Chromosome segregation errors generate a diverse spectrum of simple and complex genomic rearrangements

Peter Ly [1,9]*, Simon F. Brunner[2], Ofer Shoshani[1], Dong Hyun Kim[1], Weijie Lan[1], Tatyana Pyntikova[3], Adrienne M. Flanagan[4,5], Sam Behjati[2,6], David C. Page [3,7,8], Peter J. Campbell[2] and Don W. Cleveland [1]*

Cancer genomes are frequently characterized by numerical and structural chromosomal abnormalities. Here we integrated a centromere-specific inactivation approach with selection for a conditionally essential gene, a strategy termed CEN-SELECT, to systematically interrogate the structural landscape of mis-segregated chromosomes. We show that single-chromosome mis-segregation into a micronucleus can directly trigger a broad spectrum of genomic rearrangement types. Cytogenetic profiling revealed that mis-segregated chromosomes exhibit 120-fold-higher susceptibility to developing seven major categories of structural aberrations, including translocations, insertions, deletions, and complex reassembly through chromothripsis coupled to classical non-homologous end joining. Whole-genome sequencing of clonally propagated rearrangements identified random patterns of clustered breakpoints with copy-number alterations resulting in interspersed gene deletions and extrachromosomal DNA amplification events. We conclude that individual chromosome segregation errors during mitotic cell division are sufficient to drive extensive structural variations that recapitulate genomic features commonly associated with human disease.

Alterations in chromosome structure are pervasive in human cancers[1,2] and define a disease group known as genomic disorders[3–5]. These abnormalities can comprise a continuum of complexity, ranging from simple arm-level deletions to intricate networks of rearrangements connecting multiple chromosomes[6–10]. Tumors frequently harbor one or more forms of structural genomic rearrangements, which can underlie disease pathogenesis by altering DNA copy numbers[11], reshuffling linear DNA sequences[10,12–15], and/or perturbing regulatory architecture[5,16–20]. Cancer genome sequencing has recently uncovered complex structural changes that previously escaped detection by conventional cytogenetics, including chromothripsis—localized rearrangements from the catastrophic shattering of individual chromosomes and subsequent restitching in haphazard order[21].

Chromosomes are constantly under assault from intrinsic and extrinsic sources of DNA double-strand breaks (DSBs)[7,22–25], although these challenges are often counteracted by DNA-damage-repair mechanisms to maintain genomic integrity. Errors in mitotic cell division have emerged as a potent source of endogenous DSBs inflicted onto mis-segregated chromosomes[26]. Improper chromosome segregation causes numerical aneuploidy and has recently been linked to structural anomalies. For instance, mis-segregated chromosomes can generate unbalanced translocations when damaged in the cytokinetic furrow[27], acquire arm-level segmental imbalances[28,29], or, alternatively, become encapsulated into abnormal structures called micronuclei.

Micronuclei act as a spatially restricted compartment for DSB accumulation during interphase[30,31], with fragmentation of the micronucleated chromosome during the following mitosis[30,32,33]. Although the establishment of a genetically heritable, clonally rearranged human chromosome derived from these events has not been experimentally achieved, chromosomes within micronuclei can acquire localized rearrangement junctions in the next cell cycle[34] resembling chromothriptic-like signatures from cancer genomes[21]. Examples of chromothripsis involving multiple chromosomes and in combination with translocations have been identified in somatic[21,35–37] and inherited[5,38–42] cases, demonstrating that chromothriptic alterations can include events more diverse than rearrangements restricted to an individual chromosome. Despite its prevalence in disease, whether the spectrum of simple and complex genomic rearrangements can be initiated by a unifying mechanism remains unknown.

We previously established a somatic human cell-based approach to selectively inactivate the Y chromosome centromere by using a gene replacement strategy to rapidly exchange the centromere-specific histone H3 variant CENP-A with a loss-of-function chimeric mutant, which supports ongoing function of all centromeres except for the Y chromosome[33]. Centromere inactivation caused chromosome-specific mis-segregation into micronuclei, triggering whole-chromosome shattering during the subsequent mitosis coupled with efficient chromosome loss[33]. To explore the structural rearrangement landscape of mis-segregated chromosomes, we now develop CEN-SELECT—a strategy combining centromere inactivation-induced chromosome shattering with a drug selection marker engineered into the Y chromosome, which otherwise lacks somatically essential genes. The resulting chromosomal byproducts are

[1]Ludwig Institute for Cancer Research, Department of Cellular and Molecular Medicine, University of California San Diego School of Medicine, La Jolla, CA, USA. [2]Wellcome Sanger Institute, Hinxton, UK. [3]Whitehead Institute for Biomedical Research, Cambridge, MA, USA. [4]University College London Cancer Institute, London, UK. [5]Department of Histopathology, Royal National Orthopaedic Hospital NHS Trust, Stanmore, UK. [6]Department of Paediatrics, University of Cambridge, Cambridge, UK. [7]Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Cambridge, MA, USA. [8]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. [9]Present address: Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX, USA. *e-mail: peter.ly@utsouthwestern.edu; dcleveland@ucsd.edu

therefore unbiased from constraints typically imposed by autosomal gene-dosage effects. Implementing this approach with comprehensive cytogenetic and whole-genome sequencing (WGS) analyses, we show that mitotic segregation errors directly generate a broad spectrum of chromosomal aberrations that reconstruct the complex structural features of cancer genomes.

Results

Development of the CEN-SELECT system in diploid human cells.
To induce mitotic errors and identify cells maintaining the mis-segregated chromosome, we combined a Y centromere–specific inactivation strategy[33] with a Y chromosome–encoded selection marker, an approach we call CEN-SELECT (Fig. 1a). We started with p53-inactivated DLD-1 colorectal cancer cells that are chromosomally stable (46,XY) and carry one disrupted *CENPA* allele with the second allele modified to encode CENP-A fused to an EYFP-linked auxin-inducible degron (AID) tag (CENP-A$^{EYFP-AID/-}$)[33]. Addition of the auxin hormone indole-3-acetic acid (IAA) activates a constitutively expressed E3 ubiquitin ligase (osTIR1) to induce rapid CENP-A$^{EYFP-AID}$ degradation[43,44]. Clustered regularly interspaced palindromic repeats (CRISPR)–Cas9-mediated genome editing was used to insert a neomycin-resistance gene (*Neo^R*) (Supplementary Fig. 1a,b), which confers survival in the presence of the antibiotic G418, at the Y chromosome AZFa locus located ~2.5 Mb from the centromere on proximal Yq11.221 (ref. [45]). Correctly targeted clones were identified (Supplementary Fig. 1b,c), and a single-copy, doxycycline (DOX)-inducible wild-type CENP-A (CENP-A^{WT}) or chimeric mutant carrying a histone H3 carboxy-terminal tail substitution (CENP-A^{C-H3}) was integrated at a defined FRT site. Replacement of CENP-A$^{EYFP-AID}$ with chimeric CENP-A^{C-H3} supports functional kinetochore assembly at the centromere of all chromosomes except the Y[33], which lacks CENP-B-binding sequences[46]. Genetic modifications and cell cloning steps are outlined in Supplementary Fig. 1a.

Over half (59%) of cells underwent expected loss of the Y chromosome (LOY) (Fig. 1b) within 3 d of centromere inactivation induced by DOX/IAA addition, as determined by interphase fluorescence in situ hybridization (FISH) with Y centromere-targeting DNA probes (CEN-Y). In agreement with increased mitotic errors, the Y chromosome was encapsulated within micronuclei in 56% of cells that retained it—a ~100-fold increase over untreated cells (Fig. 1c,d).

Despite loss of kinetochore function, Y centromere identity is epigenetically preserved by CENP-A^{C-H3}-assembled chromatin[33]. We next determined whether inactivated Y centromeres could be reactivated following the removal of DOX/IAA. Endogenous CENP-A$^{EYFP-AID}$ returned to its initial steady-state level within 3–4 d of drug washout and was accompanied by a corresponding decline in CENP-A^{C-H3} synthesis (Supplementary Fig. 2a). Reaccumulated CENP-A$^{EYFP-AID}$ incorporated into centromeric chromatin in interphase and mitotic chromosomes, including the Y chromosome (Supplementary Fig. 2b-d), consistent with Y centromere reactivation, and the proportion of cells containing micronuclei approached baseline.

We next used transient Y centromere inactivation coupled with selection for *Neo^R* gene retention (Fig. 1e). While transient replacement of CENP-A$^{EYFP-AID}$ with CENP-A^{C-H3} alone had no effect on clonogenic growth from single cells, its combination with G418 selection sharply reduced survival by ~89% (Fig. 1f,g). Inspection of G418-resistant populations by FISH revealed that >90% of recovered cells scored positive for retention of the Y chromosome, the vast majority (98%) of which were positioned within the main nucleus (Fig. 1h,i), a finding consistent with centromere reactivation. Additionally, there was no effect on the viability of control cells expressing CENP-A^{WT} treated in parallel (Fig. 1h,i). Thus, CEN-SELECT can experimentally identify cells carrying chromosomes previously exposed to high-frequency mis-segregation into micronuclei.

Release of damaged micronuclear chromatin during mitosis. Micronuclei are intrinsically prone to acquiring DSBs[30,31,33]. Consistent with this, micronuclei produced by Y centromere inactivation recruited components involved in the DNA-damage response, including the MRE11-RAD50–NBS1 (MRN) complex (Supplementary Fig. 3a), which is critical for DSB recognition and signaling. Two ATM kinase substrates, H2AX[47] and NBS1 (refs. [48,49]), were also readily phosphorylated (S139 and S343, respectively) within most micronuclei examined. By contrast, whereas the 53BP1 DNA-repair factor[50,51] was rapidly recruited to DSB sites induced by chemotherapeutic agents in the major nucleus, micronuclear DSBs failed to actively recruit or retain 53BP1 (Supplementary Fig. 3b), in accordance with previous findings[52], suggesting that micronuclear DNA damage developing during interphase can persist unrepaired into the next mitosis.

Live-cell imaging of nuclear and micronuclear membranes fluorescently labeled with lamin A–green fluorescent protein (GFP) revealed that 6.4% of micronuclei ($n=140$, pooled from 3 experiments) exhibited abnormal nuclear envelope breakdown (NEBD) throughout mitosis (Supplementary Videos 1 and 2). Micronuclear envelope disassembly kinetics were next assessed by time-lapse microscopy of semipermeabilized interphase cells treated with mitotic *Xenopus* egg extracts to trigger NEBD[53,54] (Supplementary Fig. 4a-c). The extract-induced NEBD kinetics were identical between nuclei and micronuclei in most cells (Supplementary Fig. 4d,e), while in a subset (14%, $n=51$) the micronuclear envelope only partially disassembled and at a reduced rate (Supplementary Fig. 4e-g). Thus, whereas partially disassembled micronuclei can persist for another cell cycle, most micronuclei enter mitosis with timely NEBD, promoting further chromosome mis-segregation[55] and/or the reincorporation of damaged chromatin into daughter cell nuclei after mitotic exit[30].

Mitotic errors drive widespread structural rearrangements. To test whether chromosome segregation errors directly provoke gross chromosomal rearrangements, we hybridized Y chromosome paint probes to metaphase spread preparations. A majority (59 ± 8%, 92 of 153 metaphases) of cells examined from CEN-SELECT harbored at least one type of Y-specific structural change (Fig. 2a), with different alterations occurring in each cell (described below). Structural rearrangements were undetectable from 94 control metaphases rescued with CENP-A^{WT} and at <1% frequency (1 of 122 metaphases) in G418-resistant cells without centromere inactivation (Fig. 2a), consistent with a low spontaneous rearrangement rate.

To confirm that rearrangements did not result from selection for rare cells with pre-existing abnormalities, we isolated three independent clones that carried a cytogenetically normal Y chromosome (Supplementary Fig. 1a). Each clone underwent efficient LOY following centromere inactivation (Supplementary Fig. 1d). Y chromosome rearrangement frequencies were elevated in each clone following CEN-SELECT, supporting that the majority of rearrangements were generated de novo rather than by selection for pre-existing alterations (Fig. 2a). Across the parental cells and clones, CEN-SELECT produced Y chromosome rearrangements at a frequency greater than two orders of magnitude higher than G418 selection without centromere inactivation (139 of 297 metaphases versus 1 of 258 metaphases, respectively) (Fig. 2a). CEN-SELECT did not affect the low rearrangement frequency (1.4%, 2 of 144 metaphases) of a control autosome (chromosome 14), as determined using chromosome paint probes, and rearrangements between the Y and chromosome 14 were observed at a frequency (4%, 1 of 24 Y translocations) expected by chance (~4.4% in diploid cells). Clone no. 1 was selected for subsequent efforts, as it was completely haploid for the Y under untreated conditions ($n=45$ metaphases) and had a Y chromosome rearrangement frequency closely matching that of the parental clone after CEN-SELECT.
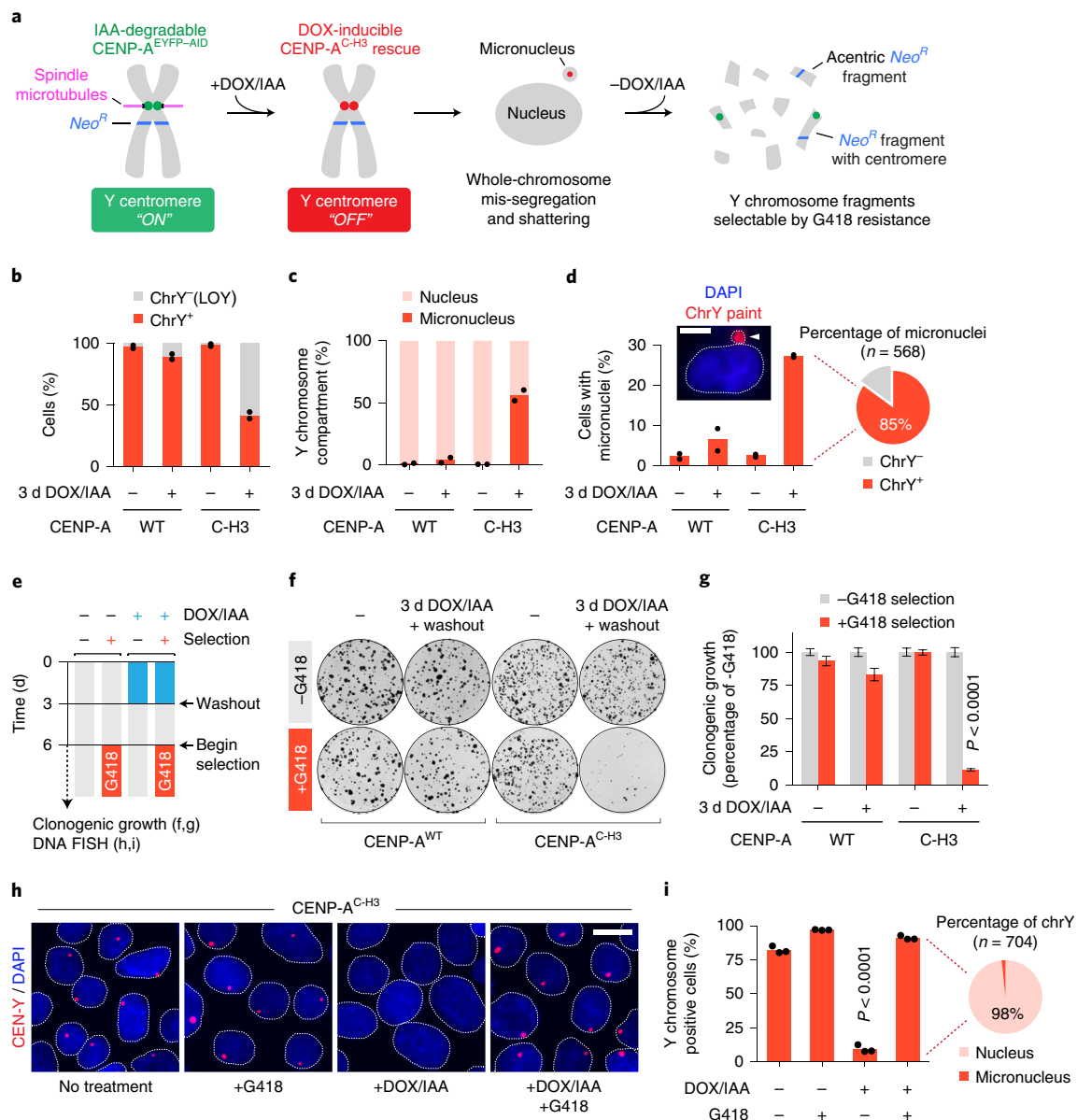
Fig. 1 | A centromere-inactivation and chromosome-selection system (CEN-SELECT) identifies cells harboring previously mis-segregated chromosomes. a, Overview of the CEN-SELECT approach, which combines a CENP-A replacement strategy to induce Y chromosome mis-segregation and micronucleus-mediated shattering with a CRISPR–Cas9-integrated neomycin-resistance gene (Neo^R). **b–d,** Engineered DLD-1 cells carrying a CENP-A^{WT} or CENP-A^{C-H3} rescue gene were treated with DOX/IAA and quantified by interphase FISH for LOY (**b**), the compartment of the Y chromosome (chrY) in chrY-positive cells (**c**), and the proportion of cells with micronuclei (**d**). Pie chart shows the fraction of micronuclei carrying a chrY-positive or chrY-negative micronucleus from CENP-A^{C-H3} rescued cells treated with DOX/IAA. Data represent the mean of $n = 2$ independent experiments from 1,081–2,085 cells (**b,d**) or 861–1,242 Y chromosomes (**c**). Scale bar, 5 μm. **e,** Experimental schematic for **f–i**. **f,** Representative colony formation plate scans. **g,** Quantification of **f**. Bar graph represents the mean ± s.e.m. of $n = 9$ biological replicates pooled from 3 independent experiments; P value derived from two-tailed Student's t-test compared with untreated cells. **h,** Representative interphase FISH images of CENP-A^{C-H3} rescue cells. **i,** Quantification of **h**. Scale bar, 10 μm. Bar graph represents the mean ± s.e.m. of $n = 3$ independent experiments from 833–918 cells; P value derived from two-tailed Student's t-test compared with untreated cells. Pie chart represents the fraction of Y chromosomes compartmentalized within the nucleus or micronucleus following DOX/IAA treatment and G418 selection.

Rearrangements recapitulate a diverse structural variant landscape.

To visualize the types of structural rearrangements generated by CEN-SELECT, we designed a metaphase FISH-based assay utilizing two DNA probe sets that distinctly labeled both halves of the Y chromosome (Fig. 2b). A set of green paint probes hybridized to the euchromatic portion of the male-specific region of the Y (MSY) spanning Yp and proximal Yq[56], and another set of red probes (41H10) hybridized to the repetitive, heterochromatic array on distal Yq (YqH). These probe sets produced a characteristic morphology when hybridized to control metaphases, enabling routine discrimination between normal (Fig. 2c) and derivative Y chromosomes.

Cytogenetic inspection of 200 Y chromosomes exhibiting microscopic abnormalities revealed seven major categories of diverse inter- and intrachromosomal structural alterations (Figs. 2d and 3a), including end fusions, translocations, insertions, complex rearrangements, inversions (Supplementary Fig. 5), isodicentrics, and deletions (see Supplementary Note for detailed descriptions of each
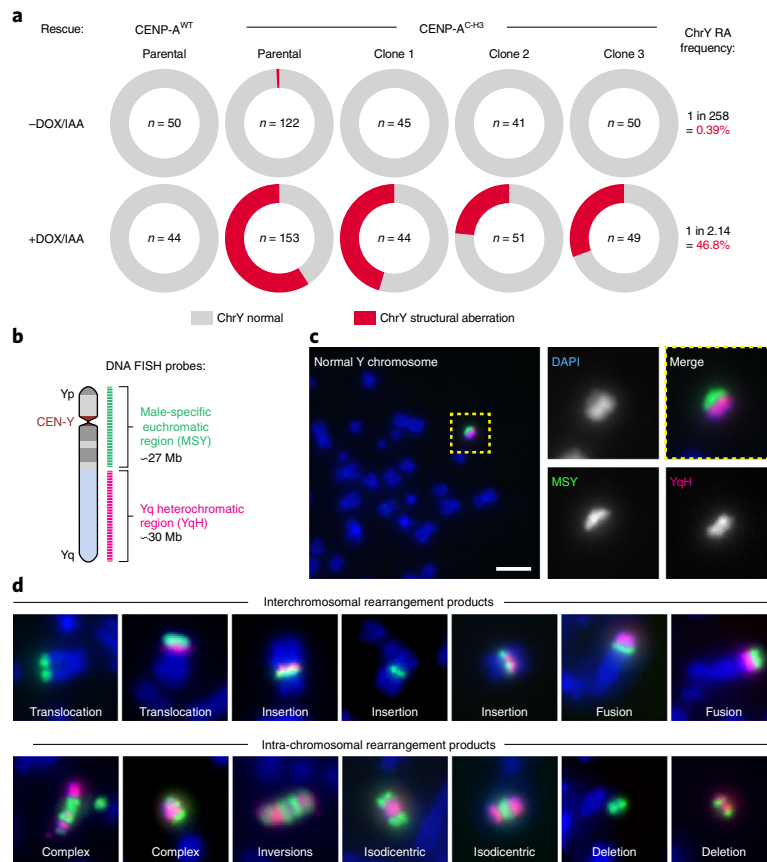
Fig. 2 | Mis-segregated chromosomes acquire a broad spectrum of structural genomic rearrangement types. a, Measurements of Y chromosome rearrangement frequencies in parental cells and 3 independent clonal lines following 0-d or 3-d DOX/IAA treatment and G418 selection. Metaphase spreads were subjected to DNA FISH using Y chromosome–specific paint probes; n = number of metaphase spreads examined. Parental CENP-A[C-H3] frequencies were pooled from three independent experiments. RA, rearrangement. **b**, Schematic of multicolored DNA FISH probes used to characterize structural anomalies of the Y chromosome. **c,d**, Representative FISH images of a normal Y chromosome without visible defects from control metaphases (scale bar, 5 μm) (**c**) and examples of derivative Y chromosomes from 3-d DOX/IAA-treated, G418-resistant metaphases (**d**). See Supplementary Note for a description of each rearrangement type.

type). Derivative chromosomes consisting of multiple rearrangement types were also observed, as exemplified by seven cases of complex aberrations combined with translocations (Fig. 3a).

To determine whether chronic centromere inactivation would influence the spectrum of rearrangement types, we continuously passaged cells with G418 alone or under CEN-SELECT conditions (Supplementary Fig. 6a). Whereas cell-doubling rates remained steady under G418 selection, CEN-SELECT cells displayed reduced growth rates during early passages (Supplementary Fig. 6b,c), reflecting competition between proliferating Y chromosome–positive cells and LOY cells undergoing cell death. Proliferation rate accelerated after p4, approaching baseline growth as Y chromosome–positive cells began outcompeting LOY cells (Supplementary Fig. 6c,d). FISH analysis revealed a passage-dependent reduction in Y chromosome fragmentation with a corresponding enrichment in Y chromosome rearrangements that eventually relocated to the nucleus (Supplementary Fig. 6d,e). At the sixth passage, >97% of metaphases (n = 43) harbored one to three copies of a normal or rearranged Y chromosome. Cytogenetic profiling of 201 derivative chromosomes obtained from this final passage detected heterogeneous rearrangement types similar to transient CEN-SELECT (Fig. 3a), but with a notable shift towards more interchromosomal rearrangements (Fig. 3b). A possible ring chromosome was also detected (1 of 401 derivative chromosomes), as well as another uncharacterized rearrangement (Supplementary Fig. 6f). Thus, mitotic chromosome segregation errors trigger a wide spectrum of rearrangements that frequently recapitulate cancer-associated structural alterations.

Rearrangement frequency and specificity through NHEJ repair.
We next compared rearrangements generated by CEN-SELECT with two established sources of structural aberrations: DNA damage induced by ionizing radiation (Supplementary Fig. 7a,b) and random chromosome segregation errors (Supplementary Fig. 7c,d) caused by spindle assembly checkpoint inactivation[27,57]. Following the G418 selection procedure, frequencies of de novo structural abnormalities affecting four autosomes or both sex chromosomes were measured from 1,968 metaphases (Supplementary Fig. 7e-g). Ionizing radiation caused an expected dose- and chromosome-length-dependent increase in rearrangements (Fig. 4a and Supplementary Fig. 7h), whereas spindle assembly checkpoint inactivation produced modest levels of structural aberrations that weakly correlated with chromosome size (Fig. 4a). In contrast, CEN-SELECT generated Y chromosome–specific rearrangements at a frequency comparable to that of rearrangements on a single autosome from sublethal ionizing radiation doses (Fig. 4a), demonstrating that CEN-SELECT achieves efficient rearrangements specifically on chromosomes undergoing increased mis-segregation rates.

With the exception of terminal deletions, the observed rearrangement products required at least one DNA joining event between two broken chromosome ends. To test the involvement of the two distinct DSB end-joining pathways (classical non-homologous end joining (NHEJ) and alternative end joining (alt-EJ)), each was transiently suppressed by RNA-interference-mediated codepletion of LIG4/DNA-PK or LIG3/PARP1, respectively[26] (Fig. 4b).
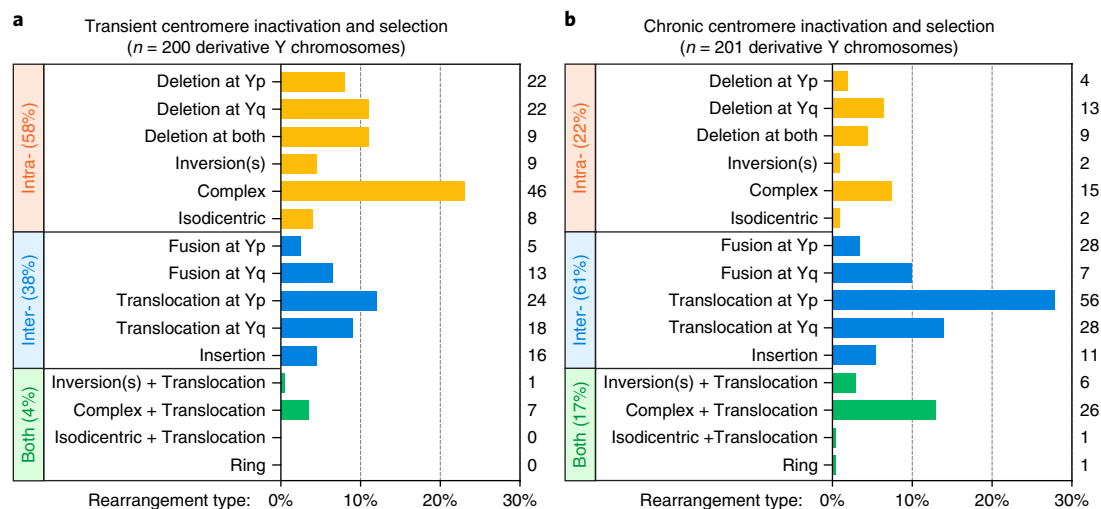
Fig. 3 | Systematic classification of the structural rearrangement landscape. a,b, The distribution of structural rearrangement types quantified from metaphase spreads by using MSY/YqH FISH probes following transient centromere inactivation induced by 3-d DOX/IAA treatment, washout, and G418 selection (**a**), or prolonged centromere inactivation induced by continuous passage in DOX/IAA and G418 (detailed in Supplementary Fig. 6) (**b**). The number of each case detected is depicted on the right of each graph.

Cytogenetic assessment revealed that inhibition of NHEJ, but not alt-EJ, markedly reduced Y chromosome rearrangement frequencies in the majority of G418-resistant cells (Fig. 4c)—94% of which harbored a Y chromosome without visible defects (149 of 159 metaphases). A complementary approach confirmed this finding, with NHEJ-deficiency conferring clonogenic sensitivity to G418 selection (Fig. 4d), consistent with the failure to reassemble a functional *Neo^R*-encoding Y chromosome. These data support classical NHEJ as the major DSB-repair mechanism facilitating rearrangement formation on mis-segregated chromosomes in human cells.

Sequencing of stable and complex genomic rearrangement events. The structural rearrangement landscape of derivative chromosomes produced by CEN-SELECT was next determined at high resolution by WGS. Single-cell-derived subclones were isolated (Fig. 5a) from sustained CEN-SELECT populations that harbored a rearranged Y chromosome within the main nucleus (Supplementary Fig. 6d–e). Metaphase FISH revealed that 22 of 40 clones carried distinct, clonal Y chromosome rearrangements (Fig. 5b) spanning the diverse spectrum observed in the bulk population (Fig. 3b). Five clones had interchromosomal translocations involving the Y with complex patterns of intrachromosomal rearrangements exhibiting overlapping FISH signals between the MSY/YqH regions (Fig. 5c). Metaphase FISH determined that >95% of cells within each clone propagated the same derivative chromosome even after long-term continuous passage without G418 (Fig. 5d,e), demonstrating that, in the absence of selective pressure, these chromosomes were stably heritable without further changes.

To identify de novo structural variants and DNA copy-number alterations, the parental DLD-1 clone and 20 derived clones (Fig. 5b) were subjected to paired-end WGS (35.7× mean depth). A control clone carrying a normal Y chromosome (PD37303a) is shown in Fig. 6a, and a summary of cytogenetics and WGS data for each clone is provided in Supplementary Table 1. The frequency of Y chromosome–specific structural variants per Mb sequence was enriched ~21-fold compared with the mean of individual autosomes, and the total number of structural variants otherwise scaled by chromosome size (Supplementary Fig. 8a,b). We note that these measurements probably underestimate the number of Y chromosome structural variants, since rearrangements involving reference assembly gaps, including the repetitive YqH region, were excluded from structural

variant analyses, although such junctions were evident by MSY copy-number changes without an associated structural variant call. The rearrangement partner was identified by WGS for seven clones—each validated by metaphase FISH (Supplementary Fig. 9).

Reconstructing 71 Y-specific intrachromosomal rearrangement breakpoints revealed that the majority of junctions lacked microhomology beyond a frequency expected by chance from random simulations (Supplementary Fig. 8c and see Methods). Among these junctions, seven contained 3 base pairs (bp) of microhomology, and two had ≥4 bp (Supplementary Fig. 8d), consistent with predominant NHEJ-dependent ligation with minimal alt-EJ involvement. Junctions with ≥2 bp of non-templated insertions were also observed within 32% of junctions (23 of 71) with a median length of 13.5 bp and maximum length of 94 bp (Supplementary Fig. 8d). Junction sequences with microhomology tracts and long insertions are provided in Supplementary Table 2. Additionally, measurements of intra-DSB distance between consecutive breakpoints indicated that a majority (76%, $n = 228$) of chromosome fragments were <1 Mb in size with a median length of 300 kb (Supplementary Fig. 8e).

Chromothripsis coupled to interchromosomal rearrangements. WGS identified 13 clones with intrachromosomal rearrangements of the Y chromosome, at least nine of which harbored the canonical signatures of chromothripsis (Fig. 6b–d and Supplementary Fig. 10), including oscillating copy-number states and clustered breakpoint patterns randomly distributed across the four possible orientations (Supplementary Fig. 8a). Similar complex rearrangements were not observed on other chromosomes. Super-resolution visualization of metaphase chromosomes revealed that complex rearrangements produced intricately linked FISH patterns with interspersed euchromatic and heterochromatic segments (Fig. 6b). Copy-number profiling further identified short deletions interspersed along the chromosome, a hallmark of chromosome shattering followed by the loss of DNA fragments, and copy-number increases at multiple loci (Fig. 6b–d and Supplementary Fig. 10), indicative of DNA replication errors and/or cofragmentation of sister chromatids. Oscillations across three, or sometimes more, copy-number levels were also observed, patterns suggesting the simultaneous involvement of both sister chromatids in the chromothriptic event[58]. Indeed, as previously shown[33], centromere-inactivated chromosomes occasionally
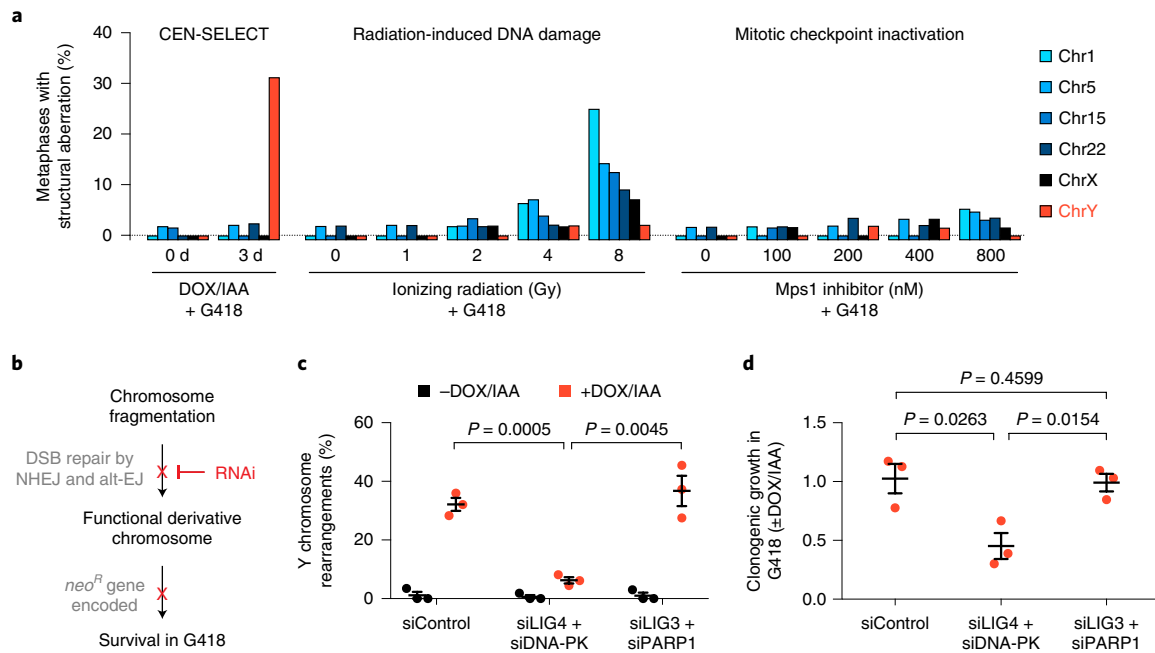
Fig. 4 | Chromosome rearrangements develop with high frequency and specificity through classical NHEJ repair. a, DLD-1 cells carrying a CENP-A^{C-H3} rescue gene were treated as indicated in Supplementary Fig. 7e, and this was followed by metaphase spread preparation and hybridization to the indicated chromosome paint probes. Metaphases were examined for structural rearrangements affecting each chromosome. Bar graph represents $n = 42$–65 metaphases scored per chromosome per condition, analyzing a total of 1,968 metaphase spreads (exact sample sizes provided in Supplementary Fig. 7g). **b**, Schematic of experimental hypothesis tested in **c** and **d**. **c,d**, DLD-1 CENP-A^{C-H3} rescue cells were treated with or without 3-d DOX/IAA and transfected with the indicated siRNAs simultaneously with DOX/IAA washout for an additional 3 d. Cells were then replated into G418 medium for 10-d selection followed by metaphase FISH using MSY/YqH probes (102–159 metaphase spreads per condition) (**c**) or 14 d at single-cell density for colony formation assays (**d**). Data in **c** and **d** represent the mean ± s.e.m. of $n = 3$ independent experiments; P values were derived from two-tailed Student's t-test comparing groups as indicated. RNAi, RNA interference.

underwent non-dysjunction during mitosis and entrapped both chromatids into the same micronucleus.

Sequencing of two distinct clones resolved independent simple or complex events involving interstitial insertion of Y chromosome fragments into different chromosome 1 arms (Fig. 6e,f). FISH analysis revealed that one (PD37306a) had distinct separation between Y and chromosome 1p sequences, as indicated by an apparent gap using a chromosome 1 painting probe (Fig. 6g). By contrast, another (PD37313a) carried colocalized chromosome 1 and MSY FISH signals (Fig. 6h). PD37313a was determined by WGS to contain a series of shuffled Y chromosome sequences juxtaposed between a de novo duplicated segment located on chromosome 1q (Fig. 6f), consistent with chromothriptic fragments inserted into a tandem duplication site, which could arise from stalled replication forks[59].

Hypermutation hotspots were also observed near five chromothriptic junctions (Supplementary Fig. 8f), indicative of kataegis co-occurring with chromosomal rearrangements[60]. These regions comprised strictly C>T mutations with an intermutation distance of <1 kb. Kataegic events occurred in proximity to Y–Y breakpoints associated with a structural variant call (130 bp to 6.5 kb) or binned copy-number change (61–109 kb) and were otherwise not observed elsewhere throughout the genome. Interphase DNA bridges[61] containing Y chromosome sequences were undetectable across all CEN-SELECT passages (Supplementary Fig. 6d), supporting micronucleus-mediated origins underlying these kataegis signatures.

We further identified complex Y chromosome rearrangements from an independent cohort of 3 of 18 male osteosarcomas interrogated by WGS[62], including chromothripsis in a 10-year-old child (PD7193a) (Supplementary Fig. 11a-c). The architectures of these rearrangements resembled those generated by CEN-SELECT,

including complex intra- and interchromosomal rearrangements accompanied by oscillating copy-number patterns. Similar molecular features were also observed at reconstructed breakpoint junctions (Supplementary Fig. 11d–f). These examples provide proof-of-principle evidence that such rearrangements occur in the context of human cancer, although the degree to which mitotic errors contribute to somatic rearrangements in different cancer types warrants further investigation.

Consequences of chromothriptic DNA copy-number alterations. We next examined how structural rearrangements provoked gene inactivation by mapping gene copy-number states along the Y chromosome. Analysis of 213 well-mapped Y-linked genes and pseudogenes revealed widespread deletions scattered across regions of the initially mis-segregated chromosome, including five clones with >65% gene loss and eight harboring less frequent (0.5–37%) deletions (Fig. 7a,b). The genomic loci of two Y-encoded genes broadly expressed in somatic cells, *KDM5D* and *EIF1AY*, were inspected at greater resolution. Among the 20 sequenced clones, five carried a structural deletion in either gene, and three exhibited codeletions in both (Supplementary Fig. 12a,b). Reverse-transcription PCR (RT–PCR) analysis verified that these genes were not expressed in the corresponding clones (Supplementary Fig. 12c) and provided further evidence that these derivative chromosomes were transcriptionally functional beyond Neo^R expression. Thus, in contrast to whole-chromosome aneuploidy, which alters gene-dosage across the mis-segregated chromosome, complex rearrangements arising from chromosome mis-segregation delete a subset of interspersed genes through copy-number alterations (Fig. 7c).

Clone PD37310a was resistant to G418 selection despite lacking a cytogenetically detectable Y chromosome when hybridized
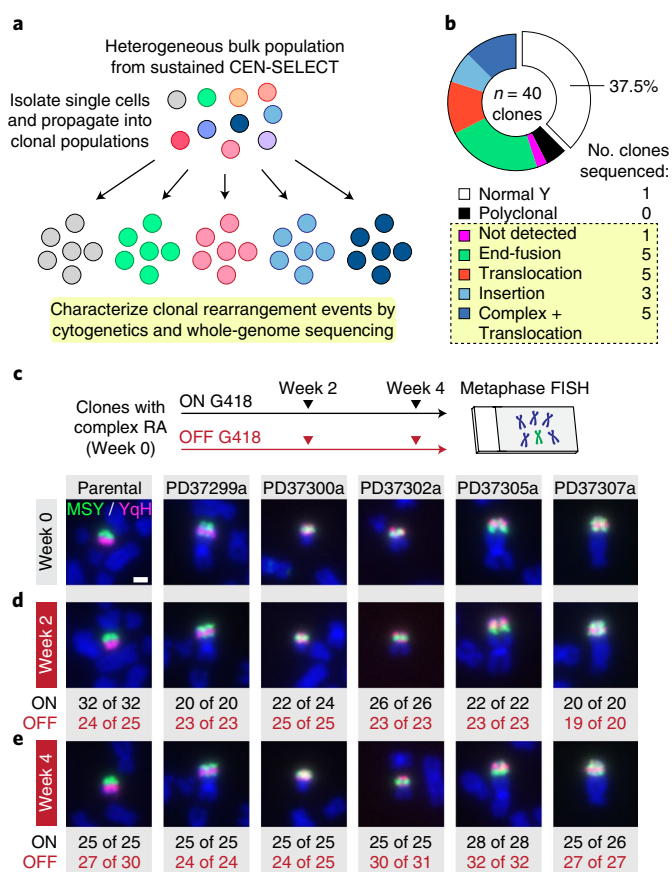
Fig. 5 | Isolation and propagation of single-cell-derived clones with genetically stable derivative chromosomes. a, Schematic of approach used to generate clonally propagated Y chromosome rearrangements from a bulk cell population. **b**, Frequencies of Y chromosome rearrangement types obtained from single-cell-derived clones. The boxed section indicates clonal rearrangements, and the number of clones subjected to WGS is shown on the right. **c,d**, Experimental schematic and representative metaphase FISH images from the indicated clones, which were passaged in parallel cultures with (ON) or without (OFF) G418 selection for 0 weeks (**c**), 2 weeks (**d**), and 4 weeks (**e**). Scale bar, 2 μm. Values below the image represent the number of metaphases positive for the depicted derivative chromosome over the total number of metaphases examined.

to painting probes. WGS revealed that the majority of the Y was lost except for an ~0.9-Mb segment surrounding the Neo^R marker at Yq11.221 (Fig. 7d). Cytogenetics with FISH probes targeting this locus (RP11–113K10) showed that this region spatially manifested as small fragments of extrachromosomal DNA (ecDNA) (Fig. 7e) reminiscent of cancer-associated double-minute chromosomes[63,64]. PD37310a cells, but not parental controls, heterogeneously maintained multiple copies of this sequence (Fig. 7f), demonstrating that chromosome shattering induced by mitotic errors could elicit ecDNA production—consistent with observations in chromothriptic tumors[21,65]. The head-to-head and tail-to-tail orientation of the two rearrangement junctions (Fig. 7d) suggested that the ecDNA structure was produced from three DSBs that generated two fragments (509 kb and 385 kb in length), followed by an inversion of one fragment and subsequent re-ligation between the four available ends (Fig. 7g,h).

Discussion

Here we show that chromosomal instability can initiate stable and diverse structural rearrangements that span the spectrum of

complexity observed in cancers and genomic disorders, expanding the catalog of possible genomic rearrangement types caused by cell-division errors. Mis-segregated chromosomes are frequently encapsulated into micronuclei that assemble dysfunctional nuclear envelopes[31,66]. The ensuing loss of nucleocytoplasmic compartmentalization throughout interphase triggers micronucleus-specific DSBs that persist unrepaired into mitosis. These DSBs subsequently resolve into highly fragmented chromatin during mitotic entry[30,32,33] that reintegrates into daughter cell genomes after mitotic exit[30,34].

Following the completion of mitosis, inherited chromosome fragments are deposited into a functional G1-phase nucleus to initiate the reassembly process by activating the DNA-damage response. Because NHEJ is temporally active during early cell-cycle phases, we propose that the NHEJ machinery engages these fragments for re-ligation, triggering both intra- and interchromosomal rearrangements depending on the availability and/or spatial proximity of compatible DNA ends. We show that inhibiting NHEJ suppresses chromosome rearrangement formation (Fig. 4c), and the breakpoint junctions of re-ligated fragments do not utilize sequence homology beyond that expected from random chance (Supplementary Fig. 8c,d). Microhomology-mediated repair through alt-EJ, however, may be involved in processing remaining DSBs that persist into S/G2 phases when end-resection is operative[67]. Whereas recombination-dependent DSB repair preserves sequence integrity at breakpoint junctions, end joining-mediated repair can generate mutagenic indels. Although the mechanisms linking rearrangements with kataegis remain unclear, our data suggest that some breakpoints (Supplementary Fig. 8f) are probably reassembled through an intermediate step consisting of exposed single-stranded DNA processed by APOBEC enzymes[68].

Use of CEN-SELECT has generated functional and genetically stable chromosomes that have undergone chromothripsis and that persist over multiple cell cycles even without continued selection. Importantly, apart from the requirement during selection to retain Neo^R, the sparse gene content and lack of essential genes on the Y chromosome eliminate constraints on the possible 'solution space' of structural variant configurations that can arise, allowing for an unbiased view of the rearrangement landscape driven by mitotic errors. For instance, complex reassembly events retained multiple segments of Y sequences despite no selective pressure for its maintenance (Fig. 6 and Supplementary Figure 10). At a minimum, the reconstitution of a stable derivative chromosome requires the re-stitching of fragments that include an epigenetically active centromere with two end-capping telomeres. Because the majority of DNA fragments produced from micronucleus-mediated chromosome shattering are acentric[33], retention of any essential chromosomal element(s) may be facilitated through rearrangements with a normally segregating, non-homologous chromosome harboring a spontaneous DSB, thereby reacquiring the capacity for its faithful inheritance.

Here we exploited unique features of the Y chromosome to reconstruct the mechanisms and consequences of genomic aberrations[69]. Although a contribution from underlying experimental factors (Y-specific features, resistance gene position) or genetic background (*TP53* status, mismatch-repair deficiency) cannot be formally excluded, we anticipate that the observations reported here are generalizable across chromosomes. We provide evidence that chromothriptic rearrangements can focally delete genes located on mis-segregated chromosomes (Fig. 7a,b), findings with implications for tumor suppressor gene loss in cancer. Such deletions probably arise through the unequal distribution of DNA fragments during mitosis[34] and from inherited fragments that either failed to participate in the chromosome rearrangement event (Fig. 7c) or stochastically reassembled into an unstable configuration. Conversely, copy-number increases can be generated through DNA replication
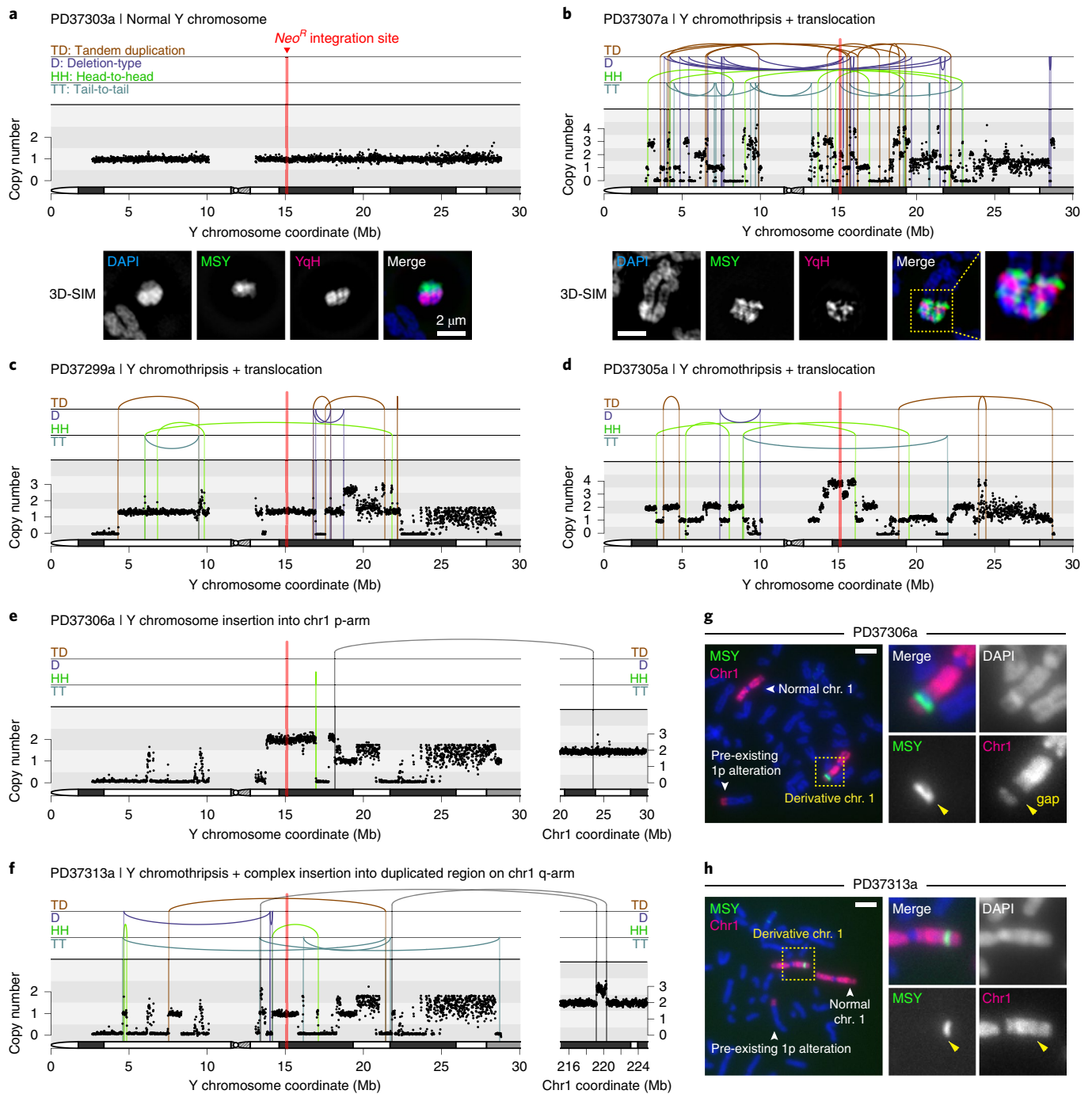
Fig. 6 | Whole-genome sequencing reveals complex rearrangements that include the hallmark signatures of chromothripsis. a,b, DNA copy-number profiles from WGS (top) and representative metaphase FISH images hybridized to the indicated probes captured by super-resolution microscopy (bottom) from clone PD37303a with a normal Y chromosome (**a**) and clone PD37307a with 83 breakpoints detected across the mappable Y chromosome region (**b**). 3D-SIM, three-dimensional structured illumination microscopy. **c–f,** DNA copy-number profiles of additional clones carrying a chromothriptic Y chromosome coupled to translocations (**c,d**), a simple insertion into chromosome 1p (**e**), or a complex insertion at a duplicated region on chromosome 1q (**f**). The x axes of Y chromosome plots are clipped at 30 Mb to exclude the Yq heterochromatic region. **g,h,** Metaphase FISH images using MSY and chromosome 1 painting probes on clones PD37306a (**g**) and PD37313a (**h**). Scale bar, 5 μm.

errors (for example, template switching) or, in more extreme cases, the ligation of fragment ends to produce autonomously replicating, circular ecDNA structures lacking centromeres and telomeres, which act as a frequent mode of oncogene amplification in cancer[63,64]. Indeed, the ecDNA carried by clone PD37310a co-amplified 14 additional genes within the vicinity of the selection marker (Fig. 7h). The short fragments generated by chromosome shattering

(Supplementary Fig. 8e) may be conducive to ecDNA production to drive gene amplification events associated with tumorigenesis and/ or chemotherapeutic resistance.

Rearrangements arising from chromosome shattering are likely to be frequent in chromosomally unstable cells exhibiting high rates of chromosome segregation errors, especially those deficient in the p53 checkpoint[28,29,65]. The majority of random chromosome
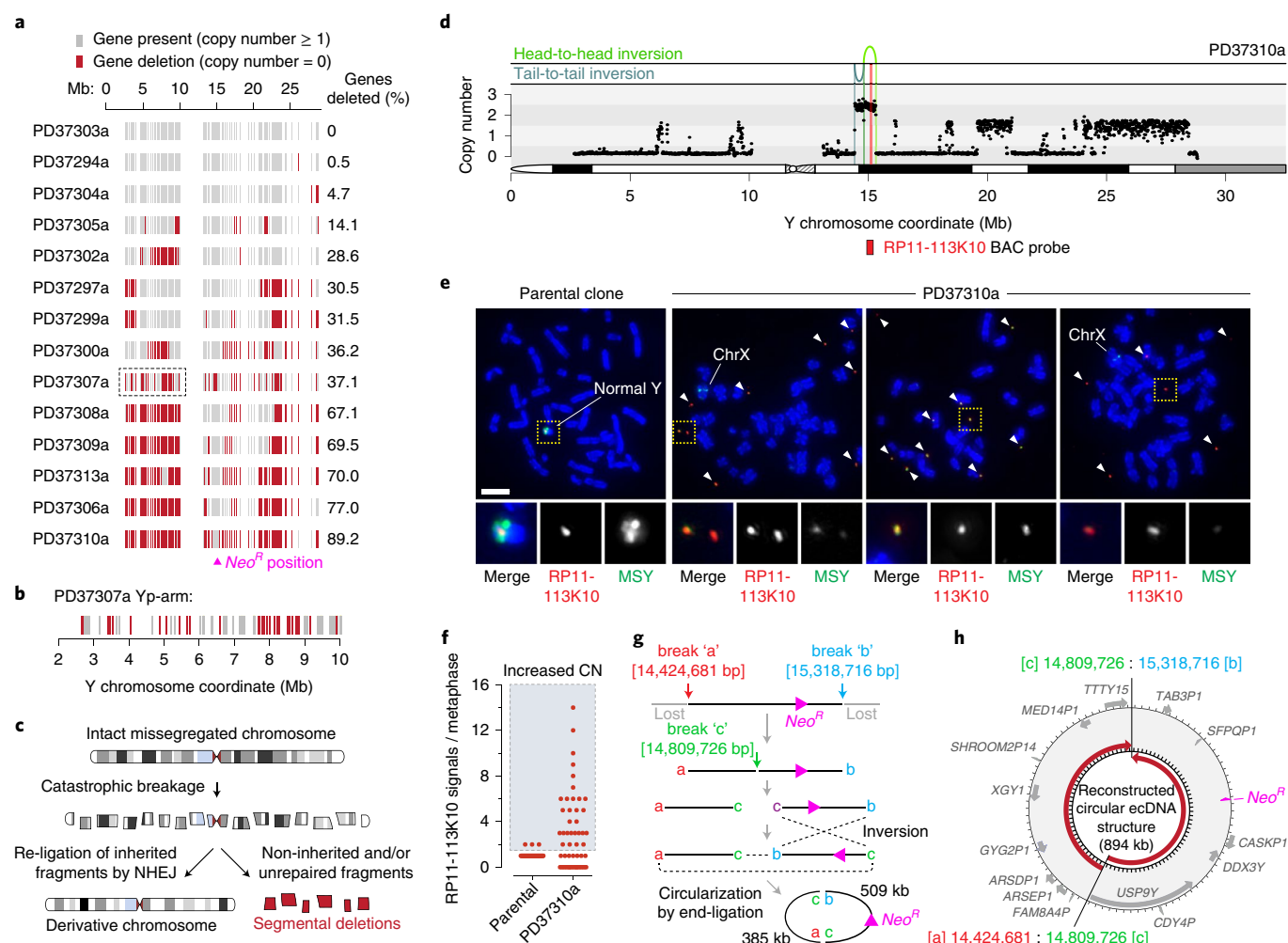
Fig. 7 | Gene disruption and extrachromosomal DNA amplification from chromosome mis-segregation-induced rearrangements. a, Each gray vertical line represents an individual gene or pseudogene depicted at its chromosomal start position, and red lines represent a copy number of zero. Clones are ranked from fewest to most gene deletions. **b**, Magnification of clone PD37307a (boxed region in **a**) exhibiting oscillating patterns of gene retention and deletion within an 8-Mb segment. **c**, Schematic of chromosome shattering and reassembly events resulting in a derivative chromosome harboring rearrangements with interspersed deletions. **d**, DNA copy-number profile of clone PD37310a, showing extensive LOY except for the region harboring the selection marker accompanied by two inversions. **e**, Images of metaphase spreads prepared from the parental or PD37310a clone and hybridized to MSY (green) and RP11–113K10 BAC (red) probes recognizing the region shown in **d**. Arrows denote ecDNA fragments hybridizing to the RP11–113K10 probe, and regions of the X chromosome partially hybridize to MSY probes, due to X–Y sequence homology. Scale bar, 5 μm. **f**, Quantification of **e**. Each data point represents an individual metaphase spread derived from the parental clone ($n = 48$) or PD37310a clone ($n = 56$). CN, copy number. **g**, Schematic depicting the predicted steps leading to the generation of the ecDNA element through the circular reassembly of two broken DNA fragments. **h**, Reconstructed ecDNA sequence from WGS. Genes and pseudogenes in the corresponding region are shown, and red arrows depict 5′ to 3′ orientation.

products formed from these events, however, are probably deleterious or confer no positive advantage. In rarer instances, chromosome shattering may trigger a series of rearrangements in which a 'perfect storm' of genetic lesions is created through DNA copy-number alterations, disruptive breakpoints, and/or mutations to enable clonal and oncogenic selective advantages[70]. Such catastrophic events can operate as a punctuated mutational mechanism to rapidly remodel the cancer genome[26,71].

URLs. CRISPR Design Tool, http://crispr.mit.edu; BACPAC Resources Program, CHORI, https://bacpacresources.org.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0360-8.

References
1. Campbell, P. J. et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat. Genet.* **40**, 722–729 (2008).
2. Stephens, P. J. et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **462**, 1005–1010 (2009).
3. Carvalho, C. M. & Lupski, J. R. Mechanisms underlying structural variant formation in genomic disorders. *Nat. Rev. Genet.* **17**, 224–238 (2016).
4. Liu, P. et al. An organismal CNV mutator phenotype restricted to early human development. *Cell* **168**, 830–842.e7 (2017).
5. Redin, C. et al. The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nat. Genet.* **49**, 36–45 (2017).
6. Bass, A. J. et al. Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. *Nat. Genet.* **43**, 964–968 (2011).
7. Yang, L. et al. Diverse mechanisms of somatic structural variations in human cancer genomes. *Cell* **153**, 919–929 (2013).

8. Baca, S. C. et al. Punctuated evolution of prostate cancer genomes. *Cell* **153**, 666–677 (2013).

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

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

11. Cogen, P. H., Daneshvar, L., Metzger, A. K. & Edwards, M. S. Deletion mapping of the medulloblastoma locus on chromosome 17p. *Genomics* **8**, 279–285 (1990).

12. Nowell, P. C. The minute chromosome (Phl) in chronic granulocytic leukemia. *Blut* **8**, 65–66 (1962).

13. Rowley, J. D. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* **243**, 290–293 (1973).

14. de Klein, A. et al. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. *Nature* **300**, 765–767 (1982).

15. Gao, Q. et al. Driver fusions and their implications in the development and treatment of human cancers. *Cell Rep.* **23**, 227–238 e3 (2018).

16. Northcott, P. A. et al. Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature* **511**, 428–434 (2014).

17. Zhang, X. et al. Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. *Nat. Genet.* **48**, 176–182 (2016).

18. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).

19. Weischenfeldt, J. et al. Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. *Nat. Genet.* **49**, 65–74 (2017).

20. Dixon, J. R. et al. Integrative detection and analysis of structural variation in cancer genomes. *Nat. Genet.* **50**, 1388–1398 (2018).

21. Stephens, P. J. et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011).

22. Lombard, D. B. et al. DNA repair, genome stability, and aging. *Cell* **120**, 497–512 (2005).

23. Lin, C. et al. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* **139**, 1069–1083 (2009).

24. Henssen, A. G. et al. PGBD5 promotes site-specific oncogenic mutations in human tumors. *Nat. Genet.* **49**, 1005–1014 (2017).

25. Garaycoechea, J. I. et al. Alcohol and endogenous aldehydes damage chromosomes and mutate stem cells. *Nature* **553**, 171–177 (2018).

26. Ly, P. & Cleveland, D. W. Rebuilding chromosomes after catastrophe: emerging mechanisms of chromothripsis. *Trends Cell Biol.* **27**, 917–930 (2017).

27. Janssen, A., van der Burg, M., Szuhai, K., Kops, G. J. & Medema, R. H. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. *Science* **333**, 1895–1898 (2011).

28. Soto, M. et al. p53 prohibits propagation of chromosome segregation errors that produce structural aneuploidies. *Cell Rep.* **19**, 2423–2431 (2017).

29. Santaguida, S. et al. Chromosome mis-segregation generates cell-cycle-arrested cells with complex karyotypes that are eliminated by the immune system. *Dev. Cell* **41**, 638–651.e5 (2017).

30. Crasta, K. et al. DNA breaks and chromosome pulverization from errors in mitosis. *Nature* **482**, 53–58 (2012).

31. Hatch, E. M., Fischer, A. H., Deerinck, T. J. & Hetzer, M. W. Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell* **154**, 47–60 (2013).

32. Kato, H. & Sandberg, A. A. Chromosome pulverization in human cells with micronuclei. *J. Natl Cancer Inst.* **40**, 165–179 (1968).

33. Ly, P. et al. Selective Y centromere inactivation triggers chromosome shattering in micronuclei and repair by non-homologous end joining. *Nat. Cell Biol.* **19**, 68–75 (2017).

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

35. Garsed, D. W. et al. The architecture and evolution of cancer neochromosomes. *Cancer Cell* **26**, 653–667 (2014).

36. 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).

37. Cortés-Ciriano, I. et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. Preprint at https://www.biorxiv.org/content/early/2018/05/30/333617 (2018).

38. Kloosterman, W. P. et al. Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum. Mol. Genet.* **20**, 1916–1924 (2011).

39. Kloosterman, W. P. et al. Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms. *Cell Rep.* **1**, 648–655 (2012).

40. Chiang, C. et al. Complex reorganization and predominant non-homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. *Nat. Genet.* **44**, 390–397 (2012).

41. Weckselblatt, B., Hermetz, K. E. & Rudd, M. K. Unbalanced translocations arise from diverse mutational mechanisms including chromothripsis. *Genome Res.* **25**, 937–947 (2015).

42. Cretu Stancu, M. et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nat. Commun.* **8**, 1326 (2017).

43. Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. & Kanemaki, M. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* **6**, 917–922 (2009).

44. Holland, A. J., Fachinetti, D., Han, J. S. & Cleveland, D. W. Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *Proc. Natl Acad. Sci. USA* **109**, E3350–E3357 (2012).

45. Sun, C. et al. An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. *Nat. Genet.* **23**, 429–432 (1999).

46. Masumoto, H., Masukata, H., Muro, Y., Nozaki, N. & Okazaki, T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell. Biol.* **109**, 1963–1973 (1989).

47. Burma, S., Chen, B. P., Murphy, M., Kurimasa, A. & Chen, D. J. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J. Biol. Chem.* **276**, 42462–42467 (2001).

48. Zhao, S. et al. Functional link between ataxia-telangiectasia and Nijmegen breakage syndrome gene products. *Nature* **405**, 473–477 (2000).

49. Wu, X. et al. ATM phosphorylation of Nijmegen breakage syndrome protein is required in a DNA damage response. *Nature* **405**, 477–482 (2000).

50. Schultz, L. B., Chehab, N. H., Malikzay, A. & Halazonetis, T. D. p53 binding protein 1 (53BP1) is an early participant in the cellular response to DNA double-strand breaks. *J. Cell Biol.* **151**, 1381–1390 (2000).

51. Anderson, L., Henderson, C. & Adachi, Y. Phosphorylation and rapid relocalization of 53BP1 to nuclear foci upon DNA damage. *Mol. Cell. Biol.* **21**, 1719–1729 (2001).

52. Terradas, M., Martin, M., Tusell, L. & Genesca, A. DNA lesions sequestered in micronuclei induce a local defective-damage response. *DNA Repair* **8**, 1225–1234 (2009).

53. Miake-Lye, R. & Kirschner, M. W. Induction of early mitotic events in a cell-free system. *Cell* **41**, 165–175 (1985).

54. Muhlhausser, P. & Kutay, U. An in vitro nuclear disassembly system reveals a role for the RanGTPase system and microtubule-dependent steps in nuclear envelope breakdown. *J. Cell Biol.* **178**, 595–610 (2007).

55. Soto, M., Garcia-Santisteban, I., Krenning, L., Medema, R. H. & Raaijmakers, J. A. Chromosomes trapped in micronuclei are liable to segregation errors. *J. Cell Sci.* **131**, jcs214742 (2018).

56. Skaletsky, H. et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**, 825–837 (2003).

57. Santaguida, S., Tighe, A., D'Alise, A. M., Taylor, S. S. & Musacchio, A. Dissecting the role of MPS1 in chromosome biorientation and the spindle checkpoint through the small molecule inhibitor reversine. *J. Cell Biol.* **190**, 73–87 (2010).

58. Li, Y. et al. Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. *Nature* **508**, 98–102 (2014).

59. Willis, N. A. et al. Mechanism of tandem duplication formation in BRCA1-mutant cells. *Nature* **551**, 590–595 (2017).

60. Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).

61. Maciejowski, J., Li, Y., Bosco, N., Campbell, P. J. & de Lange, T. Chromothripsis and kataegis induced by telomere crisis. *Cell* **163**, 1641–1654 (2015).

62. Behjati, S. et al. Recurrent mutation of IGF signalling genes and distinct patterns of genomic rearrangement in osteosarcoma. *Nat. Commun.* **8**, 15936 (2017).

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

64. deCarvalho, A. C. et al. Discordant inheritance of chromosomal and extrachromosomal DNA elements contributes to dynamic disease evolution in glioblastoma. *Nat. Genet.* **50**, 708–717 (2018).

65. Rausch, T. et al. Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. *Cell* **148**, 59–71 (2012).

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

67. Truong, L. N. et al. Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc. Natl Acad. Sci. USA* **110**, 7720–7725 (2013).

68. Taylor, B. J. et al. DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. *eLife* **2**, e00534 (2013).

69. Ly, P. & Cleveland, D. W. Interrogating cell division errors using random and chromosome-specific missegregation approaches. *Cell Cycle* **16**, 1252–1258 (2017).

70. Mardin, B. R. et al. A cell-based model system links chromothripsis with hyperploidy. *Mol. Syst. Biol.* **11**, 828 (2015).
71. Leibowitz, M. L., Zhang, C. Z. & Pellman, D. Chromothripsis: a new mechanism for rapid karyotype evolution. *Annu. Rev. Genet.* **49**, 183–211 (2015).

Author contributions

P.L. and D.W.C. conceived the project, designed the experiments, and wrote the manuscript. P.L., O.S., and D.H.K. conducted the experiments and analyzed the data. S.F.B. and P.J.C. performed and analyzed the sequencing experiments. P.L. and W.L. prepared *Xenopus* egg extracts. T.P. and D.C.P. provided critical FISH reagents. S.B. and A.M.F. contributed osteosarcoma samples. All authors provided input on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0360-8.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to P.L. or D.W.C.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Cell culture and reagents. T-REx Flp-In DLD-1 cells were cultured in DMEM (Thermo Fisher) containing 10% tetracycline-free fetal bovine serum (FBS) (Omega Scientific) and 100 U ml⁻¹ penicillin–streptomycin. hTERT-immortalized RPE-1 cells were cultured in DMEM/F12 (Thermo Fisher) supplemented with 10% FBS, 100 U ml⁻¹ penicillin-streptomycin, 2 mM L-glutamine, and 0.348% sodium bicarbonate. All cells were maintained at 37 °C under 5% CO₂ and atmospheric oxygen.

DOX and IAA were purchased from Sigma, dissolved in cell culture-grade water, and used at 1 µg ml⁻¹ and 500 µM, respectively. Where indicated, DOX/IAA washout was performed by two washes in PBS followed by addition of growth medium. The Mps1 inhibitor reversine (Santa Cruz) was dissolved in dimethylsulfoxide and used at the indicated concentrations. Geneticin (G418 Sulfate) and hygromycin (Thermo Fisher Scientific) were used at selection concentrations of 300 µg ml⁻¹ and 200 µg ml⁻¹, respectively. DLD-1 short interfering RNA (siRNA) transfections were conducted with Lipofectamine RNAiMAX reagent (Thermo Fisher Scientific) using characterized siRNAs (GE Healthcare Dharmacon) as previously described[33].

Genome engineering and modifications. T-REx Flp-In DLD-1 cells, which were previously modified to carry a heterozygous CENP-A EYFP-AID fusion tag (CENP-A^EYFP-AID/−) and stably express osTIR1[33], were used as the starting material for all DLD-1 experiments. To integrate a selectable marker into the Y chromosome by CRISPR editing, target sequences for guide RNAs were selected using CRISPR Design Tool (see URLs). An oligonucleotide (5′-AACACTTCTCTAGCACGATT-3′) encoding a guide RNA targeting the Y chromosome AZFa locus (target coordinate 12,997,776 using GRCh38 reference assembly) was cloned into the Bbs1 restriction site of the pSpCas9(BB)-2A-GFP vector (PX458, Addgene 48138) and cotransfected into DLD-1 CENP-A^EYFP-AID/− cells with a synthesized pUC57-based donor vector containing an EF1α promoter driving the neomycin-resistance gene (Neo^R) flanked by 450-bp left and right AZFa homology arms (GenScript). After transfection using X-tremeGENE 9 DNA Transfection Reagent (Sigma-Aldrich), cells were selected with G418 for 20 d and plated by limiting dilution into 96-well plates. G418-resistant, single-cell-derived clones were expanded and screened by PCR for targeted insertion of Neo^R into the AZFa locus by using a forward primer located outside the left homology arm and a reverse primer located within the EF1α promoter. A forward primer sitting inside the left donor homology arm with the same reverse primer were used as a control, as well as primers amplifying the sY1250 STS. All PCR primer sequences are provided in Supplementary Table 3. Clones identified by PCR were confirmed to harbor a single copy of a morphologically normal-appearing Y chromosome by using DNA FISH on metaphase spreads.

After Y chromosome editing, CENP-A wild type or C–H3 rescue constructs were integrated into the Flp-In locus, as previously described[33]. Briefly, cells were cotransfected with pOG44 and pcDNA5 constructs carrying the indicated CENP-A rescue gene by using X-tremeGENE 9 DNA Transfection Reagent (Sigma-Aldrich), and successfully targeted cells with correct integration were selected with hygromycin.

To generate cells with fluorescently labeled nuclear envelopes, we transduced DLD-1 and RPE-1 H2B–mRFP cells with retroviruses packaged in 293GP cells to stably express pBABE-puro-GFP-wt-lamin A (Addgene 17662). Fluorescent cells were isolated by flow cytometry on a BD FACSAria II Cell Sorter (University of California San Diego Human Embryonic Stem Cell Core Facility).

Cell growth measurements. For colony formation assays, trypsinized cells were pelleted, diluted as needed, and seeded at low density into 10-cm² or six-well plates for up to 2 weeks. Cells were fed with fresh growth medium each week. Cells were fixed with 100% methanol at room temperature, incubated in 0.5% crystal violet, 25% methanol solution with gentle agitation for 30 min, and washed with water. To account for the expected approximately tenfold differences in clonogenic growth in Fig. 4d, all DOX/IAA-treated conditions were seeded at a tenfold-higher cell density than control non-DOX/IAA-treated cells to yield a comparable number of colonies at time of fixation. For quantification of cell proliferation rates, 1 × 10⁶ cells were seeded into 10-cm² plates, and total cell numbers were counted using a hemocytometer every 5 d. Then, 1 × 10⁶ cells were reseeded into new 10-cm² plates, and the process repeated for up to 30 d.

Nuclear envelope disassembly assay using *Xenopus* egg extracts. Interphase and cytostatic factor (CSF)-arrested extracts were prepared from *Xenopus laevis* eggs as described[72]. Crude extracts were supplemented with 250 mM sucrose, and aliquots were snap frozen and stored at −80 °C. All animal protocols complied with ethical regulations set forth and approved by University of California San Diego and the Institutional Animal Care and Use Committee. For generation of RPE-1 cells with micronuclei by nocodazole washout, RPE-1 H2B–mRFP and lamin A–GFP cells were arrested in mitosis for 6 h with 100 ng ml⁻¹ nocodazole (Sigma), harvested by mitotic shake-off, and released into fresh growth medium overnight in four-well chamber slides. For nuclear envelope disassembly assays, a slightly modified protocol adapted from refs. [54] and [73] was closely followed. Briefly, chamber slides were incubated on ice for 5 min and washed with cold permeabilization buffer

(20 mM Hepes, pH 7.4, 110 mM KoAc, 5 mM MgOAc, 0.5 mM EGTA, and 250 mM sucrose) for 5 min. Cells were then incubated on ice for 5 min in permeabilization buffer containing 40 µg µl⁻¹ digitonin (Promega) while frozen egg extracts were allowed to thaw on ice. Cells were washed three times with cold permeabilization buffer for 5 min each. Thawed egg extracts (500 µl per chamber) were added, and cells were imaged on a DeltaVision RT (GE Healthcare) microscope system using a ×60 objective with 20 × 0.2-µm z-sections at intervals of 4–6 min. Time-lapse images were deconvolved, and maximum-intensity quick projections videos were generated using the softWoRx (Applied Precision) program. GFP fluorescent intensities were measured using ImageJ.

Metaphase spread preparation. Subconfluent cell cultures were incubated with 100 ng ml⁻¹ KaryoMAX (Thermo Fisher Scientific) diluted in fresh growth medium for 4 h. Cells were trypsinized, pelleted, and slowly resuspended in 0.075 M KCl solution dropwise with gentle vortexing, followed by incubation at 37 °C for 6 min and addition of 1:6 total volume of prechilled Carnoy's fixative (3:1 methanol/acetic acid). Cells were then pelleted, resuspended in prechilled Carnoy's fixative, and centrifuged. Cells were resuspended in prechilled Carnoy's fixative and stored at −20 °C. For preparation of metaphase spreads, 10 µl of fixed cell suspensions was dropped onto glass slides and air-dried.

Interphase and metaphase DNA FISH. Chromosome paint and centromere enumeration probes were purchased from MetaSystems. Yq heterochromatin (YqH, 41H10) and RP11–113K10 BAC probes were obtained from the BACPAC Resources Program, CHORI (see URLs). 41H10 BAC DNA was isolated from 200-ml bacterial cultures by using a Plasmid Midi Kit (Qiagen) and labeled with Cy3 label using the nick-translation protocol for DNA Polymerase I/DNase I (Life Technologies). RP11–113K10 BAC DNA was isolated from 50-ml bacterial cultures by using the BACMAX DNA Purification Kit (Epicentre), sonicated (Branson Ultrasonics Sonifier S-450), and labeled with tetramethyl (TM)-rhodamine using a Label IT Nucleic Acid Labeling Kit (Mirus Bio). Labeled BAC probes were resuspended in fluorescein isothiocyanate (FITC)-labeled Y chromosome paint probes (Metasystems).

For FISH on interphase cells, unsynchronized cells cultured in chamber slides were washed with PBS and fixed with Carnoy's fixative for 15 min at room temperature. Slides were dehydrated in 80% ethanol and air-dried.

FISH probes were combined as indicated, applied to interphase cells or metaphase spreads on glass slides, and sealed with a coverslip. Samples and probes were co-denatured at 75 °C for 2 min, followed by sealing with rubber cement and overnight hybridization at 37 °C in a humidified chamber. Slides were washed in 0.4× SSC at 72 °C for 2 min and in 2× SSC, 0.05% Tween-20 at room temperature for 30 s. Slides were rinsed in water, stained with DAPI, and mounted in ProLong Gold antifade mounting solution.

For quantification of chromosomal rearrangement frequencies, randomly selected metaphases were imaged, manually inspected using the indicated DNA FISH probes, and scored for de novo structural chromosomal abnormalities when compared with the parental DLD-1 karyotype.

Immunofluorescence and immuno-FISH. For immunofluorescence, cells seeded on coverslips or chamber slides were fixed with 4% formaldehyde diluted in PBS for 10 min and washed with PBS. Cells were then permeabilized with 0.3% Triton X-100, PBS for 5 min, washed with PBS, and blocked with Triton Block (0.1% Triton X-100, 2.5% FBS, 0.2 M glycine, PBS). Primary antibodies were diluted in Triton Block and applied to cells for 1 h at room temperature or overnight at 4 °C, followed by three 10-min washes with 0.1% Triton X-100, PBS. The primary antibodies to the following were used at 1:1,000 dilution in Triton Block: CENP-A (ab13939, Abcam), GFP (A10262, Invitrogen), phospho-S139 γH2AX (clone JBW301, 05–636, Millipore), and 53BP1 (NB100–304, Novus Biologicals). The primary antibodies to the following were obtained from GeneTex and used at 1:500 dilution in Triton Block: MRE11 (GTX70212), RAD50 (GTX70228), NBS1 (GTX70222), and phospho-S343 NBS1 (GTX61779). Fluorescence-conjugated secondary antibodies (Thermo Fisher Scientific) were diluted 1:500 in Triton Block and applied to cells for 1 h at room temperature, followed by 3 10-min washes with 0.1% Triton X-100, PBS and staining with DAPI. Cells were mounted in ProLong Gold antifade mounting solution and imaged as described.

For immunofluorescence combined with DNA FISH, the immunofluorescence procedure was performed as described followed by fixation in Carnoy's fixative for 10 min at room temperature and dehydration in 80% ethanol for 5 min. Samples were then processed for FISH as described.

Fixed and live-cell microscopy. Metaphase FISH images were captured on a DeltaVision Elite (GE Healthcare) microscope system with a ×60 objective and 5 × 0.5-µm z-sections. Maximum-intensity quick projections were generated using softWoRx. Immunofluorescence images were captured at 30 × 0.2-µm z-sections, followed by deconvolution and maximum-intensity quick projections using softWoRx. Structured illumination microscopy imaging was performed on metaphase spreads by using a DeltaVision OMX SR (GE Healthcare) microscope system with a ×100 objective (Olympus PlanApo 1.4 NA) and EMCCD 10-MHz camera mode. Raw images were reconstructed and aligned for super-resolution

using softWoRx with the following Wiener filters: 525 nm, 0.020; 442 nm, 0.020; 600 nm, 0.015.

For live-cell imaging experiments, 1×10^4 cells per well were seeded into CELLSTAR μClear 96-well plates (Greiner Bio-One) and imaged on a Yokogawa CQ1 confocal image cytometer with a ×40 objective. To determine mitotic timing and the formation of micronuclei with mitotic checkpoint inactivation, cells were seeded 16 h before staining with 500 nM SiR-DNA (Spirochrome) and treatment with reversine (Santa Cruz) at the indicated concentrations, followed by image acquisition at 5-min intervals with 5×2-μm z-sections per field. To monitor NEBD in cells stably expressing lamin A–GFP, micronuclei were induced by 3-d DOX/IAA treatment, followed by image acquisition at 5- or 6-min intervals with 6×1.5-μm z-sections per field. Mean intensity projections were created using ImageJ.

Immunoblotting. Whole-cell extracts were collected in SDS sample buffer and boiled for 10 min. Samples were resolved by SDS–PAGE, transferred to polyvinylidenedifluoride, and blocked with 5% milk in PBST (PBS, 0.1% Tween-20). Primary antibodies against CENP-A (2186, Cell Signaling) and GAPDH (14C10, Cell Signaling) were diluted 1:1,000 and 1:2,000, respectively, in PBST followed by incubation and three 10-min washes in PBST. Membranes were incubated with horseradish peroxidase-conjugated secondary antibodies diluted 1:4,000 in 5% milk in PBST followed by three 10-min washes in PBST and exposure to film.

RT–PCR. Total RNA was extracted by using a NucleoSpin RNA kit (Machery-Nagel) according to the manufacturer's instructions. Complementary DNA libraries were created from 4 μg total RNA by using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) with random primers. PCR was conducted using Q5 High-Fidelity polymerase (NEB) with the following reaction conditions: 98 °C for 3 min; 30 cycles of 98 °C for 1 min, 63 °C for 1 min, and 72 °C for 1 min; followed by 72 °C extension for 5 min. PCR products were resolved by electrophoresis on a 2% agarose gel and visualized with ethidium bromide. All RT–PCR primer sequences are provided in Supplementary Table 3.

Next-generation DNA sequencing and mapping. Genomic DNA was prepared by using a Quick-DNA Plus Kit (Zymo) according to manufacturer instructions. Sequencing libraries were generated from genomic DNA on robots according to manufacturer protocols. DNA was sequenced on an Illumina HiSeq X Ten platform to a targeted whole human genome coverage of 30×. Cluster generation and sequencing were performed by using the manufacturers' pipelines. Average sequence coverage achieved across all samples was 35.7× (range, 31.4–39.8×) with 151-bp paired-ends and insert size of ~500 bp. Mapping to the human genome was performed using the BWA algorithm[74] BWA-MEM (Version 0.7.8) and genome build hs37d5.

Copy-number and structural variant calling. Copy-number was called using the ascatNgs algorithm[75]. The variant calling pipeline of the Cancer Genome Project, Wellcome Sanger Institute was used to call somatic mutations[76]. The following

algorithms with standard settings, and no additional postprocessing were used: CaVEMan for substitutions and BRASS for rearrangements.

We note that clones harboring Y chromosome end-to-end fusions without additional rearrangements identified by cytogenetics proved uninformative, due to the sequence junctions spanning repetitive telomeric sequences. We also note that X–Y translocation calls arose in three clones due to mapping errors, most probably caused by sequence similarities between regions of the X and Y chromosomes. All of these calls mapped to the distal end of Xp, and were confirmed by FISH to be false-positive signals and subsequently removed from the reported sequencing data.

Simulation of random microhomology at break sites. To simulate the expected frequency of microhomology lengths at break sites, we used a randomization approach to mimic the sequence context at the actual break sites identified in our samples. The set of all rearrangement calls was replicated 100×, and a random number within the range −5,000 to 5,000 was added to each break-site position. The length of homology at the junctions of these simulated break-site sequences was then measured.

Statistical analysis and reproducibility. Statistical analyses were performed using GraphPad Prism 7.0 software with the tests as described in the figure legends; all P values were derived from experiments repeated independently at least three times. Error bars represent standard error of the mean of at least three independent experiments unless indicated. Figures with representative images were reproduced and obtained from two or more independent experiments.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The data that support the findings of this study are available from the corresponding authors upon request. Whole-genome-sequencing data have been deposited in the European Genome-phenome Archive under accession number EGAD00001004163.

References
72. Murray, A. W. Cell cycle extracts. *Methods Cell Biol.* **36**, 581–605 (1991).
73. Shankaran, S. S., Mackay, D. R. & Ullman, K. S. A time-lapse imaging assay to study nuclear envelope breakdown. *Methods Mol. Biol.* **931**, 111–122 (2013).
74. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
75. Raine, K. M. et al. ascatNgs: identifying somatically acquired copy-number alterations from whole-genome sequencing data. *Curr. Protoc. Bioinformatics* **56**, 15 9 1–15 9 17 (2016).
76. Nik-Zainal, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47–54 (2016).

natureresearch

Corresponding author(s):   Peter Ly and Don W. Cleveland

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Microscopy images were collected on DeltaVision RT, Elite, and OMX microscope systems (GE Healthcare) with softWoRx or a Yokogawa CQ1 confocal image cytometer. Whole-genome DNA sequencing was conducted using an Illumina HiSeq X Ten platform.
Data analysis	Mapping to the human genome was performed using the BWA algorithm BWA-MEM (Version 0.7.8) and genome build hs37d5. Copy number was called using the ascatNgs algorithm. The variant calling pipeline of the Cancer Genome Project, Wellcome Sanger Institute was used to call somatic mutations. The following algorithms with standard settings and without additional post-processing were used: CaVEMan for substitutions and the BRASS algorithm for rearrangements. Statistical analyses were performed with GraphPad Prism 7.0 software using the tests described in the figure legends.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

> Data reported in this study are available upon request. Whole-genome sequencing data have been deposited at the European Genome-phenome Archive (EGA) under accession number EGAD00001004163.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were not predetermined.
Data exclusions	Distal Xp-Y structural variant mapping errors caused by sequence similarities between the human Y chromosome and distal short-arm of the human X chromosome were excluded from three DLD-1 clones. In all three cases, these excluded calls were confirmed to be false by metaphase DNA fluorescence in situ hybridization. This exclusion is described in the Methods section under "Copy number and structural variant calling".
Replication	All experiments were independently reproduced multiple times, as described in the figure legends. All P-values were derived from measurements obtained from experiments conducted independently at least three times. Figures with representative images were reproduced and obtained from two or more independent experiments.
Randomization	Samples were not allocated into random experimental groups.
Blinding	The investigators were not blinded to group allocation during data collection and/or analysis.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	CENP-A (ab13939, Abcam), CENP-A (2186, Cell Signaling), GAPDH (14C10, Cell Signaling), GFP (A10262, Invitrogen), phospho-S139 γH2AX (clone JBW301, 05-636, Millipore), 53BP1 (NB100-304, Novus Biologicals), MRE11 (GTX70212, GeneTex), RAD50 (GTX70228, GeneTex), NBS1 (GTX70222, GeneTex), phospho-S343 NBS1 (GTX61779, GeneTex)
Validation	All primary antibodies used in this study are commercially available and have been validated by multiple, previously published reports. DNA damage response antibodies were validated by exposing cells to DNA damage-inducing agents as described.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	DLD-1 human colorectal cancer cells; RPE-1 human retinal pigment epithelial cells; 293T human embryonic kidney cells
Authentication	DLD-1 cells were authenticated by SNP array analysis, karyotyping, and whole-genome DNA sequencing.
Mycoplasma contamination	All cell lines used in this study were confirmed to be free of mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Oocyte-positive female Xenopus laevis frogs (Nasco, LM00531MX) were used in this study to produce egg extracts. All procedures were conducted in accordance to UCSD guidelines and regulations.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve field-collected samples.