Review

Molecular Cancer Research

Intratumor Heterogeneity: Novel Approaches for Resolving Genomic Architecture and Clonal Evolution

Ravi G. Gupta and Robert A. Somer

Abstract

High-throughput genomic technologies have revealed a remarkably complex portrait of intratumor heterogeneity in cancer and have shown that tumors evolve through a reiterative process of genetic diversification and clonal selection. This discovery has challenged the classical paradigm of clonal dominance and brought attention to subclonal tumor cell populations that contribute to the cancer phenotype. Dynamic evolutionary models may explain how these populations grow within the ecosystem of tissues, including linear, branching, neutral, and punctuated patterns. Recent evidence in breast cancer favors branching and punctuated evolution driven by genome instability as well as nongenetic sources of heterogeneity, such as epigenetic variation, hierarchal tumor cell organization, and subclonal cell–cell interactions. Resolution of the full mutational landscape of tumors could help reconstruct their phylogenetic trees and trace the subclonal origins of therapeutic resistance, relapsed disease, and distant metastases, the major causes of cancer-related mortality. Real-time assessment of the tumor subclonal architecture, however, remains limited by the high rate of errors produced by most genome-wide sequencing methods as well as the practical difficulties associated with serial tumor genotyping in patients. This review focuses on novel approaches to mitigate these challenges using bulk tumor, liquid biopsies, single-cell analysis, and deep sequencing techniques. The origins of intratumor heterogeneity and the clinical, diagnostic, and therapeutic consequences in breast cancer are also explored. *Mol Cancer Res; 15(9); 1127–37.* *©2017 AACR.*

Introduction

Breast cancers are heterogeneous both in their molecular features and clinical behaviors. With the elucidation of the relationship between estrogen receptor expression and response to endocrine therapy, breast cancer became one of the first solid tumors for which the clinical implications of intertumor heterogeneity were realized (1). A landmark investigation by Perou and colleagues showed that tumors could be further stratified into subtypes according to gene expression patterns (2). Later studies by Sorlie and colleagues established the clinical significance of these mutational signatures, which are now used to stratify breast cancer patients for therapy (3). The advent of high-throughput sequencing has enabled deeper analysis of tumor genomes on a large scale and revealed an even wider landscape of breast cancer subgroups and clinical phenotypes (4, 5).

As early as the 1970s, it was recognized that genetic heterogeneity can also be found within individual breast tumors (6). Since then, intratumor heterogeneity has been characterized in both spatial and temporal terms. In a conceptual breakthrough reported in 2009, Shah and colleagues used next-generation sequencing analysis of matched primary tumor and metastatic sites in a breast cancer patient and showed that the mutational makeup of metastases diverged significantly from the primary tumor, suggesting that clonal selection occurs during disease progression (7). Nik-Zainal and colleagues further defined the timescale of breast cancer evolution and showed how deep sequencing analysis of bulk tumor samples can reveal the presence of both clonal mutations present in most cells of the tumor (likely reflecting early events in tumorigenesis), and low-frequency subclonal mutations present only in minor subpopulations (8). Multiregion sequencing of breast tumors has also provided insights into spatial heterogeneity and shown that subclonal populations can carry driver mutations, suggesting that therapies directed against tumor subpopulations could potentially prevent disease progression (9). These findings are consistent with a recent large-scale prospective study of intratumor heterogeneity in lung cancer, which showed that subclonal diversification is associated with decreased recurrence-free survival (10). Emerging techniques, such as liquid biopsy analysis and single-cell sequencing, may yield a more comprehensive portrait of intratumor heterogeneity, but these methods currently remain limited by high costs and technical challenges, such as sequencing artifacts introduced by genome amplification.

Individual breast cancers are thus defined by both a small number of well-characterized genomic abnormalities and a much larger number of other changes that are rare or unique to the individual patient (11). The extent of subclonal heterogeneity likely remains underestimated, as low-frequency genetic variants fall below the limit of detection of conventional genomic analysis. Furthermore, most clinical trials do not assess intratumor heterogeneity, which

Department of Hematology and Oncology, MD Anderson Cancer Center at Cooper, Camden, New Jersey.

Corresponding Author: Ravi G. Gupta, MD Anderson Cancer Center at Cooper, 2 Cooper Plaza, Camden, NJ 08103. Phone: 856-735-6060; Fax: 856-735-6454; E-mail: guptar4@vcu.edu

doi: 10.1158/1541-7786.MCR-17-0070

could potentially give insights into the prognostic significance of subclonal diversity. Here, we review novel techniques for resolving tumor subclonal architecture and emerging methods for inferring the evolutionary patterns of tumors. We also explore the origins of intratumor heterogeneity and the clinical, diagnostic, and therapeutic implications in breast cancer.

Origins of Intratumor Heterogeneity

Cancer genome instability

Heterogeneity in cancer arises from the introduction of genetic and epigenetic alterations by genomic instability and different patterns of clonal evolution (Fig. 1A). Genomic instability entails processes operating at multiple scales, ranging from single nucleotides to whole chromosomes (12). Most studies of nucleotide-level instability have focused on defects in the DNA repair system, including nucleotide excision repair, base excision repair, mismatch repair, and the proofreading activity of replicative DNA polymerases. Genome-wide sequencing has recently uncovered evidence of mutations that are densely clustered in short DNA segments in a manner that cannot be explained by DNA repair defects alone. In breast cancer, for example, Nik-Zainal and colleagues found long coordinated stretches of single-nucleotide substitutions, a complex mutational process known as kataegis (8). This phenomenon has been shown to colocalize with chromosomal rearrangements occurring during the repair of DNA breaks, which exposes single-stranded DNA to the action of the APOBEC family of enzymes (13, 14).

Chromosomal rearrangements have been recognized in cancer for over 100 years. Whether these lesions actively drive the early stages of carcinogenesis or simply reflect the byproducts of tumor evolution remains the subject of debate. In a large pan-cancer analysis, Davoli and colleagues showed compelling evidence that chromosomal instability directly contributes to tumorigenesis by providing a mechanism through which cancer cells acquire additional copies of oncogenes (e.g., triplosensitivity) and lose copies of tumor suppressor genes (e.g. haploinsufficiency; ref. 15). These findings are consistent with data from yeast models, cancer cell lines, xenografts, and primary tumors, which have established a mechanistic linkage between aneuploidy and genomic instability (16, 17). Several molecular processes have been implicated in chromosomal instability, including oncogene-induced replication stress (18), defective mitosis (19), whole-genome doubling (20), and telomere attrition (21). Among the most intriguing findings of high-throughput sequencing analysis has been the recent discovery of localized complex structural rearrangements, such as chromothripsis, which causes punctuated evolution of the cancer genome on a large scale in only a single or a few cell divisions (22). The molecular basis of this phenomenon is poorly understood, but its tendency to involve highly transcribed regions and open chromatin suggests a mechanistic linkage with transcription.

Nongenetic sources of intratumor heterogeneity

Phenotypic and functional heterogeneity among tumor cells, the primary substrate of selection, arises not only from clonal evolution but also from epigenetic instability (23) and heterogeneity in the

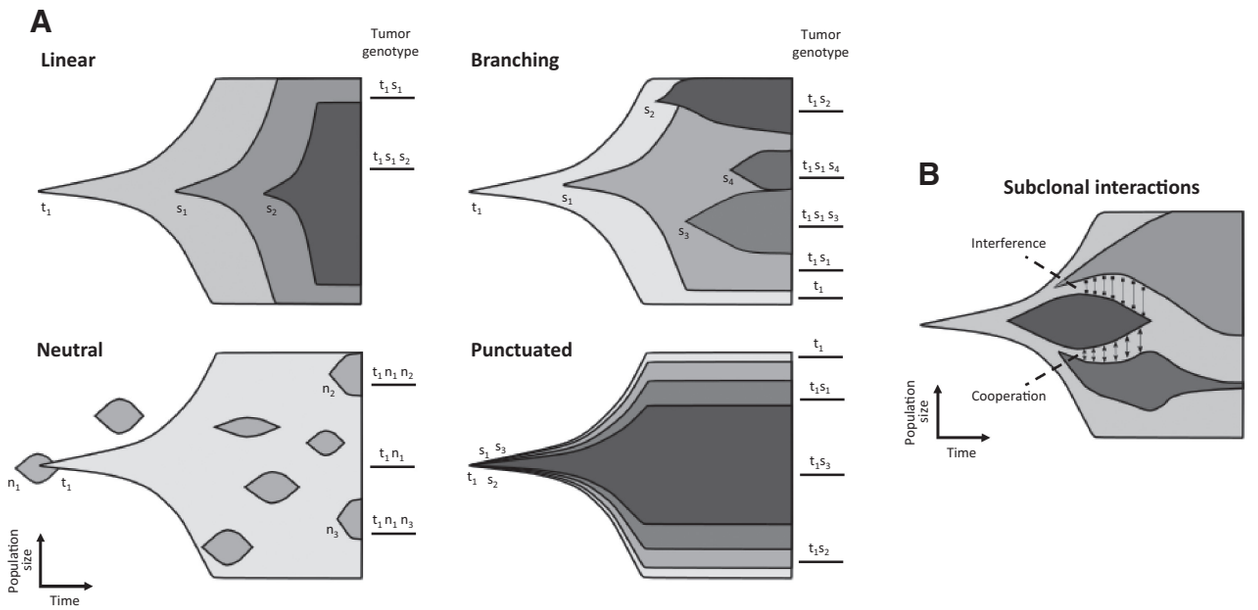

Figure 1.

A, Evolutionary mechanisms of intratumor heterogeneity. High-throughput sequencing has revealed evidence of variegated clonal architectures in breast cancer and other malignancies, including linear, branching, neutral, and punctuated evolution. These evolutionary models are not mutually exclusive and may coexist within tumors. Linear evolution is defined by sequential clonal succession but can generate intratumor heterogeneity if clonal sweeps are incomplete as shown. In branching evolution, in contrast, clonal diversity results from divergent subclones that emerge independently. Neutral evolution is defined by the absence of selection, wherein heterogeneity is a byproduct of tumorigenesis that results from stochastic processes. In the punctuated model, the majority of detectable subclonal alterations occur early in cancer evolution and tumors grow as a single clonal expansion. Subsequent subclonal expansions are rare and fail to homogenize the tumor. *t*, founder (trunk) alteration; *s*, subclonal (branch) alterations; *n*, neutral mutations. **B,** Recent evidence indicates that both competition and cooperation may occur between subclones, which may also influence the evolutionary dynamics of tumor cell subpopulations. These interactions are governed by non–cell-autonomous signaling and the tumor microenvironment.

microenvironment, such as metabolite gradients (24). A number of cancer types, including breast cancer, may also contain a hierarchy of tumorigenic and nontumorigenic subpopulations, commonly referred to as the cancer stem cell hypothesis (25). These archetypes are not mutually exclusive and can coexist within individual tumors. The stem cell model, for example, predicts that clonal evolution and tumorigenic stem cell differentiation can jointly or independently contribute to intratumor heterogeneity (26). However, the relative contribution of different sources of heterogeneity to clinical behaviors, such as drug resistance and metastasis, remains unclear. Deep sequencing studies of breast cancer and other solid tumors have revealed pervasive genetic diversification that could be the main driver of phenotypic heterogeneity. Nonetheless, the true extent of nongenetic sources of heterogeneity is likely underestimated due to the limitations of current experimental techniques, underscoring the need for novel methods for mapping intratumor epigenetic and transcriptional patterns.

Epigenetic alterations in cancer, including posttranslational histone modifications, DNA methylation, transcription factors, and noncoding RNA, have been studied for nearly 30 years. In breast cancer, epigenetic mediators, such as the histone methyl transferase EZH2 (27) and linker histone H1.0 (28), are now recognized to have key roles in tumorigenesis, including the maintenance of tumor-initiating cancer stem cells and have emerged as potential therapeutic targets. Multi-omics analysis of the genome and epigenome in other malignancies, such as brain tumors (29), prostate cancer (30), and chronic lymphocytic leukemia (31), has revealed that evolutionary phylogenies inferred from the somatic mutation and methylation profiles of tumors are highly concordant, suggesting that aberrant epigenetic states may either promote genomic instability or result from genetic changes that select for certain epigenetic patterns. These results warrant the application of a multi-omics approach to investigating the origins and functional consequences of intratumor heterogeneity in breast cancer.

Evolutionary dynamics in cancer

In the 1970s, Peter Nowell first proposed an evolutionary framework for carcinogenesis and described tumor progression in terms of the evolution of a single increasingly aggressive clone (32). This model gained support from later work by Vogelstein and colleagues in colorectal cancer, which advanced the view of linear step-wise clonal succession (33). Most data supporting linear evolution are derived from single-gene studies that likely underestimate clonal diversity. Unbiased genome-wide sequencing has recently demonstrated more complex clonal architectures that evolve during disease progression (34). The application of phylogenetic inference to deep sequencing data has also shed light on the evolutionary history of tumors and revealed different patterns of cancer evolution, including branching, neutral, and punctuated evolution.

Branching evolution is defined by the gradual accumulation of new driver mutations in subclonal populations, which expand in parallel or converge upon the same molecular pathway if positively selected. In neutral evolution, random mutations accumulate over time without selection, leading to extensive intratumor heterogeneity that is a byproduct rather than an active driver of tumorigenesis. Punctuated evolution, in contrast, posits that genomic events occur in short bursts of time, and intratumor heterogeneity is generated during the early stages of cancer. In this so-called "Big Bang" model (35), tumorigenesis is initiated by a catastrophic genomic event, such as chromothripsis, and tumors are "predetermined" to become invasive, metastatic, or drug resistant.

Current evidence of neutral evolution in cancer is derived primarily from low-depth exome data prone to sequencing errors (36), and its prevalence across cancer types remains controversial. Branching evolution, in contrast, has been evidenced by the presence of subclonal driver mutations found in deep sequencing studies of breast cancer (9) as well as other malignancies such as kidney cancer (37), lung cancer (38), and melanoma (39). One of the first observations of punctuated evolution in breast cancer was reported by Hicks and colleagues, who described "firestorms" of localized chromosomal rearrangements associated with poor outcomes (40). Most recently, punctuated evolution has been supported by deep sequencing and single-cell studies that have examined copy number variation and other chromosomal phenomena, such as chromothripsis (41).

Subclonal cell–cell interactions and clonal interference

Increasing evidence suggests that clonal evolution may not be a stochastic process but rather shaped by paracrine and cell–cell interactions. Communication between tumor cells and cells in the tumor microenvironment has been well studied in breast cancer and recognized as an important determinant of tumor progression and treatment response. Attention has recently shifted to communication between coexisting subpopulations within tumors, including subclonal interactions and clonal interference (Fig. 1B). In a mouse model of breast cancer, for example, Marusyk and colleagues demonstrated evidence of non-cell-autonomous signaling wherein tumor growth is driven by minor subpopulations of cells that do not have higher selective fitness themselves but stimulate the growth of all other clones by secreting tumor-promoting factors such as IL11 in the microenvironment (42). Other groups have identified further evidence of non–cell-autonomous signaling between tumor-initiating cancer stem cells and their progeny (43). These interactions are thought to drive tumor growth through positive feedback loops mediated by signaling molecules, such as Wnt family members and IL6, which have recently emerged as novel therapeutic targets for breast cancer. In total, these data imply that targeting only the most abundant clonal population in tumors may be a suboptimal therapeutic strategy. They further highlight the need for deep sequencing techniques to resolve the tumor subclonal architecture and provide justification for the development of treatment paradigms that inhibit cooperativity between tumor subpopulations.

Detecting Rare and Subclonal Mutations in Bulk Tumor Tissue

Massively parallel sequencing or "next-generation sequencing," a form of digital PCR that allows high-throughput analysis of the genome, has revolutionized the detection of somatic mutations in cancer. However, this approach cannot be used to detect rare mutations due to the high rate of errors that arise from library preparation and genome amplification during the sequencing process. Computational modeling by Maley and colleagues suggests that subclonal populations carrying driver mutations may be present in tumors at frequencies as low as 10^{-7} (44). Base-calling algorithms can improve the error rate of DNA sequencing but instrument-driven errors still preclude the detection of mutations that occur at frequencies less than 10^{-4} (45). Two related approaches for preparing

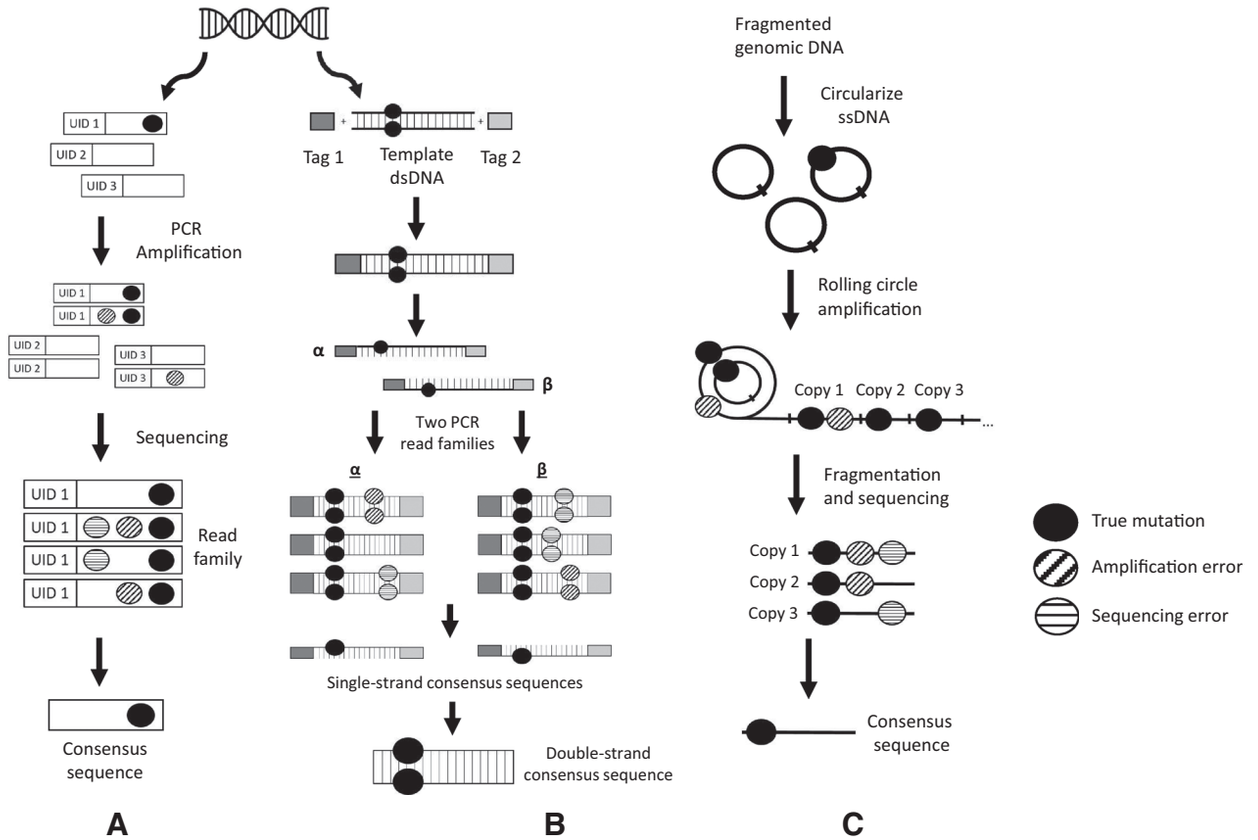

Figure 2.

Workflow of Safe Sequencing (**A**), duplex sequencing (**B**), and circle sequencing (**C**). Safe sequencing (46) and the duplex method (48) utilize molecular barcoding to label DNA templates and create "read families" that allow discrimination of true mutations from sequencing and amplification errors. Read families must contain at least three members to generate a consensus sequence. Circle sequencing (49) eliminates the need for barcoding, offering a more cost-effective alternative for identifying rare genetic variants. These techniques can identify SNVs with great accuracy but have limited resolution of structural events owing to short read lengths and low depth of sequence coverage. dsDNA, double-stranded DNA; ssDNA, single-stranded DNA.

sequencing libraries called Safe Sequencing and duplex sequencing were recently introduced to resolve this problem.

Safe Sequencing involves ligation of a barcode or "unique identifier" (UID) to each DNA template molecule followed by amplification, which results in a large number of identical daughter molecules that contain the barcode as well as any mutations present in the original template (46). Consensus analysis of UID read families allows discrimination of sequencing errors from true somatic mutations (Fig. 2A). However, this technique still produces an error rate of about two false mutations per 10,000 nucleotides, which arise during the assignment of barcodes and PCR amplification (47). In duplex sequencing, in contrast, both strands of the DNA duplex are barcoded and sequenced, and mutations are scored only if they are present at the same position in both strands (Fig. 2B; ref. 48). Duplex sequencing eliminates the effects of sequencing artifacts and detects mutations with an error rate of less than one false mutation per billion nucleotides, allowing the identification of rare genetic variants with unprecedented accuracy.

Although Safe Sequencing and the duplex method have significantly improved detection of subclonal mutations, they both suffer from short read lengths and low coverage of the genome. As a result, these relatively inefficient techniques are currently only useful for interrogating small genomic targets. One alternative to barcoding is circle sequencing, wherein genomic DNA is fragmented, circularized by ligating fragment ends, and amplified using a rolling-circle polymerase (49). Each circularized DNA fragment gives rise to about three linear copies that can be used to deduce a consensus sequence (Fig. 2C). In this approach, all copies of the template are physically linked and barcoding is not required to identify read families, which eliminates the need for redundant PCR amplification. Circle sequencing improves specificity and results in one false mutation per million nucleotides but at a cost to overall sensitivity due to loss of some template DNA during circularization (49).

The technical complexity of Safe Sequencing and circle sequencing has limited their widespread utilization in cancer biology. In contrast, several groups have reported diverse applications for duplex sequencing since its invention, including identification of drug-resistant mutations in patients with chronic myelogenous leukemia (50), early detection of ovarian cancer by sequencing of TP53 in Pap smear DNA (51), and analysis of mitochondrial DNA mutations in breast stem cells (52). Most recently, the duplex method has been applied to

significantly improve the sensitivity of circulating tumor DNA (ctDNA) analysis (53) and single-cell genome sequencing (54). Another important benefit of duplex sequencing is that it can provide accurate measurement of allelic frequencies, and thus, the mutation rate of tumors. The question of whether malignant cells exhibit an elevated rate of mutation compared with normal cells has been the subject of debate since the mutator phenotype hypothesis was first proposed as a driving force of tumor progression (55).

Emerging Approaches for Interrogating Subclonal Architecture

Liquid biopsies and serial tumor genotyping

The advent of blood-based analysis of solid tumors using intact circulating tumor cells (CTC) and cell-free ctDNA has introduced a noninvasive approach for tumor genotyping that is readily amenable to serial sampling and thus convenient for investigating temporal patterns of intratumor heterogeneity. Previous studies have applied amplicon-based methods (56) as well as whole-exome and whole-genome techniques (57) for sequencing ctDNA, but these approaches remain costly and lack sensitivity for detecting low-frequency alleles. One promising alternative called CAPP-Seq was introduced by Newman and colleagues in 2014 to address these limitations (Fig. 3). CAPP-Seq lowers the cost of ctDNA sequencing to less than $500 per patient by targeting only specific loci in the genome that are recurrently mutated for a given cancer (58). This method can resolve all major classes of mutations, including single-nucleotide variants (SNV), copy number alterations (CNA), and structural rearrangements. The authors recently integrated duplex sequencing into the CAPP-Seq workflow to further improve recovery of ctDNA from blood samples and reduce sequencing errors (53). In 30 patients with lung cancer whose tumors had been previously genotyped, this approach achieved 92% sensitivity and 100% specificity for detecting known *EGFR* mutations and could identify mutations present at frequencies as low as 0.004% in ctDNA, the most sensitive rate reported to date.

A key advantage of cell-free ctDNA analysis is the relative ease with which plasma can be collected for sequencing without the need for costly tools to enrich rare and fragile CTCs. Genetic analysis of CTCs remains hindered by several technical challenges, including low signal-to-noise ratios, amplification bias, and polymerase errors. High-throughput analysis of ctDNA is thus likely to remain the preferred clinical biomarker for interrogating genomic architecture. However, this strategy limits its analysis to SNVs, CNAs, structural rearrangements, and DNA methylation changes. Intact CTCs, in contrast, can be characterized at the level of DNA, RNA, and protein. Most notably, CTCs can also be used for functional *in vitro* and *in vivo* models to assess their potential roles in metastasis.

Several groups have demonstrated how analysis of gene expression in CTCs can provide insights into intratumor heterogeneity. Powell and colleagues, for example, reported one of the first studies of transcriptional heterogeneity in CTCs, using multiplexed quantitative PCR analysis of 87 cancer genes in individual CTCs derived from patients and single cells isolated from breast cancer cell lines (59). These experiments showed that the transcriptional profiles of CTCs differed significantly from breast cancer cell lines, suggesting that CTC profiling could be more useful than cell line data for identifying putative therapeutic targets in patients with advanced disease. Other groups have recently applied whole-exome sequencing of single CTCs in patients with prostate cancer (60) and lung cancer (61). These studies have shown that CTCs carry a large fraction of the mutations present in primary tumor and metastatic sites but also harbor many CTC-specific mutations that remain difficult to validate. Improvements in single-cell analysis methods could enable genome-wide analysis of CTCs and provide further insights into the evolutionary relationship between primary tumor, metastatic sites, and CTCs.

Single-cell analysis of bulk tumor tissue

Single-cell analysis offers the most definitive approach for reconstructing the clonal architecture and phylogenetic trees of cancer but is complicated by low input material for molecular profiling studies, such as DNA sequencing. Initial efforts for amplifying single-cell genomes used PCR-based methods, which are effective for measuring certain classes of mutations, such as CNAs, but offer low genomic coverage and limited resolution of SNVs (62). This was later followed by the introduction of multiple displacement amplification (MDA) in 2001, which utilizes a DNA polymerase with high replicative fidelity and offers much higher genomic coverage with fewer errors than PCR (63). Although potentially prone to certain biases, such as preferential amplification and allelic dropout, MDA has been widely applied for investigating genetic heterogeneity in cancer. Wang and colleagues, for example, recently introduced a technique for single-cell analysis known as Nuc-seq, which utilizes MDA in combination with duplex sequencing and flow cytometry to isolate single tumor cells for sequencing (54). In patients with breast cancer, the authors applied Nuc-seq to measure intratumor genomic heterogeneity and found evidence of punctuated clonal evolution (Fig. 4).

One alternative method for single-cell whole-genome amplification called MALBAC emerged in 2012 to mitigate amplification bias and significantly reduce the rate of sequencing errors (64). This approach minimizes amplification artifacts by utilizing quasi-linear amplification that only produces copies of the original DNA template rather than exponential copies of copies as in PCR-based methods and MDA. MALBAC currently offers the lowest false-negative rate for detecting SNVs but results in higher false positives compared with MDA owing to a lower fidelity polymerase. MALBAC has been successfully applied to measure genomic heterogeneity in CTCs of patients with lung cancer (61) and mutation rate in colon cancer cells (64).

Most studies utilizing MDA and MALBAC have examined only tens to hundreds of cells due to the high cost of single-cell genome-wide analysis and thus provide a limited representation of large tumor cell populations. Zahn and colleagues recently introduced a novel method called direct library preparation (DLP), which eliminates amplification prior to sequencing and prepares libraries directly from single-cell DNA (65). In this approach, unamplified DNA is barcoded and sequenced at low depth to generate a copy number profile for each single-cell genome. Cells with shared profiles are then used to derive a high-depth consensus genome for each clone from which SNVs can be identified. Using breast cancer xenograft tumor cells, the authors were able to resolve tumor subpopulations and clonal evolution between serial passages. At a cost of only $0.50 per cell for library preparation and a sensitivity for low-frequency subclones of about 0.05%, DLP represents a significant

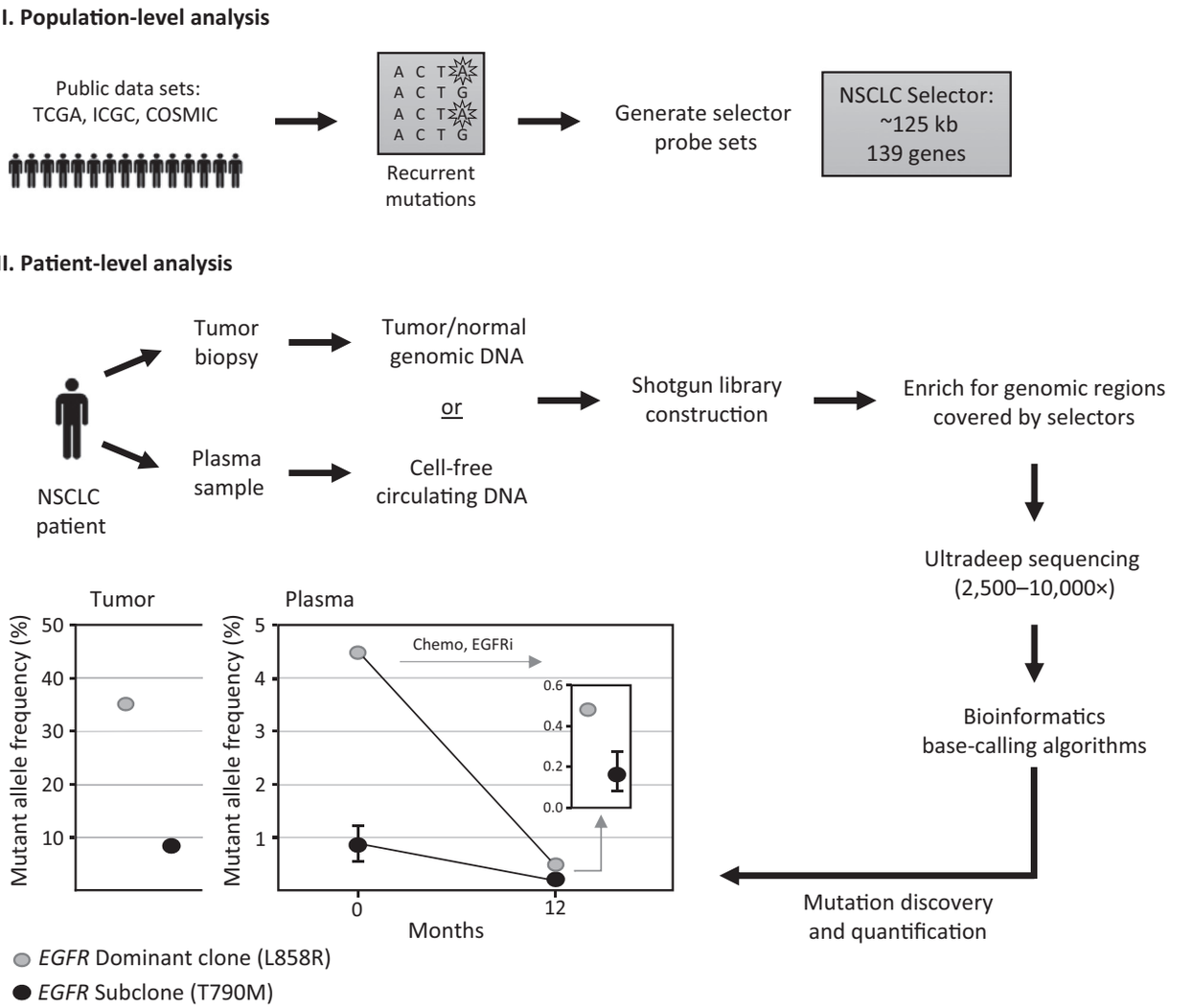

Figure 3.

Assessing subclonal architecture using cell-free ctDNA. The CAPP-Seq technique leverages novel bioinformatics methods and library construction techniques optimized for low amounts of input DNA (58). Publicly available data from genome-wide sequencing efforts are first analyzed to identify the smallest genomic region that can capture the most mutations within a given subject with cancer, which is termed a "selector." For example, Newman and colleagues created a selector covering about 125 kb of the genome and containing 139 genes in patients with NSCLC. These regions of interest are then interrogated in individual patients using tissue samples from bulk tumor and plasma during the natural history of disease. In one patient with metastatic NSCLC treated with an EGFR inhibitor, CAPP-Seq allowed resolution of resistance mutation patterns in plasma. Of note, clonal and subclonal resistance mutations in the *EGFR* gene were unexpectedly suppressed after treatment with erlotinib. The authors recently integrated the duplex method for library preparation to further increase the sensitivity of CAPP-Seq, allowing the detection of rare alleles in ctDNA with higher sensitivity and specificity (53). These findings were concordant with sequencing data from bulk tumor samples, suggesting that ctDNA analysis could potentially be applied for biopsy-fee tumor genotyping. Graph adapted from Newman and colleagues (see ref. 58). TCGA, The Cancer Genome Atlas; ICGC, International Cancer Genome Consortium; COSMIC, Catalogue of Somatic Mutations in Cancer.

advancement in throughput and cost-effectiveness for single-cell analysis.

The clearest advantage of single-cell analysis over bulk tumor assays lies in functional studies that extend beyond mutation detection, for example, epigenetic and gene expression studies to define cell states and gene regulation in cancer. Single-cell studies preserve information that is lost in bulk tumor gene expression assays, which cannot compartmentalize data by cell type and thus are unable to distinguish changes in expression that arise from gene regulation versus shifts in the cellular composition of a mixed sample. Immunophenotyping has recently shown promise as one approach for addressing this issue and assessing heterogeneity in bulk tumor biopsies (66) but may potentially mask distinct subpopulations that share common cell surface marker profiles. Recent advances in single-cell genomics, including high-throughput single-cell RNA sequencing (reviewed in detail by Kolodziejczyk and colleagues; see ref. 67) and epigenetic profiling (reviewed in detail by

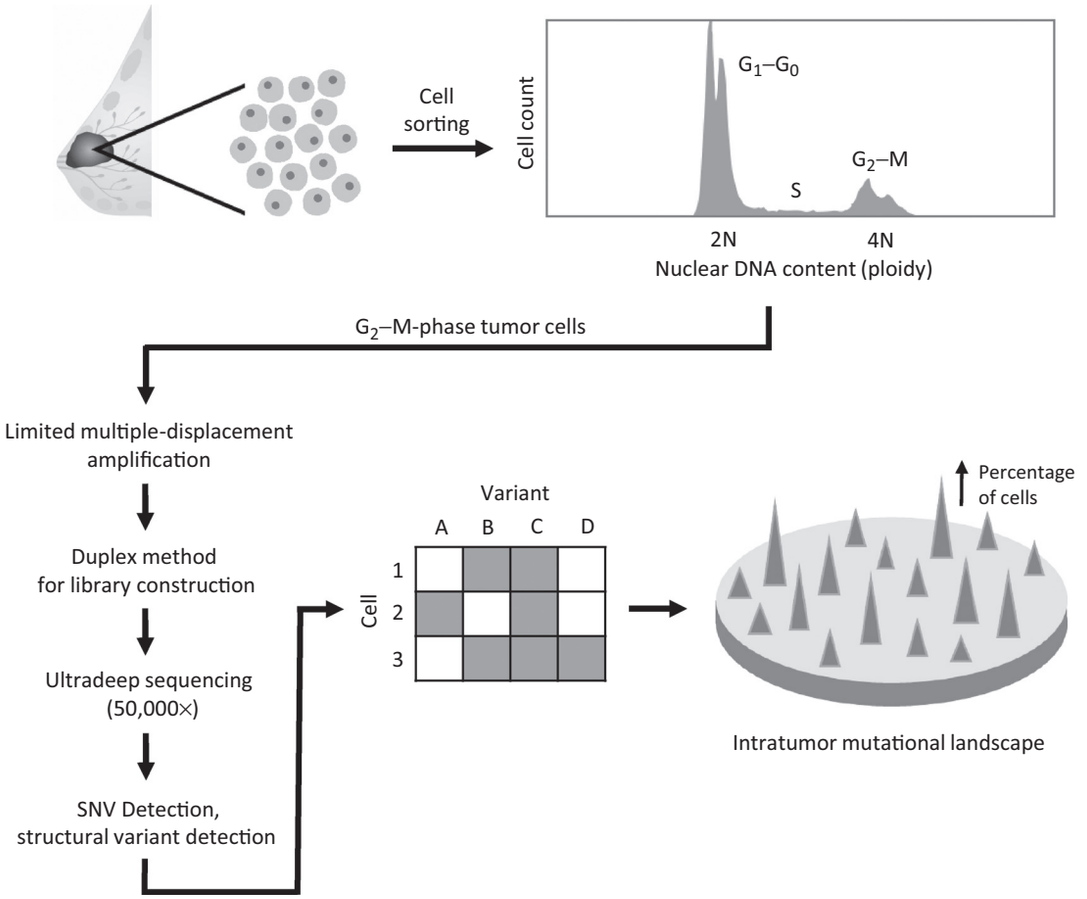

Figure 4.

Measuring clonal evolution using single-cell sequencing. Wang and colleagues recently introduced a novel cost-efficient approach to single-cell analysis that entails cell sorting to isolate G_2–M-phase tumor cells for sequencing (54). The authors applied duplex barcoding for library preparation, which allowed detection of low-frequency mutations occurring in less than 10% of the tumor mass as well as estimation of mutation rate. Single-nucleotide and structural variants were detected using bioinformatics algorithms. In two patients with ER-positive and triple-negative ductal carcinoma, the authors found aneuploid rearrangements were present early in tumor evolution and remained largely stable, whereas point mutations evolved gradually and generated extensive intratumor diversity.

Schwartzman and colleagues; see ref. 68) as well as integrated analysis of the genome, epigenome, and transcriptome at the single-cell level (69), have ushered a turning point in cancer biology, allowing functional profiling of thousands of individual cells in a single experiment. These studies can be performed on samples containing mixed cellular populations without the need for experimental purification, which could allow unbiased characterization of novel cell types and cellular states. Perhaps the most exciting promise of these techniques is their potential application in predictive modeling of cellular phenotypes and dynamics in cancer. Single-cell multi-omics methods might also be used to annotate mutation-based tumor phylogenies with functional cell state information and thus disclose, for example, the gene expression profiles of emerging subclones.

A fundamental barrier facing single-cell analysis, however, is validation of genetic and transcriptional data, which are potentially prone to errors such as false positives (due to sequencing errors) and false negatives (due to allelic dropout, amplification artifacts, and germline variants). This raises the question of whether changes in DNA and RNA detected by single-cell methods reflect true biological heterogeneity or just extensive technical errors. Several groups have shown that low-throughput techniques, such as single-cell quantitative PCR and single-molecule RNA FISH, are useful for validating new transcripts discovered by high-throughput single-cell RNA sequencing (70). Few studies, in contrast, have reported robust methods for validating subclonal mutations, which are present in less than 1% of the tumor mass. One notable exception entails comparison of single-cell mutational data with ultradeep sequencing analysis of bulk tumor DNA. This approach requires the application of sequencing techniques that can resolve the subclonal architecture with high accuracy such as the duplex method, which has been successfully applied to validate single-cell genomic data in breast cancer (54).

Inferring Clonal Evolution from Deep Sequencing Data

As the sampling depth of sequencing techniques increases, so too will the multiplicity of genomic changes found in

cancer. One particular challenge facing analysis of these data is clonal deconvolution or the identification of distinct clonal populations and their relative populations in bulk tumor samples that contain an admixture of cancer cells as well as stromal and immune cells. This has led to the development of bioinformatics tools such as the THetA (71) and TITAN (72) algorithms, which enable inference of tumor subpopulations directly from DNA sequencing data. A drawback of THetA and TITAN, however, is that they can only identify subpopulations containing CNAs and cannot distinguish subpopulations defined only by SNVs. Numerous other tools have thus been introduced that focus on the latter. Shah and colleagues, for example, developed the PyClone tool for identifying subclonal tumor populations in a large cohort of breast cancer patients and demonstrated widespread variability in the number of subpopulations present within individual tumors (73). Clonal deconvolution algorithms can be used either as a preprocessing step for sequencing data analysis or integrated into software tools, such as SCHISM (74), which enable automated reconstruction of the evolutionary trajectories (e.g., phylogenetic trees) of tumors.

Phylogenetic inference has introduced a novel approach for investigating fundamental problems in cancer, such as the separation of driver mutations from passenger variants (75) and elucidation of the timing and order of driver mutations during disease progression (76, 77). Complex and often contradictory evolutionary patterns have emerged from the application of phylogenetic inference to cancer genomes, including natural selection versus neutral evolution and linear versus branching phylogenies. At least part of this variation likely arises from differences in the markers used for phylogenetic inference (such as SNVs, CNAs, and methylation patterns) as well as the mathematical models used for reconstruction of evolutionary trees (such as maximum parsimony, minimum evolution, maximum likelihood, and Bayesian sampling; ref. 78). Few studies have compared how changes in study design affect phylogenetic inference. One important exception is a recent study by Zhao and colleagues, which drew inferences only when multiple phylogenetic methods produced the same tree topology (79). In their analysis of the clonal origins of metastases in 13 cancer types, including breast cancer, the authors showed that some metastatic lineages branch from the primary tumor early and in parallel (often well before initial diagnosis), whereas others have a single late origin. These results support both the branching and punctuated models of tumor evolution and may reconcile conflicting conclusions reached by earlier studies.

Phylogenetic tools are prone to misuse, and successful application of these methods to cancer genetics requires careful selection of the most appropriate models and algorithms for analyzing a particular data source. A computational model intended for inferring the phylogenetic history of CNAs, for example, may produce false trees if applied to SNV data. Several outstanding problems in cancer phylogenetics also remain unresolved, including a lack of models for inferring phylogenies from more complex genomic data, such as chromothripsis and kataegis as well as single-cell sequencing data. Furthermore, most studies of evolutionary patterns in cancer thus far have relied on phylogenetic models originally developed for studying Darwinian evolution of species driven by competitive selection, but whether these are adequate for characterizing clonal evolution in tumors remains unknown, particularly given recent evidence of clonal cooperation (42) and coevolution with the microenvironment (80). There is an emerging need for more sophisticated computational tools that can accommodate these subtleties as well as methods of validating the results of phylogenetic inference in individual patients.

Clinical and Therapeutic Implications of Intratumor Heterogeneity

Intratumor heterogeneity poses significant diagnostic and therapeutic challenges in the clinical setting. Spatial heterogeneity, for example, may confound the accuracy of single-region biopsies and gene expression profile tests used for prognostic risk stratification, such as Oncotype Dx (81). The clonal composition of metastatic sites may also branch from the primary tumor and thus limit the efficacy of therapies for relapsed disease that are guided by the original molecular profile of the primary tumor. This has led to the development of mathematical tools for quantifying and mapping intratumor heterogeneity, such as the Shanon index, an ecologic measure of clonal diversity that can integrate genomic data with protein expression and histopathologic tumor grade. Polyak and colleagues, for example, have applied the Shannon index in breast cancer and shown that intratumor heterogeneity in genomic traits, such as chromosome 8q24 copy number, is strongly associated with tumor grade and predictive of *in situ*-to-invasive breast carcinoma transition (82). Ecologic indices can also be used to investigate the genomic and phenotypic topography of metastatic disease and have recently shown evidence of evolutionary bottlenecks in metastasis (83). These results imply that although single-region biopsies of a primary tumor are likely prone to sampling bias, clinical decision-making based upon biopsy of single distant metastatic site might be reasonable.

Molecularly guided therapies promise to revolutionize cancer treatment, but drug resistance remains nearly universal in metastatic disease, and durable clinical benefit remains elusive, a phenomenon that may be attributable in part to clonal evolution. Several evolutionary-based strategies have thus been proposed to overcome this problem, including therapies that target early driver events shared by all tumor cells (e.g., the phylogenetic "trunk" of a tumor; ref. 84) or therapies directed against convergent phenotypic evolution (e.g., common molecular pathways affected by distinct mutations in the phylogenetic "branches" of a tumor; ref. 85). In lung cancer and melanoma, for example, patients with a higher ratio of trunk (clonal) to branch (subclonal) mutation-associated neoantigens have been found to derive greater survival benefit from immune-based therapies, such as PD-1 and CTLA-4 blockade (86). These results suggest that novel immunotherapies directed against multiple clonal neoantigens could be an effective approach to address the clinical challenges introduced by intratumor heterogeneity. Alternatively, adaptive treatment strategies that utilize dose-skipping algorithms to maintain drug-sensitive tumor subpopulations rather than eradicate disease have also shown promise in preclinical trials of breast cancer (87). This approach, which aims to prevent the competitive release of drug-resistant subclones, is particularly appealing, as it can be achieved using conventional chemotherapeutic drugs and may be applicable to existing targeted therapies.

Tumor heterogeneity, cell state plasticity, and drug resistance can also result from alterations in epigenetic control, a

possibility with profound therapeutic consequences given that correction of stable somatic mutations is substantially more difficult than reversible epigenetic changes. Until recently, epigenetic alterations were thought to be rare in solid tumors and generally restricted to hematologic malignancies, childhood cancers, and highly aggressive solid tumors, such as glioblastoma. Increasing evidence in other cancer types, such as breast cancer, however, has illuminated the importance of epigenetic regulators in functional heterogeneity, including the emergence of drug-resistant cell states (88). Epigenetic therapy has also recently shown promise as a priming treatment that sensitizes tumors, which are otherwise drug-resistant, to conventional and targeted therapies (89, 90). These results highlight the emerging need for integrated analysis of the genome, epigenome, and transcriptome, which may advance our understanding of the relationship between cell state plasticity and treatment response in cancer.

Conclusions

Intratumor heterogeneity has profound diagnostic and therapeutic implications for patients with breast cancer. Clonal diversity may confound the prognostic utility of molecular biopsies that are spatially and temporally restricted or derived from the primary tumor alone. Targeted genomic methods currently offer the most practical approach for investigating the subclonal architecture of tumors, but reductions in the cost of whole-genome sequencing and single-cell analysis could eventually shift this paradigm. Multi-omics approaches will yield deeper insights into the mechanisms of clonal selection and help resolve the functional consequences of intratumor heterogeneity. There is also a growing need to elucidate the molecular mechanisms of genetic and epigenetic instability in breast cancer, which may inform the design of novel treatment strategies, such as adaptive and immune-based therapies that target the evolutionary dynamics of cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors wish to thank Lawrence Loeb for critically reading the manuscript and providing expert guidance.

Received February 5, 2017; revised March 29, 2017; accepted June 5, 2017; published OnlineFirst June 8, 2017.

References

1. Hawkins RA, Roberts MM, Forrest AP. Oestrogen receptors and breast cancer: current status. Br J Surg 1980;67:153–69.
2. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000; 406:747–52.
3. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001; 98:10869–74.
4. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346–52.
5. The Cancer Genome Atlas Research Network. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61–70.
6. Dexter DL, Kowalski HM, Blazar BA, Fligiel Z, Vogel R, Heppner GH. Heterogeneity of tumor cells from a single mouse mammary tumor. Cancer Res 1978;38:3174–81.
7. Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, Burleigh A, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature 2009;461:809–13.
8. Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, et al. The life history of 21 breast cancers. Cell 2012;149:994–1007.
9. Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med 2015;21:751–9.
10. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the evolution of non-small-cell lung cancer. N Engl J Med 2017;376:2109–21.
11. Vogelstein B, Papdopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science 2013;339:1546–58.
12. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. Nature 1998;396:643–9.
13. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, et al. Mutational processes molding the genomes of 21 breast cancers. Cell 2012;149:979–93.
14. Roberts SA, Sterling J, Thompson C, Harris S, Mav D, Shah R, et al. Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. Mol Cell 2012;46:424–35.
15. Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, Park PJ, et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. Cell 2013;155:948–62.
16. Sheltzer JM, Blank HM, Pfau SJ, Tange Y, George BM, Humpton TJ, et al. Aneuploidy drives genomic instability in yeast. Science 2011;333: 1026–30.
17. Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, Harris BT, et al. Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science 2011;333:1039–43.
18. Zeman MK, Cimprich KA. Causes and consequences of replication stress. Nat Cell Biol 2014;16:2–9.
19. Gordon DJ, Resio B, Pellman D. Causes and consequences of aneuploidy in cancer. Nat Rev Genet 2012;13:189–203.
20. Dewhurst SM, McGranahan N, Burrell RA, Rowan AJ, Grönroos E, Endesfelder D, et al. Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. Cancer Discov 2014;4:175–85.
21. Roger L, Jones RE, Heppel NH, Williams GT, Sampson JR, Baird DM. Extensive telomere erosion in the initiation of colorectal adenomas and its association with chromosomal instability. J Natl Cancer Inst 2013;105: 1202–11.
22. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 2011;144:27–40.
23. Plass C, Pfister SM, Lindroth AM, Bogatyrova O, Claus R, Lichter P. Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. Nat Rev Genet 2013;14: 765–80.
24. Carmona-Fontaine C, Deforet M, Akkari L, Thompson CB, Joyce JA, Xavier JB. Metabolic origins of spatial organization in the tumor microenvironment. Proc Natl Acad Sci U S A 2017;114:2934–2939.
25. Shackleton M, Quintana E, Fearon ER, Morrison SJ. Heterogeneity in cancer: cancer stem cells versus clonal evolution. Cell 2009;138:822–9.
26. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. Nature 2013;501:328–37.
27. Kim KH, Roberts CW. Targeting EZH2 in cancer. Nat Med 2016;22: 128–34.
28. Torres CM, Biran A, Burney MJ, Patel H, Henser-Brownhill T, Cohen AS, et al. The linker histone H1.0 generates epigenetic and functional intratumor heterogeneity. Science 2016;353:.
29. Mazor T, Pankov A, Johnson BE, Hong C, Hamilton EG, Bell RJ, et al. DNA methylation and somatic mutations converge on the cell cycle and define similar evolutionary histories in brain tumors. Cancer Cell 2015;28:307–17.

30. Brocks D, Assenov Y, Minner S, Bogatyrova O, Simon R, Koop C, et al. Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. Cell Rep 2014;8:798–806.
31. Landau DA, Clement K, Ziller MJ, Boyle P, Fan J, Gu H, et al. Locally disordered methylation form the basis of intratumor methylome variation in chronic lymphocytic leukemia. Cancer Cell 2014;26:813–25.
32. Nowell PC. The clonal evolution of tumor cell populations. Science 1976;194:23–8.
33. Fearon ER, Hamilton SR, Vogelstein B. Clonal analysis of human colorectal tumors. Science 1987;238:193–7.
34. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21.
35. Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, et al. A Big Bang model of human colorectal tumor growth. Nat Genet 2015;47: 209–16.
36. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. Nat Genet 2016;48:238–44.
37. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012;366:883–92.
38. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science 2014;346:251–6.
39. Harbst K, Lauss M, Cirenajwis H, Isaksson K, Rosengren F, Törngren T, et al. Multiregion whole-exome sequencing uncovers the genetic evolution and mutational heterogeneity of early-stage metastatic melanoma. Cancer Res 2016;76:4765–74.
40. Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu E, et al. Novel patterns of genome rearrangement and their association with survival in breast cancer. Genome Res 2006;16:1465–79.
41. Leibowitz ML, Zhang CZ, Pellman D. Chromothripsis: a new mechanism for rapid karyotype evolution. Annu Rev Genet 2015;49:183–211.
42. Marusyk A, Tabassum DP, Altrock PM, ALmendro V, Michor F, Polyak K. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. Nature 2014;514:54–8.
43. Zhang M, Tsimelzon A, Chang CH, Fan C, Wolff A, Perou CM, et al. Intratumoral heterogeneity in a Trp53-null mouse model of human breast cancer. Cancer Discov 2015;5:520–33.
44. Etchings J. Strategies in biomedical data science. 1st ed. Hoboken (NJ): John Wiley & Sons, Inc.; 2017.
45. Robasky K, Lewis NE, Church GM. The role of replicates for error mitigation in next-generation sequencing. Nat Rev Genet 2014;15:55–62.
46. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. Proc Natl Acad Sci U S A 2011;108:9530–5.
47. Yu C, Mannan AM, Yvone GM, Ross KN, Zhang YL, Marton MA, et al. High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. Nat Biotechnol 2016; 34:419–23.
48. Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, Loeb LA. Detection of ultra-rare mutations by next-generation sequencing. Proc Natl Acad Sci U S A 2012;109:14508–13.
49. Lou DI, Hussmann JA, McBee RM, Acevedo A, Andino R, Press WH, et al. High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. Proc Natl Acad Sci U S A 2013;110: 19872–7.
50. Schmitt MW, Fox EJ, Prindle MJ, Reid-Bayliss KS, True LD, Radich JP, et al. Sequencing small genomic targets with high efficiency and extreme accuracy. Nat Methods 2015;12:423–5.
51. Krimmel JD, Schmitt MW, Harrell MI, Agnew KJ, Kennedy SR, Emond MJ, et al. Ultra-deep sequencing detects ovarian cancer cells in peritoneal fluid and reveals somatic TP53 mutations in noncancerous tissues. Proc Natl Acad Sci U S A 2016;113:6005–10.
52. Ahn EH, Hirohata K, Kohrn BF, Fox EJ, Chang CC, Loeb LA. Detection of ultra-rare mitochondrial mutations in breast stem cells by duplex sequencing. PLoS One 2015;10:e0136216.
53. Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, Scherer F, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. Nat Biotechnol 2016;34:547–55.
54. Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 2014;512:155–60.
55. Loeb LA. A mutator phenotype in cancer. Cancer Res 2001;61:3230–9.
56. Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med 2012;4: 136ra68.
57. Murtaza M, Dawson SJ, Pogrebniak K, Rueda OM, Provenzano E, Grant J, et al. Multifocal clonal evolution characterized using circulating tumor DNA in a case of metastatic breast cancer. Nat Commun 2015; 6:8760.
58. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. Nat Med 2014;20:548–54.
59. Powell AA, Talasaz AH, Zhang H, Coram MA, Reddy A, Deng G, et al. Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines. PLoS One 2012;7:e33788.
60. Lohr JG, Adalsteinsson VA, Cibulskis K, Choudhury AD, Rosenberg M, Cruz-Gordillo P, et al. Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nat Biotechnol 2014;32:479–84.
61. Ni X, Zhuo M, Su Z, Duan J, Gao Y, Wang Z, et al. Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc Natl Acad Sci U S A 2013;110:21083–8.
62. Telenius H, Carter NP, Bebb CE, Nordenskjold M, Ponder BA, Tunnacliffe A. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. Genomics 1992;13:718–25.
63. Dean FB, Nelson JR, Giesler TL, Lasken RS. Rapid amplification of plasmid and phage DNA using phi29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res 2001;11:1095–1099.
64. Zong C, Lu S, Chapman AR, Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science 2012;338:1622–6.
65. Zahn H, Steif A, Laks E, Eirew P, VanInsberghe M, Shah SP, et al. Scalable whole-genome single-cell library preparation without preamplification. Nat Methods 2017;14:167–73.
66. Snowden E, Porter W, Hahn F, Ferguson M, Tong F, Parker JS, et al. Immunophenotyping and transcriptomic outcomes in PDX-derived TNBC tissue. Mol Cancer Res 2017;15:429–38.
67. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The technology and biology of single-cell RNA sequencing. Mol Cell 2015; 58:610–20.
68. Schwartzman O, Tanay A. Single-cell epigenomics: techniques and emerging applications. Nat Rev Genet 2015;16:716–26.
69. Macaulay IC, Ponting CP, Voet T. Single-cell multiomics: multiple measurements from single cells. Trends Genet 2017;33:155–168.
70. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. Nat Rev Genet 2015;16:133–45.
71. Oesper L, Mahmoody A, Raphael BJ. THetA: inferring intra-tumor heterogeneity from high-throughput DNA sequencing data. Genome Biol 2013;14:R80.
72. Ha G, Roth A, Khattra J, Ho J, Yap D, Prentice LM, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. Genome Res 2014;24:1881–93.
73. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 2012;486:395–9.
74. Niknafs N, Beleva-Guthrie V, Naiman DQ, Karchin R. SubClonal Hierarchy Inference from Somatic Mutations: Automatic reconstruction of cancer evolutionary trees from multi-region next generation sequencing. PLoS Comput Biol 2015;11:e1004416.
75. Kumar RD, Swamidass SJ, Bose R. Unsupervised detection of cancer driver mutations with parsimony-guided learning. Nat Genet 2016;48;1288–94.
76. McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. Sci Transl Med 2015;7:283ra54.
77. Beerenwinkel N, Greenman CD, Lagergren J. Computational cancer biology: an evolutionary perspective. PLoS Comput Biol 2016;12:e1004717.
78. Schwartz R, Schaffer AA. The evolution of tumour phylogenetics: principles and practice. Nat Rev Genet 2017;18:213–229.

79. Zhao ZM, Zhao B, Bai Y, Iamarino A, Gaffney SG, Schlessinger J, et al. Early and multiple origins of metastatic lineages within primary tumors. Proc Natl Acad Sci U S A 2016;113:2140–5.
80. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. Science 2015;348:880–6.
81. Gyanchandani R, Lin Y, Lin HM, Cooper K, Normolle DP, Brufsky A, et al. Intratumor heterogeneity affects gene expression profile test prognostic risk stratification in early breast cancer. Clin Cancer Res 2016;22:5362–9.
82. Park SY, Gonen M, Kim HJ, Michor F, Polyak K. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. J Clin Invest 2010;120:636–44.
83. Almendro V, Kim HJ, Cheng Y-K, Gönen M, Itzkovitz S, Argani P, et al. Genetic and phenotypic diversity in breast tumor metastases. Cancer Res 2014;74:1338–48.
84. Bozic I, Reiter JG, Allen B, Antal T, Chatterjee K, Shah P, et al. Evolutionary dynamics of cancer in response to targeted combination therapy. Elife 2013;2:e00747.
85. Voss MH, Hakimi AA, Pham CG, Brannon AR, Chen YB, Cunha LF, et al. Tumor genetic analyses of patients with metastatic renal cell carcinoma and extended benefit from mTOR inhibitor therapy. Clin Cancer Res 2014;20:1955–64.
86. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 2016;351:1463–9.
87. Enriquez-Navas PM, Kam Y, Das T, Hassan S, Silva A, Foroutan P, et al. Exploiting evolutionary principles to prolong tumor control in preclinical models of breast cancer. Sci Transl Med 2016;8:327ra24.
88. Brown R, Curry E, Magnani L, Wilhelm-Benartzi CS, Borley J. Poised epigenetic states and acquired drug resistance in cancer. Nat Rev Cancer 2014;14:747–53.
89. Appleton K, Mackay HJ, Judson I, Plumb JA, McCormick C, Strathdee G, et al. Phase I and pharmacodynamics trial of the DNA methyltransferase inhibitor decitabine and carboplatin in solid tumors. J Clin Oncol 2007;25:4603–9.
90. Munster PN, Thurn KT, Thomas S, Raha P, Lacevic M, Miller A, et al. A phase II study of the histone deacetylase inhibitor vorinostat combined with tamoxifen for the treatment of patients with hormone therapy-resistant breast cancer. Br J Cancer 2011;104:1828–35.