Mechanisms and Consequences of Cancer Genome Instability: Lessons from Genome Sequencing Studies

June-Koo Lee,[1] Yoon-La Choi,[2,3] Mijung Kwon,[4,5] and Peter J. Park[6]

[1]Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea; email: junekoo_lee@kaist.ac.kr

[2]Department of Pathology and Translational Genomics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, South Korea

[3]Department of Health Sciences and Technology, Samsung Advanced Institute of Health Sciences and Technology (SAIHST), Sungkyunkwan University School of Medicine, Seoul 06351, South Korea; email: yla.choi@samsung.com

[4]Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115

[5]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115; email: mijung_kwon@dfci.harvard.edu

[6]Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts 02115; email: peter_park@hms.harvard.edu

Annu. Rev. Pathol. Mech. Dis. 2016. 11:283–312

First published online as a Review in Advance on February 22, 2016

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

This article's doi:
10.1146/annurev-pathol-012615-044446

Keywords

chromosomal instability, aneuploidy, chromosomal rearrangement, microsatellite instability, hypermutation, cancer evolution

Abstract

During tumor evolution, cancer cells can accumulate numerous genetic alterations, ranging from single nucleotide mutations to whole-chromosomal changes. Although a great deal of progress has been made in the past decades in characterizing genomic alterations, recent cancer genome sequencing studies have provided a wealth of information on the detailed molecular profiles of such alterations in various types of cancers. Here, we review our current understanding of the mechanisms and consequences of cancer genome instability, focusing on the findings uncovered through analysis of exome and whole-genome sequencing data. These analyses have shown that most cancers have evidence of genome instability, and the degree of instability is variable within and between cancer types. Importantly, we describe some recent evidence supporting the idea that chromosomal instability could be a major driving force in tumorigenesis and cancer evolution, actively shaping the genomes of cancer cells to maximize their survival advantage.

INTRODUCTION

Since the seminal discovery by Theodor Boveri of the presence of chromosomal abnormalities in cancer cells (1), many hypotheses have been put forth on how the unstable nature of the cancer genome is associated with tumorigenesis (2)—whether it is a cause or a consequence of tumor evolution (3) and what the underlying biological mechanisms are for this instability (4). Genome instability can be defined as an increased tendency of the genome to acquire mutations, typically conferred by dysfunctional genome maintenance processes. Here we describe various aspects of genomic instability as a cardinal feature of cancer that drives tumor evolution and facilitates other cancer hallmarks (5).

Although cancer is a disease that results from the accumulation of somatic mutations in the genome (6), analysis of its nucleotide sequence was limited for decades by our inability to sequence more than a few hundred nucleotides at a time. Development of high-throughput sequencing (HTS) technology in the past 10 years, however, has provided a pivotal turning point, enabling genome-wide, nucleotide-level analysis of genomic alterations. With the new sequencing technology, we have made significant progress on many key questions, such as how many genomic alterations are present in a cancer genome (7), what kinds of genes or pathways are frequently altered across cancer types (8), how heterogeneous cancer cells within a tumor sample are, and how the patterns of alterations evolve over time. Strictly speaking, instability refers to the rate of change (9), and the presence of a large number of somatic mutations in a cancer genome (status) does not provide precise information on the degree of instability. Nonetheless, the sequencing datasets allow us to systematically analyze the impact of instability on the entire genome and to infer the biological mechanisms underlying certain genomic lesions on the basis of the traces of DNA damage and its repair process left on the nucleotide sequence. Furthermore, comparative analysis of serially acquired or spatially separated samples provides snapshots of clonal dynamics and intratumoral heterogeneity for insights into cancer genome instability (10–12).

After a brief description of genome variation, we describe present understanding of the biochemical mechanisms in some key cancer genome instability processes, starting with large-scale (chromosome-level) instability and moving to small-scale (nucleotide-level) instability. We then present genomic consequences of such instabilities as observed in sequencing studies and new insights into their mechanisms gained through these studies. Our examples illustrate the complexities of various driving forces in tumorigenesis and the promise of a data-driven approach for exploring the pathogenic mechanisms and consequences of cancer genome instability.

TYPES OF LESIONS FOUND IN THE CANCER GENOME

Sequencing of a large number of exomes and genomes has led to detailed characterization of many types of alterations. The initial, groundbreaking studies examined about two dozen genomes (13, 14), but these were soon followed by larger-scale studies, notably those conducted by the multi-institutional consortium called The Cancer Genome Atlas (TCGA). TCGA alone has generated multidimensional genomic data including ~10,000 exome and ~1,000 high-coverage whole-genome pairs (tumor and matched normal) across more than 20 tumor types. More recently, as the cost of sequencing has come down, similar studies have proliferated, resulting in the accumulation of tens of thousands of profiles.

Major types of genomic alterations and the patterns of paired-end sequencing data that reflect such alterations are shown in **Figure 1**. Distinct genomic instability processes are involved in the generation of these alterations. In this review, we describe two major genomic instability processes

operating at different scales: chromosomal instability and nucleotide-level genomic instability. The following categories of genomic alterations are closely related to the two processes.

Single Nucleotide Variations and Small Insertion/Deletions

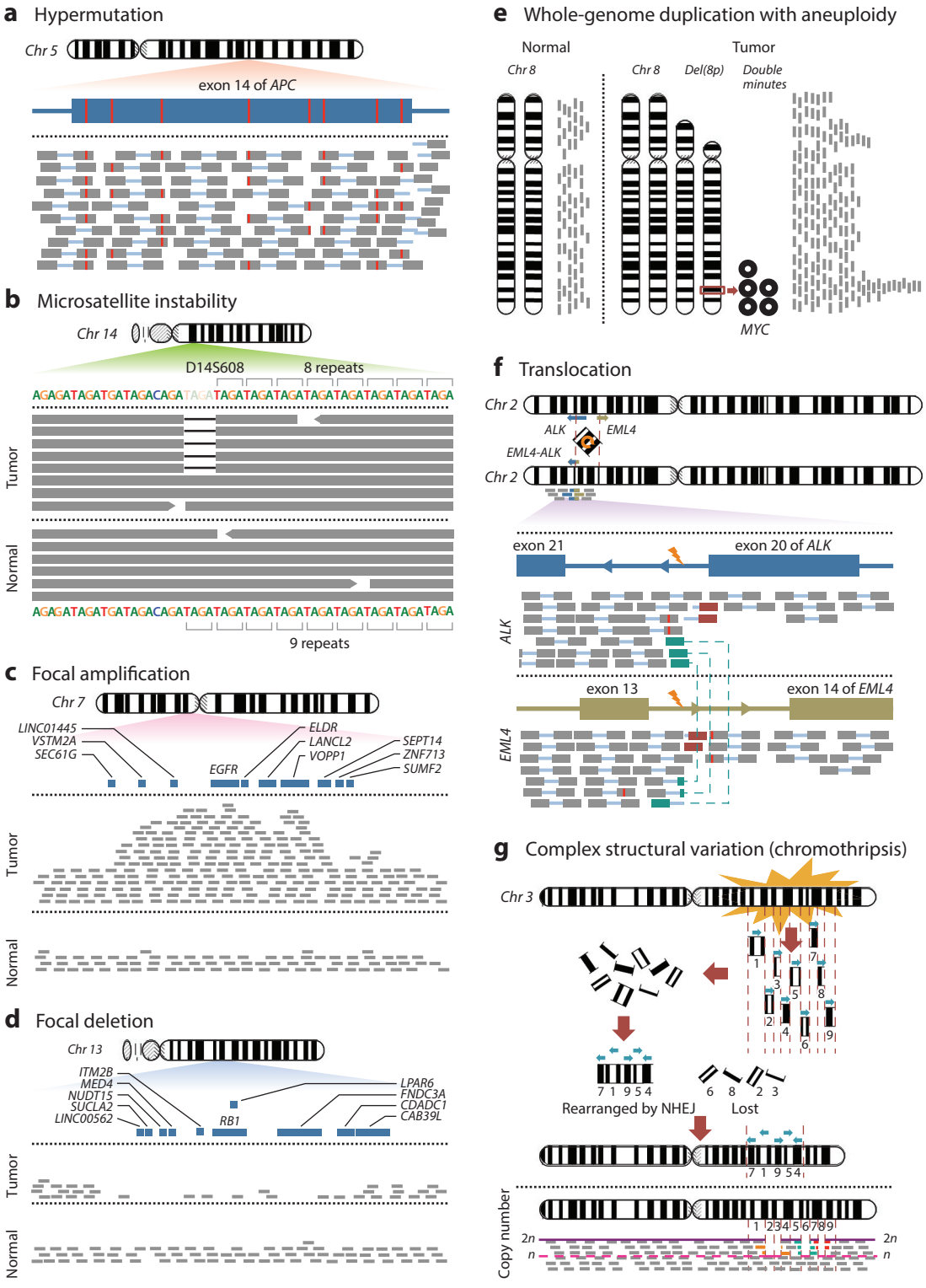
Sequencing datasets allow detection of somatic single nucleotide variations (SNVs) after germline variations (from a matched normal sample and/or a collection of normals) are removed. To detect SNVs with low variant allele fractions (a small number of cells in a population carrying the mutation), high sequencing coverage is needed. The number of somatic mutations per cancer genome can vary by orders of magnitude depending on the primary site, tumor type, and patient age (7, 15, 16). For example, some cases of posterior fossa ependymomas in infants have only a few SNVs in the entire genome; DNA methylation profiling of this tumor has revealed CpG island methylator phenotype, indicating that epigenetic mechanisms play a crucial role for tumorigenesis in this case (17). In contrast, microsatellite-unstable colon cancers can have thousands of mutations per exome owing to their defective mismatch repair system (18). A recent study on the frequency of different base substitutions and their adjacent nucleotide context uncovered more than 20 mutational signatures in the human cancers, some of which can be attributed to well-known carcinogens or DNA repair defects (16).

Copy Number Alterations

Genomic regions affected by somatic copy number alterations (CNAs) vary widely in size. Traditionally, the terms gain and loss of copy number have been used to describe large-scale CNAs such as chromosomal arm-level changes (19), whereas amplification and deletion have been used to refer to focal CNAs. But the usage of these terms is not uniform. For a long time, large-scale CNAs such as aneuploidy have been investigated with cytogenetic methods. Although fluorescence in situ hybridization and other molecular techniques are widely used for detection of predefined focal CNA lesions, introduction of array-based platforms has enabled unbiased identification of CNAs with higher resolution, on the order of ~10–100 kb (20, 21). With sequencing data, CNAs can be detected at higher resolution, sometimes even at the exon level and with base-pair resolution of the CNA boundaries. Furthermore, the same data can often provide estimates of tumor purity and ploidy, leading to absolute quantification of copy number in an allele-specific manner.

Structural Variations

The term structural variations (SVs) refers to a wide range of genomic rearrangements, including translocations, inversions, tandem duplications, transposable element insertions, and other complex rearrangements, as well as simple insertions and deletions (except the ones that are very small, e.g., <30 bp, which are typically referred to as indels). In a broad sense, copy number alterations comprise a category of SVs. CNAs and SVs are the major consequences of chromosomal instability (CIN). Since the discovery of the Philadelphia chromosome in chronic myeloid leukemia (22), conventional cytogenetic studies have identified stereotypical translocations in various types of hematologic cancers, some of which confer sensitivity to specific targeted therapies (23). Whole-genome sequencing (WGS) has transformed our ability to study SVs. Whereas array-based approaches did not allow for detection of copy-neutral changes such as inversions and translocations, WGS now allows identification of a wide range of rearrangements. Studies have revealed that the frequency and complexity of SV events vary within and across cancer types (24). Recent studies have also provided insights into the generating mechanism (25) and temporal dynamics of such lesions (26).



CHROMOSOMAL INSTABILITY

An increased rate of chromosomal change compared with normal cells, referred to as CIN, is a widespread form of genomic instability in human cancers (9, 27). Profiling with array-based platforms in the past several years has produced a molecular portrait of chromosomal aberrations across several thousand samples (28–30), although analysis of aggregated data remains difficult due to the varying levels of resolution across different generations of array platforms (30). It has become clear that nearly all cancers have some amount of chromosomal aberration, in number (aneuploidy), structure (e.g., translocation, inversion, or duplication), or both. Some of these genomic lesions are recurrent, suggesting positive selection conferred by at least one of the loci contained in the lesion, whereas others are nonrecurrent events that are more likely to be passengers than drivers. One of the most important debates related to CIN has been whether it is actively driving tumorigenesis with evolutionary selection for aneuploidy, or whether the observed alterations are simply by-products of cancer evolution (3).

Role of Chromosomal Instability in Tumorigenesis

Origins of CIN have been explored through numerous in vitro experiments and various genetically engineered mouse models (3). One likely mechanism is the dysfunction of the mitotic checkpoint (also called a spindle assembly checkpoint)—a cellular surveillance mechanism that ensures accurate segregation of metaphase chromosomes (31). If the chromosomes are not correctly attached to the spindle at the kinetochores, the mitotic checkpoint delays the start of anaphase to safeguard the proper segregation of chromosomes. Ablation or repression of mitotic checkpoint molecules (e.g., Mad2, Bub1, CenpE) in mouse models has resulted in CIN phenotypes (32–34). These models have provided important insights into the role of CIN in relation to tumorigenesis. Although the CIN phenotype accelerated tumorigenesis in many cases, the acquisition of another oncogenic alteration seemed to be necessary for development of cancers. For example, many mice with CIN that did not have increased rate of spontaneous tumorigenesis exhibited accelerated tumorigenesis when they were exposed to chemical carcinogens (35, 36). Furthermore, in this CIN background, introduction of oncogenic mutation or inactivation of tumor suppressors also increased the rate of tumorigenesis (34, 37). These findings indicate that, although aneuploidy may not directly result in tumor development, it could contribute to the promotion of tumorigenesis. In other cases, however, the CIN phenotype may be subject to negative selection. Decreased incidence of tumor development has been observed in some mouse models of CIN under various oncogenic backgrounds, compared with those without CIN (34, 37, 38). Although the molecular basis of this paradox is still unclear, it seems that CIN can be deleterious to survival and fitness of normal cells. For example, yeast or primary mouse cells with aneuploidy show impaired proliferation compared

Figure 1

Examples of genomic alterations associated with cancer genome instability. Paired-end sequencing reads from a DNA fragment are illustrated as gray blocks connected by a line (except in zoomed-out views, where the fragment is a single block). (*a*) In hypermutated cancers, there is a high rate of somatic mutation throughout the genome, possibly even with multiple mutations in an exon. In this example, exon 14 of *APC* was targeted by eight point mutations (*red lines*). (*b*) In this microsatellite-unstable cancer, a tetranucleotide repeat was deleted by erroneous replication from microsatellite D14S608 on chromosome 14. (*c*) Focal amplification targeting the *EGFR* locus, as evidenced by increased read density. In many cases, multiple adjacent genes are included in the amplicon. (*d*) Focal deletion targeting the *RB1* locus. (*e*) Complex aneuploidy, with doubling of chromosome 8 by whole-genome duplication followed by deletion of portions of 8p in two chromosomes. The *MYC* locus was amplified via formation of double minutes. (*f*) Translocation in chromosome 2p resulting in *EML4-ALK* rearrangement, linking exon 13 of *EML4* with exon 21 of *ALK*. (*g*) Chromosome 3q was shattered from a catastrophic event and then stitched to form a complex genomic rearrangement. The resultant chromothripsis lesion shows a copy number pattern that alternates between two states. Abbreviation: NHEJ, nonhomologous end joining.

with euploid counterparts (39, 40). Cells may not tolerate excessive genomic alterations (37) or proteotoxic stresses in these cases (39). In cancer, the growth disadvantages of aneuploidy can be alleviated by genome duplication (discussed below) or overcome by acquisition of genetic alterations resulting in aneuploidy tolerance, e.g., the inactivation of p53 (41, 42). In summary, these studies reveal that CIN is a double-edged sword that can be detrimental to cell survival but whose adverse effects can be tolerated in specific cellular contexts such as tumorigenesis.

Recent genome-wide studies have identified strong evidence that supports the active role of CIN in tumorigenesis. For example, the functional role of hemizygous deletions had been largely uncharacterized, in contrast to recurrent homozygous deletions that typically target classical tumor suppressor loci. It was unclear whether recurrent hemizygous deletions are precursors to homozygous deletions or hemizygous deletions themselves can confer a selective advantage to cells by haploinsufficiency (wherein loss of one copy of a gene produces phenotype) of the encompassing genes. However, a functional genomics study using an shRNA library to target 19,011 genes revealed that negative regulators of cell proliferation (named STOP genes, short for suppressors of tumorigenesis and/or proliferation) were significantly enriched within recurrent hemizygous deletions, whereas positive regulators of cell proliferation (called GO genes, short for growth enhancers and oncogenes) were depleted (43). This suggested that a cumulative dosage effect could bestow proliferative advantages (43). In a related study, a large number of cancer driver genes (~250 oncogenes and ~320 tumor suppressor genes) were predicted from the SNV mutation patterns in >8,000 tumors (44). With this expanded gene set, the researchers examined whether the distribution and potency of cancer driver genes could explain the patterns of CNAs. Surprisingly, a score that combined the density of tumor suppressors and oncogenes with their potency for each chromosome arm predicted the frequency of deletion and amplification of chromosome arms and whole chromosomes. These results indicate the role of cumulative haploinsufficiency and triplosensitivity (wherein an additional copy of a gene produces phenotype) of genes in maximizing the proliferative advantage of cancer cells (44). The findings therefore support the hypothesis that CIN plays an active role in tumorigenesis and that the complex patterns of aneuploidy frequently observed in cancers reflect their evolutionary history.

Molecular Basis of Chromosomal Instability

The mechanisms by which the cancer genome undergoes large-scale genomic alterations under CIN have been investigated through various in vitro and in vivo models as well as through molecular profiling of human samples. Several models below have identified factors that contribute to molecular pathogenesis of CIN.

Oncogene-induced replication stress. Replication stress—the impairment of DNA replication that results in stalling or collapse of replication forks (45, 46)—is induced by activation of oncogenes (e.g., *KRAS* and *CCND1*) or inactivation of tumor suppressor genes (e.g., *RB1* and *APC*) (4). Both events turn on the signaling pathways for cell proliferation, driving the cells to enter a hyper-replicative phase. However, this can give rise to stalling and, if not repaired in time, to a collapse of replication forks (reviewed in 47), resulting in DNA double-strand breakage (DSB) (48, 49). In normal cells or precancerous lesions with intact p53, apoptosis or senescence is induced as part of the DNA damage response. However, cancer cells can escape these processes due to their frequent inactivation of p53 (50).

Traces of DSBs resulting from DNA replication stress are frequently observed in cancer genomes. The best example is the common fragile sites, which are especially sensitive to replication stress (51); these regions are frequently targeted by recurrent homozygous deletions or loss

of heterozygosity events in premalignant lesions and cancers (45, 48, 49, 52). Tandem duplications are also attributable to replication fork collapse and the associated DNA repair mechanism, break-induced replication (BIR) (53). BIR is a variant of homologous recombination that can repair fork collapse–associated single-end DSB lesions in conjunction with a nearby replication fork. The BIR process can produce segmental duplications with microhomology junctions, which have frequently been observed in many types of cancers, including ovarian and breast cancers (54). Recent studies have also provided evidence that DNA replication stress can precipitate not only structural chromosomal abnormalities but also missegregation of chromosomes (55, 56), possibly due to copy number loss of CIN suppressor genes (55) or compromised stability of centromeres and kinetochores (56).

The oncogene-induced replication stress model is in accordance with the stepwise carcinogenesis model as described for colorectal cancer evolution (57). Mutations in oncogenes or in tumor suppressor or caretaker genes (e.g., *MSH2* and *MLH1*) are the early events in tumorigenesis, and these may lead to genomic instability. Cancer cells may then escape apoptosis or senescence in the presence of *TP53* mutation, which makes the cells permissive of genome instability (45, 50). Recent sequencing studies have also demonstrated this to be the case. In 112 esophageal adenocarcinomas and their precancerous lesions, including Barrett's esophagus and high-grade dysplasia (58), *TP53* mutation was preceded by mutations in cancer-related genes, including *ARID1A* and *SMARCA4*. Similar findings were identified in bladder cancers and their corresponding precancerous lesions (59).

Defective mitosis. Impairment in mitosis directly results in aneuploidy. Therefore, various defects in mitosis have been studied for their contributions to CIN.

- ***Mitotic checkpoint dysfunction.*** Although experimental models have shown that defects in the mitotic checkpoint can effectively recapitulate CIN phenotype, many CIN cancer cells have intact mitotic checkpoints (60, 61). Although somatic mutation in *BUB1*, a gene encoding serine/threonine kinase whose role is critical in mitotic checkpoint function, was identified in CIN colorectal cancer patients (62), mutations in mitotic checkpoint genes are rarely observed in human cancers (3).
- ***Defective cohesion.*** Recent studies have revealed that defects in sister chromatid cohesion may also contribute to CIN (63, 64). In line with these studies, mutations in *STAG2*, a gene encoding a subunit protein of the cohesin complex, have been observed in multiple types of human cancers exhibiting CIN phenotype (65, 66).
- ***Merotelic attachment.*** A lagging chromosome is frequently observed in the anaphase of CIN cancer cells (67). This is typically caused by abnormal chromosome attachment called merotelic attachment, when a single kinetochore is attached to microtubules emanating from different spindle poles. Because it is not detected by the spindle assembly checkpoint, merotelic attachment and subsequent lagging chromosomes are believed to be an important cause of CIN (68). Two major mechanisms have been suggested to explain the abnormal attachment (27, 68). First, dynamic interactions between spindle microtubule and kinetochore are hyperstabilized in cancer cells (69), causing difficulty in the resolution of the chromosome attachment errors (70). Second, supernumerary centrosomes that are frequently observed in human cancers can cause merotelic attachment (71). Although cancer cells can cope with catastrophic multipolar divisions by clustering their amplified centrosomes into two poles to form pseudobipolar spindles (72), the centrosome clustering results in an increased rate of merotelic attachment.
- ***Whole-genome duplication.*** Mitotic defects can also cause whole-genome duplication (WGD) (68), which can serve as a precursor to the CIN cancers (73). Several pieces of

evidence indicate that WGD is an early event in tumorigenesis. First, tetraploid cells are found more frequently in early-stage lesions of human cancers (74–76). Second, activation of several oncogenes or inactivation of tumor suppressor genes can produce tetraploid cells in their early passages (77, 78). Third, analysis of copy number data from colorectal cancers shows that copy number losses take place after the WGD event in a majority of genome-doubled samples (79). What, then, is the role of WGD in early-stage tumorigenesis? Ploidy has substantial impact on the evolutionary adaptation of cells. Tetraploid cells exhibit higher numbers of chromosomal abnormalities per cell, though not per chromosome (79), conferring greater adaptability. A recent study using isogenic yeast cells showed that tetraploid cells undergo a more rapid adaptation compared with diploid and haploid counterparts (80). In vitro evolution experiments showed that this rapid adaptation was due to both higher rates of beneficial mutations and larger fitness effects of new mutations. Furthermore, experimentally induced WGD showed increased tolerance to subsequent CIN processes (79), and tetraploid cells have shown higher tumorigenicity in mice (81). These studies demonstrate that WGD enables a more rapid adaptation to the environment and plays an important role in the early stages of tumor progression.

Telomere attrition. Dysfunctional telomeres have been proposed as a major mechanism of genomic instability, especially in epithelial cancers (82, 83). Shortening of the telomere frequently occurs in the early stages of tumorigenesis (84, 85). A study of human breast tissues by fluorescence in situ hybridization showed that CNA events were substantially more frequent at the ductal carcinoma in situ (DCIS) stage, reflecting the telomere crisis that occurs during the transition from ductal hyperplasia to DCIS (86). Shortening of telomeres activates DNA damage response, but acquisition of mutations in *TP53* during their later stages of tumorigenesis allows the cancer cells to escape apoptosis or senescence (82).

Telomere attrition–induced CIN may account for a diverse set of chromosomal aberrations. A breakage-fusion-bridge cycle (87), which is triggered by an eroded telomere, may induce CNAs and SVs ranging from focal amplifications or deletions to whole-chromosomal changes (88). Telomerase-deficient mouse models have shown abundant focal CNA events as well as nonreciprocal translocations (89–91), which are frequent in human cancers, in contrast to the conventional oncogene-activated or tumor suppressor–inactivated mice, in which whole-chromosome-level CNA events are more prevalent. Moreover, prolonged DNA damage response to dysfunctional telomeres could bypass mitosis and force the cell to enter the second S phase and induce WGD (92). This could explain why both telomere shortening and WGD are commonly observed in early-stage tumors. Furthermore, telomere attrition may account for a higher frequency of WGD in cancers among elderly patients (93).

In late-stage tumors, the length of telomeres is maintained by reactivation of telomerase (more than 90% of cases) or by the alternative lengthening of telomere (ALT) pathway (82). The genomes of these tumors display traces of telomere-associated crisis. A recent study of 22 cell lines whose telomere is maintained through the ALT pathway revealed prevalence of subtetraploid karyotype, extensive genomic rearrangement, and defective DNA damage response (94). These lesions may be attributable to telomere dysfunction in the earlier stages of tumorigenesis.

Footprints of Chromosomal Instability

Recent sequencing-based studies have provided a detailed view of chromosomal alterations in large cohorts, and comparative analysis of these profiles across many different cancer types has led

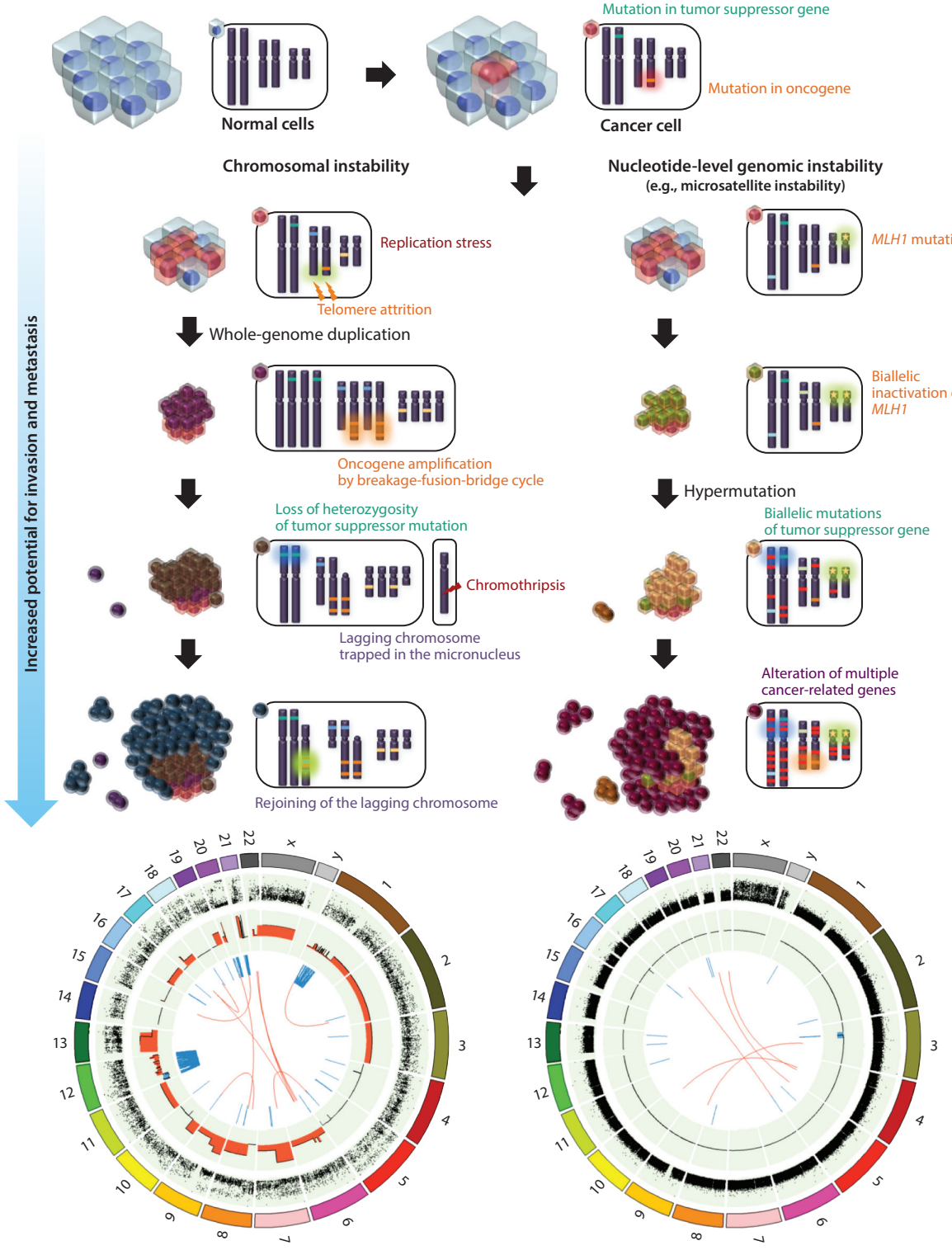
to identification of pan-cancer as well as tumor type–specific features. These data have also led to the discovery of novel mechanisms that contribute to CIN and may drive tumorigenesis.

Landscape of CNAs in cancer genomes. Accumulation of a large number of copy number profiles has allowed for a panoramic view of CNA events across thousands of tumor specimens (28–30, 95). In a recent pan-cancer analysis of 4,934 cancers in the TCGA cohort, WGD was estimated to have occurred in more than one-third of all cancers (37%), highlighting its major role in tumorigenesis (95). Moreover, these cancers with WGD exhibited a greater number of CNAs when compared with near-diploid cancers, suggesting that WGD could predispose cancer cells to CIN or make cancer cells more tolerant to subsequent CIN processes. The average ploidy of cancers with WGD was 3.31, which indicates that a large number of copy-loss events took place after the WGD (95). These cancers with WGD could represent the full-blown phenotype of CIN (illustrated in **Figure 2**); they differ from aneuploid cancers that have near-diploid genomes. We note that the estimates of purity and ploidy above were derived from array-based copy number variation data and may be unstable; analysis of WGS datasets will give more accurate predictions.

Among the many CNA lesions, focal CNAs are the most common, followed by arm-level and chromosome-level CNAs. Large-scale events, including arm- and chromosome-level events, show low amplitude (e.g., one copy gain or loss), in contrast to focal CNAs, which generally have higher amplitude. The focal CNAs in past studies have encompassed three or four genes on average (95), but this estimate was largely dependent on the resolution of the array platforms and the algorithms used to identify such CNAs. As WGS data allow a much higher spatial resolution (96), smaller focal CNAs are likely to be detected.

To identify the driver CNAs, the most common approach is to search for recurrently altered regions across a large number of samples. Frequently amplified regions typically encompass classical oncogenes known to be activated by gene amplifications, such as *CCND1*, *EGFR*, and *MYC* (95). In the amplified regions without classical oncogenes, chromatin modification–related genes (e.g., *BRD4*, *KDM2A*, *KDM5A*) are among the enriched (95), highlighting their role as a new cancer hallmark (45). Previous studies have indicated that 7–17% of homozygous deletions targeted recessive tumor suppressor genes, including *CDKN2A/B*, *STK11*, and *PTEN*, suggesting that these deletions are footprints of positive selection (28, 29, 95). Identification of the deletion targets and subsequent functional experiments continue to result in discovery of novel tumor suppressors. For example, *PARK2*, a gene encoding E3 ubiquitin ligase that targets cyclin D and cyclin E, was recently identified as a tumor suppressor, frequently targeted by homozygous deletion (97); *QKI* is also targeted by homozygous deletions or loss-of-function mutations in glioblastoma, suggesting its role as a tumor suppressor (98). In many cases, homozygous deletions encompass the largest genes in the human genome. These genes are prone to DNA damage because their transcription cannot be completed by the start of S phase, leading to a conflict between transcription and replication (45). In these cases, the homozygous deletions are passenger events that reflect the replication stress–associated genomic instability (29).

Different patterns of genomic instability are found depending on the tissue origin (93, 95). The WGD events are common in epithelial cancers: More than 50% of samples display WGD in lung squamous cell carcinoma, lung adenocarcinoma, and bladder cancer. By contrast, hematologic cancers very rarely contain WGD events, and only 11% of glioblastomas show WGD. Furthermore, cancer types from similar developmental lineages often share the same pattern of arm-level or focal CNAs (99). For example, similar amplification and deletion patterns have been observed among squamous cell carcinomas (head and neck carcinoma and lung squamous cell carcinoma) and among



female reproductive cancers (ovary, uterine, cervix, and breast) (95). These observations indicate that CIN drives tumorigenesis in close relation to the developmental context of tissue origin.

Identification of structural variations in cancer genomes. Genomic rearrangements have long been recognized as a cardinal feature of cancer. Many oncogenic rearrangements have been described (23, 100–102), with some fusion oncoproteins successfully targeted by small-molecule inhibitors (103, 104). However, the number of fusions with therapeutic relevance has been relatively small, partly due to the lack of technologies to comprehensively identify SVs in the genome. With HTS, this situation has changed dramatically. RNA sequencing is often performed to quantify gene expression levels, but it also allows for identification of fusion transcripts, with single nucleotide resolution of the breakpoint when it is contained within a sequenced read (105–107). Major limitations of this approach, however, are that the fusion transcripts must be expressed at a sufficiently high level to be detected and that only gene-gene fusions can be captured. The more comprehensive approach is to utilize WGS data. First applied to identifying germline SVs in the human genome (108), this approach was quickly adapted for cancer genome SV detection. One of the first papers on this topic, published in 2008, analyzed two lung cancer genomes (109); with the cost of sequencing lowered, the number of samples with WGS has grown quickly and includes >1,000 cases in TCGA alone. Although the bioinformatics methods for identification of SVs based on transcriptome or WGS data have improved greatly over the years, their detection sensitivity is not satisfactory, especially due to the complexity of many SV events. SV analysis becomes even more challenging for those SVs present in a small fraction of cells or when the sequencing coverage is not high.

Systematic searches for oncogenic SV events using these technologies have identified a large number of cancer-associated rearrangements, as reported in numerous recent papers. Given that the majority of these rearrangements are located in noncoding regions, many of them are likely to be passenger events that reflect a genomic instability process. A close examination of SVs and their breakpoints across a large number of samples has revealed biological mechanisms that generate SVs (24, 110). Yang et al. (24) analyzed 140 WGS cases from 10 cancer types and found an average of 185 somatic SV events per tumor, with a large variation among samples. The number of SVs was also variable between the cancer types (e.g., high in breast cancer and lung squamous cell carcinomas versus low in clear cell renal cell carcinomas), indicating that a different spectrum of instability mechanisms is present in each tissue of origin. It will be informative to investigate

Figure 2

Two major processes of cancer genome instability. The left column illustrates an example of CIN. Oncogene-induced replication stress can generate DNA double-strand breaks, and error-prone repair of these lesions can result in various types of genomic alterations, including loss of heterozygosity, deletions, and tandem duplications. Cancers with CIN frequently undergo WGD, after which chromosomal loss or aneuploidy modulates copy numbers of beneficial mutations to maximize survival advantage. Chromosomal missegregation often results in lagging chromosomes, which can be trapped in the micronucleus. Chromosomal catastrophe in the micronucleus can result in complex rearrangements such as chromothripsis. The right column illustrates nucleotide-level instability with MSI. Biallelic inactivation of *MLH1* results in dysfunctional mismatch repair, leading to the hypermutation phenotype affecting microsatellites. A high mutational load on cancer-related genes could drive tumorigenesis, while near-diploid karyotypes are maintained. In the bottom portion of the figure, Circos plots were generated from two colorectal cancer samples with corresponding genomic instability processes. From the outside in, SNVs (allele fractions), CNAs (*red* for gains, *blue* for losses), and SVs (*red* for interchromosomal SVs, *blue* for intrachromosomal SVs) are shown. The left plot shows a CIN cancer with WGD, with clustered intrachromosomal rearrangements and alternating copy number states in chromosome 1 corresponding to chromothripsis. The right plot shows a microsatellite-unstable cancer with numerous SNVs and few CNAs. Abbreviations: CIN, chromosomal instability; CNA, copy number alteration; MSI, microsatellite instability; NHEJ, nonhomologous end joining; SNV, single nucleotide variation; SV, structural variation; WGD, whole-genome duplication.

the ploidy as well as the relative timing of the SV events (on the basis of the fraction of cells carrying each event, as estimated using the number of sequencing reads that support the event) to understand the genomic instability process shaping the genomes of these cancers.

Analysis of sequence homology at SV breakpoints gives some insights into the mechanism that gave rise to the SV (24). Not surprisingly, there is a notable difference between the mechanisms found in germline and somatic SVs, corresponding to different driving forces in normal and cancer cells. For instance, transposable element insertion was the dominant mechanism in nontumor genomes, whereas nonhomologous end joining and alternative end joining were dominant in tumor genomes. Another noticeable difference in cancer genomes is the higher percentage of microhomology-mediated break-induced replication (MMBIR), a replication-dependent mechanism that can repair a collapsed replication fork (111), reflecting replication stress–associated DNA damage in cancers. Moreover, a detailed analysis of glioblastoma cases showed that multiple mechanisms may be involved in the generation of a lesion, e.g., *EGFR* gain through a replication-based mechanism and *CDKN2A/B* loss through nonhomologous end joining repair of DSBs. This illustrates one of the many complexities in the forces that generate somatic rearrangements in the cancer genome.

Comprehensive analysis of SVs in cancer genomes has also elucidated the extent of retrotransposition insertions in cancer (112–115). Retrotransposons are widespread, potentially mobile genomic elements in the human germline, with their activity normally suppressed by various epigenetic mechanisms. Although previous studies suggested that their somatic activation could contribute to tumorigenesis (116, 117), the prevalence of such somatic retrotransposition events had been largely unknown. In the first retrotransposition analysis based on WGS data, Lee et al. (112) found ~200 insertions in 43 cases, with an average of 28 insertions in colorectal genomes. The authors also found that the genes affected by retrotransposition tended to be targeted by point mutations and their mRNA expression levels were significantly downregulated, suggesting a potential role of retrotransposition in cancer. A subsequent study across 200 WGS cases as well as 767 exomes found numerous events, including 35 insertions into exonic regions (114), whereas a study of 290 WGS cases concluded that these events were enriched in gene deserts and heterochromatin and had no general effects on transcription levels of genes at the insertion points (115). All studies found that the amount of retrotransposition activity is highly variable among different tumor types, with especially high rates in lung, head and neck, and colorectal cancers (112–115). Although their overall contribution to tumorigenesis may not be substantial, retrotranspositions are likely to play an etiologic role in specific tumors, and so further elucidation of their role in these cancers is warranted.

Chromosome shattering and reassembly. Among the most interesting findings of SV analysis is the discovery of a localized, massive genomic rearrangement, now termed chromothripsis (*thripsis* means "shattering" in Greek) (26). Typical characteristics of chromothripsis are a large number of rearrangements within a restricted area of chromosome, oscillation of copy numbers between a small number (typically two) of copy number states, and retention of heterozygosity in the high-copy number state, with loss of heterozygosity in the low-copy number state (118). As gradual addition of massive genomic rearrangements would be accompanied by a large variation in copy number states due to the error-prone DNA damage repair mechanisms (119), the presence of only two copy number states indicates a short time frame during which these complex genomic lesions are generated. Moreover, the striking pattern of clustered rearrangements suggests that these lesions are formed by mitotic errors and the missegregation of a single chromosome (120). The initial report found high prevalence of chromothripsis in osteosarcoma (3 out of 9) and chordoma

(2 out of 11), which commonly originate in bone tissues (26). Subsequent reports have revealed that this phenomenon is also observed in many other types of cancers. A large survey of array-based copy number profiles across various cancer types has estimated, using stringent criteria, that >1.5% of 8,227 cancers exhibited copy number oscillations characteristic of chromothripsis (30). Although correlations with clinical parameters have not been reported yet, chromothripsis may contribute to tumorigenesis by affecting cancer-related genes. In a small-cell lung cancer cell line, a double-minute chromosome harboring *MYC* was generated as a result of chromothripsis, and it acted as a substrate for further amplification while conferring selective advantages to the clone (26). Many cancer-related genes, including classical oncogenes (e.g., *MYCN*, *EGFR*, and *CCND1*) and tumor suppressor genes (e.g., *CDKN2A*, *PTEN*, and *ARID1A*), have been reported to be affected by chromothripsis (26, 30).

Multiple mechanisms have been suggested for the pathogenesis of chromothripsis, including ionizing radiation, telomere attrition, and aborted apoptosis (120). Among these, the most likely explanation comes from experiments on micronuclei and the associated chromosome pulverization (121). Cells with defective mitosis frequently exhibit anaphase-lagging chromosomes due to merotelic attachments (3); these chromosomes sometimes fail to join the segregated chromosomal mass at the poles and are subsequently sequestered in a structure called the micronucleus. The chromosome in the micronucleus may undergo defective DNA replication and nuclear envelope rupture, resulting in DNA damage and eventually leading to massive DNA fragmentation (121, 122). Some of these micronuclei harboring damaged chromosomes may persist in daughter cells or else be reincorporated into daughter nuclei after nuclear envelope breakdown. The many fragments after pulverization of the chromosome in the micronucleus could be stitched by an end-joining process, consistent with the lack of homology at the breakpoints that has been reported previously (26). By using live-cell imaging and single-cell sequencing, Zhang et al. have shown that chromosomes missegregated to micronuclei can generate chromothripsis-like genome rearrangements (123). At the same time, replication stress and fork collapse in the micronucleus may make complex genomic lesions by a replication-based process such as MMBIR. This could explain the apparently contradictory finding from a study of germline DNA from a patient with developmental delay, in which complex breakpoint structures with many copy number changes and microhomology sequences were reported (124). Notably, generation of chromothripsis lesions is frequently preceded by defects in the DNA damage pathway, which permits the propagation of massive DNA damage into daughter cells. Enrichment of chromothripsis events among the *TP53*-mutant samples in acute myeloid leukemia reflects this finding (125).

Impact of chromatin features on the CNA landscape. To explain the patterns of CNAs, correlations with various genomic and epigenomic factors have been explored. Application of HTS to the characterization of the epigenetic features and three-dimensional structure of chromosome folding has provided a variety of datasets that can be compared with CNA datasets. A comprehensive set of chromatin interactions can be profiled with techniques such as Hi-C [based on chromosome conformation capture techniques (126)] and ChIA-PET [chromatin interaction analysis by paired-end tag sequencing (127)]. Comparison of CNA landscape with Hi-C experimental data indicates that two breakpoints of many CNA lesions are likely to be located in spatial proximity to each other in the interphase nuclei (128). It is plausible that physical proximity is a prerequisite for the formation of structural genomic lesions, as exemplified in the cases of recurrent chromosomal translocations (129, 130). The distribution of CNA lengths is well explained by the fractal globule model (128), a model of chromosomal conformation in the interphase nuclei that enables maximal packing as well as easy folding and unfolding (131). Replication timing is

also associated with formation of CNA lesions, as indicated by overrepresentation of rearrangement breakpoints mapping to early replicating chromosome bands in neuroblastomas (132) and confirmed by large-scale analysis (110, 133, 134). Genomic domains that are replicated simultaneously are also spatially clustered in the nucleus to facilitate efficient DNA replication (135), and the interactions between the two DNA breaks from each boundary increase the likelihood of CNAs in these regions (111).

NUCLEOTIDE-LEVEL INSTABILITY

Genomic instability at the nucleotide level is frequently manifested in a hypermutation phenotype. Representative examples are cancers with defects in the repair system of DNA replication errors (136), as illustrated in **Figure 2**. In normal cells, the extraordinary fidelity of human DNA replication is achieved by two major systems: polymerase proofreading and mismatch repair (MMR). Cancers with a defect in these systems typically exhibit a hypermutation phenotype, and recent sequencing-based studies revealed their genome-wide mutational spectrum (137). Although a causal relationship between somatic mutation and cancer underpins our understanding of tumorigenesis, whether a large number of mutations is advantageous to cancer cells is a matter of controversy (138, 139).

A number of studies have examined the functional consequences of hypermutation phenotypes in tumorigenesis. First, there are cancers with a very high number of either mutations or CNAs but not both, suggesting that a cancer with a large number of mutations does not require CNAs for tumor progression. These outliers also result in an inverse correlation between the number of mutations and the number of CNAs in an aggregate analysis of >3,000 exomes (140). A recent exome sequencing study of chemical-induced and *Kras* oncogene–induced mouse cancers (141) is also consistent with the idea that each instability process may independently drive tumorigenesis. Second, analysis of hypermutated cancers has revealed that cancer-related genes are frequently affected by the instability process (137, 142, 143), which confers selective advantages to the affected cells. Third, a recent study has suggested that there is an upper limit to the possible number of mutations in a cancer cell. This study on pediatric patients with inherited biallelic MMR defects showed that the mutational load in cancers with both MMR and polymerase proofreading defects did not exceed 2×10^4 mutations per exome (144). This threshold suggests a balance between nucleotide-level genomic instability and selective pressure, with a higher mutational load being detrimental to cancer and subject to negative selection.

Although the high mutational load and frequent branched evolution in these cancers can be a hurdle for effective treatment, the hypermutation phenotype could also be used as a therapeutic target. Hypermutated cancers can present a large number of neoantigens, triggering anticancer immune responses. The immune response may then be suppressed by high expression of immune checkpoint molecules, including PD-1 and PD-L1 (145). Remarkably, a recent Phase II clinical trial has shown promising efficacy of anti-PD-1 monoclonal antibody pembrolizumab in MMR-deficient colorectal cancers (146). Although the result of this small study needs to be validated in Phase III trials, it is evident that nucleotide-level genomic instability can confer selective vulnerability to treatment, suggesting a new therapeutic avenue for this subset of cancers.

Below, we discuss major genome instability processes operating at the nucleotide level. Many of them originate from defects in the DNA repair system. Although this subject has been studied for decades, recent analyses using exome and WGS data have provided a more detailed characterization of the genome-wide consequences of these instability processes as well as their contributions to tumorigenesis.

Mismatch Repair Defect and Microsatellite Instability

Microsatellites—short tandem repeat sequences in the human genome—are prone to strand slippage during DNA replication. Microsatellite instability (MSI) refers to the hypermutation phenotype involving a large number of indel (insertion/deletion) mutations in microsatellites, acquired as a result of defective MMR. The MMR system normally functions to correct mismatched nucleotides and insertion-deletion loops that occur during replication (147). Briefly, *MSH2* encodes the protein that recognizes the mismatched bases or insertion-deletion loops by formation of a heterodimer with MSH6 or MSH3. The MSH complex then interacts with the MLH1-PMS2 heterodimer, the interplay protein that can recruit other MMR-related proteins to make nicks in both sides of the DNA lesion. When the lesion is removed, replicative polymerase fills the gap, and the repair process is completed by ligation (147). First described in familial and sporadic colorectal cancers in 1993 (148–150), MSI has been a prototype of nucleotide-level genomic instability. MSI arises when there is biallelic inactivation of MMR genes (e.g., *MSH2*, *MLH1*, *MSH6*, and *PMS2*) by mutation or promoter hypermethylation (151). In hereditary nonpolyposis colorectal cancer, the most common type of hereditary colorectal cancer syndrome (152), one defective copy of an MMR gene is inherited. Inactivation of the residual normal copy by a second hit or epigenetic silencing can lead to a malignant transformation.

MSI is observed in approximately 15% of colorectal, 20% of stomach, and 30% of endometrial cancers (18, 153, 154). In colorectal cancer, MSI status has several clinical implications. Patients with MSI colorectal cancers are generally considered to be in a favorable prognosis group (155). Importantly, they do not benefit from 5-fluorouracil-based adjuvant chemotherapy (156), possibly because the MMR system, which is critical for induction of apoptosis in response to 5-fluorouracil-incorporated DNA lesions, is inoperative (157). Therefore, accurate diagnosis of MSI status from clinical samples has been a long-standing issue.

The current gold standard test is the Bethesda panel (158), a PCR-based analysis of the fragment length variation for a small, predefined set of microsatellites. These markers are monomorphic or quasi-monomorphic microsatellites, which means that their lengths are uniform in >99% of the European population (159). Patients having two or more positive markers are diagnosed as MSI-high, those having one positive marker as MSI-low, and those having zero positive markers as microsatellite stable. Immunohistochemical staining for MMR proteins is also used as a screening method for MSI detection (151). For detection of *MLH1* promoter methylation, methylation-specific PCR is the current standard (160). These low-throughput methods, however, provide only a partial view of a genome-wide phenomenon (161). This makes it challenging to distinguish the cases that are borderline between MSI-high, MSI-low, and microsatellite stable. Moreover, the genes impacted by MSI are known to vary across cancer types (162). Thus, application of the Bethesda panel to endometrial cancers, for example, is not optimal because the panel was originally designed for the diagnosis of MSI in hereditary nonpolyposis colorectal cancer (163).

Application of HTS to MSI analysis has led to new insights into the genome-wide consequences of MMR defects. In a recent analysis of colorectal and endometrial exomes/genomes from TCGA, a computational method was used to identify microsatellites whose lengths were altered between the tumor and matched normal pairs (137). Although sequencing technologies have a higher error rate in homopolymer regions, the computational approach identified a large number of MSI loci accurately, with those on the Bethesda panel among the most significant. Across ~280 samples, the concordance between exome-based and Sanger-based Bethesda results was remarkably high.

Exome analysis has provided a more detailed view of several characteristics of MSI events. First, overrepresented among the recurrent MSI events in coding sequences were known cancer-related genes, with high frameshift-to-inframe ratios that suggest strong positive selection. Second, the

analysis produced a nearly complete list of genes targeted by MSI, with different genes targeted in colorectal (*TGFBR2* and *ACVR2A*) and endometrial (*JAK1* and *TFAM*) cancers. Although the cancer-type specificity of MSI impact was well known, there had been no genome-wide comparisons previously. For cancer types in which MSI is less frequent, this exome-based approach in future studies will easily generate a list of most affected genes. Third, expression of alleles harboring frameshift MSI was frequently reduced, likely reflecting RNA surveillance by nonsense-mediated decay. Fourth, the analysis delineated the boundaries between the three MSI categories. In particular, whether MSI-low can be a separate category (161) has been a subject of debate, given that some levels of MSI could sometimes be explained by background mutations or normal replication errors (151). Lastly, correlative analysis using whole-genome data established the relationship between the rate of SNVs and chromatin structure. Previous studies showed that SNVs are more frequent in heterochromatic regions (164) and late-replicating regions (15) of cancer genomes. By contrast, MSI events were found to be enriched in euchromatin regions and early-replicating DNA segments (137). Another recent analysis also indicated that mutations acquired after the inactivation of the MMR pathway were no longer restricted in the late-replicating heterochromatins (165), suggesting a critical role of the MMR pathway in the suppression of mutations in the early-replicating euchromatin regions, which encompass functionally important, actively transcribed genes.

Mutation in the Exonuclease Domain of Replicative DNA Polymerases

With their proofreading activities, eukaryotic replicative polymerases, including polymerases ε and δ, enable DNA replication with extreme fidelity in human cells. The 3′ to 5′ exonuclease activity is critical for their proofreading function; mutational disruption of the exonuclease domain could result in inaccurate replication of the genome, leading to a hypermutation phenotype (166). Mutations in the exonuclease domains of the *POLE* and *POLD1* genes encoding these polymerases have been reported in sporadic cases and cell lines of colorectal cancer (167, 168). Recent TCGA studies also showed that approximately 3% of colorectal cancers and 9% of endometrial cancers have somatic mutations in *POLE*, often coexisting with MMR gene mutations (18, 153). Furthermore, germline mutations in the exonuclease domain of *POLE* and *POLD1* genes were associated with susceptibility to colorectal and endometrial cancers (169).

Analysis of WGS data has greatly aided our understanding of the genomic instability process caused by *POLE* or *POLD1* exonuclease domain mutations (142, 169). These cancers have a very high number of somatic mutations in their genomes (often >100 per Mb). In particular, several hotspot mutations in the catalytic residues of the *POLE* gene are consistently associated with extreme hypermutation (142, 144). This ultramutator phenotype generally exhibits even more mutations than MSI cancers. These cancers are also typically microsatellite stable even in the presence of MMR gene mutation (137) and are devoid of CNAs (144), indicating that this ultramutator phenotype may independently drive tumorigenesis. It is important to note that endometrial cancers harboring *POLE* mutations carry a favorable prognosis (170), similar to colorectal cancers with MSI (155), underscoring the prognostic impact of hypermutation on clinical presentations. Another feature of cancers with the *POLE* mutation is the strand-specific distribution of characteristic mutation signatures (142). Studies have reported an extreme frequency of TCT>TAT or TCG>TTG mutations when the catalytic residues are mutated (16, 142, 144). Analysis of the individual sequence reads has indicated that these mutations are selectively enriched in the leading strands, consistent with the role of DNA polymerase ε in leading strand synthesis in eukaryotes (142, 171, 172).

The role of the *POLD1* exonuclease domain mutation is less well defined. One study reported a driver role of *POLD1* mutation as well as a related idiosyncratic mutation signature in ultramutated cancers with biallelic MMR deficiency (144). However, another study reported

that *POLD1* mutations in sporadic adult cancers are mostly passenger events and that cancers harboring this mutation often exhibit MSI (142).

Base-Excision Repair Defect

The structure of DNA bases is continuously assaulted by various chemical reactions, including oxidation, deamination, and alkylation, which could ultimately result in single-base lesions (173). For example, production of reactive oxygen species (ROS) could oxidize the guanine base into 8-oxoguanine (8-oxoG), which prefers to pair with adenine (A) by formation of a Hoogsteen base pair rather than with cytosine. If this 8-oxoG:A pair is not properly repaired, DNA replication of the damaged base could result in G>T transversion, which is a well-known signature of smoking-associated mutation (174). Recognition, excision, and repair of the damaged site are performed by base-excision repair (BER), an evolutionarily conserved DNA repair system. Three enzymes are the key components of the human BER system: OGG1, MUTYH, and MTH1 (encoded by *NUDT1*) (175). OGG1 and MUTYH are DNA glycosylases that can remove the damaged bases, including 8-oxoG, from the DNA helix. OGG1 is the primary defender against 8-oxoG by excising it to make an abasic site. If the unremoved 8-oxoG pairs with A, MUTYH excises the mismatched A base to prevent G>T transversion. MTH1 works on the free nucleotide pools, which are even more susceptible to oxidation by ROS (176), by hydrolyzing the oxidized deoxyguanosine triphosphates (dGTPs) to prevent their incorporation into the genome.

Association between the defects of the BER system and tumorigenesis has been described in colorectal cancers. This can be partly explained by the importance of ROS in colorectal tumorigenesis, as colonic epithelium is continuously exposed to ROS produced by microbiota (177). Direct evidence of BER defect–associated tumorigenesis was first described in the discovery of *MUTYH*-associated polyposis in a British family in 2002 (178). In three family members with autosomal recessive colonic polyposis, the researchers identified a compound heterozygote germline mutation of *MUTYH*. This loss-of-function variant in *MUTYH* resulted in defective BER, which caused somatic G>T mutation in both copies of *APC*. *KRAS* was also affected by this G>T transversion signature, frequently showing G12C mutation (179). Tumors from patients with *MUTYH*-associated polyposis exhibit neither microsatellite instability nor CIN, indicating that BER defect is an independent genomic instability pathway. A germline variant of *OGG1* is more frequently observed in patients with advanced colorectal cancer compared with normal subjects (180), but its association with susceptibility to colorectal cancer is controversial (181). *MTH1* variants have not been found to be associated with cancer susceptibility.

Because the BER pathway is involved in the repair of single-base damage caused by diverse biological processes, cancer cells seem to be dependent on the BER system, as exemplified by the elevated expression of MTH1 in human cancers (182). Therefore, the BER pathway can be a promising therapeutic target. Recent studies have demonstrated that cancer cells are selectively vulnerable to inhibition of MTH1 compared with normal cells, as this enzyme is functionally important only in cancer cells (183, 184). Inhibition of MTH1 increased incorporation of 8-oxoG, leading to DNA damage response and apoptosis in a variety of cancer cell lines and xenograft models. Although this novel treatment concept should be validated in further studies, treatment strategies targeting generalized dependence of cancer cells (cancer phenotypic lethality) on the BER pathway have opened a new prospect in cancer therapeutics (185).

Nucleotide-Excision Repair Defect

Nucleotide-excision repair (NER) is a nonspecific DNA repair system that recognizes and corrects relatively large, helix-distorting lesions, which are frequently induced by exogenous mutagens

(186). These lesions include UV-induced cyclobutane pyrimidine dimers and DNA adducts induced by various chemicals including cisplatin and aflatoxin. Two major branches operate in the NER pathway. The first one is transcription-coupled repair, which is initiated by the stalling of RNA polymerase II after recognition of DNA lesions. The second class is global genome repair, which is activated by the xeroderma pigmentosum C (XPC) complex that recognizes DNA lesions more globally. These two pathways share the XPA-RPA complex, which enables the opening of the DNA helix by XPB and XPD helicases. Structure-specific nucleases XPF-ERCC1 and XPG are recruited to this DNA lesion by XPA; two incisions are made, one on each side of the DNA lesion, to produce a gap of 27–30 nucleotides. This gap is restored by replicative polymerases with the opposite undamaged strand as a template, and DNA ligation completes the NER process.

The role of the NER system has been intensely studied in the context of UV-associated DNA damage. A well-known example is xeroderma pigmentosum (XP), a rare autosomal recessive cancer syndrome whose carriers exhibit a nearly 10,000-fold increase in the incidence of UV-associated, early-onset skin cancer relative to the general population (187). Skin fibroblasts from XP patients exhibit defects in the NER system, and these individuals have germline mutations in NER-associated genes, including *XPC* (188). Mutations in *POLH* (encoding translesion synthesis polymerase η) are also observed in patients with a variant form of XP (187). With a defective NER system, UV-induced pyrimidine dimers are not effectively repaired. As a result, the unrepaired C-C dimers are paired with A-A by translesion synthesis during DNA replication, resulting in C-C>T-T transition, which is the typical signature of UV-associated mutation. Studies have shown that cancer-related genes including *TP53* and *PTCH* are frequently affected by this mutation signature in skin cancers from XP patients (189, 190).

Recent cancer genome sequencing efforts have further highlighted the critical contribution of the NER pathway in genome maintenance. Early studies using WGS revealed that somatic mutation frequency is significantly lower in highly expressed genes (191, 192), and this finding was recently validated in a larger cohort (15). Moreover, mutations are selectively depleted in the transcribed strand compared with the nontranscribed strand (191, 192), clearly suggesting the active role of the transcription-coupled repair process. In another study using melanoma genomes, investigators found depletion of mutations including C>T transition in open chromatin regions, which are marked by DNase I hypersensitivity (193), indicating that UV-associated mutations are more efficiently repaired in these regions. Experimental evidence indicating limited access of the NER complex to the dense chromatin regions also supports this finding (194). Moreover, melanoma genomes harboring NER gene mutations showed relatively higher mutation rates in the open chromatin regions (193). This supports the active involvement of the NER pathway in repairing mutations in regulatory sequences.

Localized Hypermutation by AID/APOBEC Family Cytidine Deaminases

Active DNA repair by the genome maintenance pathways mentioned above may explain the nonrandom distribution of somatic mutations in cancer cells, particularly in association with chromatin states (164, 165). However, sequencing studies have also revealed that in several types of cancers, mutations are densely clustered in short DNA segments in a way that cannot be explained by a DNA repair defect. These localized C>T/G>A hypermutation clusters, termed kataegis by one group (195), were described in 2012 (195, 196). A mutation cluster usually spans up to several megabases, with one to several hundred bases between mutations. Notably, examination of individual sequence reads revealed a strand-specific and coordinated mutation pattern: Adjacent mutations that were close enough to be contained in a read were always located on the same read, and the same type of mutation (for example, C>T) continued in a stretch before switching to the

opposite type (in this case, G>A). Furthermore, kataegis was frequently colocalized with genomic rearrangements. Therefore, kataegis is likely to occur in long stretches of single-stranded DNA, introduced during the repair of DNA breaks that also produce genomic rearrangements (195, 196). A recent study proposed that BIR, which exposes single-stranded DNA intermediates during the repair process, may be the precipitating condition for kataegis (197).

Analysis of the mutational signature of kataegis also pointed to enzymes responsible for this process. The predominant signature was C>T or C>G at TCX trinucleotide sites; this strongly suggests the role of apolipoprotein B editing complexes (APOBECs) in the pathogenesis of kataegis (195, 196). APOBEC family enzymes are cytidine deaminases, which can convert cytosine into uracil by deamination (198). Under normal conditions, APOBEC family enzymes play a role in innate defense against retroviruses by editing DNAs or RNAs. Given their ability to insert mutations into DNAs, it has long been speculated that APOBEC family enzymes may be associated with tumorigenesis (199).

APOBEC-associated mutation is the second-most common mutation signature in a pan-cancer analysis (present in 14.4% of all cases) (16) and is especially frequent in bladder, cervix, breast, lung, and head and neck cancers (200, 201). Cancer-related genes can be affected by APOBEC-associated mutagenesis—a recent analysis of TCGA head and neck cancer exome data revealed that *PIK3CA* helical domain mutations can be attributable to APOBEC activity in human papillomavirus–negative cases (143). Multiple factors are linked with the APOBEC-associated mutation signature, including high expression of APOBEC family enzymes (202, 203) and germline deletion polymorphism spanning the *APOBEC3A-APOBEC3B* loci (204). It is interesting that carriers of this deletion polymorphism have been shown to be at increased risk for breast cancer (205).

Recent studies of diffuse large B cell lymphoma and multiple myeloma revealed that activation-induced deaminase (AID) can also generate kataegis-like mutation clusters in the context of B cells (206, 207). This enzyme generates a mutational signature distinct from the APOBEC-associated signature: In these studies, C>T hypermutations occurred at C nucleotides preceded by purines. In the affected samples, hypermutation clusters were recurrently observed in well-known AID-associated regions, including *IGH-MYC* and *IGH-CCND1* fusions and transcribed promoters. Subsequent studies indicated that these AID target regions were in proximity in the nucleus due to the formation of superenhancers (208). Recruitment of AID to these large enhancer elements may explain the recurrent off-target hypermutation in many B cell lineage–specific genes (206).

GENOMIC INSTABILITY AND CANCER EVOLUTION

Cancer evolves continuously by addition of genomic alterations and the selective advantages conferred on a subset of clones by such alterations. Advances in single-cell sequencing have allowed a more detailed characterization of intratumoral heterogeneity, including identification of gene expression signatures belonging to different subclones (209). A better understanding of intratumoral heterogeneity and the evolutionary dynamics of the subclones is a major challenge in cancer therapeutics (210), as researchers attempt to explain the mixed treatment responses as well as primary resistance to anticancer agents. Because genomic instability is the source of a wide range of alterations in the cancer genome, modification to this fundamental process during cancer evolution may be beneficial to a clone in escaping from the various selection pressures, including anticancer treatments (**Figure 3**). Previous studies have shown several remarkable examples of genomic instability variation during cancer evolution.

As described above, telomere attrition–associated chromosomal fusion and breakage often occur in the early stages of tumorigenesis. In the later stages, reactivation of telomerase or ALT contributes to tumor progression (82). In line with this biological process, an analysis of pancreatic

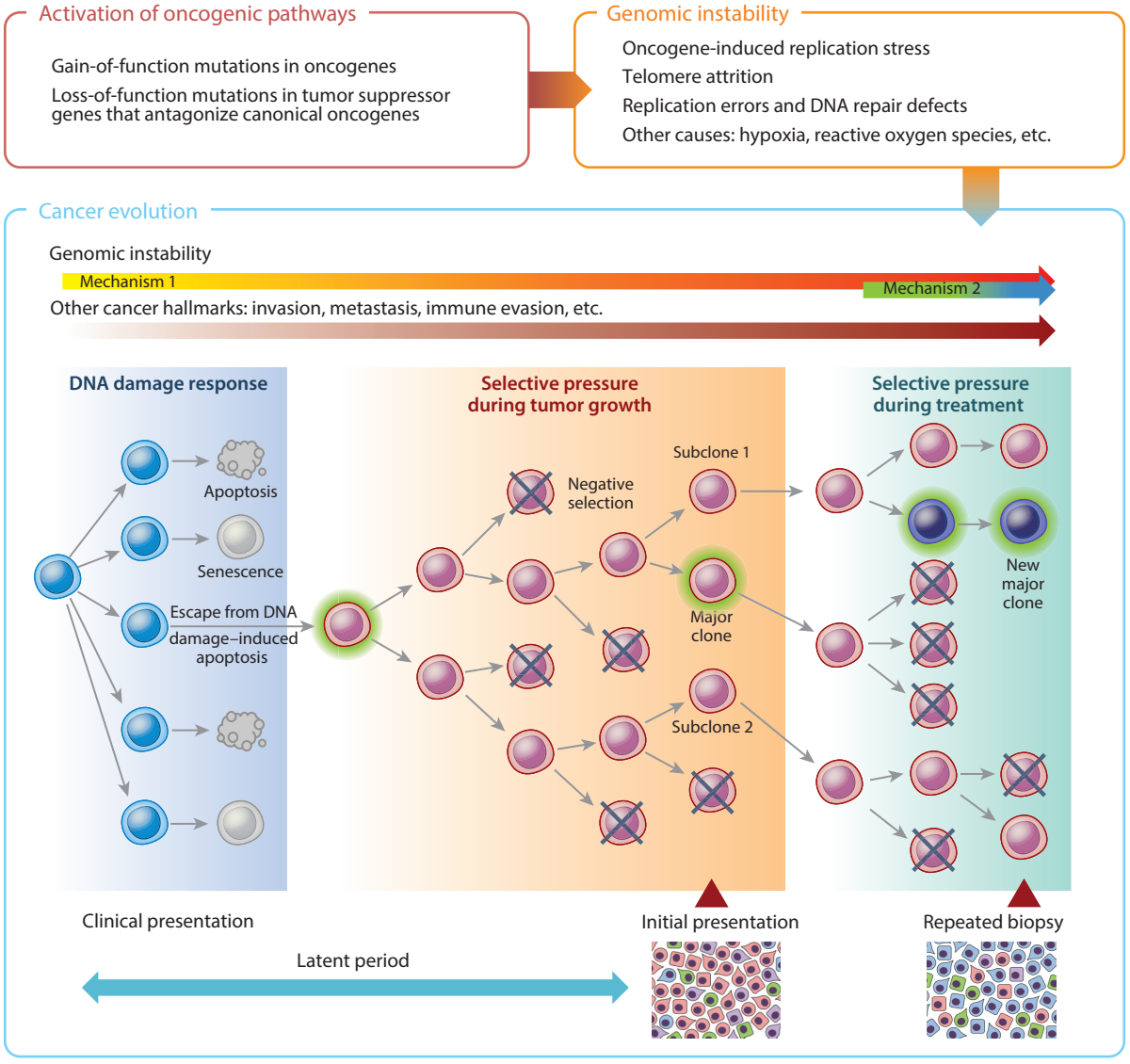


Figure 3

Mechanisms and consequences of cancer genome instability. Cancer progression is an evolutionary process driven by somatic alterations. In the early stages of tumorigenesis, frequent DNA damage by genomic instability results in the activation of DNA damage response pathways; the cells with significant DNA damage then undergo apoptosis or senescence, for example, by activating p53. However, some cells may escape this surveillance process by acquisition of genetic alterations such as mutations in *TP53*, thus avoiding apoptosis and accumulating further genetic alterations as they grow. Genomic instability may drive branched evolution of cancer cells. Clones with mutations that are not advantageous for survival will be subject to negative selection, whereas clones with higher invasiveness or metastatic potential will be subject to positive selection. Anticancer treatments such as chemotherapy will exert selective pressure, resulting in clonal repopulation. Comparison of histology and genomic alterations at repeated biopsies will be needed to reconstruct the complex evolutionary history of a cancer.

cancer genomes revealed a distinctive pattern of lesions indicating break-fusion-bridge cycles (fold-back inversions) in all metastatic lesions in a patient (211). In contrast, other signatures of genomic lesions were continuously added; the genomic instability process generating these lesions seemed to be persistent after the break-fusion-bridge cycles.

Two other studies provided evidence that mutational signatures of non-small-cell lung cancer can change during its evolution (11, 12). Contribution of smoking to the early genomic lesions in non-small-cell lung cancers is obvious, with the increase in the characteristic G>T/C>A transversions. However, in some cases, the relative fraction of smoking-associated mutations decreased during tumor progression, while the APOBEC-associated signature affecting several cancer-related genes increased in the subclones. Furthermore, lung cancer from a patient who stopped smoking 20 years prior to the diagnosis exhibited a WGD event generated in the context of a smoking-associated mutation signature (11). This highlights the long latency of non-small-cell lung cancer before clinical detection, and the very early occurrence of WGD in clinical cancer samples.

Genomic instability can be affected by anticancer treatments. Some of the agents are mutagenic, whereas others induce resistant mechanisms that alter DNA damage repair pathways. In recurrent gliomas, researchers have addressed the mutational impact of temozolomide, a frequently used alkylating agent for adjuvant treatment after surgical resection (212). They sequenced exomes of initial and recurrent glioma pairs to identify the mutational difference between the two time points, as well as to compare the mutational signatures of temozolomide-treated and -untreated patients. For low-grade gliomas with *IDH1* mutation evolving into high-grade gliomas, temozolomide treatment resulted in a temozolomide-associated hypermutation phenotype (C>T transition in the CpC and CpT context) as well as more oncogenic aberrations and progression to grade IV. Furthermore, *MLH1* mutation and *MGMT* promoter hypermethylation are observed only in recurrent samples, and this finding is compatible with previous observations (213). MMR defects may contribute to the hypermutation phenotype in conjunction with temozolomide treatment. The impact of other widely used chemotherapeutic agents on the mutational signature of the cancer genome is as yet largely unknown. As genome profiling based on a smaller amount of tissue or even a liquid biopsy becomes more accessible, it will be possible to obtain serial samples from the same patient to better analyze the impact of different therapeutic agents on the mutational landscape. Information on how distinct classes of anticancer agents impact the heterogeneous cancer cell population differently will provide valuable insights into the development of novel combination treatments as well as into the optimal timing and sequence of treatments.

CONCLUDING REMARKS

Cancer genome sequencing studies have provided a wealth of new insights into the role of cancer genome instability in driving tumorigenesis. We now have genome-wide mutational data on thousands of cancer genomes, from single nucleotide variants and indels to copy number and structural variants spanning a wide length scale. However, current models are still insufficient to explain genomic instability fully. For example, epigenetic mechanisms including altered chromatin remodeling and higher-order chromatin organization may modulate the function of cancer-related genes in conjunction with genomic instability processes, but we have relatively little genome-wide cancer epigenome data at this point. Further studies that utilize new technologies—e.g., single-cell RNA and DNA sequencing, chromatin accessibility and other epigenetic assays, and serial tracking of tumor progression in the same patient using less invasive biopsies—will enable a higher-resolution view of how genomic instability impacts the different subpopulations of a tumor.

These studies will lay the foundation for novel therapeutic approaches that exploit the selective vulnerability of cancers conferred by their unstable genomes.

DISCLOSURE STATEMENT

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

ACKNOWLEDGMENTS

J.L. is supported by a fellowship (NRF-2013H1A2A1032691) from the National Research Foundation of Korea, and P.J.P. is supported in part by the Harvard Ludwig Center. The authors thank Drs. Teresa Davoli and Lixing Yang for their comments on this manuscript.

LITERATURE CITED

1. Boveri T. 1914. *Zur frage der entstehung maligner tumoren*. Jena, Ger.: Gustav Fischer
2. Sieber OM, Heinimann K, Tomlinson IPM. 2003. Genomic instability—the engine of tumorigenesis? *Nat. Rev. Cancer* 3(9):701–8
3. Holland AJ, Cleveland DW. 2009. Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. *Nat. Rev. Mol. Cell Biol.* 10(7):478–87
4. Negrini S, Gorgoulis VG, Halazonetis TD. 2010. Genomic instability—an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* 11(3):220–28
5. Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. *Cell* 144(5):646–74
6. Cooper GM. 1982. Cellular transforming genes. *Science* 217(4562):801–6
7. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. 2013. Cancer genome landscapes. *Science* 339(6127):1546–58
8. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, et al. 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505(7484):495–501
9. Lengauer C, Kinzler KW, Vogelstein B. 1998. Genetic instabilities in human cancers. *Nature* 396(6712):643–49
10. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, et al. 2012. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366(10):883–92
11. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, et al. 2014. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 346(6206):251–56
12. Zhang J, Fujimoto J, Zhang J, Zhang J, Wedge DC, et al. 2014. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* 346(6206):256–59
13. Sjöblom T, Jones S, Wood LD, Parsons DW, Lin J, et al. 2006. The consensus coding sequences of human breast and colorectal cancers. *Science* 314(5797):268–74
14. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, et al. 2007. The genomic landscapes of human breast and colorectal cancers. *Science* 318(5853):1108–13
15. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, et al. 2013. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499(7457):214–18
16. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, et al. 2013. Signatures of mutational processes in human cancer. *Nature* 500(7463):415–21
17. Mack S, Witt H, Piro R, Gu L, Zuyderduyn S, et al. 2014. Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. *Nature* 506(7489):445–50
18. Cancer Genome Atlas Network. 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487(7407):330–37
19. Shaffer LG, McGowan-Jordan J, Schmid M. 2013. *ISCN 2013: An International System for Human Cytogenetic Nomenclature (2013)*. Basel, Switz.: Karger

20. Kallioniemi A, Kallioniemi O-P, Sudar D, Rutovitz D, Gray JW, et al. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258(5083):818–21
21. Mei R, Galipeau PC, Prass C, Berno A, Ghandour G, et al. 2000. Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. *Genome Res.* 10(8):1126–37
22. Nowell PC, Hungerford DA. 1960. Chromosome studies on normal and leukemic human leukocytes. *J. Natl. Cancer Inst.* 25(1):85–109
23. Rowley JD. 2001. Chromosome translocations: dangerous liaisons revisited. *Nat. Rev. Cancer* 1(3):245–50
24. Yang L, Luquette LJ, Gehlenborg N, Xi R, Haseley PS, et al. 2013. Diverse mechanisms of somatic structural variations in human cancer genomes. *Cell* 153(4):919–29
25. Stephens PJ, McBride DJ, Lin M-L, Varela I, Pleasance ED, et al. 2009. Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* 462(7276):1005–10
26. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, et al. 2011. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 144(1):27–40
27. Gordon DJ, Resio B, Pellman D. 2012. Causes and consequences of aneuploidy in cancer. *Nat. Rev. Genet.* 13(3):189–203
28. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, et al. 2010. The landscape of somatic copy-number alteration across human cancers. *Nature* 463(7283):899–905
29. Bignell GR, Greenman CD, Davies H, Butler AP, Edkins S, et al. 2010. Signatures of mutation and selection in the cancer genome. *Nature* 463(7283):893–98
30. Kim TM, Xi R, Luquette LJ, Park RW, Johnson MD, Park PJ. 2013. Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes. *Genome Res.* 23(2):217–27
31. Kops GJ, Weaver BA, Cleveland DW. 2005. On the road to cancer: aneuploidy and the mitotic checkpoint. *Nat. Rev. Cancer* 5(10):773–85
32. Michel LS, Liberal V, Chatterjee A, Kirchwegger R, Pasche B, et al. 2001. MAD2 haplo-insufficiency causes premature anaphase and chromosome instability in mammalian cells. *Nature* 409(6818):355–59
33. Jeganathan K, Malureanu L, Baker DJ, Abraham SC, van Deursen JM. 2007. Bub1 mediates cell death in response to chromosome missegregation and acts to suppress spontaneous tumorigenesis. *J. Cell Biol.* 179(2):255–67
34. Rao CV, Yang YM, Swamy MV, Liu T, Fang Y, et al. 2005. Colonic tumorigenesis in $BubR1^{+/-}Apc^{Min/+}$ compound mutant mice is linked to premature separation of sister chromatids and enhanced genomic instability. *PNAS* 102(12):4365–70
35. Babu JR, Jeganathan KB, Baker DJ, Wu X, Kang-Decker N, van Deursen JM. 2003. Rae1 is an essential mitotic checkpoint regulator that cooperates with Bub3 to prevent chromosome missegregation. *J. Cell Biol.* 160(3):341–53
36. Dai W, Wang Q, Liu T, Swamy M, Fang Y, et al. 2004. Slippage of mitotic arrest and enhanced tumor development in mice with *BUBR1* haploinsufficiency. *Cancer Res.* 64(2):440–45
37. Weaver BA, Silk AD, Montagna C, Verdier-Pinard P, Cleveland DW. 2007. Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell* 11(1):25–36
38. Chesnokova V, Kovacs K, Castro AV, Zonis S, Melmed S. 2005. Pituitary hypoplasia in $Pttg^{-/-}$ mice is protective for $Rb^{+/-}$ pituitary tumorigenesis. *Mol. Endocrinol.* 19(9):2371–79
39. Torres EM, Sokolsky T, Tucker CM, Chan LY, Boselli M, et al. 2007. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science* 317(5840):916–24
40. Williams BR, Prabhu VR, Hunter KE, Glazier CM, Whittaker CA, et al. 2008. Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science* 322(5902):703–9
41. Thompson SL, Compton DA. 2010. Proliferation of aneuploid human cells is limited by a p53-dependent mechanism. *J. Cell Biol.* 188(3):369–81
42. Torres EM, Dephoure N, Panneerselvam A, Tucker CM, Whittaker CA, et al. 2010. Identification of aneuploidy-tolerating mutations. *Cell* 143(1):71–83
43. Solimini NL, Xu Q, Mermel CH, Liang AC, Schlabach MR, et al. 2012. Recurrent hemizygous deletions in cancers may optimize proliferative potential. *Science* 337(6090):104–9
44. Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, et al. 2013. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* 155(4):948–62

45. Macheret M, Halazonetis TD. 2015. DNA replication stress as a hallmark of cancer. *Annu. Rev. Pathol.* 10:425–48
46. Zeman MK, Cimprich KA. 2013. Causes and consequences of replication stress. *Nat. Cell Biol.* 16(1):2–9
47. Petermann E, Orta ML, Issaeva N, Schultz N, Helleday T. 2010. Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol. Cell* 37(4):492–502
48. Bartkova J, Rezaei N, Liontos M, Karakaidos P, Kletsas D, et al. 2006. Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* 444(7119):633–37
49. Di Micco R, Fumagalli M, Cicalese A, Piccinin S, Gasparini P, et al. 2006. Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature* 444(7119):638–42
50. Halazonetis TD, Gorgoulis VG, Bartek J. 2008. An oncogene-induced DNA damage model for cancer development. *Science* 319(5868):1352–55
51. Durkin SG, Glover TW. 2007. Chromosome fragile sites. *Annu. Rev. Genet.* 41:169–92
52. Gorgoulis VG, Vassiliou L-VF, Karakaidos P, Zacharatos P, Kotsinas A, et al. 2005. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* 434(7035):907–13
53. Costantino L, Sotiriou SK, Rantala JK, Magin S, Mladenov E, et al. 2014. Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science* 343(6166):88–91
54. McBride DJ, Etemadmoghadam D, Cooke SL, Alsop K, George J, et al. 2012. Tandem duplication of chromosomal segments is common in ovarian and breast cancer genomes. *J. Pathol.* 227(4):446–55
55. Burrell RA, McClelland SE, Endesfelder D, Groth P, Weller M-C, et al. 2013. Replication stress links structural and numerical cancer chromosomal instability. *Nature* 494(7438):492–96
56. Coschi CH, Ishak CA, Gallo D, Marshall A, Talluri S, et al. 2014. Haploinsufficiency of an RB-E2F1-Condensin II complex leads to aberrant replication and aneuploidy. *Cancer Discov.* 4(7):840–53
57. Fearon ER, Vogelstein B. 1990. A genetic model for colorectal tumorigenesis. *Cell* 61(5):759–67
58. Weaver JMJ, Ross-Innes CS, Shannon N, Lynch AG, Forshew T, et al. 2014. Ordering of mutations in preinvasive disease stages of esophageal carcinogenesis. *Nat. Genet.* 46(8):837–43
59. Nordentoft I, Lamy P, Birkenkamp-Demtroder K, Shumansky K, Vang S, et al. 2014. Mutational context and diverse clonal development in early and late bladder cancer. *Cell Rep.* 7(5):1649–63
60. Tighe A, Johnson VL, Albertella M, Taylor SS. 2001. Aneuploid colon cancer cells have a robust spindle checkpoint. *EMBO Rep.* 2(7):609–14
61. Gascoigne KE, Taylor SS. 2008. Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs. *Cancer Cell* 14(2):111–22
62. Cahill DP, Lengauer C, Yu J, Riggins GJ, Willson JKV, et al. 1998. Mutations of mitotic checkpoint genes in human cancers. *Nature* 392(6673):300–3
63. Barber TD, McManus K, Yuen KW, Reis M, Parmigiani G, et al. 2008. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *PNAS* 105(9):3443–48
64. Manning AL, Yazinski SA, Nicolay B, Bryll A, Zou L, Dyson NJ. 2014. Suppression of genome instability in pRB-deficient cells by enhancement of chromosome cohesion. *Mol. Cell* 53(6):993–1004
65. Solomon DA, Kim T, Diaz-Martinez LA, Fair J, Elkahloun AG, et al. 2011. Mutational inactivation of *STAG2* causes aneuploidy in human cancer. *Science* 333(6045):1039–43
66. Balbás-Martínez C, Sagrera A, Carrillo-de-Santa-Pau E, Earl J, Márquez M, et al. 2013. Recurrent inactivation of *STAG2* in bladder cancer is not associated with aneuploidy. *Nat. Genet.* 45(12):1464–69
67. Thompson SL, Compton DA. 2008. Examining the link between chromosomal instability and aneuploidy in human cells. *J. Cell Biol.* 180(4):665–72
68. Holland AJ, Cleveland DW. 2012. Losing balance: the origin and impact of aneuploidy in cancer. *EMBO Rep.* 13(6):501–14
69. Bakhoum SF, Thompson SL, Manning AL, Compton DA. 2009. Genome stability is ensured by temporal control of kinetochore-microtubule dynamics. *Nat. Cell Biol.* 11(1):27–35
70. Bakhoum SF, Genovese G, Compton DA. 2009. Deviant kinetochore microtubule dynamics underlie chromosomal instability. *Curr. Biol.* 19(22):1937–42
71. Ganem NJ, Godinho SA, Pellman D. 2009. A mechanism linking extra centrosomes to chromosomal instability. *Nature* 460(7252):278–82

72. Quintyne NJ, Reing JE, Hoffelder DR, Gollin SM, Saunders WS. 2005. Spindle multipolarity is prevented by centrosomal clustering. *Science* 307(5706):127–29
73. Storchova Z, Kuffer C. 2008. The consequences of tetraploidy and aneuploidy. *J. Cell Sci.* 121(Pt. 23):3859–66
74. Levine DS, Rabinovitch PS, Haggitt RC, Blount PL, Dean PJ, et al. 1991. Distribution of aneuploid cell populations in ulcerative colitis with dysplasia or cancer. *Gastroenterology* 101(5):1198–210
75. Galipeau PC, Cowan DS, Sanchez CA, Barrett MT, Emond MJ, et al. 1996. 17p (p53) allelic losses, 4N (G2/tetraploid) populations, and progression to aneuploidy in Barrett's esophagus. *PNAS* 93(14):7081–84
76. Olaharski AJ, Sotelo R, Solorza-Luna G, Gonsebatt ME, Guzman P, et al. 2006. Tetraploidy and chromosomal instability are early events during cervical carcinogenesis. *Carcinogenesis* 27(2):337–43
77. Fodde R, Kuipers J, Rosenberg C, Smits R, Kielman M, et al. 2001. Mutations in the *APC* tumour suppressor gene cause chromosomal instability. *Nat. Cell Biol.* 3(4):433–38
78. Wang X, Zhou YX, Qiao W, Tominaga Y, Ouchi M, et al. 2006. Overexpression of aurora kinase A in mouse mammary epithelium induces genetic instability preceding mammary tumor formation. *Oncogene* 25(54):7148–58
79. Dewhurst SM, McGranahan N, Burrell RA, Rowan AJ, Gronroos E, et al. 2014. Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Discov.* 4(2):175–85
80. Selmecki AM, Maruvka YE, Richmond PA, Guillet M, Shoresh N, et al. 2015. Polyploidy can drive rapid adaptation in yeast. *Nature* 519(7543):349–52
81. Fujiwara T, Bandi M, Nitta M, Ivanova EV, Bronson RT, Pellman D. 2005. Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* 437(7061):1043–47
82. Maser RS, DePinho RA. 2002. Connecting chromosomes, crisis, and cancer. *Science* 297(5581):565–69
83. Artandi SE, DePinho RA. 2010. Telomeres and telomerase in cancer. *Carcinogenesis* 31(1):9–18
84. Engelhardt M, Drullinsky P, Guillem J, Moore MA. 1997. Telomerase and telomere length in the development and progression of premalignant lesions to colorectal cancer. *Clin. Cancer Res.* 3(11):1931–41
85. Roger L, Jones RE, Heppel NH, Williams GT, Sampson JR, Baird DM. 2013. Extensive telomere erosion in the initiation of colorectal adenomas and its association with chromosomal instability. *J. Natl. Cancer Inst.* 105(16):1202–11
86. Chin K, de Solorzano CO, Knowles D, Jones A, Chou W, et al. 2004. In situ analyses of genome instability in breast cancer. *Nat. Genet.* 36(9):984–88
87. McClintock B. 1938. The production of homozygous deficient tissues with mutant characteristics by means of the aberrant mitotic behavior of ring-shaped chromosomes. *Genetics* 23(4):315–76
88. De Lange T. 2005. Telomere-related genome instability in cancer. *Cold Spring Harb. Symp. Quant. Biol.* 70:197–204
89. Artandi SE, Chang S, Lee SL, Alson S, Gottlieb GJ, et al. 2000. Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. *Nature* 406(6796):641–45
90. O'Hagan RC, Chang S, Maser RS, Mohan R, Artandi SE, et al. 2002. Telomere dysfunction provokes regional amplification and deletion in cancer genomes. *Cancer Cell* 2(2):149–55
91. Else T, Trovato A, Kim AC, Wu Y, Ferguson DO, et al. 2009. Genetic p53 deficiency partially rescues the adrenocortical dysplasia phenotype at the expense of increased tumorigenesis. *Cancer Cell* 15(6):465–76
92. Davoli T, Denchi EL, de Lange T. 2010. Persistent telomere damage induces bypass of mitosis and tetraploidy. *Cell* 141(1):81–93
93. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, et al. 2012. Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* 30(5):413–21
94. Lovejoy CA, Li W, Reisenweber S, Thongthip S, Bruno J, et al. 2012. Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLOS Genet.* 8(7):e1002772
95. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, et al. 2013. Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* 45(10):1134–40

96. Xi R, Hadjipanayis AG, Luquette LJ, Kim TM, Lee E, et al. 2011. Copy number variation detection in whole-genome sequencing data using the Bayesian information criterion. *PNAS* 108(46):E1128–36
97. Gong Y, Zack TI, Morris LG, Lin K, Hukkelhoven E, et al. 2014. Pan-cancer genetic analysis identifies PARK2 as a master regulator of G1/S cyclins. *Nat. Genet.* 46(6):588–94
98. Brennan CW, Verhaak RGW, McKenna A, Campos B, Noushmehr H, et al. 2013. The somatic genomic landscape of glioblastoma. *Cell* 155(2):462–77
99. Myllykangas S, Himberg J, Bohling T, Nagy B, Hollmen J, Knuutila S. 2006. DNA copy number amplification profiling of human neoplasms. *Oncogene* 25(55):7324–32
100. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, et al. 2005. Recurrent fusion of *TMPRSS2* and ETS transcription factor genes in prostate cancer. *Science* 310(5748):644–48
101. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, et al. 2007. Identification of the transforming *EML4-ALK* fusion gene in non-small-cell lung cancer. *Nature* 448(7153):561–66
102. Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, et al. 2012. RET, ROS1 and ALK fusions in lung cancer. *Nat. Med.* 18(3):378–81
103. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, et al. 2001. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N. Engl. J. Med.* 344(14):1031–37
104. Kwak EL, Bang YJ, Camidge DR, Shaw AT, Solomon B, et al. 2010. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N. Engl. J. Med.* 363(18):1693–703
105. Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, et al. 2009. Transcriptome sequencing to detect gene fusions in cancer. *Nature* 458(7234):97–101
106. Ju YS, Lee WC, Shin JY, Lee S, Bleazard T, et al. 2012. A transforming *KIF5B* and *RET* gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res.* 22(3):436–45
107. Singh D, Chan JM, Zoppoli P, Niola F, Sullivan R, et al. 2012. Transforming fusions of *FGFR* and *TACC* genes in human glioblastoma. *Science* 337(6099):1231–35
108. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, et al. 2007. Paired-end mapping reveals extensive structural variation in the human genome. *Science* 318(5849):420–26
109. Campbell PJ, Stephens PJ, Pleasance ED, O'Meara S, Li H, et al. 2008. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat. Genet.* 40(6):722–29
110. Drier Y, Lawrence MS, Carter SL, Stewart C, Gabriel SB, et al. 2013. Somatic rearrangements across cancer reveal classes of samples with distinct patterns of DNA breakage and rearrangement-induced hypermutability. *Genome Res.* 23(2):228–35
111. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. 2009. Mechanisms of change in gene copy number. *Nat. Rev. Genet.* 10(8):551–64
112. Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, et al. 2012. Landscape of somatic retrotransposition in human cancers. *Science* 337(6097):967–71
113. Solyom S, Ewing AD, Rahrmann EP, Doucet T, Nelson HH, et al. 2012. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res.* 22(12):2328–38
114. Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M. 2014. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. *Genome Res.* 24(7):1053–63
115. Tubio JMC, Li Y, Ju YS, Martincorena I, Cooke SL, et al. 2014. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. *Science* 345(6196):1251343
116. Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, et al. 1992. Disruption of the *APC* gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* 52(3):643–45
117. Iskow RC, McCabe MT, Mills RE, Torene S, Pittard WS, et al. 2010. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell* 141(7):1253–61
118. Korbel JO, Campbell PJ. 2013. Criteria for inference of chromothripsis in cancer genomes. *Cell* 152(6):1226–36
119. Forment JV, Kaidi A, Jackson SP. 2012. Chromothripsis and cancer: causes and consequences of chromosome shattering. *Nat. Rev. Cancer* 12(10):663–70
120. Holland AJ, Cleveland DW. 2012. Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. *Nat. Med.* 18(11):1630–38

121. Crasta K, Ganem NJ, Dagher R, Lantermann AB, Ivanova EV, et al. 2013. DNA breaks and chromosome pulverization from errors in mitosis. *Nature* 482(7383):53–58
122. Hatch EM, Fischer AH, Deerinck TJ, Hetzer MW. 2013. Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell* 154(1):47–60
123. Zhang CZ, Spektor A, Cornils H, Francis JM, Jackson EK, et al. 2015. Chromothripsis from DNA damage in micronuclei. *Nature* 522(7555):179–84
124. Liu P, Erez A, Nagamani SCS, Dhar SU, Kołodziejska KE, et al. 2011. Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell* 146(6):889–903
125. Rausch T, Jones DTW, Zapatka M, Stütz AM, Zichner T, et al. 2012. Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with *TP53* mutations. *Cell* 148(1–2):59–71
126. Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. 2012. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* 58(3):268–76
127. Zhang J, Poh HM, Peh SQ, Sia YY, Li G, et al. 2012. ChIA-PET analysis of transcriptional chromatin interactions. *Methods* 58(3):289–99
128. Fudenberg G, Getz G, Meyerson M, Mirny LA. 2011. High order chromatin architecture shapes the landscape of chromosomal alterations in cancer. *Nat. Biotechnol.* 29(12):1109–13
129. Wijchers PJ, de Laat W. 2011. Genome organization influences partner selection for chromosomal rearrangements. *Trends Genet.* 27(2):63–71
130. Zhang Y, McCord RP, Ho YJ, Lajoie BR, Hildebrand DG, et al. 2012. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* 148(5):908–21
131. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326(5950):289–93
132. Schleiermacher G, Janoueix-Lerosey I, Combaret V, Derre J, Couturier J, et al. 2003. Combined 24-color karyotyping and comparative genomic hybridization analysis indicates predominant rearrangements of early replicating chromosome regions in neuroblastoma. *Cancer Genet. Cytogenet.* 141(1):32–42
133. Barlow JH, Faryabi RB, Callen E, Wong N, Malhowski A, et al. 2013. Identification of early replicating fragile sites that contribute to genome instability. *Cell* 152(3):620–32
134. Pedersen BS, De S. 2013. Loss of heterozygosity preferentially occurs in early replicating regions in cancer genomes. *Nucleic Acids Res.* 41(16):7615–24
135. Kitamura E, Blow JJ, Tanaka TU. 2006. Live-cell imaging reveals replication of individual replicons in eukaryotic replication factories. *Cell* 125(7):1297–308
136. Preston BD, Albertson TM, Herr AJ. 2010. DNA replication fidelity and cancer. *Semin. Cancer Biol.* 20(5):281–93
137. Kim T-M, Laird PW, Park PJ. 2013. The landscape of microsatellite instability in colorectal and endometrial cancer genomes. *Cell* 155(4):858–68
138. Tomlinson IPM, Novelli MR, Bodmer WF. 1996. The mutation rate and cancer. *PNAS* 93(25):14800–3
139. Fox EJ, Prindle MJ, Loeb LA. 2013. Do mutator mutations fuel tumorigenesis? *Cancer Metastasis Rev.* 32(3–4):353–61
140. Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. 2013. Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* 45(10):1127–33
141. Westcott PMK, Halliwill KD, To MD, Rashid M, Rust AG, et al. 2014. The mutational landscapes of genetic and chemical models of *Kras*-driven lung cancer. *Nature* 517(7535):489–92
142. Shinbrot E, Henninger EE, Weinhold N, Covington KR, Goksenin AY, et al. 2014. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Res.* 24(11):1740–50
143. Henderson S, Chakravarthy A, Su X, Boshoff C, Fenton TR. 2014. APOBEC-mediated cytosine deamination links *PIK3CA* helical domain mutations to human papillomavirus-driven tumor development. *Cell Rep.* 7(6):1833–41
144. Shlien A, Campbell BB, de Borja R, Alexandrov LB, Merico D, et al. 2015. Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. *Nat. Genet.* 47(3):257–62

145. Llosa NJ, Cruise M, Tam A, Wicks EC, Hechenbleikner EM, et al. 2015. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov.* 5(1):43–51
146. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, et al. 2015. PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* 372(26):2509–20
147. Peltomaki P. 2003. Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J. Clin. Oncol.* 21(6):1174–79
148. Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, et al. 1993. Clues to the pathogenesis of familial colorectal cancer. *Science* 260(5109):812–16
149. Thibodeau SN, Bren G, Schaid D. 1993. Microsatellite instability in cancer of the proximal colon. *Science* 260(5109):816–19
150. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. 1993. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363(6429):558–61
151. de la Chapelle A, Hampel H. 2010. Clinical relevance of microsatellite instability in colorectal cancer. *J. Clin. Oncol.* 28(20):3380–87
152. Lynch HT, de la Chapelle A. 2003. Hereditary colorectal cancer. *N. Engl. J. Med.* 348(10):919–32
153. Cancer Genome Atlas Research Network. 2013. Integrated genomic characterization of endometrial carcinoma. *Nature* 497(7447):67–73
154. Cancer Genome Atlas Research Network. 2014. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 513(7517):202–9
155. Popat S, Hubner R, Houlston R. 2005. Systematic review of microsatellite instability and colorectal cancer prognosis. *J. Clin. Oncol.* 23(3):609–18
156. Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, et al. 2003. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N. Engl. J. Med.* 349(3):247–57
157. Tajima A, Hess MT, Cabrera BL, Kolodner RD, Carethers JM. 2004. The mismatch repair complex hMutSα recognizes 5-fluorouracil-modified DNA: implications for chemosensitivity and resistance. *Gastroenterology* 127(6):1678–84
158. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, et al. 1998. A national cancer institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 58(22):5248–57
159. Kim TM, Park PJ. 2014. A genome-wide view of microsatellite instability: old stories of cancer mutations revisited with new sequencing technologies. *Cancer Res.* 74(22):6377–82
160. Herman JG, Umar A, Polyak K, Graff JR, Ahuja N, et al. 1998. Incidence and functional consequences of *hMLH1* promoter hypermethylation in colorectal carcinoma. *PNAS* 95(12):6870–75
161. Laiho P, Launonen V, Lahermo P, Esteller M, Guo M, et al. 2002. Low-level microsatellite instability in most colorectal carcinomas. *Cancer Res.* 62(4):1166–70
162. Kuismanen SA, Moisio AL, Schweizer P, Truninger K, Salovaara R, et al. 2002. Endometrial and colorectal tumors from patients with hereditary nonpolyposis colon cancer display different patterns of microsatellite instability. *Am. J. Pathol.* 160(6):1953–58
163. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, et al. 2004. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J. Natl. Cancer Inst.* 96(4):261–68
164. Schuster-Bockler B, Lehner B. 2012. Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* 488(7412):504–7
165. Supek F, Lehner B. 2015. Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature* 521:81–84
166. Heitzer E, Tomlinson I. 2014. Replicative DNA polymerase mutations in cancer. *Curr. Opin. Genet. Dev.* 24:107–13
167. da Costa LT, Liu B, el-Deiry W, Hamilton SR, Kinzler KW, et al. 1995. Polymerase delta variants in RER colorectal tumours. *Nat. Genet.* 9(1):10–11

168. Flohr T, Dai JC, Buttner J, Popanda O, Hagmuller E, Thielmann HW. 1999. Detection of mutations in the DNA polymerase δ gene of human sporadic colorectal cancers and colon cancer cell lines. *Int. J. Cancer* 80(6):919–29
169. Palles C, Cazier J-B, Howarth KM, Domingo E, Jones AM, et al. 2013. Germline mutations affecting the proofreading domains of *POLE* and *POLD1* predispose to colorectal adenomas and carcinomas. *Nat. Genet.* 45(2):136–44
170. Church DN, Stelloo E, Nout RA, Valtcheva N, Depreeuw J, et al. 2015. Prognostic significance of *POLE* proofreading mutations in endometrial cancer. *J. Natl. Cancer Inst.* 107(1):402
171. Pursell ZF, Isoz I, Lundström E-B, Johansson E, Kunkel TA. 2007. Yeast DNA polymerase ε participates in leading-strand DNA replication. *Science* 317(5834):127–30
172. McElhinny SAN, Gordenin DA, Stith CM, Burgers PM, Kunkel TA. 2008. Division of labor at the eukaryotic replication fork. *Mol. Cell* 30(2):137–44
173. Krokan HE, Bjørås M. 2013. Base excision repair. *Cold Spring Harb. Perspect. Biol.* 5(4):a012583
174. Hecht SS. 1999. Tobacco smoke carcinogens and lung cancer. *J. Natl. Cancer Inst.* 91(14):1194–210
175. David SS, O'Shea VL, Kundu S. 2007. Base-excision repair of oxidative DNA damage. *Nature* 447(7147):941–50
176. Topal MD, Baker MS. 1982. DNA precursor pool: a significant target for *N*-methyl-*N*-nitrosourea in C3H/10T1/2 clone 8 cells. *PNAS* 79(7):2211–15
177. Huycke MM, Gaskins HR. 2004. Commensal bacteria, redox stress, and colorectal cancer: mechanisms and models. *Exp. Biol. Med.* 229(7):586–97
178. Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, et al. 2002. Inherited variants of *MYH* associated with somatic G:C→T:A mutations in colorectal tumors. *Nat. Genet.* 30(2):227–32
179. Lipton L, Halford SE, Johnson V, Novelli MR, Jones A, et al. 2003. Carcinogenesis in MYH-associated polyposis follows a distinct genetic pathway. *Cancer Res.* 63(22):7595–99
180. Smith CG, West H, Harris R, Idziaszczyk S, Maughan TS, et al. 2013. Role of the oxidative DNA damage repair gene *OGG1* in colorectal tumorigenesis. *J. Natl. Cancer Inst.* 105(16):1249–53
181. Kinnersley B, Buch S, Castellvi-Bel S, Farrington SM, Forsti A, et al. 2014. Re: Role of the oxidative DNA damage repair gene *OGG1* in colorectal tumorigenesis. *J. Natl. Cancer Inst.* 106(5):dju086
182. Speina E, Arczewska KD, Gackowski D, Zielińska M, Siomek A, et al. 2005. Contribution of hMTH1 to the maintenance of 8-oxoguanine levels in lung DNA of non-small-cell lung cancer patients. *J. Natl. Cancer Inst.* 97(5):384–95
183. Gad H, Koolmeister T, Jemth A-S, Eshtad S, Jacques SA, et al. 2014. MTH1 inhibition eradicates cancer by preventing sanitation of the dNTP pool. *Nature* 508(7495):215–21
184. Huber KVM, Salah E, Radic B, Gridling M, Elkins JM, et al. 2014. Stereospecific targeting of MTH1 by (*S*)-crizotinib as an anticancer strategy. *Nature* 508(7495):222–27
185. Helleday T. 2014. Cancer phenotypic lethality, exemplified by the non-essential MTH1 enzyme being required for cancer survival. *Ann. Oncol.* 25(7):1253–55
186. de Laat WL, Jaspers NG, Hoeijmakers JH. 1999. Molecular mechanism of nucleotide excision repair. *Genes Dev.* 13(7):768–85
187. DiGiovanna JJ, Kraemer KH. 2012. Shining a light on xeroderma pigmentosum. *J. Investig. Dermatol.* 132(3 Pt. 2):785–96
188. Cleaver JE. 1968. Defective repair replication of DNA in xeroderma pigmentosum. *Nature* 218(5142):652–56
189. Dumaz N, Drougard C, Sarasin A, Daya-Grosjean L. 1993. Specific UV-induced mutation spectrum in the p53 gene of skin tumors from DNA-repair-deficient xeroderma pigmentosum patients. *PNAS* 90(22):10529–33
190. Bodak N, Queille S, Avril MF, Bouadjar B, Drougard C, et al. 1999. High levels of patched gene mutations in basal-cell carcinomas from patients with xeroderma pigmentosum. *PNAS* 96(9):5117–22
191. Lee W, Jiang Z, Liu J, Haverty PM, Guan Y, et al. 2010. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. *Nature* 465(7297):473–77
192. Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, et al. 2010. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 463(7278):184–90

193. Polak P, Lawrence MS, Haugen E, Stoletzki N, Stojanov P, et al. 2014. Reduced local mutation density in regulatory DNA of cancer genomes is linked to DNA repair. *Nat. Biotechnol.* 32(1):71–75
194. Yasuda T, Sugasawa K, Shimizu Y, Iwai S, Shiomi T, Hanaoka F. 2005. Nucleosomal structure of undamaged DNA regions suppresses the non-specific DNA binding of the XPC complex. *DNA Repair* 4(3):389–95
195. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, et al. 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell* 149(5):979–93
196. Roberts SA, Sterling J, Thompson C, Harris S, Mav D, et al. 2012. Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. *Mol. Cell* 46(4):424–35
197. Sakofsky CJ, Roberts SA, Malc E, Mieczkowski PA, Resnick MA, et al. 2014. Break-induced replication is a source of mutation clusters underlying kataegis. *Cell Rep.* 7(5):1640–48
198. Conticello SG. 2008. The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* 9(6):229
199. Harris RS, Petersen-Mahrt SK, Neuberger MS. 2002. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol. Cell* 10(5):1247–53
200. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, et al. 2013. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat. Genet.* 45(9):970–76
201. Burns MB, Temiz NA, Harris RS. 2013. Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat. Genet.* 45(9):977–83
202. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, et al. 2013. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 494(7437):366–70
203. Taylor BJ, Nik-Zainal S, Wu YL, Stebbings LA, Raine K, et al. 2013. DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. *eLife* 2:e00534
204. Nik-Zainal S, Wedge DC, Petljak M, Butler AP, Bolli N, et al. 2014. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nat. Genet.* 46(5):487–91
205. Long J, Delahanty RJ, Li G, Gao Y-T, Lu W, et al. 2013. A common deletion in the *APOBEC3* genes and breast cancer risk. *J. Natl. Cancer Inst.* 105(8):573–79
206. Qian J, Wang Q, Dose M, Pruett N, Kieffer-Kwon KR, et al. 2014. B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. *Cell* 159(7):1524–37
207. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, et al. 2014. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat. Commun.* 5:2997
208. Meng FL, Du Z, Federation A, Hu J, Wang Q, et al. 2014. Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell* 159(7):1538–48
209. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, et al. 2014. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344(6190):1396–401
210. Bedard PL, Hansen AR, Ratain MJ, Siu LL. 2013. Tumour heterogeneity in the clinic. *Nature* 501(7467):355–64
211. Campbell PJ, Yachida S, Mudie LJ, Stephens PJ, Pleasance ED, et al. 2010. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* 467(7319):1109–13
212. Johnson BE, Mazor T, Hong C, Barnes M, Aihara K, et al. 2013. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* 343(6167):189–93
213. Hunter C, Smith R, Cahill DP, Stephens P, Stevens C, et al. 2006. A hypermutation phenotype and somatic *MSH6* mutations in recurrent human malignant gliomas after alkylator chemotherapy. *Cancer Res.* 66(8):3987–91

Annual Review of Pathology: Mechanisms of Disease

Volume 11, 2016

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Pathology: Mechanisms of Disease* articles may be found at http://www.annualreviews.org/errata/pathol