nature genetics

Review article

https://doi.org/10.1038/s41588-024-01916-2

Aneuploidy as a driver of human cancer

Received: 23 April 2024

Accepted: 20 August 2024

Published online: 02 October 2024

Check for updates

Eran Sdeor[1], Hajime Okada[1], Ron Saad[1,2], Tal Ben-Yishay[1,2] & Uri Ben-David[1]

Aneuploidy, an abnormal chromosome composition, is a major contributor to cancer development and progression and an important determinant of cancer therapeutic responses and clinical outcomes. Despite being recognized as a hallmark of human cancer, the exact role of aneuploidy as a 'driver' of cancer is still largely unknown. Identifying the specific genetic elements that underlie the recurrence of common aneuploidies remains a major challenge of cancer genetics. In this Review, we discuss recurrent aneuploidies and their function as drivers of tumor development. We then delve into the context-dependent identification and functional characterization of the driver genes underlying driver aneuploidies and examine emerging strategies to uncover these driver genes using cancer genomics data and cancer models. Lastly, we explore opportunities for targeting driver aneuploidies in cancer by leveraging the functional consequences of these common genetic alterations.

Aneuploidy, a genomic state characterized by an abnormal number of chromosomes or chromosome arms, is widely recognized as a hallmark of cancer[1]. Such large copy number alterations (CNAs) affect hundreds or thousands of genetic elements at once, leading to changes in gene dosage and altered gene expression profiles. These alterations thus have profound effects on cell behavior and function[2], which may either suppress or promote tumorigenesis, depending on the context[1,3]. Approximately 90% of solid tumors are aneuploid[4], and aneuploidy is strongly correlated with a poor prognosis, underscoring the urgency of assessing and integrating it in clinical care decisions[1]. Therefore, understanding the driving role of aneuploidy in tumor formation and progression is essential for developing therapeutic approaches to fight this hallmark trait of cancer.

Many factors shape the aneuploidy landscape, including tumor stage, cell of origin, cellular microenvironment and genomic context (that is, co-occurring genetic alterations). Consequently, many of the cellular consequences of aneuploidy depend on the specific altered chromosome and the specific cellular context. These factors, among others, collectively contribute to shaping the eventual aneuploidy landscapes of human tumors, as reviewed before[1,2].

Cells may missegregate their chromosomes during cell division, leading to various possible outcomes, including the generation of aneuploid daughter cells[5,6]. The selection for and against specific karyotypes under specific cellular circumstances is believed to be driven by the combined effect of the oncogenes and the tumor-suppressor genes (TSGs) residing within each chromosomal fragment. It is thought that the overall consequence of the dysregulation of genes, whether promoting or suppressing cell survival and proliferation, dictates the net adaptive values of the cellular karyotypes[7]. A comprehensive pan-cancer analysis of aneuploidy across >10,000 tumors has revealed that each cancer type exhibits a unique pattern of aneuploidy. Tumors of similar tissue origins tend to share similar karyotypes[8]. For example, gastrointestinal tumors tend to acquire gains of chromosome arms 8q and 13q and chromosome 20, while gynecological tumors are all characterized by the gain of chromosome arm 1q[8,9]. Another large-scale analysis suggested that tissue-specific gene expression of oncogenic proliferation signals determines which copy number changes and aneuploidies would be selected for during tumorigenesis, thus explaining the characteristic tissue-specificity patterns of aneuploidy[9].

Moreover, it has been suggested that tissue-specific proliferation-promoting genes tend to be enriched in chromosome arms that are often aneuploid in tumors originating from the respective tissue. For example, proliferation-promoting genes in human mammary epithelial cells were found to be enriched in chromosomal gains that are frequently acquired in breast cancer[9,10]. Another study proposed that specific aneuploidies amplify the chromosome arm-wide gene

[1]Department of Human Molecular Genetics and Biochemistry, Faculty of Medical and Health Sciences, Tel Aviv University, Tel Aviv, Israel. [2]The Blavatnik School of Computer Science, Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. e-mail: ubendavid@tauex.tac.ac.il

https://doi.org/10.1038/s41588-024-01916-2

BOX 1

Aneuploidy detection methods and recurrence analyses

Methods for aneuploidy detection

Beginning with SNP arrays of cell populations[101], we have now advanced to the point where we can perform whole-exome sequencing and whole-genome sequencing (WGS) of single cells. This not only enables us to study the dominant clones within a cancer cell population but also allows for exploring karyotypic heterogeneity and clonal dynamics within the population[102,103]. Genomic sequencing has enabled us to infer CNAs and characterize the aneuploidy profiles of cancer cells[8,104]. Furthermore, in addition to genomic methods, transcriptomics has become increasingly popular for inferring both focal CNAs and aneuploidies[105–107]. Today, using techniques such as single-cell RNA sequencing, CNAs in general, and aneuploidies in particular, can be identified by comparing chromosome-wide gene expression imbalances to diploid chromosome-wide gene expression patterns or by identifying alterations in allele ratios using haplotype-aware analyses[108–115]. After sequencing or SNP array genotyping, various tools can be employed to generate segmented copy number calls, depending on the sequencing method used (Fig. 1a). Among them, Rawcopy is suitable for SNP arrays[116], CNVkit is applicable for WGS, whole-exome sequencing or targeted sequencing[117], AneuFinder is suitable for single-cell WGS[118] and both InferCNV and Numbat are suitable for single-cell RNA sequencing[114,119]. Other powerful tools are algorithms, such as ABSOLUTE[120] and PureCN[121], which estimate tumor purity and ploidy from DNA sequencing data and allow the identification of focal CNAs and aneuploidies that are specific to the cancer component of the tumors. These tools also consider subclonal CNAs, providing insights into the contribution of aneuploidy to clonal dynamics. Subsequently, the segmented copy number data can be converted into arm-level copy number data using tools such as ASCETS (arm-level somatic copy number events in targeted sequencing)[122]. Arm-level copy number data can then be visualized and clustered, and statistical analyses can assist in the identification of recurrent focal CNAs and aneuploidies (Fig. 1a). Overall, the technological advances allow us to easily infer aneuploidy, based on DNA and RNA sequencing platforms.

Aneuploidy recurrence analyses

To identify potential driver aneuploidies, recurrence analysis commonly involves examining samples from large cohorts of patients with cancer, with the assumption that driver aneuploidies are those that are detected more frequently than what would be expected based on random or uniform distributions. Consequently, several tools have been developed to identify recurrent aneuploidy patterns in the cancer genome. One widely used tool is GISTIC2.0 (ref. 61), a statistical algorithm that detects recurrent CNAs in cancer genomes by evaluating the frequency and the amplitude of the observed events. GISTIC2.0 is designed to pinpoint genomic regions that are frequently gained or lost across a set of cancer samples, providing insights into potential driver aneuploidies in tumorigenesis (Fig. 1a). This tool, among others, has considerably contributed to aneuploidy profiling across cancer genomes, aiding in the exploration of tumorigenesis complexities. It is important to note, however, that recurrence analyses are inevitably confounded by the granularity of the genomic analyses, that is, by which subset of tumors is being considered as the background for the analysis. In other words, the identification of recurrent aneuploidies largely depends on the denominator. For example, a recurrent aneuploidy can be associated with a particular tissue type, as exemplified by squamous cancers that are characterized by chromosome arm 3p loss and chromosome arm 3q gain, regardless of the specific organ in which they arose[8]. Therefore, while the conventional approach seeks to identify recurrent aneuploidies specific to a particular cancer type, other tumor-associated features, such as molecular subtypes and even specific genetic alterations, may greatly affect the prevalence of aneuploidy. For example, when considering breast cancer as a whole, we would miss recurrent aneuploidies that are specific to only one of the breast cancer molecular subtypes[21,123–126] (Fig. 1b) or those that are associated with a specific altered oncogenic pathway[64,127]. Additionally, tumor stage could be associated with aneuploidy recurrence. Indeed, some tumor-specific aneuploidies, such as chromosome arm 20q gain in GBM, tend to arise later in tumorigenesis than others, such that their identification as 'recurrent' would depend on the tumor stage of the studied patient cohort[11,20]. Ultimately, similar to all genomic analyses of human cancer, it is imperative to consider the appropriate level of granularity when profiling aneuploidy patterns in an attempt to identify 'recurrent' aneuploidies. As more and more genomic data from patients with cancer become available, this will ensure a comprehensive understanding of the intricate complexities within cancer genomes. This in turn could resolve potential ambiguities that might arise from a more generalized perspective that does not consider the remarkable context sensitivity of aneuploidy.

expression patterns of the normal tissue of tumor origin[11]. This observation implies that normal tissue cells and tumors derived from the same tissue share similar transcriptional programs, many of which are not shared with other tissues[9]. This suggests that specific aneuploidies serve as a genetic means to 'fix' such transcriptional programs. We recently found that tissue-specific and chromosome-specific features can largely explain the aneuploidy landscapes observed across human cancers[12]. These findings collectively suggest that recurrence patterns of aneuploidy highly depend on the tissue of origin, which must be considered when exploring the driving role of a recurrent aneuploidy or the genes driving such recurrence.

In this Review, we first define driver aneuploidies and explore their emergence and the approaches used for their exposure. We then discuss the driver genes underlying these aneuploidies, describing the complexity of their identification and the strategies used for this purpose, ranging from genomic analyses to the latest research models in the field of aneuploidy. Finally, we briefly discuss strategies to target both driver and nondriver genes to selectively kill aneuploid cancer cells.

Driver aneuploidies

Defining driver aneuploidies

Owing to its complexity and the existence of different definitions, it is important to establish a clear definition of aneuploidy before identifying driver aneuploidies. For both conceptual and practical reasons, we adhere to our previously proposed definition[1], in which aneuploidy is defined as CNAs that encompass whole chromosomes or entire chromosome arms (excluding the short arms of acrocentric chromosomes). We note that large CNAs that encompass as many (or more) genes as small chromosome arms can also be included in this definition[1].

The high recurrence of a given aneuploidy in a specific tumor type, in and of itself, does not necessarily justify classifying it as a driver

aneuploidy. Determining whether an observed aneuploidy qualifies as a driver should be considered in a specific context. Here, 'driver aneuploidies' would actively contribute to the initiation, progression or maintenance of cancer and be under positive selection due to a phenotypic advantage (such as increased survival, proliferation or migration) that they confer on the cancer cells under the relevant conditions. For example, driver aneuploidies may promote rapid tumor growth[13], therapy resistance[14,15] or the ability to adapt to changing microenvironments[1,16–18]. By contrast, 'passenger aneuploidies' do not actively contribute to tumorigenesis and instead often result from the genomic instability of cancer cells; they are tolerated as they do not impose a substantial fitness cost, allowing them to evade negative selection[1,12,19].

Identifying driver aneuploidies: recurrence analyses and beyond

Over the years, a major goal of cancer research has been to understand the structure of the cancer genome. In addition to more traditional cytogenetic approaches, numerous advanced genomic analysis methods have been recently developed to decipher the genetic content of cancer cells. These advancements have led to innovative methods for aneuploidy detection and for recurrence analyses (Box 1). The systematic mapping of aneuploidy landscapes can help nominate potential driver aneuploidies (Fig. 1a). Considering the cellular context, for example, tissue type, tumor stage or molecular subtype (Fig. 1b), is crucial in such analyses, as it greatly impacts the identification of recurrent aneuploidies.

As mentioned above, recurrence alone proves insufficient in the pursuit for driver aneuploidies and it is crucial to go beyond merely identifying recurrent alterations. The distinction between driver aneuploidies and passenger aneuploidies is essential, and statistical association does not prove causation. Recurrent aneuploidies that emerge and expand early on during tumorigenesis[20] as well as those that already pre-exist in rare cells in the normal tissues of patients[21] make for stronger driver candidates. However, to determine whether an aneuploidy actively contributes to tumorigenesis, one must assess its impact on gene expression. The differentially expressed genes can then be subjected to gene set enrichment analysis (Fig. 1c), as recurrent aneuploidies are expected to be associated with altered signaling pathways[1] and consequently affect cellular behavior (Fig. 1d).

The same principles that apply to the distinction between recurrent and driver mutations can provide further insights. For example, even when encountering a frequently recurring mutation, understanding whether it is indeed a driver mutation requires consideration of factors, such as gene size, genomic location and nucleotide composition, among others[22]. Furthermore, it is essential to determine whether this mutation consistently co-occurs or is mutually exclusive with other mutations, as this can teach us about its potential role in tumorigenesis[8,23,24].

Overall, the comprehensive evaluation of potential driver aneuploidies demands a multifaceted approach that considers potential genomic confounders (contextual associations) and includes the assessment of their functional consequences.

Shaping aneuploidy through chromosome missegregation and selection

Evidence for chromosome-specific rates of chromosome missegregation. Several studies suggest that chromosome missegregation is not entirely random (Fig. 2). One study reported that chromosome segregation fidelity is related to centromeric cohesion forces, as chromosomes 1 and 2 were found to be particularly prone to a weakening of centromeric cohesion, resulting in an increased rate of their missegregation[25]. In addition, for most human chromosomes, chromosome segregation favors centromeres with a high abundance of centromeric DNA sequences and centromere proteins[26]. The larger the chromosome and the smaller its centromere, the more prone it is to missegregation[26]. Furthermore, the nuclear location of chromosomes during interphase has been suggested to determine their rate of missegregation, such that peripheral chromosomes are more likely to missegregate[27]. This study indicates that the location of a chromosome relative to the spindle poles largely determines the likelihood for its missegregation[27]. Together, these studies suggest that both intrinsic features of the chromosome and the spatial positioning of the chromosome within the nucleus play a role in chromosome missegregation, contributing to the biased emergence of specific aneuploidies in specific contexts. In other words, not all chromosomes are equally prone to chromosome missegregation.

Evidence for positive and negative selection. Upon its emergence, aneuploidy has a substantial impact on numerous genes and, consequently, global gene expression[1,3,28]. These changes often lead to imbalanced protein production, resulting in proteotoxic stress that disrupts cellular function[29,30], thereby posing a substantial challenge for cellular and organismal survival during development[1,31]. Induction of aneuploidy is known to result in several types of cellular stresses (mitotic, replicative, proteotoxic and metabolic)[29]. Nonetheless, aneuploidy is well tolerated in human cancer, suggesting that cancer cells must overcome these fitness costs. A recurrent driver aneuploidy is thus expected to confer a fitness advantage, offsetting the detrimental consequences of widespread gene dysregulation.

A study of mouse embryonic fibroblasts showed that gaining a single chromosome did not result in cellular transformation and actually led to reduced proliferation of the cells compared to their diploid counterparts[32]. However, throughout extensive passaging, the cells acquired additional aneuploidies that eventually facilitated an increase in proliferation. This suggests that, although a single trisomy reduces the proliferation rate, karyotypic evolution eventually selects for additional aneuploidies that rescue the cell proliferation deficiency[32]. Importantly, even aneuploidy of a single chromosome or chromosome arm could function as a tumor promoter under different circumstances, for example, in human embryonic stem cells[33] or in metastasis[34,35]. We recently induced chromosome missegregation in human retinal pigment epithelial (RPE1) cells and followed their karyotypic evolution[36]. Although the initial aneuploidies following chromosome missegregation were consistent with previously described chromosome-specific missegregation rates[27], the eventual karyotypes of the clones differed substantially. Similarly, karyotypically heterogeneous aneuploid cells were generated in renal proximal tubular epithelial cells and human mammary epithelial cells[10], and, although the initial pool of aneuploid cells represented all chromosomes, after a few days in culture, only specific aneuploidies were selected for, which were similar to common aneuploidies in renal and breast cancers, respectively. This suggests that, at least in this context, nonrandom chromosome missegregation does not determine the eventual clonal karyotypes[36]. Moreover, the characteristic aneuploidy patterns across human tumors do not follow reported prevalences of chromosome missegregation either[8]. These findings collectively indicate that, while aneuploidy has a detrimental effect on cells, specific aneuploidies (which we define as driver aneuploidies) may confer a fitness advantage under particular conditions. This underscores the critical role of selection in shaping the aneuploidy landscape of human cancers (Fig. 2).

Two contrasting yet intertwined forces determine the aneuploidy landscape: positive selection, which favors aneuploidies that acquire traits promoting their survival or proliferation, and negative selection, which acts to eliminate aneuploid cells unable to adapt to their genomic alteration. To date, most studies have focused on important roles for positive selection of aneuploidies[7,9,11,37]. For example, chromosome arms with an enriched density of oncogenes and TSGs appear to coincide with regions that are susceptible to gain or loss, respectively, an observation that could be attributed to the cumulative impact of

https://doi.org/10.1038/s41588-024-01916-2

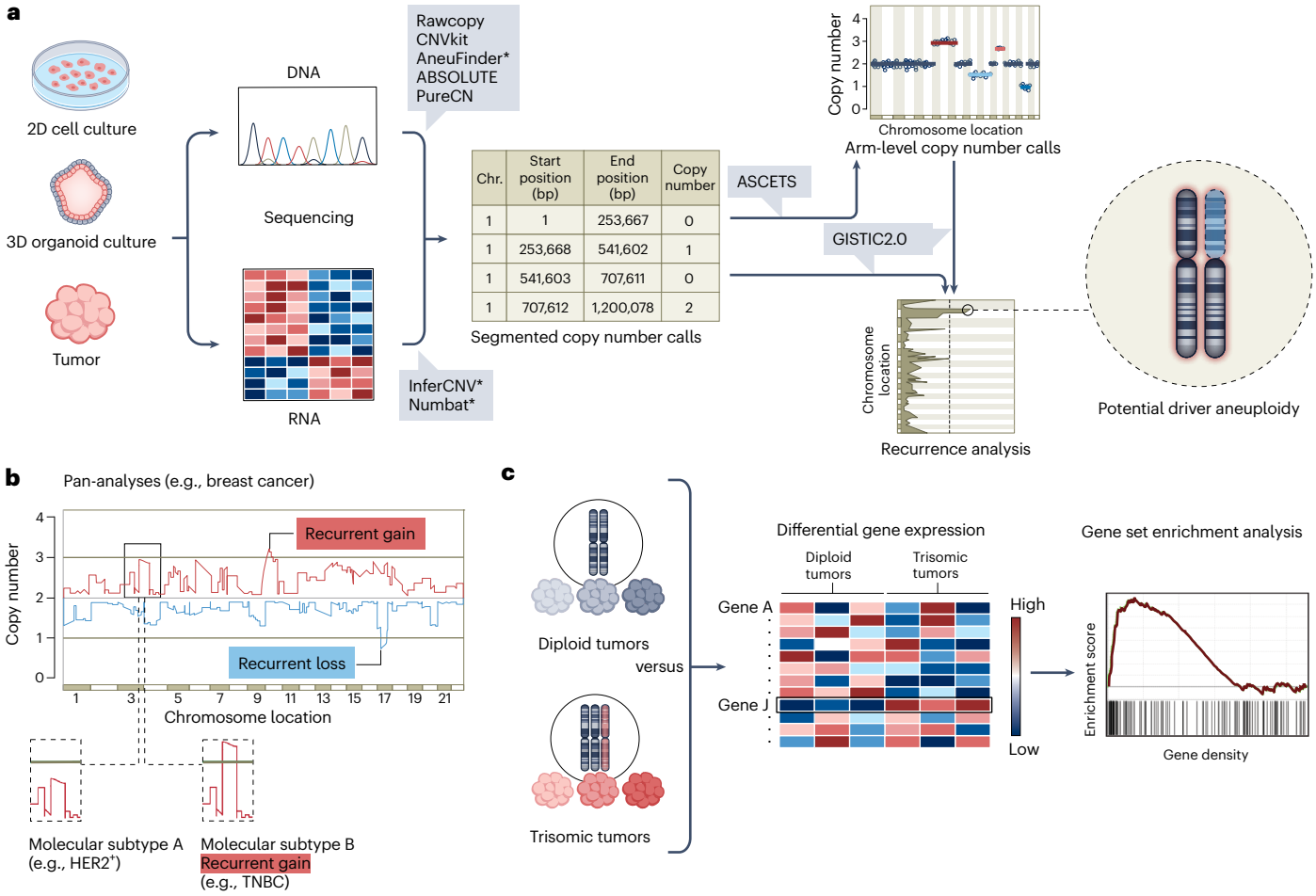

Fig. 1 | Strategies to identify driver aneuploidies. a, A driver aneuploidy is expected to be a recurrent event in the genomic landscape of cancer cells. Strategies to identify recurrent aneuploidy patterns involve using various computational tools to convert sequencing or array genotyping data into copy number information (Box 1). The copy number data are segmented based on the positions of the CNAs on the chromosome (Chr.). These data can then be visualized as arm-level copy number changes and subjected to recurrence analyses. 2D, two dimensional; 3D, three dimensional; asterisks represent algorithms for copy number inference from single-cell DNA or RNA sequencing. **b**, Many factors shape the aneuploidy landscapes of human tumors and should be taken into consideration when attempting to identify recurrent aneuploidies. For example, the recurrent aneuploidy pattern of a specific cancer type may differ based on its molecular subtype: a given aneuploidy may not reach a recurrence threshold in breast cancer as a whole, as it is recurrent in one molecular subtype but not in another (shown in the insets). TNBC, triple-negative breast cancer. **c**, Aneuploidy recurrence per se is insufficient to designate an aneuploidy as a driver event of tumorigenesis. For aneuploidy to be a true driver, it should affect gene expression, which could be evaluated by gene set enrichment analysis. **d**, Consequently, driver aneuploidies are expected to promote cellular tumorigenic phenotypes, such as increased proliferation, increased cell migration or increased cell survival following drug exposure, whereas passenger aneuploidies do not.

modifying such genes[7]. Moreover, proliferation-promoting genes and proliferation-arresting genes vary among tissues, making them tissue type specific, thus contributing to the recurrence of tissue-specific driver aneuploidies[9]. Furthermore, a recent study suggested that somatic copy number amplifications, including aneuploidies, are positively selected in cancer evolution to buffer against the subsequent gene inactivation through loss-of-function point mutations[37].

Importantly, however, recent research has highlighted the importance of negative selection in shaping the aneuploidy landscape, which has been underappreciated until now[12,19]. A novel method called BISCUT (breakpoint identification of significant cancer undiscovered targets) identifies loci in tumor genomes that are more prone to breaking, thus leading to the forming of aneuploidies[19]. This approach found more evidence for loci under negative selection than for those under positive selection. Furthermore, by applying a machine learning approach, we recently revealed that the impact of TSG density in predicting chromosome gains is larger than that of oncogene density and vice versa for oncogene density and predicting chromosome losses, highlighting the importance of negative selection in shaping the aneuploidy landscape of human cancers[12].

https://doi.org/10.1038/s41588-024-01916-2

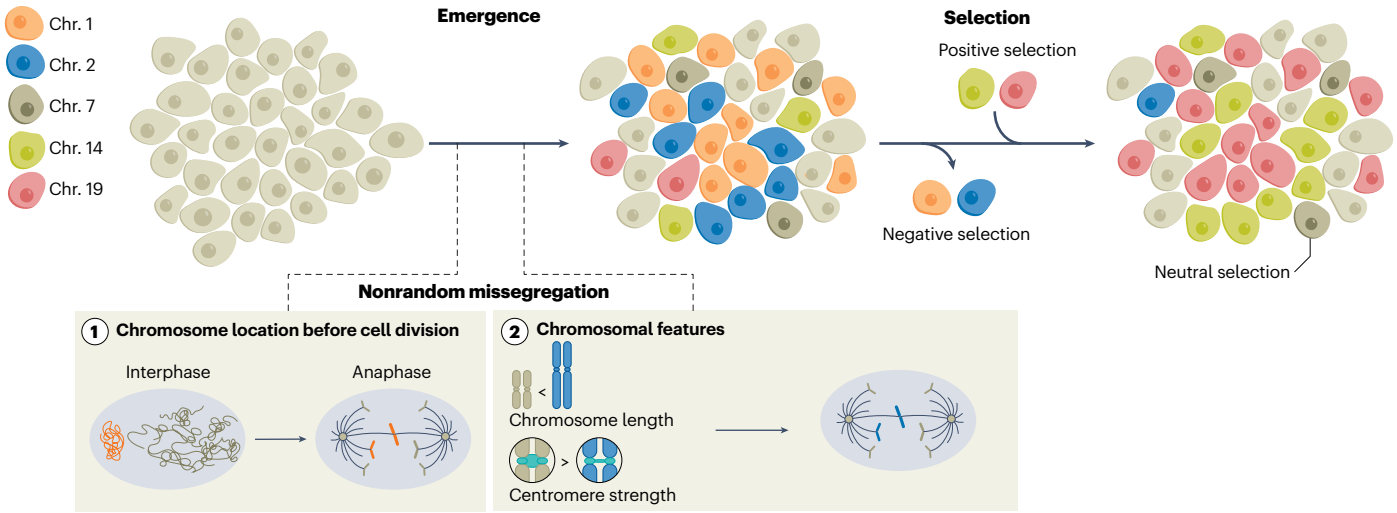

Fig. 2 | Emergence and selection of driver aneuploidies. Chromosome missegregation has been proposed as a nonrandom event influenced by (1) the spatial organization within the nucleus before cell division and (2) chromosome features such as centromeric properties and chromosome length. Peripheral chromosomes during interphase are more prone to missegregate, as exemplified by chromosome 1 (orange). Additionally, longer chromosomes with weaker centromeres (shown in turquoise) are also reported to missegregate more frequently, as exemplified by chromosome 2 (blue). The detrimental effects of aneuploidy on cellular function are profound, and cells need to overcome these challenges to ensure their survival. Upon emergence, cells undergo karyotypic evolution driven by selective pressures, both positive and negative, through which cells select for driver aneuploidies that confer advantages, exemplified here by chromosomes 14 and 19, which promote survival and offset the fitness costs of emerging aneuploidies. Consequently, an adapted karyotype that is distinct from the initial emerging karyotype is selected for.

Together, these studies suggest that, unlike the evolution of cancer point mutations, which are thought to be shaped mostly by positive selection[38,39], the evolution of cancer aneuploidy is shaped by both positive and negative selection. Positive selection promotes the dysregulation of key genes that contribute to tumor progression, whereas negative selection acts as a safeguard against the strong detrimental effects of aneuploidy.

Driver genes underlying driver aneuploidies

Defining driver genes

In the context of aneuploidy, 'driver genes' are specific genes residing within the driver aneuploidies and underlying their recurrence. These genes play a pivotal role in initiating, progressing or maintaining recurrent aneuploidies. Dysregulation of the expression of such genes is essential for the development or the persistence of specific recurrent aneuploidies within cancer cells. Alterations in these genes are positively selected for during cancer evolution due to their ability to confer a selective advantage, promoting cell survival, proliferation or adaptation to the tumor microenvironment under specific conditions. As in the case of driver aneuploidies, the concept of driver genes is context dependent. In theory, different driver genes may even contribute to the same aneuploidy at various stages of tumorigenesis, under different environmental conditions or during different drug treatments, although empirical evidence for this is currently lacking.

By definition, most of the genes residing within a driver aneuploidy are passenger genes and are not the primary cause of that specific aneuploidy. These genes either retain normal expression despite the alteration of their copy number, owing to compensation mechanisms, or change their expression without detrimental or beneficial consequences for the cancer cells[30,40–46]. Here, we define two categories of driver genes: a 'strong' driver gene, which individually provides a substantial fitness advantage to the cells and is therefore a major contributor to a driver aneuploidy, and a 'weak' driver gene, which provides only a minor fitness advantage to the cells, such that multiple such genes must genetically interact to collectively provide a substantial fitness advantage.

Driver genes are not limited to protein-coding genes, as genes that regulate gene expression, such as microRNAs (miRs) or long noncoding RNAs (lncRNAs), may also drive cancer[47–53]. For example, miR-3676 downregulates the T cell leukemia–lymphoma 1 (*TCL1*) oncogene[54]. This miR resides on chromosome arm 17p and is co-deleted with *TP53* in B cell chronic lymphocytic leukemia, resulting in the overexpression of *TCL1*, which contributes to tumor progression[54]. Another example is the lncRNA PVT1, which resides on chromosome arm 8q and is cogained with the *MYC* oncogene in almost all *MYC*-amplified cancers[55]. Gaining PVT1 is required for high MYC protein levels in 8q-gained human cancer cells, and removing PVT1 from the *MYC*-driven HCT116 colon cancer cells reduced its tumorigenic potency[55]. These findings suggest that noncoding genes can also function as driver genes, demonstrating the need to expand the search for aneuploidy drivers beyond the protein-coding space.

One or few strong drivers versus the cumulative effect of many weak drivers

There are several strong driver genes known to be major contributors to, if not the sole drivers of, recurrent aneuploidies. One such strong driver is *MYC*, which drives chromosome arm 8q gain[56]. In mice, *Myc* resides on chromosome 15 and drives its gain in T cell lymphoma[17]. The driving role of this gene was convincingly demonstrated by genetically engineering *MYC* expression from murine chromosome 6, which led to the common appearance of chromosome 6 gain in T cell lymphoma-bearing mice[17]. These results established *MYC* as a bona fide strong driver of recurrent trisomy 15 in mice. *TP53* is another strong driver gene that is considered to underlie the recurrence of chromosome arm 17p loss in human cancers[57], as is *PTEN*, which is associated with chromosome arm 10q loss[58].

However, evidence shows that, even in these cases in which strong driver genes could be identified, they are usually not the sole drivers, and aneuploidy recurrence is usually driven by more than a single gene (Fig. 3a). Evidence for this comes from genomic analyses of strong driver genes that need to be biallelically inactivated in cancer. For example, loss of chromosome arm 17p is a common way to inactivate one *TP53* allele, both in hematopoietic cancers[57] and in breast cancer[35]. However, this loss is also common without evidence for biallelic inactivation of *TP53* (refs. 35,57), suggesting that there might be other

https://doi.org/10.1038/s41588-024-01916-2

genes involved in driving this aneuploidy. Similarly, *PTEN* is biallelically inactivated in 40% of patients with glioblastoma (GBM), whereas chromosome arm 10q is lost in about 80% of patients[19].

Stronger evidence for the involvement of multiple genes in driving recurrent aneuploidies is coming from functional studies[59]. A study of *Trp53* in mice demonstrated that a heterozygous deletion of the region syntenic to human 17p13.1, combined with a deletion of the other *Trp53* allele, had a more severe effect on the development of hematopoietic cancers than the effect of *Trp53* homozygous deletion alone[57]. This finding suggests that additional genes residing on chromosome arm 17p may function as TSGs and contribute to tumor development in this context[57]. Another recent study suggested that *MDM4*, a strong p53 inhibitor, drives chromosome arm 1q gain in some cancer types[18]. However, overexpression of *MDM4* did not fully recapitulate the effect of this gain, and upregulation of the anti-apoptotic gene *BCL9* was shown to also contribute to it[18]. This suggests that, alongside *MDM4*, *BCL9* is likely a co-driver of this gain[18]. Similarly, a study in mice demonstrated that *Rad21* contributes to the prevalence of trisomy 15 in T cell lymphomas, suggesting that *Myc* is not the only driver gene underlying this trisomy[17]. Consistently, in human Ewing sarcoma, the common gain of chromosome arm 8q involves more than just an additional copy of *MYC*[60], as an extra copy of *RAD21*, which also resides on this chromosome arm, alleviates oncogene-induced replication stress in this disease[60].

For many recurrent aneuploidies, strong driver genes have not been identified, suggesting that their prevalence could be driven by the combined effect of multiple weak driver genes. In line with this notion, the characteristic aneuploidy patterns observed in human cancers were suggested to be shaped by the genomic distribution and the potency of TSGs, oncogenes and essential genes[7,9]. For a gene to contribute to aneuploidy, its copy number-driven dosage imbalance should affect its gene expression levels, consequently impacting the protein level (in the case of coding genes). Therefore, the genes driving somatic CNA and aneuploidy have been proposed to be haploinsufficient and triplosensitive genes that reside on the lost or gained chromosomes or chromosome arms, respectively[7]. Furthermore, recurrently gained chromosome arms were shown to be enriched for oncogenes and genes associated with proliferation, whereas recurrently lost arms were enriched for TSGs[7,9] (Fig. 3a). For example, *WRN* was recently suggested to be a haploinsufficient TSG, which contributes to the loss of chromosome arm 8p in human cancers[19].

It has also been suggested that aneuploidy could 'hard-wire' chromosome-wide gene expression biases that pre-exist in the cancer tissue of origin[11]. We recently reproduced this finding in machine learning-based modeling of aneuploidy patterns[12], in which we found that chromosome arms that tend to be gained exhibit higher tissue-specific chromosome arm expression levels than other chromosome arms, whereas chromosome arms that tend to be lost exhibit lower tissue-specific expression levels. Importantly, the genomic associations of recurrent aneuploidies with genomic features, such as gene density and gene expression, do not contradict the potential presence

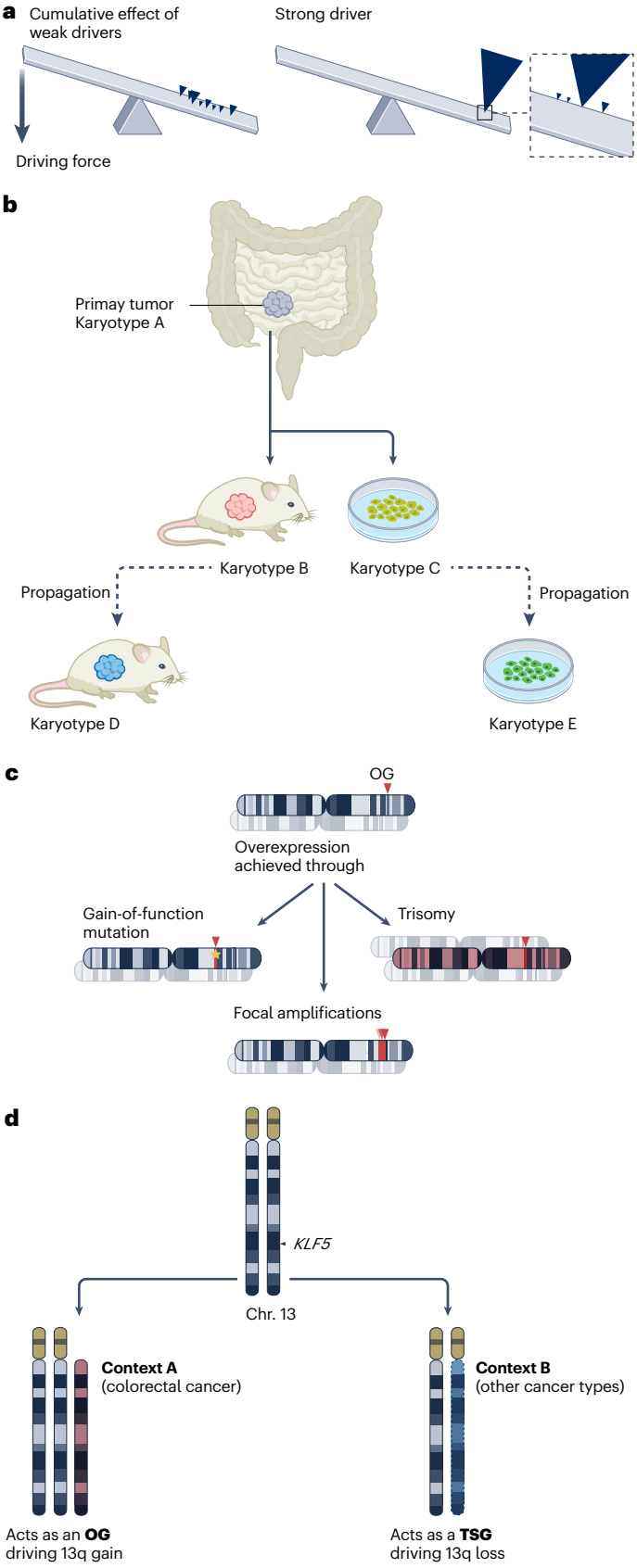

Fig. 3 | Challenges in identifying the driver genes underlying driver aneuploidies. a, Seesaw analogy illustrating the driving forces underlying driver aneuploidies. Identifying driver genes underlying a driver aneuploidy is challenging, as it can be driven by a cumulative effect of many genes (left). Even with a strong driver gene, additional genes on the aneuploid chromosome can contribute to the beneficial effect of the aneuploidy, and strong drivers can mask the weaker ones (right). **b**, Distinct cellular microenvironments provide a selective advantage to different karyotypes (represented by different colors). For example, the same cell type may present different aneuploidies in three versus two dimensions (karyotypes B and C, respectively), or whenever cancer models are propagated (karyotypes D and E). **c**, Identifying an oncogene (OG) as a driver of a trisomy could be masked by the fact that its overexpression may result from multiple mechanisms, such as gain-of-function mutations and focal amplifications, in addition to trisomy. The multiple alternative ways to activate an oncogene can also complicate its identification as a bona fide driver when comparing tumors with and without the trisomy. **d**, The tissue specificity of a gene acting as an oncogene or a TSG, exemplified here with *KLF5*, can result in its driving opposing aneuploidies in cancer types that originate from different tissues.

of a few genes with strong effects. On the other hand, as mentioned above, the mere presence of oncogenes or TSGs within recurrent aneuploidies is not sufficient to define them as aneuploidy drivers; further gene-specific investigations are required to confirm their role as driver genes and distinguish them from passenger genes.

In sum, it becomes evident that aneuploidies are frequently shaped by the interplay between multiple genes, rather than by a single strong driver gene. This intricate network of genetic interactions must be considered in systematic attempts to identify the complex driving role(s) of recurrent aneuploidies, as single-gene-focused approaches (for example, most knockout and overexpression screens) would likely be limited for this purpose.

Strategies for identifying driver genes

Genomic analyses and their limitations. The identification of driver genes underlying driver aneuploidies must overcome multiple inherent challenges associated with genomic analyses (Box 2 and Fig. 3b–d). First, it is expected that driver genes are located within the minimal recurrently aberrant region[1]. Therefore, the use of genomic tools such as GISTIC2.0 ('genomic identification of significant targets in cancer'; Box 1) for the identification of significantly altered chromosomal regions proves useful[61]. Another recently developed tool mentioned already above is BISCUT[19], which allows the identification of changes in breakpoint frequencies of partial somatic CNAs. Using BISCUT could help narrow down the region that undergoes selection, potentially identifying driver genes within it. However, these tools heavily depend on the quality and the comprehensiveness of the available genomic databases; if a relevant driver gene or aneuploidy is not well represented in these databases, it may be missed. Another potent method in this context is synteny analysis to compare recurrent aneuploidies across species. Owing to incomplete synteny between the mouse and human genomes, comparing the minimal recurrent region between species can help narrow down regions of interest[60,62–64]. Once the region of interest has been narrowed down, further genomic analyses can elucidate specific gene candidates. Synteny analysis is, of course, limited by the various degrees of synteny between humans and other species as well as the fact that some driver genes may be species specific and not necessarily evolutionarily conserved.

Combining aneuploidy patterns with mutation and focal CNA patterns can point to candidate driver genes. First, examining the mutation and gene-level CNA frequency could help identify a gene as a suspected cancer gene, increasing the likelihood that it is an aneuploidy driver. Furthermore, the specific types of mutations (for instance, missense or nonsense) can further support a gene as a putative driver[65]. Once a suspected driver gene is identified, co-occurrence and mutual exclusivity analyses with CNAs could provide further evidence that it is, indeed, a driver gene of aneuploidy. For example, known oncogenes were shown to be mutually exclusive with certain recurrent aneuploidies: *KRAS* mutations are mutually exclusive with chromosome arm 18q gains in pancreatic cancer, and *BRAF* mutations are mutually exclusive with chromosome arm 20q gains in colorectal cancer[18]. For an analysis of TSGs, it is important to distinguish genes that are biallelically inactivated, in which the aneuploidy could affect one allele and the other allele is perturbed in a different manner, from genes that are monoallelically affected, in which the other allele remains unperturbed. It is also important to note that mutual exclusivity and co-occurrence analyses are mostly valuable for identifying strong driver genes, rather than the potentially long tail of weaker drivers.

Differential gene expression analysis is another important means for identifying driver genes. Given that driver genes are assumed to be transcriptionally dysregulated, they are expected to be differentially expressed in aneuploid tumors in comparison to the diploid normal tissue of origin. To focus on changes primarily driven by CNAs in tumor cells, it is essential to integrate gene expression data with copy number data. Genes showing consistent and correlated copy number and gene expression changes are strong candidates for driver genes. By employing this method, *MCM4*, for example, was suggested as a driver gene of chromosome arm 8q gain in osteosarcoma[66]. Another approach involves investigating genes residing within a driver aneuploidy and exploring their expression patterns in relation to clinical outcomes of patients. Using this method, *EIF4EBP1* was suggested as a driver gene of trisomy 8 in Ewing sarcoma[67]. However, it is important to note that gene expression can change over time as tumors evolve. A gene may act as a driver at one stage but not at another, and this temporal aspect can be overlooked when analyzing static snapshots of gene expression data.

In vitro models and their limitations. Following genomic association analyses, the next crucial step is to experimentally validate the suspected driver genes through the utilization of functional models. Human cancer cell lines serve as one of the main and most powerful in vitro tools in cancer research[68]. Despite their value, cancer cell lines lack the heterogeneity characteristic of tumors and may not fully represent the complex in vivo conditions. Nonetheless, they remain an essential resource in cancer research, including for the research of aneuploidy. A comparison of cancer cell lines with and without a given aneuploidy can shed light on the role of the aneuploidy and on the genes that may drive it. For example, a comparison of multiple myeloma cell lines with and without chromosome arm 1q gain unveiled genes that reside on this arm and are more essential when the chromosome arm is gained[69].

Comparison of cancer cell lines with and without a given aneuploidy is confounded by other genomic differences between the cell lines. Therefore, isogenic models of human cancer before or after the introduction of aneuploidy are valuable (Fig. 4a). Such isogenic models can be established by isolating clones of cell lines that spontaneously acquired aneuploidy[70], by inducing random aneuploidy and isolating clones with a specific aneuploidy[10,30,36,71] or by direct aneuploidy induction using chromosome-specific engineering methods. The current chromosomal engineering methods are mostly based on the CRISPR–Cas9 system and were recently thoroughly reviewed[72,73]. In brief, current approaches for directed aneuploidy induction include restoring disomy in aneuploid cells using CRISPR targeting (ReDACT)[18] and molecular alteration of chromosomes with engineered tandem elements (MACHETE), two Cas9-based tools to engineer large genomic deletions[74]. Additional methods involve linking a dead Cas9 protein to either a plant kinesin[75], the centromere protein T (CENP-T)[76] or a mutant kinetochore scaffold 1 (KNL1) protein[59] to induce targeted chromosome missegregation.

The above techniques for generating aneuploidies vary in effectiveness and require single-cell cloning, which is resource intensive and laborious, limiting the scalability of experiments[73]. While there is still progress to be made, these methods have nevertheless greatly improved our ability to engineer specific aneuploidies. Once isogenic cells are generated, they can be subjected to systemic genomic, transcriptomic and proteomic characterization[10,30,32,36,77] as well as to genetic perturbation and pharmacological screens[18,30,36] to identify candidate aneuploidy driver genes. Such isogenic systems can also be used to test specific hypotheses[78,79], thereby allowing confirmation of candidate driver genes. Another advantage of engineered cell lines is their ability to functionally demonstrate the causal role of candidate driver genes using evolution experiments. For example, *RAD21* and *EIF4EBP1* were shown to contribute to the growth-promoting properties of trisomy 8 in Ewing sarcoma[60,67] and *MDM4* and *BCL9* were shown to contribute to the tumorigenic phenotypes of chromosome arm 1q gain[18] by applying gain-of-function and loss-of-function genetic manipulations and assessing their effect on aneuploidy-driven phenotypes.

In vivo models and their limitations. While in vitro models offer valuable insights, in vivo models provide a more comprehensive understanding of how candidate driver genes function in living organisms,

https://doi.org/10.1038/s41588-024-01916-2

BOX 2

Challenges in identifying driver genes

Identifying driver genes underlying driver aneuploidies is a challenging mission. First, by definition, aneuploidy affects hundreds to thousands of genes at once[1]. Second, as described above, in most cases, there is more than one single driver gene, and the interplay between several strong driver genes and multiple additional genes with a weaker effect determines the fitness advantage of each driver aneuploidy. Third, the cellular microenvironment shapes the constraints on karyotype evolution, making it difficult to dissect the functional role(s) of aneuploidy and the underlying genes in a cell-autonomous manner (Fig. 3b). For example, the aneuploidy landscapes differ between human tumors and human cancer cell lines. Therefore, while cancer cell lines can still be very useful for modeling aneuploidy, the selection pressures that shape aneuploidy in vitro and in vivo are not identical[1,70,81]. Furthermore, even within the same cell line, evolution through clonal selection leads to karyotypic heterogeneity[70]. Fourth, the existence of a common oncogene and/or TSG within a driver aneuploidy does not automatically indicate that this is a driver gene. Established cancer driver genes are not necessarily the causative factor behind aneuploidy recurrence. For example, consider the epidermal growth factor receptor (EGFR) and its overexpression in HER2$^+$ breast cancer[128]. Although initially presumed to drive the common gain of chromosome 7 in some tumor types, it was recently suggested that EGFR overexpression, but not high *EGFR* copy number, is a poor prognostic factor in HER2$^+$ primary breast cancer[128]. Other mechanisms, such as transcriptional regulation, post-translational modifications or genomic alterations that extend beyond changes in copy number (for example, point mutations), could potentially be the main contributors to its expression (Fig. 3c). In GBM, *EGFR* overexpression is often driven by its amplification, and trisomy of chromosome 7 is a common event, but the aneuploidy appears to be somewhat independent of gene amplification[129]. These findings imply that, although alterations in *EGFR* expression may influence cancer development, they might be disassociated from trisomy of chromosome 7, indicating that *EGFR*, although driving cancer, is not necessarily a driver gene of this aneuploidy. Furthermore, the tissue specificity of driver genes is another important challenge. While there are pan-cancer TSGs and oncogenes that drive recurrent aneuploidies, there are also tissue-specific TSGs and oncogenes that contribute to specific aneuploidies[7]. It is possible that distinct genes drive the same aneuploidy in different tumor types, and the same gene may even drive opposite aneuploidies. For example, although *KLF5* functions as a TSG in some cancer types, it plays a distinct oncogenic role in colorectal cancer, potentially explaining why chromosome 13 is recurrently gained in colorectal cancer despite its common loss in most tumor types[12] (Fig. 3d). Finally, multiple genes have been proposed as drivers of chromosome arm 1q gain across various cancer types, including *MDM4*, *BCL9* (ref. 18), *APH1A* and *NCSTN*[10]; the exact contribution of each of these genes to the recurrence of this aneuploidy appears to depend on the cancer type and also on the cellular context (for example, cellular milieu and nutrient availability). These findings align with the observation that many genes exhibit highly variable effects on cell proliferation in a cell type-specific manner[9]. The tissue specificity of driver genes thus adds another layer of complexity to the identification of driver genes underlying driver aneuploidies.

bridging the gap between the laboratory and the clinic. The most basic in vivo model is the cell line xenograft model, which involves the introduction of cell lines into mice, both with and without the specific aneuploidy of interest[18,32,34,78,79]; this allows the monitoring of tumor formation, growth, metastasis and other relevant factors. Subsequently, these cell line-derived tumors can be collected for further investigation, including DNA sequencing and gene expression analyses. Furthermore, exogenous modulation of candidate driver genes can be carried out to observe their impact on tumor features, thereby functionally interrogating their driver role[18]. A notable limitation of this approach is that it is based on cancer cell lines, with all the caveats associated with them[80].

Another model used for identifying driver genes is patient-derived xenografts (PDXs), which involve the direct transplantation of human tumors into mice (Fig. 4a). This approach can provide a more accurate representation of the human cancer genome than other mouse models. However, as we previously suggested, the selection pressures in this animal model might differ from those in human patients, and the propagation of PDXs can cause them to diverge from the primary tumors from which they originated[81–83]. Additionally, as PDX models lack an intact immune system, they may not fully recapitulate the genetics and the histology of human tissue tumors[84]. Nevertheless, PDX models offer an opportunity to identify driver genes underlying driver aneuploidies that, for example, drive drug resistance (Fig. 4a).

Genetically engineered mouse models (GEMMs) are another approach for identifying driver genes (Fig. 4a), as they can recapitulate tumor formation within the organism. GEMMs enable the development of heterogeneous tumors and are immunocompetent, allowing modeling of the complexity of tumor formation and its interaction with the immune system. We previously reported that, in breast cancer GEMMs, the prevalence of CNAs and recurrent aneuploidy patterns are primarily determined by the initiating perturbations (that is, the genetic manipulation used for tumor initiation)[64]. To directly examine mouse models of aneuploidy, in vivo chromosome missegregation can also be induced in mice, for example, by inducing chromosomal instability. This could result in tumor formation and potentially involve selection for specific karyotypes[16,17,85]. Once a suitable GEMM is obtained, comparative studies of gene expression can be performed between tumors with and without an aneuploidy of interest. Moreover, assuming that the same gene(s) underlie the recurrence of aneuploidy in both species, GEMMs can be compared to human tumors to narrow down chromosomal regions of interest and to identify potential driver genes[64,86,87]. Subsequently, functional studies can be performed to confirm the roles of candidate driver genes. For example, a candidate driver gene could be ectopically expressed on a different chromosome to investigate its role in driving chromosomal alterations[17]. Finally, recent approaches such as MACHETE allow for in vivo chromosome engineering, making it possible to investigate the role of recurrent large chromosomal aberrations in the GEMM setting[74].

In sum, in vivo models provide critical insights into the function of candidate driver genes within living organisms, and they serve as indispensable tools in the ongoing pursuit of identifying the driver genes underlying driver aneuploidies in cancer. However, there is no 'one model fits all' for aneuploidy research, and therefore we need to carefully choose the appropriate model for each question.

Targeting driver aneuploidies

Aneuploidy leads to many cellular consequences, such as chromosomal instability[1,31,88,89], cell cycle dysregulation[90], widespread gene expression changes[33,91] and various cellular stresses that include replicative,

https://doi.org/10.1038/s41588-024-01916-2

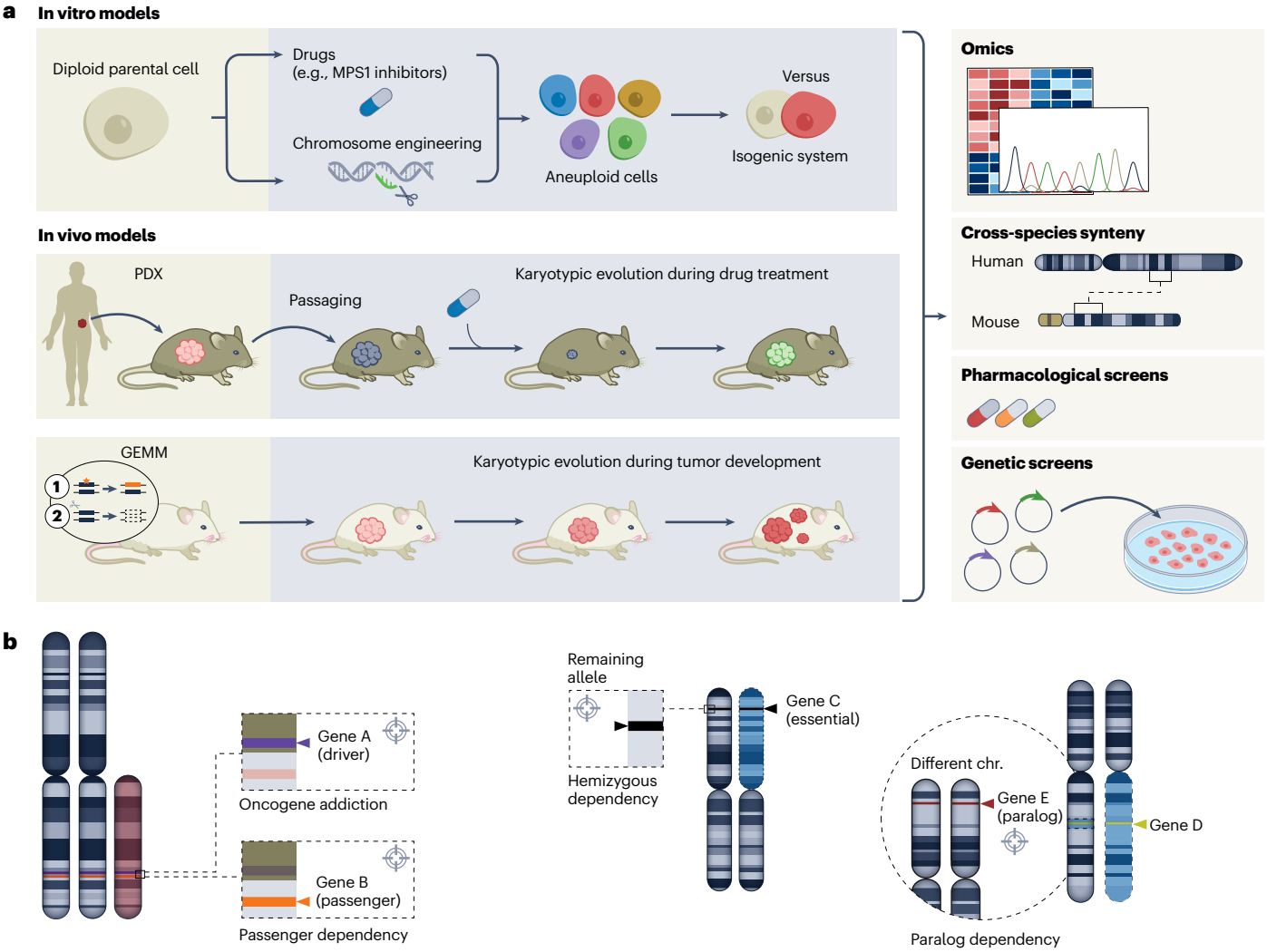

Fig. 4 | Modeling and targeting driver aneuploidies. a, Determining the cellular effects of a specific aneuploidy is necessary to identify its driver genes and can be achieved by comparing isogenic cancer models with and without the aneuploidy of interest. In vitro, aneuploidy can be induced (either in a targeted or untargeted manner), and clones with or without an aneuploidy of interest can be derived. In vivo models offer additional insights into the driving role of genes underlying driver aneuploidies. PDXs can be used to investigate the role(s) of genes underlying drug-resistant aneuploidies, thus facilitating translational research, whereas GEMMs can be used to perturb putative driver genes and study early stages of tumorigenesis. Further omics, cross-species synteny analyses and functional studies, such as drug and CRISPR screens, can be performed to confirm whether the putative driver gene(s) are indeed responsible for driving the observed phenotypes. Cell colors represent different karyotypes. MPS1, monopolar spindle 1. **b**, Aneuploid cells can create vulnerabilities to the disruption of various types of genes and thus provide opportunities to target recurrent aneuploidies. Here, driver genes are inherently designated as targets (oncogene addiction), but passenger genes in their vicinity may also offer vulnerabilities (passenger dependency) (left). Recurrent chromosome losses often result in hemizygous essential genes, rendering cells dependent on the unperturbed allele (hemizygous dependency) (center), while homozygous losses of genes can render cells dependent on paralog genes that compensate for the former's loss (paralog dependency), which could potentially be therapeutically targeted (right).

mitotic, proteotoxic and metabolic stresses (reviewed in ref. 29). Targeting these general cellular consequences of aneuploidy in highly aneuploid cells is the focus of much research and has been discussed elsewhere[1,30,31,36,92]. Here, we specifically focus on the potential therapeutic targeting of both driver and nondriver genes underlying recurrent aneuploidies (Fig. 4b).

Uncovering driver genes within recurrent aneuploidies is the mainstay of the potential therapeutic targeting of aneuploidy (Fig. 4b). For example, in a study of *Kras*-driven breast cancer in mice, the mesenchymal–epithelial transition factor (*cMet*) oncogene was suggested to drive recurrent amplifications on chromosome 6 in tumors that are resistant to therapy[93]. Accordingly, targeting these tumors with the *cMet* inhibitor tepotinib resulted in decreased proliferation and a reduction in tumor size, proving that *cMet* was indeed essential for the resistant tumors[93]. Similarly, as mentioned above, *MDM4* and *BCL9* were shown to promote chromosome arm 1q gain in ovarian cancer[18], which naturally positions them as candidate targets. Likewise, *APH1A* and *NCSTN*, encoding subunits of the γ-secretase complex, were recently suggested to drive 1q gain in breast cancer cells, rendering them more sensitive to γ-secretase inhibition[10]. Overall, the identification of driver genes within driver aneuploidies inherently designates them as candidate targets, as some of them may constitute so-called oncogenic addictions.

In addition to driver genes underlying recurrent aneuploidies, co-occurring passenger alterations within these aneuploidies may also create unique cellular vulnerabilities[1,31,46,94,95]. For chromosome arm or whole-chromosome losses, several studies have demonstrated that hemizygosity of essential genes could render the aneuploid cells

more sensitive to their perturbation (Fig. 4b). For example, in colorectal cancer, *POLR2A* is co-deleted with *TP53* as part of a chromosome arm aneuploidy[96]. Furthermore, the same aneuploidy also sensitized prostate cancer cells to perturbation of RBX1, an RNA polymerase II subunit A (POLR2A) activator[79], further illustrating this phenomenon. In addition to direct targeting of genes residing on the affected chromosome, copy number loss of genes with functional paralogs on other chromosomes can render cells more sensitive to the inhibition of these paralogs[12,97–100] (Fig. 4b).

Chromosome arm gains can also confer cellular vulnerabilities, attributed to the presence of passenger genes (Fig. 4b). For example, in ovarian cancer cells with 1q gain, upregulation of the uridine–cytidine kinase encoded by *UCK2* rendered them more sensitive to nucleotide analogs[18]. We recently exposed another class of such passenger genes, which we termed amplification-related gain-of-sensitivity genes[46]. These genes reside within recurrently gained genomic loci and are toxic to cells when overexpressed, therefore undergoing dosage compensation[46]. Targeting the dosage-compensation mechanisms can therefore selectively inhibit cells in which these genes are gained due to aneuploidy.

It is important to note that, in some cases, cellular vulnerabilities of recurrent aneuploidies could be identified even when the underlying culprit genes remain unknown. For example, loss of chromosome arm 8p was shown to result in altered lipid metabolism, which induces autophagy and increases resistance to hypoxic conditions[78], thereby rendering the cells more sensitive to autophagy inhibitors. We note that identifying such cellular vulnerabilities of driver aneuploidies can shed light on their functional consequences, which can ultimately also contribute to uncovering underlying driver genes.

In summary, the identification and understanding of the cellular consequences of driver aneuploidies as well as those of specific driver and nondriver genes within these aneuploidies, offer promising avenues for developing targeted therapies against aneuploid tumors.

Conclusions and outlook

The understanding of the functional consequences of recurrent aneuploidies requires us to first identify which aneuploidies are subject to positive selection during tumorigenesis (driver aneuploidies) and then to expose the driver genes underlying these aneuploidies. This pursuit has substantially progressed in recent years, yet several challenges remain. In particular, defining driver aneuploidies within a specific context is fundamental for precise and effective exploration of the aneuploidy driver genes underlying cancer progression.

Genomic analyses have greatly expanded our understanding of cancer aneuploidy. As tumor sequencing data become more available and computational analysis tools improve, it is increasingly feasible to identify and prioritize candidate aneuploidy driver genes based on genomic tumor data. Nonetheless, the complexity of this problem and its combinatorial nature still pose a serious challenge. Moreover, genomic analyses are not sufficient to prove causality and should be complemented with functional studies. The ongoing quest to unravel the driver genes within driver aneuploidies is multifaceted. Novel genomic tools, in vitro models and in vivo models, combined with chromosome engineering approaches and coupled to systematic approaches for genetic and pharmacological perturbations, hold great promise for the near future. Further refining and expanding genomic, transcriptomic and functional databases, improving chromosomal engineering methods and advancing in vivo models will further facilitate greater understanding of the role of aneuploidy in cancer formation, progression and response to therapy.

References

1. Ben-David, U. & Amon, A. Context is everything: aneuploidy in cancer. *Nat. Rev. Genet.* **21**, 44–62 (2020).
2. Li, R. & Zhu, J. Effects of aneuploidy on cell behaviour and function. *Nat. Rev. Mol. Cell Biol.* **23**, 250–265 (2022).
3. Vasudevan, A. et al. Aneuploidy as a promoter and suppressor of malignant growth. *Nat. Rev. Cancer* **21**, 89–103 (2021).
4. Weaver, B. A. & Cleveland, D. W. Does aneuploidy cause cancer? *Curr. Opin. Cell Biol.* **18**, 658–667 (2006).
5. Compton, D. A. Mechanisms of aneuploidy. *Curr. Opin. Cell Biol.* **23**, 109–113 (2011).
6. Santaguida, S. & Amon, A. Short- and long-term effects of chromosome mis-segregation and aneuploidy. *Nat. Rev. Mol. Cell Biol.* **16**, 473–485 (2015).
7. Davoli, T. et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155**, 948–962 (2013).
8. Taylor, A. M. et al. Genomic and functional approaches to understanding cancer aneuploidy. *Cancer Cell* **33**, 676–689 (2018).
9. Sack, L. M. et al. Profound tissue specificity in proliferation control underlies cancer drivers and aneuploidy patterns. *Cell* **173**, 499–514 (2018).
10. Watson, E. V. et al. Chromosome evolution screens recapitulate tissue-specific tumor aneuploidy patterns. *Nat. Genet.* **56**, 900–912 (2024).
11. Patkar, S. et al. Hard wiring of normal tissue-specific chromosome-wide gene expression levels is an additional factor driving cancer type-specific aneuploidies. *Genome Med.* **13**, 93 (2021).
12. Jubran, J. et al. Machine-learning analysis reveals an important role for negative selection in shaping cancer aneuploidy landscapes. *Genome Biol.* **25**, 95 (2024).
13. Stopsack, K. H. et al. Aneuploidy drives lethal progression in prostate cancer. *Proc. Natl Acad. Sci. USA* **116**, 11390–11395 (2019).
14. Ippolito, M. R. et al. Gene copy-number changes and chromosomal instability induced by aneuploidy confer resistance to chemotherapy. *Dev. Cell* **56**, 2440–2454 (2021).
15. Lukow, D. A. et al. Chromosomal instability accelerates the evolution of resistance to anti-cancer therapies. *Dev. Cell* **56**, 2427–2439 (2021).
16. Shoshani, O. et al. Transient genomic instability drives tumorigenesis through accelerated clonal evolution. *Genes Dev.* **35**, 1093–1108 (2021).
17. Trakala, M. et al. Clonal selection of stable aneuploidies in progenitor cells drives high-prevalence tumorigenesis. *Genes Dev.* **35**, 1079–1092 (2021).
18. Girish, V. et al. Oncogene-like addiction to aneuploidy in human cancers. *Science* **381**, eadg4521 (2023).
19. Shih, J. et al. Cancer aneuploidies are shaped primarily by effects on tumour fitness. *Nature* **619**, 793–800 (2023).
20. Gerstung, M. et al. The evolutionary history of 2,658 cancers. *Nature* **578**, 122–128 (2020).
21. Williams, M. J. et al. Luminal breast epithelial cells from wildtype and *BRCA* mutation carriers harbor copy number alterations commonly associated with breast cancer. Preprint at *bioRxiv* https://doi.org/10.1101/2F2024.05.01.591587 (2024).
22. Lawrence, M. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
23. Idbaih, A. et al. BAC array CGH distinguishes mutually exclusive alterations that define clinicogenetic subtypes of gliomas. *Int. J. Cancer* **122**, 1778–1786 (2008).
24. Skoulidis, F. & Heymach, J. V. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nat. Rev. Cancer* **19**, 495–509 (2019).
25. Worrall, J. T. et al. Non-random mis-segregation of human chromosomes. *Cell Rep.* **12**, 3366–3380 (2018).
26. Dumont, M. et al. Human chromosome-specific aneuploidy is influenced by DNA-dependent centromeric features. *EMBO J.* **39**, e102924 (2020).

27. Klaasen, S. J. et al. Nuclear chromosome locations dictate segregation error frequencies. *Nature* **607**, 604–609 (2022).
28. Sheltzer, J. M. & Amon, A. The aneuploidy paradox: costs and benefits of an incorrect karyotype. *Trends Genet.* **27**, 446–453 (2011).
29. Zhu, J. et al. Cellular stress associated with aneuploidy. *Dev. Cell* **44**, 420–431 (2018).
30. Ippolito, M. R. et al. Increased RNA and protein degradation is required for counteracting transcriptional burden and proteotoxic stress in human aneuploid cells. Preprint at *bioRxiv* https://doi.org/10.1101/2023.01.27.525826 (2023).
31. Gordon, D. J. et al. Causes and consequences of aneuploidy in cancer. *Nat. Rev. Genet.* **13**, 189–203 (2012).
32. Sheltzer, J. M. et al. Single-chromosome gains commonly function as tumor suppressors. *Cancer Cell* **31**, 240–255 (2017).
33. Ben-David, U. et al. Aneuploidy induces profound changes in gene expression, proliferation and tumorigenicity of human pluripotent stem cells. *Nat. Commun.* **5**, 4825 (2014).
34. Vasudevan, A. et al. Single-chromosomal gains can function as metastasis suppressors and promoters in colon cancer. *Dev. Cell* **52**, 413–428 (2020).
35. Laue, K. et al. Inactivation of p53 drives breast cancer brain metastasis by altering fatty acid metabolism. Preprint at *bioRxiv* https://doi.org/10.1101/2023.12.20.572490 (2023).
36. Zerbib, J. et al. Human aneuploid cells depend on the RAF/MEK/ERK pathway for overcoming increased DNA damage. Preprint at *bioRxiv* https://doi.org/10.1101/2023.01.27.525822 (2023).
37. Alfieri, F. et al. Cancer genomes tolerate deleterious coding mutations through somatic copy number amplifications of wild-type regions. *Nat. Commun.* **14**, 3594 (2023).
38. Martincornea, I. & Campbell, P. J. Somatic mutation in cancer and normal cells. *Science* **349**, 1483–1489 (2015).
39. Martincorena, I. et al. Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029–1041 (2017).
40. Stingele, S. et al. Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells. *Mol. Syst. Biol.* **8**, 608 (2012).
41. Liu, Y. et al. Systematic proteome and proteostasis profiling in human trisomy 21 fibroblast cells. *Nat. Commun.* **8**, 1212 (2017).
42. Gonçalves, E. et al. Widespread post-transcriptional attenuation of genomic copy-number variation in cancer. *Cell Syst.* **5**, 386–398 (2017).
43. Brennan, C. M. et al. Protein aggregation mediates stoichiometry of protein complexes in aneuploid cells. *Genes Dev.* **33**, 1031–1047 (2019).
44. Schukken, K. M. & Sheltzer, J. M. Extensive protein dosage compensation in aneuploid human cancers. *Genome Res.* **32**, 1254–1270 (2022).
45. Spealman, P. et al. Post-transcriptional mechanisms modulate the consequences of adaptive copy number variation. Preprint at *bioRxiv* https://doi.org/10.1101/2023.10.20.563336 (2023).
46. Rendo, V. et al. A compendium of amplification-related gain of sensitivity (ARGOS) genes in human cancer. Preprint at *bioRxiv* https://doi.org/10.1101/2023.12.16.571980 (2023).
47. Bueno, M. J. et al. Genetic and epigenetic silencing of microRNA-203 enhances *ABL1* and *BCR-ABL1* oncogene expression. *Cancer Cell* **13**, 496–506 (2008).
48. Kim, Y. K. et al. Functional links between clustered microRNAs: suppression of cell-cycle inhibitors by microRNA clusters in gastric cancer. *Nucleic Acids Res.* **37**, 1672–1681 (2009).
49. Tsuchida, A. et al. miR-92 is a key oncogenic component of the miR-17–92 cluster in colon cancer. *Cancer Sci.* **102**, 2264–2271 (2011).
50. Rao, E. et al. The miRNA-17~92 cluster mediates chemoresistance and enhances tumor growth in mantle cell lymphoma via PI3K/AKT pathway activation. *Leukemia* **26**, 1064–1072 (2012).
51. Yang, L. et al. lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature* **500**, 598–602 (2013).
52. Prensner, J. R. et al. The long non-coding RNA PCAT-1 promotes prostate cancer cell proliferation through cMyc. *Neoplasia* **16**, 900–908 (2014).
53. Lanzós, A. et al. Discovery of cancer driver long noncoding RNAs across 1112 tumour genomes: new candidates and distinguishing features. *Sci. Rep.* **7**, 41544 (2017).
54. Balatti, V. et al. *TCL1* targeting miR-3676 is codeleted with tumor protein p53 in chronic lymphocytic leukemia. *Proc. Natl Acad. Sci. USA* **112**, 2169–2174 (2015).
55. Tseng, Y. Y. et al. *PVT1* dependence in cancer with *MYC* copy-number increase. *Nature* **512**, 82–86 (2014).
56. Jones, L. et al. Gain of MYC underlies recurrent trisomy of the MYC chromosome in acute promyelocytic leukemia. *J. Exp. Med.* **207**, 2581–2594 (2010).
57. Liu, Y. et al. Deletions linked to *TP53* loss drive cancer through p53-independent mechanisms. *Nature* **531**, 471–475 (2016).
58. Bostrom, J. et al. Mutation of the *PTEN* (MMAC1) tumor suppressor gene in a subset of glioblastomas but not in meningiomas with loss of chromosome arm 10q. *Cancer Res.* **58**, 29–33 (1998).
59. Bosco, N. et al. KaryoCreate: a CRISPR-based technology to study chromosome-specific aneuploidy by targeting human centromeres. *Cell* **186**, 1985–2001 (2023).
60. Su, X. A. et al. *RAD21* is a driver of chromosome 8 gain in Ewing sarcoma to mitigate replication stress. *Genes Dev.* **35**, 556–572 (2021).
61. 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).
62. Herschkowitz, J. I. et al. Comparative oncogenomics identifies breast tumors enriched in functional tumor-initiating cells. *Proc. Natl Acad. Sci. USA* **109**, 2778–2783 (2012).
63. Silva, G. O. et al. Cross-species DNA copy number analyses identifies multiple 1q21–q23 subtype-specific driver genes for breast cancer. *Breast Cancer Res. Treat.* **152**, 347–356 (2015).
64. Ben-David, U. et al. The landscape of chromosomal aberrations in breast cancer mouse models reveals driver-specific routes to tumorigenesis. *Nat. Commun.* **7**, 12160 (2016).
65. Martínez-Jiménez, F. et al. A compendium of mutational cancer driver genes. *Nat. Rev. Cancer* **20**, 555–572 (2020).
66. Kuijjer, M. L. et al. Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes Chromosomes Cancer* **51**, 696–706 (2012).
67. Funk, M. C. et al. Chromosome 8 gain drives cancer progression by hijacking the translation factor 4E-BP1 sensitizing for targeted CDK4/6 inhibition. Preprint at *bioRxiv* https://doi.org/10.1101/2022.12.11.519935 (2024).
68. Ghandi, M. et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).
69. Sklavenitis-Pistofidis, R. et al. Systematic characterization of therapeutic vulnerabilities in multiple myeloma with Amp1q reveals increased sensitivity to the combination of MCL1 and PI3K inhibitors. Preprint at *bioRxiv* https://doi.org/10.1101/2023.08.01.551480 (2023).
70. Ben-David, U. et al. Genetic and transcriptional evolution alters cancer cell line drug response. *Nature* **560**, 325–330 (2018).
71. Hintzen, D. C. et al. The impact of monosomies, trisomies and segmental aneuploidies on chromosomal stability. *PLoS ONE* **17**, e0268579 (2022).
72. Truong, M. A. et al. Modeling specific aneuploidies: from karyotype manipulations to biological insights. *Chromosome Res.* **31**, 25 (2023).

73. Lakhani, A. A. et al. Aneuploidy in human cancer: new tools and perspectives. *Trends Genet.* **39**, 968–980 (2023).
74. Barriga, F. M. et al. MACHETE identifies interferon-encompassing chromosome 9p21.3 deletions as mediators of immune evasion and metastasis. *Nat. Cancer* **3**, 1367–1385 (2022).
75. Truong, M. A. et al. A kinesin-based approach for inducing chromosome-specific mis-segregation in human cells. *EMBO J.* **42**, e111559 (2023).
76. Tovini, L. et al. Targeted assembly of ectopic kinetochores to induce chromosome-specific segmental aneuploidies. *EMBO J.* **42**, e111587 (2023).
77. Kuznetsova, A. Y. et al. Chromosomal instability, tolerance of mitotic errors and multidrug resistance are promoted by tetraploidization in human cells. *Cell Cycle* **14**, 2810–2820 (2015).
78. Cai, Y. et al. Loss of chromosome 8p governs tumor progression and drug response by altering lipid metabolism. *Cancer Cell* **29**, 751–766 (2016).
79. Li, Y. et al. Heterozygous deletion of chromosome 17p renders prostate cancer vulnerable to inhibition of RNA polymerase II. *Nat. Commun.* **9**, 4394 (2018).
80. Ben-David, U. et al. Genomic evolution of cancer models: perils and opportunities. *Nat. Rev. Cancer* **19**, 97–109 (2019).
81. Ben-David, U. et al. Patient-derived xenografts undergo mouse-specific tumor evolution. *Nat. Genet.* **49**, 1567–1575 (2017).
82. Sun, H. et al. Comprehensive characterization of 536 patient-derived xenograft models prioritizes candidates for targeted treatment. *Nat. Commun.* **12**, 5086 (2021).
83. Hoge, A. C. H. et al. DNA-based copy number analysis confirms genomic evolution of PDX models. *NPJ Precis. Oncol.* **6**, 30 (2022).
84. Sajjad, H. et al. Cancer models in preclinical research: a chronicle review of advancement in effective cancer research. *Animal Model. Exp. Med.* **4**, 87–103 (2021).
85. Baslan, T. et al. Ordered and deterministic cancer genome evolution after p53 loss. *Nature* **608**, 795–802 (2022).
86. Maser, R. et al. Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. *Nature* **447**, 966–971 (2007).
87. Rennhack, J. P. et al. Integrated analyses of murine breast cancer models reveal critical parallels with human disease. *Nat. Commun.* **10**, 3261 (2019).
88. Tijhuis, A. E. et al. The emerging links between chromosomal instability (CIN), metastasis, inflammation and tumour immunity. *Mol. Cytogenet.* **12**, 17 (2019).
89. Chunduri, N. K. & Storchová, Z. The diverse consequences of aneuploidy. *Nat. Cell Biol.* **21**, 54–62 (2019).
90. Wenzel, E. S. & Singh, A. T. K. Cell-cycle checkpoints and aneuploidy on the path to cancer. *In Vivo* **32**, 1–5 (2018).
91. Sheltzer, J. M. et al. Transcriptional consequences of aneuploidy. *Proc. Natl Acad. Sci. USA* **109**, 12644–12649 (2012).
92. Cohen-Sharir, Y. et al. Aneuploidy renders cancer cells vulnerable to mitotic checkpoint inhibition. *Nature* **590**, 486–491 (2021).
93. Salgueiro, L. et al. Acquisition of chromosome instability is a mechanism to evade oncogene addiction. *EMBO Mol. Med.* **12**, e10941 (2020).
94. Nijhawan, D. et al. Cancer vulnerabilities unveiled by genomic loss. *Cell* **150**, 842–854 (2012).
95. Paolella, B. R. et al. Copy-number and gene dependency analysis reveals partial copy loss of wild-type *SF3B1* as a novel cancer vulnerability. *eLife* **6**, e23268 (2017).
96. Liu, Y. et al. *TP53* loss creates therapeutic vulnerability in colorectal cancer. *Nature* **520**, 697–701 (2015).
97. Muller, F. L. et al. Passenger deletions generate therapeutic vulnerabilities in cancer. *Nature* **488**, 337–342 (2012).
98. Viswanathan, S. R. et al. Genome-scale analysis identifies paralog lethality as a vulnerability of chromosome 1p loss in cancer. *Nat. Genet.* **50**, 937–943 (2018).
99. Ito, T. et al. Paralog knockout profiling identifies *DUSP4* and *DUSP6* as a digenic dependence in MAPK pathway-driven cancers. *Nat. Genet.* **53**, 1664–1672 (2021).
100. Huth, T. et al. Chromosome 8p engineering reveals increased metastatic potential targetable by patient-specific synthetic lethality in liver cancer. *Sci. Adv.* **9**, eadh1442 (2023).
101. Buetow, K. H. et al. High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc. Natl Acad. Sci. USA* **98**, 581–584 (2001).
102. Navin, N. et al. Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
103. Wang, Y. et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
104. Zack, T. I. et al. Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **45**, 1134–1140 (2013).
105. Serin Harmanci, A. et al. CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data. *Nat. Commun.* **11**, 89 (2020).
106. Flensburg, C. et al. Detecting copy number alterations in RNA-seq using SuperFreq. *Bioinformatics* **37**, 4023–4032 (2021).
107. Bařinka, J. et al. RNAseqCNV: analysis of large-scale copy number variations from RNA-seq data. *Leukemia* **36**, 1492–1498 (2022).
108. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
109. Griffiths, J. A. et al. Mosaic autosomal aneuploidies are detectable from single-cell RNAseq data. *BMC Genomics* **18**, 904 (2017).
110. Fan, J. et al. Linking transcriptional and genetic tumor heterogeneity through allele analysis of single-cell RNA-seq data. *Genome Res.* **28**, 1217–1227 (2018).
111. Kumar, M. et al. Single-cell analysis of copy-number alterations in serous ovarian cancer reveals substantial heterogeneity in both low- and high-grade tumors. *Cell Cycle* **19**, 3154–3166 (2020).
112. Gao, R. et al. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat. Biotechnol.* **39**, 599–608 (2021).
113. Nahmad, A. D. et al. Frequent aneuploidy in primary human T cells after CRISPR–Cas9 cleavage. *Nat. Biotechnol.* **40**, 1807–1813 (2022).
114. Gao, T. et al. Haplotype-aware analysis of somatic copy number variations from single-cell transcriptomes. *Nat. Biotechnol.* **41**, 417–426 (2023).
115. De Falco, A. et al. A variational algorithm to detect the clonal copy number substructure of tumors from scRNA-seq data. *Nat. Commun.* **14**, 1074 (2023).
116. Mayrhofer, M., Viklund, B. & Isaksson, A. Rawcopy: improved copy number analysis with Affymetrix arrays. *Sci. Rep.* **6**, 36158 (2016).
117. Talevich, E. et al. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).
118. Bakker, B. et al. Single cell sequencing reveals karyotype heterogeneity in murine and human malignancies. *Genome Biol.* **17**, 115 (2016).
119. Patel, A. P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).
120. Carter, S. et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* **30**, 413–421 (2012).
121. Riester, M. et al. PureCN: copy number calling and SNV classification using targeted short read sequencing. *Source Code Biol. Med.* **11**, 13 (2016).

122. Spurr, L. F. et al. Quantification of aneuploidy in targeted sequencing data using ASCETS. *Bioinformatics* **37**, 2461–2463 (2021).
123. Curtis, C. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
124. Weinstein, J. N. et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **45**, 1113–1120 (2013).
125. Nassar, A. et al. Correlation of HER2 overexpression with gene amplification and its relation to chromosome 17 aneuploidy: a 5-year experience with invasive ductal and lobular carcinomas. *Int. J. Clin. Exp. Pathol.* **15**, 6254–6261 (2014).
126. Kuzmin, E. et al. Evolution of chromosome-arm aberrations in breast cancer through genetic network rewiring. *Cell Rep.* **43**, 113988 (2024).
127. Gatza, M. et al. An integrated genomics approach identifies drivers of proliferation in luminal-subtype human breast cancer. *Nat. Genet.* **46**, 1051–1059 (2014).
128. Lee, H. et al. Prognostic and predictive values of EGFR overexpression and *EGFR* copy number alteration in HER2-positive breast cancer. *Br. J. Cancer* **112**, 103–111 (2015).
129. Lopez-Gines, C. et al. New pattern of *EGFR* amplification in glioblastoma and the relationship of gene copy number with gene expression profile. *Mod. Pathol.* **23**, 856–865 (2010).

Acknowledgements

We thank the other members of the Ben-David laboratory for fruitful discussions. Figures were created using https://www.biorender.com. Work in the Ben-David laboratory is supported by the Israel Cancer Research Fund Project Award, the Israel Science Foundation project grant (1805/21), the BSF project grant (2019228) and the European Research Council Starting Grant (945674). E.S. is supported by the Eugenia Gelman Estate for Cancer Research PhD Fellowship, H.O. is supported by the JSPS Overseas Research Fellowship, and T.B.-Y. and R.S. are supported by the Edmond J. Safra Center for Bioinformatics at Tel Aviv University.

Author contributions

U.B.-D. conceived the manuscript, wrote the manuscript and designed the figures together with E.S. and approved the final version of the manuscript. E.S. wrote the manuscript and designed the figures together with U.B.-D. and prepared the figures. H.O., R.S. and T.B.-Y. provided valuable input on the manuscript and participated in its editing.

Competing interests

U.B.-D. receives consulting fees from Accent Therapeutics. The other authors declare no competing interest.

Additional information

Correspondence and requests for materials should be addressed to Uri Ben-David.

Peer review information *Nature Genetics* thanks the anonymous reviewers for their contribution to the peer review of this work.