**Aneuploidy and complex genomic rearrangements in cancer evolution**

Toby M. Baker1,2,#, Sara Waise1,3,#, Maxime Tarabichi1,4 and Peter Van Loo1,2,5,\*

1The Francis Crick Institute, London, UK

2Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

3Cancer Sciences Unit, University of Southampton, Southampton, UK

4Institute for Interdisciplinary Research (IRIBHM), Université Libre de Bruxelles, Brussels, Belgium

5Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

#: these authors contributed equally

\*: correspondence: pvanloo@mdanderson.org

**Abstract**

Mutational processes that alter large genomic regions occur frequently in developing tumors. They range from simple copy number gains and losses to the shattering and reassembly of entire chromosomes. These catastrophic events, such as chromothripsis, chromoplexy and the formation of extrachromosomal DNA, affect the expression of many genes and therefore have a substantial impact on the fitness of the cells in which they arise. In this Review, we cover large genomic alterations, the mechanisms that cause them, and their impact on tumor development and evolution.

**Introduction**

Genomic alterations are a feature universal to all tumors. Among these, single nucleotide variants (SNVs), multi-nucleotide variants (MNVs) and small insertions and deletions (indels), affect only a few base pairs. Other alterations impact much larger regions and affect multiple genes. They include alterations that change the copy number of entire genomic regions, such as aneuploidy, whole genome duplications, and extrachromosomal DNA (ecDNA). Other forms of large-scale genome alterations include genomic catastrophes that cause extensive rearrangements, such as chromothripsis, chromoplexy and breakage-fusion-bridge cycles.

The frequency with which genomic alterations are observed depends on a combination of the physical factors that determine their rate of occurrence, and selective factors that determine their rate of fixation in the population, i.e. cell survival and proliferation1 (**Fig. 1**). Even if a particular alteration occurs very frequently in individual cells, it will rarely be observed at the bulk level if it confers a substantial selective disadvantage, as those cells will likely die or fail to proliferate. Conversely, an alteration that occurs very infrequently but confers a substantial selective advantage, may be observed in many tumors. As individual small alterations are numerous and affect only a limited region, both the magnitude of their selective effect and the targets under selection can be reasonably estimated. For example, among thousands of observed individual SNVs in each tumor, only a tiny fraction are thought to be under active selection2. In fact, most of the variance in SNV and indel mutation rates along the cancer genome can be explained by local tissue-specific chromatin density and mismatch repair3,4. Thus, differences in mutation frequencies along the genome are driven primarily by disparities in rates of occurrence (i.e. physical factors), rather than fitness effects.

Less is known about the fitness impact of large-scale genomic alterations in cancer. Although particular genes are recurrently impacted by large alterations5, each individual event typically affects multiple genes simultaneously, and likely has a greater average impact than smaller variants. Therefore, one could hypothesize that the frequencies of large-scale alterations are under substantially larger selection constraints than SNVs and indels. However, as the mutational processes and repair mechanisms associated with large-scale genome alterations are highly complex, physical factors must also have a substantial influence. Here we review large-scale genomic alterations in cancer and how the interplay between physical and selective effects determines their occurrence. We discuss the formation of copy number alterations, whole-genome duplications, and different genomic catastrophes, and their role in tumor initiation and subsequent evolution.

**Aneuploidy**

During the late 19th and early 20th centuries, Theodor Boveri and David von Hansemann first noted that aneuploidy, the deviation from a balanced complement of chromosomes in a cell6-8, was a feature common to human cancers, leading them to speculate that it may be the root cause of carcinogenesis9,10. We now know that cancers also commonly gain and lose smaller genomic regions, as well as whole chromosomes. The term copy number alteration (CNA) describes the complete range of gains and losses occurring in cancer genomes. Aneuploidy and CNAs are caused by chromosomal instability (CIN), a cellular state in which chromosomes, or sections of chromosomes, are gained or lost. As CIN can be episodic11,12, an aneuploid state may reflect past rather than ongoing CIN. It is therefore impossible to directly observe CIN from a copy number measurement at a single timepoint, although copy number heterogeneity within a tumor indicates recent CIN. Multi-region13 or single-cell sequencing experiments are therefore well suited to assess ongoing CIN. Numerous single-cell sequencing studies have revealed high levels of ongoing CIN in most studied cancer types, including primary tumours11,14-17, xenografts18,19 and cell lines20.

Although CNAs can be any length, they are on aggregate bimodally distributed, as both highly focal events and chromosome-/chromosome arm-level events are frequent. Nevertheless, not all underlying processes lead to bimodality. For example, processes underlying tandem duplications can generate alterations from kilobases to megabases in size21.

By encoding features such as CNA size and total copy number across tumors, sets of copy number signatures corresponding to recurrent features of CIN have been identified22,23. Many of these signatures have been linked to specific causes of CIN, such as homologous recombination deficiency and chromothripsis, which we review later. Several signatures were also linked to inactivating mutations in particular genes, thereby connecting driver mutations and the initiation of specific CIN processes22,23. Associations with response to therapy indicate that classifying CIN has clinical use23. However, many signatures have unknown origin, suggesting that multiple processes underlying CIN remain unidentified22,23.

Mechanisms that cause CIN and lead to aneuploidy and CNAs have been extensively reviewed elsewhere24,25, thus are only briefly summarized here. A major cause of CIN is mitotic defects that cause chromosome mis-segregation during anaphase, leading to daughter cells inheriting an imbalanced set of chromosomes. Mis-attachments of the mitotic spindle to kinetochores are key in generating aneuploidy. Cells ensure that all kinetochores are correctly attached to the mitotic spindle before anaphase through the spindle assembly checkpoint (SAC). Although mutations in SAC genes induce aneuploidy26 *in vitro*, they are rare *in vivo* and not thought to be a major contributor to aneuploidy in human cancers24. Instead, merotelic attachments, whereby microtubules from both spindle poles attach to a kinetochore27 causing lagging chromosomes undetected by the SAC28, are a major contributor to chromosome mis-segregation in cancer cell lines with CIN29.

Additional centrosomes can also induce CIN. Supernumerary centrosomes can cause a multipolar cell division, resulting in descendants with substantial chromosome mis-segregation. Although multipolar divisions are a source of instability in some cell lines30, particularly under conditions of external stress31, such catastrophic losses are typically unviable32. Instead, additional centrosomes tend to cluster together into two groups such that bipolar division ensues33,34. Surplus centromeres primarily drive aneuploidy by causing lagging chromosomes through an increase in merotelic attachments32. DNA replication stress has also been linked to increased levels of CIN35.

Mutations that inactivate *TP53* have been linked to higher CNA levels in human cancers36-39, suggesting that they may induce or allow greater tolerance to CIN. Such mutations are early events in most cancer types12,40-42, and may thus initiate some of the earliest CIN in tumor development. However, the cell line evidence for *TP53* inactivation directly causing CIN is mixed, underscoring that the overall context of growth conditions is important43. In addition, functional TP53 may more potently inhibit structural chromosomal alterations than chromosome-level gains or losses44.

Although gains and losses of all chromosomes are observed in most tumor types38,45,46, their frequencies vary greatly. Some aneuploid states are near-ubiquitous in particular cancers, e.g. loss of the 3p arm in clear-cell renal cell carcinomas47, 17p loss in ovarian adenocarcinomas48, and chromosome 7 gain and chromosome 10 loss in glioblastomas49. However, most individual chromosome gains or losses only occur in a minority of tumors.

Some of the differences in CNA frequency across the genome may be due to the physical characteristics of chromosomes. For example, in cells from the female Indian muntjac, a species with very few chromosomes (2n=6), chromosomes with larger kinetochores are more likely to mis-segregate50. The DNA content of the centromere on each human chromosome has also been shown to affect mis-segregation51. Chromosome distance from the nuclear center at interphase was reported to correlate with mis-segregation likelihood during mitosis52. Distance from the nuclear center correlates with chromosome size and therefore, larger chromosomes mis-segregate more frequently53. Other physical features that affect CNA frequencies remain to be identified.

Substantial evidence points to selection driving the observed frequencies of different CNAs. The measured gain and loss frequencies of chromosome arms are positively correlated with oncogene and tumor suppressor gene density, respectively54. Genes associated with proliferative potential in certain tissues are linked to recurrent CNAs in corresponding cancers55. Homozygous inactivation of tumor suppressor genes often occurs through the combination of a loss of one parental allele and an inactivating mutation on the other allele56. This could either be a germline mutation conferring familial risk (e.g. *BRCA1* in breast and ovarian cancer57), or a somatic mutation (e.g. *TP5358*). Furthermore, CNAs that arise independently multiple times over the course of a single tumor’s timeline have been identified, further implying that specific CNAs confer a selective advantage13,59. Finally, by considering the breakpoints of telomere- and centromere-bounded copy number changes across thousands of tumors, a recent study posited that selective pressures have a greater impact on the CNA landscape than physical considerations1.

Even though certain recurrent CNAs appear actively selected for in different tumor types, the experimental evidence for the tumorigenic and proliferative effects of aneuploidy and ongoing CIN is mixed6,60-62. Nevertheless, aneuploidy and CNAs more generally are associated with progression and poorer prognosis in different tumor types63-66. Tumors with intermediate CIN are associated with the worst outcomes, with tumors with the most unstable genomes having better prognosis60,67,68. Separating the impact of existing CNAs from that of ongoing CIN is challenging, despite high levels of subclonal copy number diversity – likely emerging due to recent CIN – leading to poorer outcomes in multiple cancer types6,59,69. Therefore, the effects of CNAs and CIN are highly contextual, as recently reviewed elsewhere6.

There is conflicting evidence as to whether CIN helps or hinders tumor immune responses70,71, but this may reflect disparate responses to CIN from different components of the immune system. Triple-negative breast cancer cells with CIN were shown to rely on the innate immune system through the cGAS-STING pathway for survival72.

CNAs may play a key role in early tumor development. Healthy tissues frequently contain clones with driver mutations linked to tumor development and initiation, yet few CNAs73-77, with the exception of the liver and placenta78,79. Aneuploidy in pre-malignant lesions has been linked to progression in many cancer types2.

Some CNAs appear to be recurrent initiating events in certain cancers. Gains of chromosomes 7, 19 and 20 as well as losses of chromosome 10 are estimated to occur already in childhood in glioblastomas, and potentially even earlier12. The loss of chromosome 3p combined with gain of 5q is also a very early event in clear-cell renal cell carcinoma80. However, most gains occur over a wide period during tumor evolution, though they follow cancer-type specific trends12. Breast and stomach adenocarcinomas accumulate gains throughout their evolution, whereas cancers driven by exogenous mutagens, such as melanomas and lung cancers, typically only accumulate gains after a substantial number of mutations have occurred12.

Certain CNAs also appear to be advantageous for later-stage metastasis development, such as loss of chromosome 9p in clear-cell renal cell carcinomas81, and gain of 12p in pancreatic adenocarcinomas82. Generally, metastases have more aberrant genomes than primary tumors82-84, consistent with a late stage of tumor evolution85,86. However, this may in part be due to a detection bias: CNAs that are present in the metastasis-seeding cells but at very low overall frequency in the primary tumor are much easier to detect in the metastases due to clonal expansion of the seeding cell(s)87. The presence of CIN has been found to increase expression of epithelial-mesenchymal transition genes, necessary for metastatic spread, as well as promoting an inflammatory response that enhances metastasis83.

In general, subclonal CNAs appear to be under similar selection pressures as earlier clonal alterations. As with clonal CNAs, the greater the density of oncogenes relative to tumor suppressor genes on a chromosome arm, the more likely it is to have a subclonal gain13. The same copy number events have been found to arise independently in different expanding populations within the same tumor, further indicating active selection for some CNAs13,88. However, a mathematical model based on multi-region sequencing has suggested that, in colorectal cancer, after an initial burst of CNAs, the accumulation of mutations and CNAs may be under little selective pressure89.

Although CNAs continue to accumulate over the course of tumor development, a number of copy number gains appear to occur in periodic bursts. Many clonal copy number gains were shown to arise over a much shorter period than expected by chance, across multiple cancer types12. Multiple single-cell and multi-region analyses also point to a burst of high levels of chromosomal instability, followed by lower levels later in tumor evolution11,14,90.

**Whole Genome Duplications**

Tumor cells can double their entire genome in a single event. Frequency estimates for such whole genome duplications (WGDs) range from 28-37% in pan-cancer cohorts38,39 and up to 56% in metastases84. However, the indicated higher frequency of WGD in metastases may only be true for certain cancer types39,91. WGD rates vary greatly between different cancer types, low in clear-cell renal cell carcinomas, differentiated thyroid cancers, non-Hodgkin lymphomas, and gastrointestinal neuroendocrine tumors, but very high in osteosarcomas and germ-cell tumors12,39.

WGDs can arise through cytokinesis defects that result in a tetraploid binucleated cell, leading to genome duplicated descendants after a subsequent successful mitosis92 (**Fig. 2a**). Similarly, sustained mitotic arrest by the spindle assembly checkpoint can result in the cell exiting mitosis without completion, or skipping it completely93 (**Fig. 2b**). Among the factors that can cause these cytokinesis defects and mitotic failures are mutations or aberrant expression of well-known tumor suppressor genes such as APC and BRCA245,92, as well as telomere crisis94. Cell fusion, promoted by viral infection95, can also cause WGDs (**Fig. 2c**). Cell fusion frequency in tumor development remains unclear, as it is difficult to observe, although it has been identified through the presence of donor DNA in tumors from patients who received bone marrow transplants96.

WGD has been linked to increase in CIN and subsequent aneuploidy, for instance by generating additional centrosomes, as discussed earlier. Additionally, a cell that has recently undergone WGD may have insufficient machinery to fully replicate, causing both under- and over-replicated regions in progeny cells97. In tumors with a subclonal WGD, subclones with WGD have more CNAs than non-WGD subclones, further supporting an increase in CIN caused by WGD13. The relative frequency of whole chromosome aneuploidies compared to arm-level aneuploidies is much higher for WGD than non-WGD tumors, pointing towards mis-segregation as a major consequence of increased instability arising from WGD98. WGDs also appear to increase tolerance to aneuploidy99. Indeed, the average copy number of each chromosome in genome duplicated tumors is around three, indicating substantial chromosomal loss in duplicated cells39.

As WGD doubles the DNA content to be replicated per cell cycle, it imposes a substantial burden on the cell. Therefore, the high frequency of WGD in tumors implies that it may confer a selective advantage. Consistent with this, WGD has been linked to poorer survival in multiple cancer types39. Furthermore, *TP53* knock-out cell line experiments indicate that only tetraploid cells give rise to tumors when transplanted into mice100,101.

WGD has been proposed to mitigate the effect of deleterious mutations in cancer development, by providing additional copies of essential genes102. Mutations in essential genes in tumor genomic regions with haploid loss of heterozygosity are under negative selection103, which is relieved following WGD102. Before WGD, losses are less frequent on arms with a high concentration of essential genes, but no such relationship is observed for post-WGD losses102. WGDs therefore likely change the selection landscape of copy number losses. In contrast, WGDs appear to have little impact on the fitness landscape of copy number gains. Although the frequency of specific gains differs between WGD and non-WGD tumors (such as gains of chromosomes 19 and 20 in glioblastoma, which are considerably more common in non-WGD tumors104), in most cases the event frequencies are highly correlated98,104.

Another potential advantage of genome duplications is that increased CIN post-WGD provides more opportunities for CNAs with high selective impact and therapy resistance105,106. Moreover, WGD events were reported to promote tumorigenesis through changes to chromatin organization101. Separately, computational simulations were employed to demonstrate that WGD allows cancer cells to reach an optimal triploid copy number state with lower fitness cost than through sequential independent chromosome gains107.

WGDs often appear within the clonal evolution period of cancer13, typically following initiating driver mutations and CNAs12. Real-time estimates point to WGDs occurring years and sometimes decades before diagnosis12. When they both occur in the same tumors, a *TP53* inactivating mutation almost always precedes WGD, supporting that *TP53* mutations either increase WGD occurrence, or increase tolerance to WGD39,108.

**Structural variants**

Genomic rearrangements, also called structural variants (SVs), are a key source of somatic mutations in human cancer21,109. Although some of these SVs are interlinked and share underlying causes with CNAs (**Fig. 2d**), they form a distinct semantic class of changes. These may be classified as simple SVs, occurring in isolation (*e.g.* translocations and inversions)110,111, or as part of more complex events where multiple rearrangements may occur simultaneously, or through a linked series of events112-114. Recent whole-genome sequencing studies have enabled identification of novel rearrangement patterns, and refinement of the classification of known processes21,114,115, indicating that at least some of these events have a key role in the initiation and progression of cancer113,116.

DNA replication and repair mechanisms have key roles in the generation of both SVs and CNAs117-119. For example, breakpoints in chromothripsis, ecDNA, and CNA events have been observed to show microhomology110,120,121, indicative of breaks repaired by non-homologous end joining (NHEJ, which in contrast to homologous recombination, does not require template homology122) or microhomology-mediated break-induced replication (MMBIR, repairing single-stranded DNA which shares microhomology with the 3’ end of the replication fork123). Another double-strand DNA repair mechanism, alternative end-joining, has been particularly implicated in translocations122.

Simple rearrangements usually arise from breakpoints restricted to one or two chromosomes111,124, and are common across many cancer types21. These events may be further sub-categorized according to whether they alter DNA copy number. Balanced variants, such as inversions and reciprocal translocations, result in relative conservation of copy number111, whereas unbalanced rearrangements (*e.g.* deletions, duplications, non-reciprocal translocations) lead to CNAs111. Rearrangements were first linked to oncogenesis with the identification of the *BCR*-*ABL1* gene fusion in leukemia125, and have since been implicated in the development of a number of cancer types126-130.

Two more complex processes related to simple rearrangements were identified recently: rigmaevents comprise cumulative clustered deletions occurring in *trans*. This process appears to be distinct from chromothripsis, in which clustered breakpoints are confined to one allele110,115. Pyrgo events were characterized by clustered duplications without a significant increase in copy number115. Relative to their simple counterparts, rigmas are enriched in late-replicating regions, fragile sites, and large genes. Similarly, pyrgo events occur more frequently in early-replicating regions and super-enhancers. These differences may be due to distinct selection pressures or driver processes115.

Chromoplexy is characterized by interdependent, shuffled, apparent chains of rearrangements, typically involving multiple chromosomes113,131. These events are also likely related to simple rearrangements, namely balanced translocations, which can be viewed as chromoplexy chains involving only two chromosomes21 (**Fig. 3a**). As such, chromosome translocations and chromoplexy chains are likely the same entity. Breakpoints in chromoplexy may be associated with deletion bridges, or generate gene fusions with relative conservation of copy number132. These rearrangements underlie pathogenic gene fusions and enhancer hijacking events in a range of malignancies21,109,116,119,131,133, and may result in synchronous disruption of multiple oncogenes113,132-134. Chromoplexy is particularly prevalent in prostate cancer and frequently generates *TMPRSS2-ERG* driver fusions113,116,134. To date, chromoplexy events have been inferred solely from whole genome sequencing data, and the initiating mechanism is not currently well-characterized. However, chromoplexy chains are likely formed by double-stand breaks involving multiple chromosomes, followed by incorrect repair111,135. Chromoplexy chains often co-occur with other rearrangement patterns, such as chromothripsis (discussed below), and may have a causal role in some of these events116,119,136,137.

**Chromothripsis**

Chromothripsis describes a phenomenon whereby chromosomes are shattered into fragments that are rejoined near-randomly in a single event110 (**Fig. 3b**). This generates massive, clustered rearrangements restricted to either a single chromosome, or localized regions of multiple chromosomes110,138. These events are prevalent across cancers and are particularly frequent in bone and soft tissue tumours21,116,139.

A set of criteria have been defined to distinguish chromothripsis from other rearrangement events120. First, as chromothripsis is a localized phenomenon, it leads to high density breakpoint clustering in a short genomic interval110,120. Moreover, these events are confined to one allele, and are frequently associated with loss of genomic segments110,138. Classical chromothripsis events in a diploid context should therefore show alternating segments with retention and loss of heterozygosity, leading to a characteristic profile of copy number oscillation between two states120,139. Finally, as rejoining of shattered chromosomes is a random process, breakpoints from the four classes of intrachromosomal rearrangements (tandem duplications, deletions, and head-to-head and tail-to-tail inversions) should be represented in approximately equal proportions110.

Chromothripsis events may occur in diverse genomic contexts, with localized rearrangements in otherwise quiet genomes, or within a background of more widespread genomic rearrangements116,139. The properties of chromosomal changes have been shown to differ widely across tumor types116, suggesting different underlying mechanisms or selective constraints. For example, in prostate cancer, classical single-chromosome chromothripsis is frequent, whereas in glioblastoma, small focal chromothripsis events are particularly common, and in melanoma, osteosarcoma and liposarcoma, chromothripsis events often involve multiple chromosomes116. Many of these events are followed by amplification, which we discuss in the following section.

From a mechanistic perspective, chromothripsis is one of the better-studied complex rearrangement processes. Polyploidy and abnormalities in *TP53* appear to predispose to chromothripsis in some tumor types, but it may also occur in diploid tumors with normal p53 function121,139. Chromothripsis can result from dicentric chromosome formation as part of breakage-fusion-bridge cycles (discussed below), or through DNA damage to lagging chromosomes in micronuclei140. This spatial restriction explains the clustered nature of the observed rearrangements120,139. Chromothripsis events arising through these routes show distinct genomic footprints and underlying DNA breakage mechanisms141. Events secondary to chromatin bridge breakage show highly localized breakpoints with frequent foldback inversions, whereas rearrangements following micronuclear rupture are distributed more evenly across chromosomes and show oscillation between two or three copy number states141. In chromothripsis resulting from chromosome bridge breakage, DNA breakage occurs by mechanical force140.

At mitosis, lagging chromosomes can recruit a nuclear envelope to form a micronucleus142. However, recruitment of key nuclear proteins to lagging chromosomes is inhibited by bundled spindle microtubules143, resulting in defective micronuclear function and collapse of the nuclear envelope during S phase12-15. On exposure to cytoplasm, DNA previously contained within the micronucleus is subject to damage and fragmentation144. This phenomenon seems to be a function of aberrant DNA repair: in experimental models, accumulation of DNA-RNA hybrids on chromatids within micronuclei ultimately results in induction of double strand DNA breaks, largely mediated through the base excision repair enzymes APE1 and MPG144. Experimentally induced loss of these proteins vastly reduces, but does not fully ablate, DNA damage accumulation in interphase144. These findings indicate a role for multiple processes in fragmentation of chromatids at chromothripsis, with both NHEJ and MMBIR representing possible additional candidates120,121,145.

DNA arrangements generated by a single chromothripsis event have the potential to yield multiple oncogenic alterations120. Consistent with this, the presence of chromothripsis confers a worse outcome in multiple cancer types121,146-149. Chromothripsis could promote tumorigenesis through focal deletion of tumor suppressor genes present on lagging chromosomes38,46,110,116,139,150. Rearrangements can also generate enhancer hijacking events and oncogenic gene fusions119,151. Finally, chromothripsis could result in oncogene amplification, for example through ligation of fragments to generate circular chromatids, leading to formation of ecDNA110,114,139,145,150 (discussed below). Chomothripsis-amplification has been reported in multiple tumor types, including sarcoma and glioblastoma116,139. In a recent pan-cancer study, chromothripsis followed by amplification of underlying oncogenes (likely through multiple independent physical processes) was the most prevalent consequence on cancer genes116, suggesting positive selection for such events, as discussed in the next section.

**Complex structural variants underlying high-level amplifications**

High-level amplifications are frequent oncogenic events in cancer genomes and can be caused by multiple processes. First described in maize genomes152, breakage-fusion-bridge (BFB) cycles are now a recognized mechanism of gene amplification in multiple cancer types153. These rearrangements may be initiated by various events, including dysregulation or loss of telomeres112,154, fragile site expression155,156, and DNA breaks induced by replication stress154. The key event initiating BFB cycles is the fusion of two sister chromatids, generating a dicentric chromosome157. During the subsequent anaphase, this may break at any point between the two centromeres154, resulting in double-strand breaks which may in turn trigger further cycles (**Fig. 3c**). These events may be terminated by telomere stabilization, either through capping or by fusion with another chromosome158. Repeated cycles of chromosome breakage and end fusion result in a characteristic pattern of palindromic fold-back inversions and jumps in DNA copy number112,154,158,159.

Akin to mis-segregations, BFB cycles lead to uneven inheritance of genetic material. Over multiple iterations, this process may generate high copy-number heterogeneity, facilitating selection of cell populations with tumorigenic profiles. In human cancers, BFB cycles have been implicated in the amplification of specific oncogenes, including *ERBB2* in HER2-positive breast tumors158-160, *CCND1* in head and neck squamous cell carcinoma161 and *MET* in gastric cancer162. Tumors exhibiting the structural hallmarks of BFB cycles show shorter survival in patients with head and neck, and ovarian cancers161,163. Chromothripsis is often observed together with other rearrangement patterns3, such as BFBs110,112,154. Micronucleus formation may provide a unifying mechanism linking these two processes: broken bridge chromosomes have been observed to segregate into micronuclei, leading to chromothripsis140.

Recently, three further complex rearrangement processes generating gene amplification have been inferred from whole genome sequencing data. The first category, cycles of templated insertions, were described alongside a distinct computational approach for identifying and classifying complex rearrangements21. This process is characterized by duplication of multiple segments from different reference chromosomes with insertion into one derivative chromosome and appears to be an important cause of *TERT* activation in hepatocellular carcinoma21.

Tyfonas115 and seismic amplifications114 are complex amplification events that share underlying characteristics with ecDNA. Tyfonas generate elevated copy number and, similar to BFBs, are enriched in fold-back inversions115. However, theseamplicons show higher copy number and typically contain a higher proportion of the genome than other amplification mechanisms115,164. Relative to ecDNA and BFBs, tyfonas were found to be enriched in sarcoma, melanoma and breast cancer115. These events are associated with the characteristic *MDM2*/*CDK4* amplification in dedifferentiated liposarcoma115, and may underlie the circular neochromosomes observed in this tumor type115,165.

Seismic amplifications involve chromothripsis followed by repeated circular recombination generating high-level amplifications on ecDNAs114 (**Fig. 3d**). ecDNA is often formed by the ligation of chromothriptic fragments, as ~30% of amplicons show evidence of chromothripsis164. Seismic amplifications appear to represent a combination of known mechanisms, but with distinct characteristics. For example, in contrast to tyfonas*,* fold-back inversions are lacking. These events may account for a proportion of previously unclassifiable rearrangements21,114. Other related rearrangement processes, such as BFB events, may also generate ecDNA114,166.

**ecDNAs**

ecDNAs are double-stranded circular chromatids, ranging from approximately 100 kilobases to several megabases in size167-170. These elements are distinct from neochromosomes, lacking centromeric and telomeric sequences164,165. First described in neuroblastoma in 1965171, ecDNAs are now established as common carriers of amplified oncogenes164, known to promote tumorigenesis and mediate therapeutic resistance, and associated with worse outcome in multiple cancer types164,168,172. Although ecDNA in cancer has been recently reviewed elsewhere168,173, we outline some key results on ecDNA in tumor development.

ecDNA lacks centromeric sequences and therefore segregate unevenly at mitosis. This can lead to large increases in copy number, and expansion of intra-tumor heterogeneity over relatively few cell cycles, enabling rapid adaptation to environmental pressures167,168,170,174-176. The resulting genetically diverse pool facilitates selection of cells with mutational profiles conferring a fitness advantage168,169. Genes encoded on ecDNA show increased copy number, with concomitant elevation in mRNA transcript levels168,174. Interestingly, however, this increase in transcription is not solely a function of increased copy number: oncogenes amplified on ecDNA show higher levels of transcription than copy number-matched chromosomal DNA164,167,168.

ecDNA can also act as a mobile regulatory element174. Given that it exhibits reduced nucleosome compaction, with depletion of repressive histone marks and enrichment of active histone marks167, ecDNA shows significantly enhanced chromatin accessibility relative to linear chromosomes164,167. In addition, ecDNA shows widespread long-range chromatin connectivity, with contact sites predominantly in promoter regions167. These elements are enriched for enhancer signatures, and associate particularly with transcriptionally active chromosomal genes167. Given these observations, ecDNA is proposed to increase oncogene transcription by acting in *trans* as a mobile transcriptional super-enhancer167,169.

ecDNA also enhances oncogene transcription by hijacking regulatory elements in *trans.* DNA FISH of interphase nuclei revealed clustering of 10-100 ecDNAs169. Molecules in these hubs showed higher levels of transcription than those outside, allowing ecDNAs lacking an enhancer to access enhancer elements of nearby molecules169. This spatial localization is reported to depend on the bromodomain and extra-terminal domain protein BRD4169.

**Complex rearrangements and tumor evolution**

As detailed above, complex rearrangement events may substantially alter copy number and expression of oncogenes and tumor suppressor genes, and generate oncogenic gene fusions. As such, they can be important tumor driver events110,116, generating so-called evolutionary hopeful monsters177,178 with key roles in tumor initiation, development and evolution. Multi-cellular organisms evolved multiple mechanisms to restrain the occurrence and/or fixation of these events, some of which might be unknown. For example, in the case of chromothripsis, an Aurora-B-mediated correction of kinetochore positioning avoids mis-segregation events179, and cGAS-STING-mediated clearance of micronuclei leads to autophagic degradation of DNA fragments180. Even after escaping the cell’s surveillance and clearance mechanisms, large-scale rearrangements would generally be expected to have a negative fitness effect, given the scale of their disruption. For example, chromosome shattering will break apart multiple genes and therefore, most chromothripsis occurrences would likely be lethal or lead to a selective disadvantage. The chromothripsis and other complex events observed in cancer genomes, hence likely represent a biased series of events, enriched for occurrences under positive selection.

As with CNAs, physical considerations also affect the frequency with which different complex rearrangements are observed. Larger chromosomes more frequently mis-segregate and become entrapped in micronuclei, where they may undergo chromothripsis, partially explaining the higher rates of chromothripsis observed in larger chromosomes52. Although chromosome size is predictive of chromosomal frequency of chromothripsis, correcting for chromosome size seems to reveal further trends related to their physical and selective constraints. Notably chromosome 17p has a significantly higher rate of involvement in chromothripsis events than expected from its size alone72. This could be due to both the selective advantage of disrupting *TP53* (located on 17p), and the relatively shorter telomeres of 17p181-183.

Analysis of cancer genomes provides insight into the timing of complex events in tumor evolution. Recent data suggest that chromothripsis is more commonly early and clonal rather than late and subclonal110,116,121, although there may be some exceptions139. Consistent with this, chromothripsis events are enriched in known cancer drivers and are particularly associated with driver gene amplification3. In acral melanomas, which frequently exhibit amplified chromothripsis, both chromothripsis and the subsequent amplification of the oncogene *CCND1* typically occur early in tumor evolution116. In lung adenocarcinomas, oncogene amplifications associated with complex rearrangements, including chromothripsis and chromoplexy, are also typically early events119. In contrast, although chromothripsis was found to be an early event in lung squamous cell carcinoma, when associated with *SOX2* amplification, it often appeared later in tumor development116. Although longitudinal sequencing showed that chromoplexy can occur both early and late in prostate cancers184, this generally appears to be a clonal event116. These observations are consistent with the known role of chromothripsis and chromoplexy in generating driver events or oncogenic gene fusions in multiple fusion-driven cancer types (including *EML4*-*ALK* in lung adenocarcinoma and *TMPRSS*­2-*ERG* in prostate cancer)21,109,113,119,131,133,134,137,185,186, as well as having the potential to disrupt multiple drivers at once in a single catastrophic event. Indeed, in prostate and pancreatic cancers and acral melanomas, single complex events targeting multiple driver genes have been reported113,116,187.

Treatment-induced selective pressures may also result in complex rearrangement events occurring later in tumor evolution. Pre- and post-treatment sequencing of colorectal cancer biopsies revealed chromothripsis- or BFB-mediated amplifications of mutant *BRAF*, leading to resistance to vemurafenib166. ecDNA has been particularly implicated in acquired resistance to targeted therapy *in vitro*175. Glioblastoma cells harboring mutant *EGFR* on ecDNA reduce *EGFR* copy number in response to tyrosine kinase inhibitor administration, conferring drug resistance175. Similarly, extrachromosomal amplification of the *DHFR* and *BRAF* genes confers therapeutic resistance in cervical carcinoma and melanoma models, respectively188,189. In at least some cases, this process is dynamic. In glioblastoma, mutant *EGFR* copy number rebounds following drug withdrawal175. In melanoma, there is evidence of dynamic switching between ecDNA and chromosomal amplification in response to MAPK inhibitor administration189.

Data regarding when other complex rearrangements arise during tumor evolution are sparse. It is likely that *rigma* and BFB events occur early in the development of oesophageal115 and pancreatic112 adenocarcinoma, respectively. As BFB events are closely associated with chromothripsis110,112,154, this process may represent an early event in other tumors. However, direct evidence for the contribution and timing of complex events in tumor evolution is currently lacking21,114,115.

**Conclusion and future perspectives**

Next-generation sequencing technologies have allowed the systematic and detailed identification of large-scale genome alterations in thousands of tumors. Numerous different types of complex rearrangements have been identified so far, as well as different CNA patterns. Their frequencies of occurrence result from the interplay between physical effects that determine their rate of formation, and the impact they have on cellular fitness. Both these factors are still poorly understood. A better understanding of the complex perturbations that CNAs and complex genomic rearrangements impose on cells will require the combination of experimental systems that can induce such alterations with more sophisticated computational models and analyses. By resolving these large-scale alterations at the single-cell level, we may be able to separate the influence of physical and selective factors, through a comparison of the rate of their occurrence in individual cells to the rate of their progenies’ expansion.

Large-scale genome alterations often play a crucial role in tumor initiation, development, and evolution. Timing inference from sequenced cancer genomes suggests that these complex events often are critical early driver events in tumor evolution, advocating a potential role in the initiation of some cancers. Some of these events, particularly those leading to the formation and further evolution of ecDNAs, are highly dynamic and can evolve quickly under selective pressures imposed by the tumor microenvironment or by cancer therapies. However, many unanswered questions remain. How do we more precisely time the occurrence of complex rearrangements in cancer development and evolution? Can we build statistical models that work from a single biopsy, or is longitudinal data necessary? What are the precise physical factors that determine their occurrence across cancers and cancer types? How do these changes affect phenotype, and how can we model or measure the fitness effects of complex events or large CNAs in a specific genomic context? How do we disentangle the underlying mutational processes, given that few large-scale CNAs and complex rearrangements are expected to be selectively neutral? How do we understand the interactions between such events?

Although CNAs and SVs are intimately linked, they are often considered separately, leading to limitations in their detection and characterization. Moreover, there is currently no consensus in the field on how complex SVs and their associated CNAs can be classified. A more integrated analysis of genomic rearrangements and CNAs is required to delineate their roles in tumor development. A CIN signature approach that includes both structural and CNA features22,23 would likely allow a better understanding of the processes that lead to these alterations. In addition, the detection and characterization of these events may be improved further with platforms that sequence longer reads, aiding the identification, precise reconstruction, and classification of complex rearrangement patterns.

**Figure legends**

**Figure 1. The effect of selective and physical considerations on alteration frequencies. a**, An illustration of three arbitrary alterations with different rates of formation and selective advantage. **b**,Illustration of the impact of both physical and selective effects on the observed frequencies of large-scale genomic alterations. Alterations with a high formation rate due to physical effects will occur frequently in individual tumor cells, however, a positive selective impact is needed for cells with the alterations to expand relative to non-altered tumor cells. Alteration A will occur more often than other alterations in individual tumor cells, but will not expand due to its reduced fitness. Alterations B and C arise at a similar rate, but B is able to expand much more than C due to the large fitness increase it confers.

**Figure 2. Causes of whole genome duplication and links between complex genomic events. a**, Cytokinesis failure during mitosis can result in binucleation, and genome duplicated cells in subsequent cell divisions. **b**, Mitotic slippage is a process where mitotic arrest can result in the cell skipping mitosis. **c**, Two tumor cells or a tumor and healthy cell can fuse to form a genome duplicated tumor cell. **d,** Schematic showing how a lagging chromosome event can cause aneuploidy as well as lead to the formation of micronuclei along with broken bridge chromosomes.

**Figure 3. A summary of the processes underlying four classes of large-scale genomic rearrangements.** **a**, Chromoplexy arises through translocations involving multiple chromosomes, generating apparent chains of rearrangements. **b**, In chromothripsis, chromosomes are shattered and near-randomly rejoined in a single event. **c**, Breakage-fusion-cycles are initiated by fusion of two sister chromatids, generating a dicentric chromosome. This chromosomal bridge breaks during the subsequent anaphase, resulting in double strand DNA breaks which may trigger further cycles. **d**, Seismic amplifications represent a combination of previously-described rearrangement events, whereby chromothripsis is followed by amplification through ecDNA formation.

**Acknowledgements**

The authors would like to thank Nischalan Pillay for helpful discussion on the manuscript. T.M.B., S.W., and P.V.L. are supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (CC2008), the UK Medical Research Council (CC2008), and the Wellcome Trust (CC2008). For the purpose of open access, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. T.M.B. is supported by a PhD fellowship from Boehringer Ingelheim Fonds. S.W. is a Jean Shanks/Pathological Society Clinical Lecturer. M.T. was supported as a postdoctoral researcher of the F.R.S.-FNRS. P.V.L. is a CPRIT Scholar in Cancer Research and acknowledges CPRIT grant support (RR210006).

**Competing Interests**

The authors declare no competing interests.

**Data Availability**

This manuscript contains no original research or analysis.

**References**

1. Shih, J. *et al.* Cancer aneuploidies are shaped primarily by effects on tumour fitness. *Nature* **619**, 793-800 (2023).

2. Stratton, M.R., Campbell, P.J. & Futreal, P.A. The cancer genome. *Nature* **458**, 719-24 (2009).

3. Polak, P. *et al.* Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **518**, 360-364 (2015).

4. Supek, F. & Lehner, B. Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature* **521**, 81-4 (2015).

5. Mermel, C.H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* **12**, R41 (2011).

6. Ben-David, U. & Amon, A. Context is everything: aneuploidy in cancer. *Nat Rev Genet* **21**, 44-62 (2020).

7. Boveri, T. *Zur frage der entstehung maligner tumoren*, (Fischer, 1914).

8. von Hansemann, D. Ueber asymmetrische Zelltheilung in Epithelkrebsen und deren biologische Bedeutung. *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin* **119**, 299-326 (1890).

9. Hardy, P.A. & Zacharias, H. Reappraisal of the Hansemann-Boveri hypothesis on the origin of tumors. *Cell Biol Int* **29**, 983-92 (2005).

10. Wunderlich, V. JMM---past and present. Chromosomes and cancer: Theodor Boveri's predictions 100 years later. *J Mol Med (Berl)* **80**, 545-8 (2002).

11. Gao, R. *et al.* Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat Genet* **48**, 1119-30 (2016).

12. Gerstung, M. *et al.* The evolutionary history of 2,658 cancers. *Nature* **578**, 122-128 (2020).

13. Watkins, T.B.K. *et al.* Pervasive chromosomal instability and karyotype order in tumour evolution. *Nature* **587**, 126-132 (2020).

14. Minussi, D.C. *et al.* Breast tumours maintain a reservoir of subclonal diversity during expansion. *Nature* **592**, 302-308 (2021).

15. Gao, R. *et al.* Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat Biotechnol* **39**, 599-608 (2021).

16. Baslan, T. *et al.* Novel insights into breast cancer copy number genetic heterogeneity revealed by single-cell genome sequencing. *Elife* **9**(2020).

17. Zaccaria, S. & Raphael, B.J. Characterizing allele- and haplotype-specific copy numbers in single cells with CHISEL. *Nat Biotechnol* **39**, 207-214 (2021).

18. Zahn, H. *et al.* Scalable whole-genome single-cell library preparation without preamplification. *Nat Methods* **14**, 167-173 (2017).

19. Salehi, S. *et al.* Clonal fitness inferred from time-series modelling of single-cell cancer genomes. *Nature* **595**, 585-590 (2021).

20. Wu, C.Y. *et al.* Integrative single-cell analysis of allele-specific copy number alterations and chromatin accessibility in cancer. *Nat Biotechnol* **39**, 1259-1269 (2021).

21. Li, Y. *et al.* Patterns of somatic structural variation in human cancer genomes. *Nature* **578**, 112-121 (2020).

22. Steele, C.D. *et al.* Signatures of copy number alterations in human cancer. *Nature* **606**, 984-991 (2022).

23. Drews, R.M. *et al.* A pan-cancer compendium of chromosomal instability. *Nature* **606**, 976-983 (2022).

24. Gordon, D.J., Resio, B. & Pellman, D. Causes and consequences of aneuploidy in cancer. *Nat Rev Genet* **13**, 189-203 (2012).

25. Sansregret, L. & Swanton, C. The Role of Aneuploidy in Cancer Evolution. *Cold Spring Harb Perspect Med* **7**(2017).

26. Giam, M. & Rancati, G. Aneuploidy and chromosomal instability in cancer: a jackpot to chaos. *Cell Div* **10**, 3 (2015).

27. Guerrero, A.A., Martinez, A.C. & van Wely, K.H. Merotelic attachments and non-homologous end joining are the basis of chromosomal instability. *Cell Div* **5**, 13 (2010).

28. Cimini, D. *et al.* Merotelic kinetochore orientation is a major mechanism of aneuploidy in mitotic mammalian tissue cells. *J Cell Biol* **153**, 517-27 (2001).

29. Thompson, S.L. & Compton, D.A. Examining the link between chromosomal instability and aneuploidy in human cells. *J Cell Biol* **180**, 665-72 (2008).

30. Telentschak, S., Soliwoda, M., Nohroudi, K., Addicks, K. & Klinz, F.J. Cytokinesis failure and successful multipolar mitoses drive aneuploidy in glioblastoma cells. *Oncol Rep* **33**, 2001-8 (2015).

31. Tyagi, I.S. *et al.* Intrinsic and chemically-induced daughter number variations in cancer cell lines. *Cell Cycle* **20**, 537-549 (2021).

32. Ganem, N.J., Godinho, S.A. & Pellman, D. A mechanism linking extra centrosomes to chromosomal instability. *Nature* **460**, 278-82 (2009).

33. Ring, D., Hubble, R. & Kirschner, M. Mitosis in a cell with multiple centrioles. *J Cell Biol* **94**, 549-56 (1982).

34. Quintyne, N.J., Reing, J.E., Hoffelder, D.R., Gollin, S.M. & Saunders, W.S. Spindle multipolarity is prevented by centrosomal clustering. *Science* **307**, 127-9 (2005).

35. Burrell, R.A. *et al.* Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492-496 (2013).

36. Ciriello, G. *et al.* Emerging landscape of oncogenic signatures across human cancers. *Nat Genet* **45**, 1127-33 (2013).

37. Taylor, A.M. *et al.* Genomic and Functional Approaches to Understanding Cancer Aneuploidy. *Cancer Cell* **33**, 676-689 e3 (2018).

38. Zack, T.I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nat Genet* **45**, 1134-40 (2013).

39. Bielski, C.M. *et al.* Genome doubling shapes the evolution and prognosis of advanced cancers. *Nat Genet* **50**, 1189-1195 (2018).

40. Bennett, W.P. *et al.* p53 protein accumulates frequently in early bronchial neoplasia. *Cancer Res* **53**, 4817-22 (1993).

41. Shiao, Y.H., Rugge, M., Correa, P., Lehmann, H.P. & Scheer, W.D. p53 alteration in gastric precancerous lesions. *Am J Pathol* **144**, 511-7 (1994).

42. Brentnall, T.A. *et al.* Mutations in the p53 gene: an early marker of neoplastic progression in ulcerative colitis. *Gastroenterology* **107**, 369-78 (1994).

43. Bronder, D. *et al.* TP53 loss initiates chromosomal instability in fallopian tube epithelial cells. *Dis Model Mech* **14**(2021).

44. Soto, M. *et al.* p53 Prohibits Propagation of Chromosome Segregation Errors that Produce Structural Aneuploidies. *Cell Rep* **19**, 2423-2431 (2017).

45. Storchova, Z. & Kuffer, C. The consequences of tetraploidy and aneuploidy. *J Cell Sci* **121**, 3859-66 (2008).

46. Zack, T.I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nature Genetics* **45**, 1134-1140 (2013).

47. Zbar, B., Brauch, H., Talmadge, C. & Linehan, M. Loss of alleles of loci on the short arm of chromosome 3 in renal cell carcinoma. *Nature* **327**, 721-4 (1987).

48. Foulkes, W.D., Black, D.M., Stamp, G.W., Solomon, E. & Trowsdale, J. Very frequent loss of heterozygosity throughout chromosome 17 in sporadic ovarian carcinoma. *Int J Cancer* **54**, 220-5 (1993).

49. Bigner, S.H., Mark, J. & Bigner, D.D. Cytogenetics of human brain tumors. *Cancer Genet Cytogenet* **47**, 141-54 (1990).

50. Drpic, D. *et al.* Chromosome Segregation Is Biased by Kinetochore Size. *Curr Biol* **28**, 1344-1356 e5 (2018).

51. Dumont, M. *et al.* Human chromosome-specific aneuploidy is influenced by DNA-dependent centromeric features. *EMBO J* **39**, e102924 (2020).

52. Klaasen, S.J. *et al.* Nuclear chromosome locations dictate segregation error frequencies. *Nature* **607**, 604-609 (2022).

53. Bochtler, T. *et al.* Micronucleus formation in human cancer cells is biased by chromosome size. *Genes Chromosomes Cancer* **58**, 392-395 (2019).

54. Davoli, T. *et al.* Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155**, 948-62 (2013).

55. Sack, L.M. *et al.* Profound Tissue Specificity in Proliferation Control Underlies Cancer Drivers and Aneuploidy Patterns. *Cell* **173**, 499-514 e23 (2018).

56. Ryland, G.L. *et al.* Loss of heterozygosity: what is it good for? *BMC Med Genomics* **8**, 45 (2015).

57. Merajver, S.D. *et al.* Germline BRCA1 mutations and loss of the wild-type allele in tumors from families with early onset breast and ovarian cancer. *Clin Cancer Res* **1**, 539-44 (1995).

58. Donehower, L.A. *et al.* Integrated Analysis of TP53 Gene and Pathway Alterations in The Cancer Genome Atlas. *Cell Rep* **28**, 1370-1384 e5 (2019).

59. Jamal-Hanjani, M. *et al.* Tracking the Evolution of Non-Small-Cell Lung Cancer. *N Engl J Med* **376**, 2109-2121 (2017).

60. Hoevenaar, W.H.M. *et al.* Degree and site of chromosomal instability define its oncogenic potential. *Nat Commun* **11**, 1501 (2020).

61. Sheltzer, J.M. *et al.* Single-chromosome Gains Commonly Function as Tumor Suppressors. *Cancer Cell* **31**, 240-255 (2017).

62. Rutledge, S.D. *et al.* Selective advantage of trisomic human cells cultured in non-standard conditions. *Sci Rep* **6**, 22828 (2016).

63. Choma, D., Daures, J.P., Quantin, X. & Pujol, J.L. Aneuploidy and prognosis of non-small-cell lung cancer: a meta-analysis of published data. *Br J Cancer* **85**, 14-22 (2001).

64. Xu, J., Huang, L. & Li, J. DNA aneuploidy and breast cancer: a meta-analysis of 141,163 cases. *Oncotarget* **7**, 60218-60229 (2016).

65. Stopsack, K.H. *et al.* Aneuploidy drives lethal progression in prostate cancer. *Proc Natl Acad Sci U S A* **116**, 11390-11395 (2019).

66. Xu, J. *et al.* Prognostic value of DNA aneuploidy in gastric cancer: a meta-analysis of 3449 cases. *BMC Cancer* **19**, 650 (2019).

67. Silk, A.D. *et al.* Chromosome missegregation rate predicts whether aneuploidy will promote or suppress tumors. *Proc Natl Acad Sci U S A* **110**, E4134-41 (2013).

68. Birkbak, N.J. *et al.* Paradoxical relationship between chromosomal instability and survival outcome in cancer. *Cancer Res* **71**, 3447-52 (2011).

69. van Dijk, E. *et al.* Chromosomal copy number heterogeneity predicts survival rates across cancers. *Nat Commun* **12**, 3188 (2021).

70. Senovilla, L. *et al.* An immunosurveillance mechanism controls cancer cell ploidy. *Science* **337**, 1678-84 (2012).

71. Davoli, T., Uno, H., Wooten, E.C. & Elledge, S.J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**(2017).

72. Hong, C. *et al.* cGAS-STING drives the IL-6-dependent survival of chromosomally instable cancers. *Nature* **607**, 366-373 (2022).

73. Martincorena, I. *et al.* Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880-6 (2015).

74. Martincorena, I. *et al.* Somatic mutant clones colonize the human esophagus with age. *Science* **362**, 911-917 (2018).

75. Wijewardhane, N., Dressler, L. & Ciccarelli, F.D. Normal Somatic Mutations in Cancer Transformation. *Cancer Cell* **39**, 125-129 (2021).

76. Yokoyama, A. *et al.* Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature* **565**, 312-317 (2019).

77. Lawson, A.R.J. *et al.* Extensive heterogeneity in somatic mutation and selection in the human bladder. *Science* **370**, 75-82 (2020).

78. Duncan, A.W. *et al.* Frequent aneuploidy among normal human hepatocytes. *Gastroenterology* **142**, 25-8 (2012).

79. Coorens, T.H.H. *et al.* Inherent mosaicism and extensive mutation of human placentas. *Nature* **592**, 80-85 (2021).

80. Mitchell, T.J. *et al.* Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. *Cell* **173**, 611-623 e17 (2018).

81. Turajlic, S. *et al.* Deterministic Evolutionary Trajectories Influence Primary Tumor Growth: TRACERx Renal. *Cell* **173**, 595-610 e11 (2018).

82. Nguyen, B. *et al.* Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. *Cell* **185**, 563-575 e11 (2022).

83. Bakhoum, S.F. *et al.* Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature* **553**, 467-472 (2018).

84. Priestley, P. *et al.* Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* **575**, 210-216 (2019).

85. Martinez-Jimenez, F. *et al.* Pan-cancer whole-genome comparison of primary and metastatic solid tumours. *Nature* (2023).

86. Cancer Genome Atlas Research, N. *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **45**, 1113-20 (2013).

87. Birkbak, N.J. & McGranahan, N. Cancer Genome Evolutionary Trajectories in Metastasis. *Cancer Cell* **37**, 8-19 (2020).

88. Funnell, T. *et al.* Single-cell genomic variation induced by mutational processes in cancer. *Nature* **612**, 106-115 (2022).

89. Sottoriva, A. *et al.* A Big Bang model of human colorectal tumor growth. *Nat Genet* **47**, 209-16 (2015).

90. Cross, W. *et al.* The evolutionary landscape of colorectal tumorigenesis. *Nat Ecol Evol* **2**, 1661-1672 (2018).

91. Martínez-Jiménez, F. *et al.* Pan-cancer whole genome comparison of primary and metastatic solid tumors. *bioRxiv*, 2022.06.17.496528 (2022).

92. Lens, S.M.A. & Medema, R.H. Cytokinesis defects and cancer. *Nat Rev Cancer* **19**, 32-45 (2019).

93. Ganem, N.J., Storchova, Z. & Pellman, D. Tetraploidy, aneuploidy and cancer. *Curr Opin Genet Dev* **17**, 157-62 (2007).

94. Davoli, T. & de Lange, T. Telomere-driven tetraploidization occurs in human cells undergoing crisis and promotes transformation of mouse cells. *Cancer Cell* **21**, 765-76 (2012).

95. Duelli, D.M., Hearn, S., Myers, M.P. & Lazebnik, Y. A primate virus generates transformed human cells by fusion. *J Cell Biol* **171**, 493-503 (2005).

96. Dittmar, T. & Zanker, K.S. Tissue Regeneration in the Chronically Inflamed Tumor Environment: Implications for Cell Fusion Driven Tumor Progression and Therapy Resistant Tumor Hybrid Cells. *Int J Mol Sci* **16**, 30362-81 (2015).

97. Gemble, S. *et al.* Genetic instability from a single S phase after whole-genome duplication. *Nature* **604**, 146-151 (2022).

98. Prasad, K. *et al.* Whole-Genome Duplication Shapes the Aneuploidy Landscape of Human Cancers. *Cancer Res* **82**, 1736-1752 (2022).

99. Dewhurst, S.M. *et al.* Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Discov* **4**, 175-185 (2014).

100. Fujiwara, T. *et al.* Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437**, 1043-7 (2005).

101. Lambuta, R.A. *et al.* Whole-genome doubling drives oncogenic loss of chromatin segregation. *Nature* (2023).

102. Lopez, S. *et al.* Interplay between whole-genome doubling and the accumulation of deleterious alterations in cancer evolution. *Nat Genet* **52**, 283-293 (2020).

103. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029-1041 e21 (2017).

104. Carter, S.L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol* **30**, 413-21 (2012).

105. Lukow, D.A. *et al.* Chromosomal instability accelerates the evolution of resistance to anti-cancer therapies. *Dev Cell* **56**, 2427-2439 e4 (2021).

106. Coward, J. & Harding, A. Size Does Matter: Why Polyploid Tumor Cells are Critical Drug Targets in the War on Cancer. *Front Oncol* **4**, 123 (2014).

107. Laughney, A.M., Elizalde, S., Genovese, G. & Bakhoum, S.F. Dynamics of Tumor Heterogeneity Derived from Clonal Karyotypic Evolution. *Cell Rep* **12**, 809-20 (2015).

108. Kim, J.E. *et al.* High prevalence of TP53 loss and whole-genome doubling in early-onset colorectal cancer. *Exp Mol Med* **53**, 446-456 (2021).

109. Anderson, N.D. *et al.* Rearrangement bursts generate canonical gene fusions in bone and soft tissue tumors. *Science* **361**, eaam8419 (2018).

110. Stephens, P.J. *et al.* Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. *Cell* **144**, 27-40 (2011).

111. Dahiya, R., Hu, Q. & Ly, P. Mechanistic origins of diverse genome rearrangements in cancer. *Seminars in Cell & Developmental Biology* **123**, 100-109 (2022).

112. Campbell, P.J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109-1113 (2010).

113. Baca, Sylvan C. *et al.* Punctuated Evolution of Prostate Cancer Genomes. *Cell* **153**, 666-677 (2013).

114. Rosswog, C. *et al.* Chromothripsis followed by circular recombination drives oncogene amplification in human cancer. *Nature Genetics* **53**, 1673-1685 (2021).

115. Hadi, K. *et al.* Distinct Classes of Complex Structural Variation Uncovered across Thousands of Cancer Genome Graphs. *Cell* **183**, 197-210.e32 (2020).

116. The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature* **578**, 82-93 (2020).

117. Kloosterman, Wigard P. *et al.* Constitutional Chromothripsis Rearrangements Involve Clustered Double-Stranded DNA Breaks and Nonhomologous Repair Mechanisms. *Cell Reports* **1**, 648-655 (2012).

118. Richardson, C. & Jasin, M. Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature* **405**, 697-700 (2000).

119. Lee, J.J.-K. *et al.* Tracing Oncogene Rearrangements in the Mutational History of Lung Adenocarcinoma. *Cell* **177**, 1842-1857.e21 (2019).

120. Korbel, J.O. & Campbell, P.J. Criteria for inference of chromothripsis in cancer genomes. *Cell* **152**, 1226-36 (2013).

121. Rausch, T. *et al.* Genome Sequencing of Pediatric Medulloblastoma Links Catastrophic DNA Rearrangements with TP53 Mutations. *Cell* **148**, 59-71 (2012).

122. Boboila, C., Alt, F.W. & Schwer, B. Chapter One - Classical and Alternative End-Joining Pathways for Repair of Lymphocyte-Specific and General DNA Double-Strand Breaks. in *Advances in Immunology*, Vol. 116 (ed. Alt, F.W.) 1-49 (Academic Press, 2012).

123. Hastings, P.J., Ira, G. & Lupski, J.R. A Microhomology-Mediated Break-Induced Replication Model for the Origin of Human Copy Number Variation. *PLOS Genetics* **5**, e1000327 (2009).

124. Weckselblatt, B. & Rudd, M.K. Human Structural Variation: Mechanisms of Chromosome Rearrangements. *Trends in Genetics* **31**, 587-599 (2015).

125. Kurzrock, R., Gutterman, J.U. & Talpaz, M. The Molecular Genetics of Philadelphia Chromosome–Positive Leukemias. *New England Journal of Medicine* **319**, 990-998 (1988).

126. Turc-Carel, C. *et al.* Consistent chromosomal translocation in alveolar rhabdomyosarcoma. *Cancer Genetics and Cytogenetics* **19**, 361-362 (1986).

127. Delattre, O. *et al.* Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359**, 162-165 (1992).

128. Robinson, D.R. *et al.* Identification of recurrent NAB2-STAT6 gene fusions in solitary fibrous tumor by integrative sequencing. *Nature Genetics* **45**, 180-185 (2013).

129. Soda, M. *et al.* Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. *Nature* **448**, 561-566 (2007).

130. Stransky, N., Cerami, E., Schalm, S., Kim, J.L. & Lengauer, C. The landscape of kinase fusions in cancer. *Nature Communications* **5**, 4846 (2014).

131. Neuse, C.J. *et al.* Genome instability in multiple myeloma. *Leukemia* **34**, 2887-2897 (2020).

132. Oey, H. *et al.* Whole-genome sequencing of human malignant mesothelioma tumours and cell lines. *Carcinogenesis* **40**, 724-734 (2019).

133. Rustad, E.H. *et al.* Revealing the Impact of Structural Variants in Multiple Myeloma. *Blood Cancer Discovery* **1**, 258-273 (2020).

134. Berger, M.F. *et al.* The genomic complexity of primary human prostate cancer. *Nature* **470**, 214-220 (2011).

135. Pellestor, F., Gaillard, J.B., Schneider, A., Puechberty, J. & Gatinois, V. Chromoanagenesis, the mechanisms of a genomic chaos. *Seminars in Cell & Developmental Biology* **123**, 90-99 (2022).

136. Viswanathan, S.R. *et al.* Structural Alterations Driving Castration-Resistant Prostate Cancer Revealed by Linked-Read Genome Sequencing. *Cell* **174**, 433-447.e19 (2018).

137. Mansfield, A.S. *et al.* Neoantigenic Potential of Complex Chromosomal Rearrangements in Mesothelioma. *Journal of Thoracic Oncology* **14**, 276-287 (2019).

138. Leibowitz, M.L., Zhang, C.-Z. & Pellman, D. Chromothripsis: A New Mechanism for Rapid Karyotype Evolution. *Annual Review of Genetics* **49**, 183-211 (2015).

139. Cortés-Ciriano, I. *et al.* Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nature Genetics* **52**, 331-341 (2020).

140. Umbreit, N.T. *et al.* Mechanisms generating cancer genome complexity from a single cell division error. *Science* **368**, eaba0712 (2020).

141. Bao, L., Zhong, X., Yang, Y. & Yang, L. Starfish infers signatures of complex genomic rearrangements across human cancers. *Nature Cancer* **3**, 1247-1259 (2022).

142. Hatch, Emily M., Fischer, Andrew H., Deerinck, Thomas J. & Hetzer, Martin W. Catastrophic Nuclear Envelope Collapse in Cancer Cell Micronuclei. *Cell* **154**, 47-60 (2013).

143. Liu, S. *et al.* Nuclear envelope assembly defects link mitotic errors to chromothripsis. *Nature* **561**, 551-555 (2018).

144. Tang, S., Stokasimov, E., Cui, Y. & Pellman, D. Breakage of cytoplasmic chromosomes by pathological DNA base excision repair. *Nature* **606**, 930-936 (2022).

145. Zhang, C.-Z. *et al.* Chromothripsis from DNA damage in micronuclei. *Nature* **522**, 179-184 (2015).

146. Hirsch, D. *et al.* Chromothripsis and Focal Copy Number Alterations Determine Poor Outcome in Malignant Melanoma. *Cancer Research* **73**, 1454-1460 (2013).

147. Kloosterman, W.P. *et al.* Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer. *Genome Biology* **12**, R103 (2011).

148. Magrangeas, F., Avet-Loiseau, H., Munshi, N.C. & Minvielle, S. Chromothripsis identifies a rare and aggressive entity among newly diagnosed multiple myeloma patients. *Blood* **118**, 675-678 (2011).

149. Molenaar, J.J. *et al.* Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature* **483**, 589-593 (2012).

150. Ly, P. *et al.* Chromosome segregation errors generate a diverse spectrum of simple and complex genomic rearrangements. *Nature Genetics* **51**, 705-715 (2019).

151. Parker, M. *et al.* C11orf95-RELA fusions drive oncogenic NF-kappaB signalling in ependymoma. *Nature* **506**, 451-5 (2014).

152. McClintock, B. THE STABILITY OF BROKEN ENDS OF CHROMOSOMES IN ZEA MAYS. *Genetics* **26**, 234-282 (1941).

153. Bhargava, R., Fischer, M. & O’Sullivan, R.J. Genome rearrangements associated with aberrant telomere maintenance. *Current Opinion in Genetics & Development* **60**, 31-40 (2020).

154. Tanaka, H. & Watanabe, T. Mechanisms Underlying Recurrent Genomic Amplification in Human Cancers. *Trends in Cancer* **6**, 462-477 (2020).

155. Coquelle, A., Pipiras, E., Toledo, F., Buttin, G. & Debatisse, M. Expression of Fragile Sites Triggers Intrachromosomal Mammalian Gene Amplification and Sets Boundaries to Early Amplicons. *Cell* **89**, 215-225 (1997).

156. Marotta, M. *et al.* A common copy-number breakpoint of ERBB2 amplification in breast cancer colocalizes with a complex block of segmental duplications. in *Breast cancer research : BCR* Vol. 14 R150 (2012).

157. Thomas, R., Marks, D.H., Chin, Y. & Benezra, R. Whole chromosome loss and associated breakage–fusion–bridge cycles transform mouse tetraploid cells. *The EMBO Journal* **37**, 201-218 (2018).

158. Ferrari, A. *et al.* A whole-genome sequence and transcriptome perspective on HER2-positive breast cancers. *Nature Communications* **7**, 12222 (2016).

159. Marotta, M. *et al.* Palindromic amplification of the ERBB2 oncogene in primary HER2-positive breast tumors. *Scientific Reports* **7**(2017).

160. Bignell, G.R. *et al.* Architectures of somatic genomic rearrangement in human cancer amplicons at sequence-level resolution. *Genome Research* **17**, 1296-1303 (2007).

161. Reshmi, S.C. *et al.* Inverted duplication pattern in anaphase bridges confirms the breakage-fusion-bridge (BFB) cycle model for 11q13 amplification. *Cytogenetic and Genome Research* **116**, 46-52 (2007).

162. Hellman, A. *et al.* A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell* **1**, 89-97 (2002).

163. Wang, Y.K. *et al.* Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nature Genetics* **49**, 856-865 (2017).

164. Kim, H. *et al.* Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. *Nature Genetics* **52**, 891-897 (2020).

165. Garsed, Dale W. *et al.* The Architecture and Evolution of Cancer Neochromosomes. *Cancer Cell* **26**, 653-667 (2014).

166. Shoshani, O. *et al.* Chromothripsis drives the evolution of gene amplification in cancer. *Nature* **591**, 137-141 (2021).

167. Wu, S. *et al.* Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* **575**, 699-703 (2019).

168. Wu, S., Bafna, V., Chang, H.Y. & Mischel, P.S. Extrachromosomal DNA: An Emerging Hallmark in Human Cancer. *Annual Review of Pathology: Mechanisms of Disease* **17**, 367-386 (2022).

169. Hung, K.L. *et al.* ecDNA hubs drive cooperative intermolecular oncogene expression. *Nature* **600**, 731-736 (2021).

170. Verhaak, R.G.W., Bafna, V. & Mischel, P.S. Extrachromosomal oncogene amplification in tumour pathogenesis and evolution. *Nature Reviews Cancer* **19**, 283-288 (2019).

171. Davicioni, E. *et al.* Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas. *Cancer Res* **66**, 6936-46 (2006).

172. Chapman, O.S. *et al.* The landscape of extrachromosomal circular DNA in medulloblastoma. *bioRxiv*, 2021.10.18.464907 (2021).

173. Yi, E., Chamorro Gonzalez, R., Henssen, A.G. & Verhaak, R.G.W. Extrachromosomal DNA amplifications in cancer. *Nat Rev Genet* **23**, 760-771 (2022).

174. Turner, K.M. *et al.* Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**, 122-125 (2017).

175. Nathanson, D.A. *et al.* Targeted Therapy Resistance Mediated by Dynamic Regulation of Extrachromosomal Mutant EGFR DNA. *Science* **343**, 72-76 (2014).

176. deCarvalho, A.C. *et al.* Discordant inheritance of chromosomal and extrachromosomal DNA elements contributes to dynamic disease evolution in glioblastoma. *Nature Genetics* **50**, 708-717 (2018).

177. McGranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* **168**, 613-628 (2017).

178. Markowetz, F. A saltationist theory of cancer evolution. *Nat Genet* **48**, 1102-3 (2016).

179. Sen, O., Harrison, J.U., Burroughs, N.J. & McAinsh, A.D. Kinetochore life histories reveal an Aurora-B-dependent error correction mechanism in anaphase. *Dev Cell* **56**, 3082-3099 e5 (2021).

180. Mackenzie, K.J. *et al.* cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature* **548**, 461-465 (2017).

181. Martens, U.M. *et al.* Short telomeres on human chromosome 17p. *Nat Genet* **18**, 76-80 (1998).

182. Britt-Compton, B. *et al.* Structural stability and chromosome-specific telomere length is governed by cis-acting determinants in humans. *Hum Mol Genet* **15**, 725-33 (2006).

183. Graakjaer, J. *et al.* The pattern of chromosome-specific variations in telomere length in humans is determined by inherited, telomere-near factors and is maintained throughout life. *Mech Ageing Dev* **124**, 629-40 (2003).

184. Hong, M.K. *et al.* Tracking the origins and drivers of subclonal metastatic expansion in prostate cancer. *Nat Commun* **6**, 6605 (2015).

185. Choi, J. *et al.* Integrated mutational landscape analysis of uterine leiomyosarcomas. *Proceedings of the National Academy of Sciences* **118**, e2025182118 (2021).

186. Wang, Z. *et al.* RNA sequencing of esophageal adenocarcinomas identifies novel fusion transcripts, including NPC1-MELK, arising from a complex chromosomal rearrangement. *Cancer* **123**, 3916-3924 (2017).

187. Notta, F. *et al.* A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns. *Nature* **538**, 378-382 (2016).

188. Singer, M.J., Mesner, L.D., Friedman, C.L., Trask, B.J. & Hamlin, J.L. Amplification of the human dihydrofolate reductase gene via double minutes is initiated by chromosome breaks. *Proceedings of the National Academy of Sciences* **97**, 7921-7926 (2000).

189. Song, K. *et al.* Plasticity of Extrachromosomal and Intrachromosomal BRAF Amplifications in Overcoming Targeted Therapy Dosage Challenges. *Cancer Discovery* **12**, 1046-1069 (2022).