancing toward potential clinical translation. Actionable driver events have a wide range of recurrent clonal or subclonal fractions (McGranahan et al. 2015), which suggests the need to include clonality as a dimension of clinical cancer sequence analysis. Moreover, the analysis of mutational clonality can indicate temporal shifts in the processes

that drive somatic mutation. Andor et al. (2015) recently showed that tumor heterogeneity, detected by NGS, is correlated with increased mortality. As clinical sequencing helps to formulate an improved understanding of events driving metastatic cancer (Robinson et al. 2017), analysis of tumor heterogeneity may eventually help to predict drug resistance.

CONCLUDING REMARKS

The inherent genomic complexity of cancers gives rise to a range of genomic events and signatures that are becoming increasingly relevant in patient-treatment stratification. However, the expectation that an individual's tumor will harbor previously described and functionally characterized genomic events is not certain. The successful clinical application of personalized genomic medicine therefore must rely on broad screening approaches, a conclusion that was also reached in a study comparing WES and

WGS in gastric cancer (Wang et al. 2014). A whole-genome approach is currently the most efficient way to build a comprehensive picture of the genomic variation in a tumor without resorting to multiple technical platforms. That being said, there are still significant challenges that must be overcome before approaches such as whole-genome and transcriptome analysis (WGTA) can be universally adopted. However, on the assumption that sequencing costs will follow the historical downward trend, a more gradual uptake of WGTA for more refined stratification and subtyping of rare tumors may be achievable in the short term. Furthermore, as unanticipated clinical successes from WGTA continue to permeate the field and the infrastructure to support such approaches mature, a transition away from targeted sequencing in the clinic is perhaps to be expected.

WGTA adds complexity to the analysis of genomic data, particularly given the potential for the discovery of novel variants that require
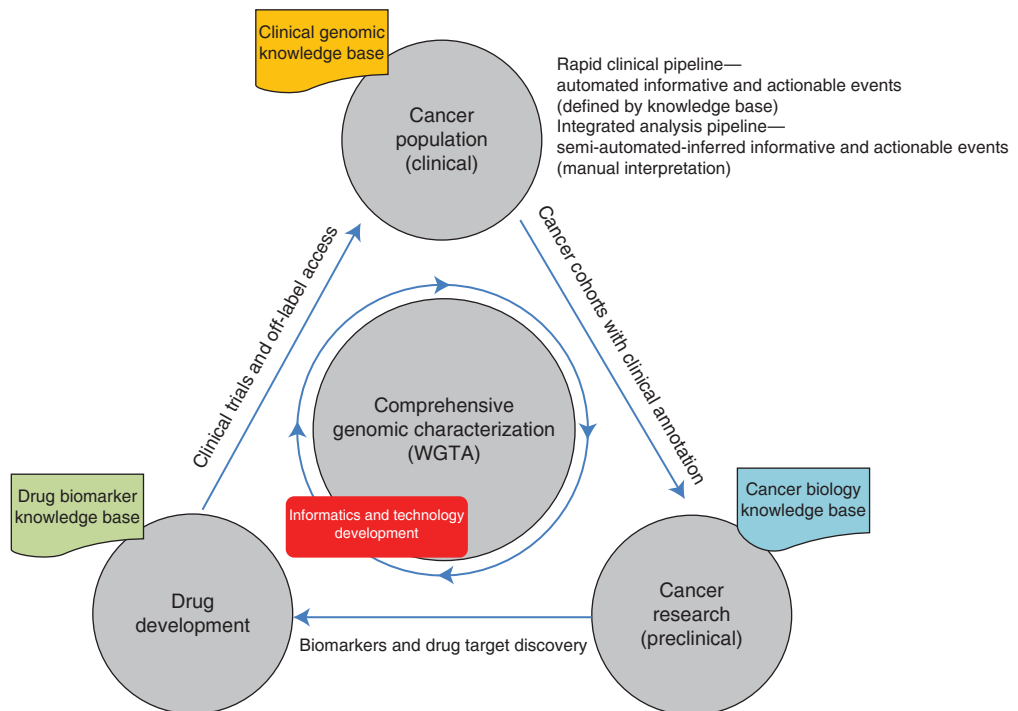

Figure 2. A model for whole-genome and transcriptome analysis (WGTA)-driven personalized genomics in the cancer clinic.

additional investigation and interpretation. Up-to-date and comprehensive knowledge bases that document current and emerging clinically relevant genomic information are therefore essential if identification and stratification of genomic results is to be achieved in a clinically relevant time frame across a broad range of clinical pipelines. To this end, a number of knowledge bases have been established in recent years, which attempt to collate, summarize, and serve up relevant genomic-therapeutic associations (for review, see Kumar-Sinha and Chinnaiyan 2018). Computational approaches leveraging natural language processing for capturing and curating new and existing knowledge are likely to be successful in increasing the efficiency of these efforts (Yim et al. 2016). Finally, algorithms that can reliably infer the potential relevance of novel genomic events and then link these events to hypothetical or inferred clinical action with limited manual intervention are needed. Their development will be critical for reducing analysis and interpretation costs, shortening turnaround time, and promoting widespread adoption of multi-omic approaches in the clinic.

For some forward-thinking jurisdictions, collating the complete genomic information for tumors coupled with extensive clinical information will provide an unprecedented research platform to understand the mechanisms underlying therapeutic response, acquired resistance, and failure. Likewise, we envisage that WGTA could be undertaken many times during the course of the disease, serially analyzing the changes accrued during various stages of treatment providing a real-time view of cancer progression and treatment response. Such studies will generate a feedback loop, which will be invaluable for the study of cancers as they relate to improved disease stratification and therapeutic intervention. Ultimately, clinical annotation of the therapeutic action taken, and outcome observed, would feed into large-scale analyses in the preclinical setting to further the understanding of the biological processes driving cancer and identifying rational biomarkers for cancer diagnostics and new drug development (Fig. 2).

REFERENCES

Alaei-Mahabadi B, Bhadury J, Karlsson JW, Nilsson JA, Larsson E. 2017. Global analysis of somatic structural genomic alterations and their impact on gene expression in diverse human cancers. *Proc Natl Acad Sci* **113:** 13768–13773.

Alexandrov LB, Stratton MR. 2014. Mutational signatures: The patterns of somatic mutations hidden in cancer genomes. *Curr Opin Genet Dev* **24:** 52–60.

Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA Jr, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale AL, et al. 2013a. Signatures of mutational processes in human cancer. *Nature* **500:** 415–421.

Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. 2013b. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep* **3:** 246–259.

Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, Ji HP, Maley CC. 2015. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat Med* **22:** 105–113.

Arun B, Bayraktar S, Liu DD, Barrera AMG, Atchley D, Pusztai L, Litton JK, Valero V, Meric-Bernstam F, Hortobagyi GN. 2011. Response to neoadjuvant systemic therapy for breast cancer in *BRCA* mutation carriers and noncarriers: A single-institution experience. *J Clin Oncol* **29:** 3739–3746.

Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, Park K, Kitabayashi N, MacDonald TY, Ghandi M, et al. 2013. Punctuated evolution of prostate cancer genomes. *Cell* **153:** 666–677.

Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, Lawrence MS, Sivachenko AY, Sougnez C, Zou L, et al. 2012. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* **486:** 405–409.

Bosdet IE, Docking TR, Butterfield YS, Mungall AJ, Zeng T, Coope RJ, Yorida E, Chow K, Bala M, Young SS, et al. 2013. A clinically validated diagnostic second-generation sequencing assay for detection of hereditary *BRCA1* and *BRCA2* mutations. *J Mol Diagn* **15:** 796–809.

Bose P, Pleasance ED, Jones M, Shen Y, Ch'ng C, Reisle C, Schein JE, Mungall AJ, Moore R, Ma Y. 2015. Integrative genomic analysis of ghost cell odontogenic carcinoma. *Oral Oncol* **51:** e71–e75.

Byrski T, Gronwald J, Huzarski T, Grzybowska E, Budryk M, Stawicka M, Mierzwa T, Szwiec M, Wiśniowski R, Siolek M. 2010. Pathologic complete response rates in young women with *BRCA1*-positive breast cancers after neoadjuvant chemotherapy. *J Clin Oncol* **28:** 375–379.

Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. 2012. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2:** 401–404.

Cerbinskaite A, Mukhopadhyay A, Plummer ER, Curtin NJ, Edmondson RJ. 2012. Defective homologous recombination in human cancers. *Cancer Treat Rev* **38:** 89–100.

Chia SK, Bramwell VH, Tu D, Shepherd LE, Jiang S, Vickery T, Mardis E, Leung S, Ung K, Pritchard KI, et al. 2012. A 50-gene intrinsic subtype classifier for prognosis and pre-

diction of benefit from adjuvant tamoxifen. *Clin Cancer Res* **18:** 4465–4472.

Chmielecki J, Crago AM, Rosenberg M, O'Connor R, Walker SR, Ambrogio L, Auclair D, McKenna A, Heinrich MC, Frank DA, et al. 2013. Whole-exome sequencing identifies a recurrent NAB2-STAT6 fusion in solitary fibrous tumors. *Nat Genet* **45:** 131–132.

Chong Z, Ruan J, Gao M, Zhou W, Chen T, Fan X, Ding L, Lee AY, Boutros P, Chen J, et al. 2017. novoBreak: Local assembly for breakpoint detection in cancer genomes. *Nat Methods* **14:** 65–67.

Davies H, Morganella S, Purdie CA, Jang SJ, Borgen E, Russnes H, Glodzik D, Zou X, Viari A, Richardson AL, et al. 2017. Whole-genome sequencing reveals breast cancers with mismatch repair deficiency. *Cancer Res* **77:** 4755–4762.

De Leeneer K, Hellemans J, De Schrijver J, Baetens M, Poppe B, Van Criekinge W, De Paepe A, Coucke P, Claes K. 2011. Massive parallel amplicon sequencing of the breast cancer genes *BRCA1* and *BRCA2*: Opportunities, challenges, and limitations. *Hum Mutat* **32:** 335–344.

Duarte T, Spencer C. 2016. Personalized proteomics: The future of precision medicine. *Proteomes* **4:** 29.

Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, Van Tine BA, Hoog J, Goiffon RJ, Goldstein TC, et al. 2012. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* **486:** 353–360.

ENCODE Project Consortium TEP. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489:** 57–74.

Farmer H, McCabe N, Lord CJ, Tutt ANJ, Johnson DA, Richardson TB, Santarosa M, Dillon KJ, Hickson I, Knights C, et al. 2005. Targeting the DNA repair defect in *BRCA* mutant cells as a therapeutic strategy. *Nature* **434:** 917–921.

Foley SB, Rios JJ, Mgbemena VE, Robinson LS, Hampel HL, Toland AE, Durham L, Ross TS. 2015. Use of whole genome sequencing for diagnosis and discovery in the cancer genetics clinic. *eBioMedicine* **2:** 74–81.

Gan C, Love C, Beshay V, Macrae F, Fox S, Waring P, Taylor G. 2015. Applicability of next generation sequencing technology in microsatellite instability testing. *Genes* **6:** 46–59.

Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, et al. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6:** pl1.

Gelmon KA, Tischkowitz M, Mackay H, Swenerton K, Robidoux A, Tonkin K, Hirte H, Huntsman D, Clemons M, Gilks B, et al. 2011. Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: A phase 2, multicentre, open-label, non-randomised study. *Lancet Oncol* **12:** 852–861.

Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, Santos A, Lopez-Bigas N. 2013. IntOGen-mutations identifies cancer drivers across tumor types. *Nat Methods* **10:** 1081–1082.

Greenman C, Pleasance E, Newman S, Yang F, Fu B, Nik-Zainal S, Jones D, Lau K, Carter N, Edwards P, et al. 2012. Estimation of rearrangement phylogeny for cancer genomes. *Genome Res* **22:** 346–361.

Griffith M, Miller CA, Griffith OL, Krysiak K, Skidmore ZL, Ramu A, Walker JR, Dang HX, Trani L, Larson DE, et al. 2015. Optimizing cancer genome sequencing and analysis. *Cell Systems* **1:** 210–223.

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: The next generation. *Cell* **144:** 646–674.

Helleday T, Eshtad S, Nik-Zainal S. 2014. Mechanisms underlying mutational signatures in human cancers. *Nat Rev Genet* **15:** 585–598.

Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A, Kadel S, Moll I, Nagore E, Hemminki K, et al. 2013. *TERT* promoter mutations in familial and sporadic melanoma. *Science* **339:** 959–961.

Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA. 2013. Highly recurrent *TERT* promoter mutations in human melanoma. *Science* **339:** 957–959.

Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. *Nature* **431:** 931–945.

Iyer G, Audenet F, Middha S, Carlo MI, Regazzi AM, Funt S, Al-Ahmadie H, Solit DB, Rosenberg JE, Bajorin DF. 2017. Mismatch repair (MMR) detection in urothelial carcinoma (UC) and correlation with immune checkpoint blockade (ICB) response. *J Clin Oncol* **35:** 4511.

Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, Jahesh G, Khan H, Coombe L, Warren RL, et al. 2017. ABySS 2.0: Resource-efficient assembly of large genomes using a Bloom filter. *Genome Res* **27:** 768–777.

Jones SJM, Laskin J, Li YY, Griffith OL, An J, Bilenky M, Butterfield YS, Cezard T, Chuah E, Corbett R. 2010. Evolution of an adenocarcinoma in response to selection by targeted kinase inhibitors. *Genome Biol* **11:** 1–12.

Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. 2017. NetMHCpan-4.0: Improved peptide–MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol* **199:** 3360–3368.

Kaufman B, Shapira-Frommer R, Schmutzler RK, Audeh MW, Friedlander M, Balmana J, Mitchell G, Fried G, Bowen K, Fielding A. 2013. Olaparib monotherapy in patients with advanced cancer and a germ-line *BRCA1/2* mutation: An open-label phase II study. *J Clin Oncol* **31:** 11024.

Kennedy RD, Quinn JE, Mullan PB, Johnston PG, Harkin DP. 2004. The role of *BRCA1* in the cellular response to chemotherapy. *J Natl Cancer Inst* **96:** 1659–1668.

Ku GY, Sanchez-Vega F, Chatila W, Margolis M, Fein C, Ilson DH, Hechtman JF, Tuvy Y, Bouvier N, Kundra R, et al. 2017. Correlation of benefit from immune checkpoint inhibitors with next gen sequencing (NGS) profiles in esophagogastric cancer (EGC) patients. *J Clin Oncol* **35:** 4025.

Kumar-Sinha C, Chinnaiyan AM. 2018. Precision oncology in the age of integrative genomics. *Nat Biotechnol* **36:** 46–60.

Kwon HJ, Fehniger TE, Marko-Varga G. 2016. Building the basis for proteomics in personalized medicine for targeted treatment. *Clin Transl Med* **5:** 19.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W,

et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409:** 860–921.

Laskin J, Jones S, Aparicio S, Chia S, Ch'ng C, Deyell R, Eirew P, Fok A, Gelmon K, Ho C, et al. 2015. Lessons learned from the application of whole-genome analysis to the treatment of patients with advanced cancers. *Cold Spring Harb Mol Case Stud* **1:** a000570.

Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, et al. 2013. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499:** 214–218.

Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, et al. 2015. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* **372:** 2509–2520.

Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, et al. 2017. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* **357:** 409–413.

Liu B, Conroy JM, Morrison CD, Odunsi AO, Qin M, Wei L, Trump DL, Johnson CS, Liu S, Wang J. 2015. Structural variation discovery in the cancer genome using next generation sequencing: Computational solutions and perspectives. *Oncotarget* **6:** 5477–5489.

Maciejowski J, Imielinski M. 2016. Modeling cancer rearrangement landscapes: From pattern to mechanism, and back. *Curr Opin Syst Biol* **1:** 54–61.

McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C. 2015. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci Transl Med* **7:** 283ra54.

Morris V, Rao X, Pickering C, Foo WC, Rashid A, Eterovic K, Kim T, Chen K, Wang J, Shaw K, et al. 2017. Comprehensive genomic profiling of metastatic squamous cell carcinoma of the anal canal. *Mol Cancer Res* **15:** 1542–1550.

Nice EC. 2016. From proteomics to personalized medicine: The road ahead. *Expert Rev Proteomics* **13:** 341–343.

Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC. 2016. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534:** 47–54.

Niu B, Ye K, Zhang Q, Lu C, Xie M, McLellan MD, Wendl MC, Ding L. 2014. MSIsensor: Microsatellite instability detection using paired tumor-normal sequence data. *Bioinformatics* **30:** 1015–1016.

Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin ML, Ordóñez GR, Bignell GR, et al. 2010a. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* **463:** 191–196.

Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, Jones D, Lin ML, Beare D, Lau KW, Greenman C, et al. 2010b. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463:** 184–190.

Raphael BJ, Hruban RH, Aguirre AJ, Moffitt RA, Yeh JJ, Stewart C, Robertson AG, Cherniack AD, Gupta M, Getz G, et al. 2017. Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell* **32:** 185–203.e13.

Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. 2012. DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28:** i333–i339.

Rheinbay E, Parasuraman P, Grimsby J, Tiao G, Engreitz JM, Kim J, Lawrence MS, Taylor-Weiner A, Rodriguez-Cuevas S, Rosenberg M, et al. 2017. Recurrent and functional regulatory mutations in breast cancer. *Nature* **547:** 55–60.

Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS. 2015. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348:** 124–128.

Robertson AG, Shih J, Yau C, Gibb EA, Oba J, Mungall KL, Hess JM, Uzunangelov V, Walter V, Danilova L, et al. 2017. Integrative analysis identifies four molecular and clinical subsets in uveal melanoma. *Cancer Cell* **32:** 204–220.e15.

Robinson DR, Wu YM, Lonigro RJ, Vats P, Cobain E, Everett J, Cao X, Rabban E, Kumar-Sinha C, Raymond V, et al. 2017. Integrative clinical genomics of metastatic cancer. *Nature* **548:** 297–303.

Robson M, Im SA, Senkus E, Xu B, Domchek SM, Masuda N, Delaloge S, Li W, Tung N, Armstrong A, et al. 2017. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. *N Engl J Med* **377:** 523–533.

Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. 2016. DeconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol* **17:** 31.

Roy R, Chun J, Powell SN. 2012. *BRCA1* and *BRCA2*: Different roles in a common pathway of genome protection. *Nat Rev Cancer* **12:** 68–78.

Rubio-Perez C, Tamborero D, Schroeder MP, Antolín AA, Deu-Pons J, Perez-Llamas C, Mestres J, Gonzalez-Perez A, Lopez-Bigas N. 2015. In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. *Cancer Cell* **27:** 382–396.

Salipante SJ, Scroggins SM, Hampel HL, Turner EH, Pritchard CC. 2014. Microsatellite instability detection by next generation sequencing. *Clin Chem* **60:** 1192–1199.

Sikov WM, Barry WT, Hoadley KA, Pitcher BN, Singh B, Tolaney SM, Kuzma CS, Pluard TJ, Somlo G, Port ER. 2015. Impact of intrinsic subtype by PAM50 and other gene signatures on pathologic complete response (pCR) rates in triple-negative breast cancer (TNBC) after neoadjuvant chemotherapy (NACT) plus/-carboplatin (Cb) or bevacizumab (Bev): CALGB 40603/150709 (Alliance). *2014 San Antonio Breast Cancer Symposium*, Abstract S4-05. San Antonio, TX, December 9–13.

Spies N, Weng Z, Bishara A, McDaniel J, Catoe D, Zook JM, Salit M, West RB, Batzoglou S, Sidow A. 2017. Genome-wide reconstruction of complex structural variants using read clouds. *Nat Methods* **14:** 915–920.

Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, et al. 2011. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144:** 27–40.

Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, Zainal SN, Martin S, Varela I, Bignell GR, et al.

2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486:** 400.

Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. 2014. OptiType: Precision HLA typing from next-generation sequencing data. *Bioinformatics* **30:** 3310–3316.

Tattini L, D'Aurizio R, Magi A. 2015. Detection of genomic structural variants from next-generation sequencing data. *Front Bioeng Biotechnol* **3:** 92.

Telli ML, Timms KM, Reid J, Hennessy B, Mills GB, Jensen KC, Szallasi Z, Barry WT, Winer EP, Tung NM. 2016. Homologous recombination deficiency (HRD) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. *Clin Cancer Res* **22:** 3764–3773.

The Cancer Genome Atlas Network. 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* **368:** 2059–2074.

Timms KM, Abkevich V, Hughes E, Neff C, Reid J, Morris B, Kalva S, Potter J, Tran TV, Chen J. 2014. Association of *BRCA1/2* defects with genomic scores predictive of DNA damage repair deficiency among breast cancer subtypes. *Breast Cancer Res* **16:** 475.

Tomasetti C, Li L, Vogelstein B. 2017. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* **355:** 1330–1334.

Tutt A, Ellis P, Kilburn L, Gilett C, Pinder S, Abraham J, Barrett S, Barrett-Lee P, Chan S, Cheang M. 2015. Abstract S3-01: The TNT trial: A randomized phase III trial of carboplatin (C) compared with docetaxel (D) for patients with metastatic or recurrent locally advanced triple negative or *BRCA1/2* breast cancer (CRUK/07/012). *Cancer Res* **75:** S3–S01.

Vinagre J, Almeida A, Pópulo H, Batista R, Lyra J, Pinto V, Coelho R, Celestino R, Prazeres H, Lima L, et al. 2013. Frequency of *TERT* promoter mutations in human cancers. *Nat Commun* **4:** 2185.

Von Minckwitz G, Hahnen E, Fasching PA, Hauke J, Schneeweiss A, Salat C, Rezai M, Blohmer JU, Zahm DM, Jackisch C. 2014. Pathological complete response (PCR) rates after carboplatin-containing neoadjuvant chemotherapy in patients with germline *BRCA* (*gBRCA*) mutation and triple-negative breast cancer (TNBC): Results from GeparSixto. *American Society of Clinical Oncology (ASCO) Annual Meeting*, Abstract 1005. Chicago, IL, May 30–June 3.

Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, Johns AL, Miller D, Nones K, Quek K, et al. 2015. Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* **518:** 495–501.

Wala J, Bandopadhayay P, Greenwald N, O'Rourke R, Sharpe T, Stewart C, Schumacher SE, Li Y, Weischenfeldt J, Yao X, et al. 2017. Genome-wide detection of structural variants and indels by local assembly. *bioRxiv* doi: 10.1101/105080.

Wang K, Yuen ST, Xu J, Lee SP, Yan HHN, Shi ST, Siu HC, Deng S, Chu KM, Law S, et al. 2014. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat Genet* **46:** 573–582.

Weinhold N, Jacobsen A, Schultz N, Sander C, Lee W. 2014. Genome-wide analysis of noncoding regulatory mutations in cancer. *Nat Genet* **46:** 1160–1165.

Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen Y-J, Makhijani V, Roth GT, et al. 2008. The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452:** 872–876.

Wrzeszczynski KO, Frank MO, Koyama T, Rhrissorrakrai K, Robine N, Utro F, Emde AK, Chen BJ, Arora K, Shah M, et al. 2017. Comparing sequencing assays and human-machine analyses in actionable genomics for glioblastoma. *Neurol Genet* **3:** e164.

Yang D, Khan S, Sun Y, Hess K, Shmulevich I, Sood AK, Zhang W. 2011. Association of *BRCA1* and *BRCA2* mutations with survival, chemotherapy sensitivity, and gene mutator phenotype in patients with ovarian cancer. *JAMA* **306:** 1557–1565.

Yim Ww, Yetisgen M, Harris WP, Kwan SW. 2016. Natural language processing in oncology. *JAMA Oncol* **2:** 797.

Zare F, Dow M, Monteleone N, Hosny A, Nabavi S. 2017. An evaluation of copy number variation detection tools for cancer using whole exome sequencing data. *BMC Bioinformatics* **18:** 286.

Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, Srinivasan P, Gao J, Chakravarty D, Devlin SM, et al. 2017. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* **23:** 703–713.

Zhang J, Mardis ER, Maher CA. 2016. INTEGRATE-neo: A pipeline for personalized gene fusion neoantigen discovery. *Bioinformatics* **33:** 555–557.

Zhao EY, Shen Y, Pleasance E, Kasaian K, Leelakumari S, Jones M, Bose P, Ch'ng C, Reisle C, Eirew P, et al. 2017. Homologous recombination deficiency and platinum-based therapy outcomes in advanced breast cancer. *Clin Cancer Res* **23:** 7521–7530.

Zolkind P, Uppaluri R. 2017. Checkpoint immunotherapy in head and neck cancers. *Cancer Metastasis Rev* **36:** 475–489.