



Supplementary Materials for
Mechanisms Generating Cancer Genome Complexity from a
Single Cell Division Error

Neil T. Umbreit, Cheng-Zhong Zhang, Luke D. Lynch, Logan J. Blaine, Anna M. Cheng,
Richard Tourdot, Lili Sun, Hannah F. Almubarak, Kim Judge, Thomas J. Mitchell,
Alexander Spektor, and David Pellman

correspondence to: david_pellman@dfci.harvard.edu, neilt_umbreit@dfci.harvard.edu,
cheng-zhong_zhang@dfci.harvard.edu

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S23
Tables S1 to S3
Captions for Movies S1 to S8

Other Supplementary Materials for this manuscript includes the following:

Movies S1 to S8

Materials and Methods

Cell culture and generation of cell lines

All cells were cultured at 37°C in 5% CO₂ atmosphere with 100% humidity. U2OS, HeLa, and telomerase-immortalized RPE-1 cells were grown in DMEM/F12 (1:1) media (Gibco) supplemented with 10% FBS, 100 IU/ml penicillin, and 100 µg/ml streptomycin. Telomerase-immortalized BJ cells were grown in DMEM/Medium 199 (4:1) supplemented with 10% FBS, 100 IU/ml penicillin, and 100 µg/ml streptomycin. For cell lines with doxycycline-inducible constructs, tetracycline-free FBS (X&Y Cell Culture) was used in all culture media. A complete description of all cell lines used in this study is provided in Table S3.

Protein and shRNA expression constructs were cloned into pLenti-CMV-Neo-DEST or pLenti-CMV-Blast-DEST (Addgene) and co-transfected with lentiviral packaging plasmids, pMD2.G and psPAX2, into 293FT cells using Lipofectamine 2000 (Life Technologies) according to the manufacturer's instructions. RPE-1, U2OS, HeLa, or BJ cells were virally transduced in the presence of 10 µg/ml polybrene for 16 hours. Transduced cells were obtained either by antibiotic selection (500 µg/ml geneticin or 10 µg/ml blasticidin) started 24 hours after transduction, or by fluorescence activated cell sorting 5-7 days after transduction.

Expression construct	Reference
GFP-BAF	This study
SNAP-BAF	This study
rtTA3	Addgene #26429
TRF2-DN (Tet-ON)	(30)
Cas9 (Tet-ON)	(79)
Chr4 subtelomere sgRNA	This study
TagRFP-T-Utrophin 261	(80)
RFP-NLS	(15)
shRNA p21 / shRNA Rb	(81)

Cell cycle synchronization and methods to generate chromosome bridges or micronuclei

To prevent cell cycle arrest (e.g. from telomere dysfunction), all experiments were performed in cells depleted of p53 using 40 nM ON-TARGETplus siRNA SMARTpool L-003329-00-0050 (Dharmacon), or in cells constitutively expressing shRNAs against p21 and Retinoblastoma (81). siRNA was transfected with Lipofectamine 3000 (Life Technologies) according to the manufacturer's instructions 24 hours before the start of the experiment.

Chromosome bridges were generated by the following procedures: (1) TRF2-DN induction with 0.1 µg/ml doxycycline for 14-16 hours; (2) Cas9 induction with 1 µg/ml doxycycline for 14-16 hours in cells constitutively expressing sgRNA targeting the Chr4 subtelomere (see details below); (3) topoisomerase II inhibition with 100 nM ICRF-193; (4) partial depletion of condensin by transfection of 1 nM SMC2 ON-TARGETplus siRNA SMARTpool L-006836-01-0005 (Dharmacon) using Lipofectamine 3000 (Life Technologies). Each method was optimized to generate chromosome bridges at a moderate frequency, ~20-30% per cell division (frequency in untreated RPE-1 and BJ

hTert cells is <<1%). For TRF2-DN and CRISPR-based methods, bridges begin forming during cell divisions occurring at least ~8 hours after washout of doxycycline, and ~1-2 days after transfection for SMC2 knockdown, whereas ICRF-193 effects are essentially immediate. HeLa and U2OS cells generated bridges at a basal frequency of ~5-10% per cell division and did not require experimental induction.

Cell synchronization and induction of micronuclei by nocodazole block and washout was performed as described (15); cells were treated with 100 ng/ml nocodazole for 6 hours, followed by mitotic-shakeoff and three washes with warm medium. To detect interphase DNA replication, cells were incubated continuously with 10 μ M EdU, added three hours after mitotic shakeoff. EdU incorporation was detected using the Click-iT Plus EdU Alex Fluor 647 Imaging Kit (Life Technologies). To block initiation of DNA replication, 10 μ M PHA-767491 or 250 nM flavopiridol was added three hours after shakeoff, together with 50 μ M Z-VAD to prevent apoptosis during prolonged arrest in cells depleted of p21 and Rb (82). For G2 synchronization, cells were treated with 9 μ M RO-3306 for 18 hours. G2-arrested cells were released into mitosis by washing seven times with warm medium. To label mitotic DNA synthesis, cells were released from G2 block into 10 μ M EdU. For generation of micronuclei by MPS1 inhibition, G2-arrested cells were released into 0.8 μ M NMS-P715.

Drug	Product details	Concentration
Doxycycline	Clontech Labs #631311	0.1-1 μ g/ml
ICRF-193	Sigma #I4659	100 nM
Latrunculin A	Life Technologies #L12370	0.5 μ M
ML7	Abcam #ab120848	20 μ M
Nocodazole	Sigma #M1404	100 ng/ml
NMS-P715	MilliporeSigma #4759495MG	0.8 μ M
5-ethynyl-2'-deoxyuridine	Life Technologies #C10640	10 μ M
RO-3306	MilliporeSigma #2176995MG	9 μ M
PHA-767491	Selleck #S2742	10 μ M
Flavopiridol	Selleck #S1230	250 nM
Z-VAD-FMK	Selleck #S7023	50 μ M
Etoposide	Selleck #S1225	10 μ M

CRISPR-mediated gene knockouts and Chr4 telomere loss

For RPE-1 cells with doxycycline-inducible Cas9 expression (79), sgRNA constructs were cloned into plasmid pLenti-Guide-Puro (Addgene) and delivered by lentiviral transduction. Guide RNA targeting sequences: TREX1, 5'-GAGAGCTTGCTACCACACG-3'; Chr4 subtelomere, 5'-TTTAGTGCCCGGCCGCAAGG-3'. Transduced cells were obtained by selection with 12 μ g/ml puromycin.

To generate CRISPR knockouts by Cas9/sgRNA transfection, TrueCut Cas9 v2 (Invitrogen) was incubated with TrueGuide modified synthetic sgRNA (Thermo Fisher) and transfected using Lipofectamine CRISPRMAX (Invitrogen) according to the manufacturer's instructions. Knockout clones were generated by single-cell flow sorting as above. Guide RNA targeting sequences: TREX1, 5'-

GCAGGTACGTACCCAACCAT-3'; SUN1, 5'-CAAGACTCGTCCAATACAGG-3'; SUN2, 5'-ACTGCATGGTGACGCCAACT-3'.

DNA isolation and Western blotting

Bulk DNA isolation was performed using PureLink Genomic DNA kit (Invitrogen) according to the manufacturer's instructions. For Western blots, cells were collected by trypsinization, pelleted, washed in PBS, and lysed by addition of an equal volume of 2X SDS lysis buffer (0.1 M Tris pH 6.8, 4% SDS, and 12% β -mercaptoethanol). Samples were denatured at 100°C for 10 minutes, passed through a 28 $\frac{1}{2}$ gauge needle to shear genomic DNA, run on NUPAGE 4-12% Bis-Tris polyacrylamide gel (Life Technologies), and transferred to PVDF membrane using iBlot (Life Technologies). Odyssey Blocking Buffer (LI-COR) was used for membrane blocking and antibody dilutions. Blots were incubated with primary antibodies for 1 hour at room temperature or overnight at 4°C. Three washes were performed with PBS-T, followed by incubation with fluorescent secondary antibodies (LI-COR) for 1 hour at room temperature, and three more PBS-T washes. Membranes were visualized using a ChemiDoc MP Imaging System (BioRad).

Antibody	Manufacturer	Dilution
Myc	Life Technologies #MA121316	1:1000
Cas9	Cell Signaling Technology # 14697S	1:1000
GAPDH	Abcam #ab8245	1:5000
γ -H2AX	MilliporeSigma #05-636-I	1:1000
SMC2	Abcam #ab10412	1:5000
α -Tubulin	Sigma #T9026	1:5000
TREX1	Abcam #ab185228	1:1000
SUN1	Abcam #ab124770	1:1000
SUN2	Abcam #ab124916	1:1000
IR680 anti-mouse	LI-COR #926-68072	1:5000
IR680 anti-rabbit	LI-COR #926-68073	1:5000
IR800 anti-mouse	LI-COR #926-32212	1:5000
IR800 anti-rabbit	LI-COR #926-32213	1:5000

Live-cell imaging

Cells were plated on #1.5 glass-bottom plates (MatTek), or on ibiTreat 24-well μ -Plates or 35 mm dishes (ibidi). Micropatterns (CYTOO) were custom-made on #1.5 coverslips and mounted in a CYTOOchamber for imaging. Substrate stiffness experiments were performed using CytoSoft Imaging plates (Advanced BioMatrix) coated with 5 μ g/ml fibronectin (Sigma). Prior to the start of imaging, SNAP-tagged proteins were labeled using SNAP-Cell 647-SiR dye (New England Biolabs) according to the manufacturer's instructions.

Widefield fluorescence microscopy was performed on a Nikon inverted microscope (Ti-E or Ti2) with Perfect Focus, outfitted with an environmental enclosure to maintain cell culture conditions (37°C and humidified 5% CO₂). Imaging was performed using a 20 \times /0.75 NA Plan Apochromat Lambda objective (Nikon); Z-stacks of

four images at 2- μ m spacing were acquired with a Zyla 4.2 sCMOS camera (Andor), at intervals of 5-10 minutes for up to 48 hours.

Actin/myosin contractility was inhibited during live-imaging experiments as follows. Cells were induced to generate chromosome bridges during the first 12-14 hours of imaging in normal growth medium. During a gap between image acquisition intervals, the medium was then replaced to introduce 0.5 μ M Latrunculin A (Life Technologies) to disrupt actin dynamics, or 20 μ M ML7 (Abcam) to inhibit myosin activation by myosin light chain kinase. Live-imaging continued in drug medium for the remainder of the experiment. Bridges that were intact at the start of drug treatment were analyzed.

Live-cell confocal microscopy to visualize chromosome bridges was performed on a Ti2 inverted microscope fitted with a CSU-W1 spinning disk system (Nikon). Z-stacks (seven images at 1- μ m spacing) were collected every 5-10 minutes for 24 hours, using a Zyla 4.2 sCMOS camera (Andor), and a 40 \times /0.95 NA Plan Apochromat Lambda objective with the correction collar set to 0.17. For micronuclear chromosomes undergoing mitotic replication, imaging was performed on a Nikon Ti-E inverted microscope fitted with a Yokogawa CSU-22 spinning disk head with the Borealis modification. Z-stacks (seven images at 1- μ m spacing) were collected every 2 minutes for 2 hours, using a CoolSnap HQ2 CCD camera (Photometrics), and a 100 \times /1.45 NA Plan Apochromat Lambda oil immersion objective (Nikon). For all live-cell confocal imaging, an environmental enclosure was used to maintain cell culture conditions (37°C and humidified 5% CO₂).

Indirect immunofluorescence microscopy

Cells were seeded on #1.5 coverslips and were washed once with PBS prior to fixation. For staining of RPA1, γ -H2AX, and LAP2, and for labeling of EdU incorporation, cells were fixed with 4% paraformaldehyde for 15 minutes. For staining of actin, myosin heavy chain, phospho-myosin light chain 2, paxillin, SUN1, and SUN2, cells were fixed with PTEMF buffer (20 mM PIPES pH 6.8, 1 mM MgCl₂, 10 mM EGTA, 0.2% Triton X-100, 4% paraformaldehyde) for 15 minutes.

Fixed samples were washed three times with PBS, permeabilized for 5 minutes at room temperature with PBS+0.5% Triton X-100, and washed three more times with PBS. Samples were then blocked in PBS+3% BSA for one hour, incubated with primary antibodies diluted in blocking buffer, washed three times with PBS+0.05% Triton X-100, incubated with secondary antibodies diluted in blocking buffer, and washed again three times in PBS+0.05% Triton X-100. After DNA staining in PBS+2.5 μ g/ml Hoechst 33342 (Life Technologies) for 20 minutes, samples were washed twice in PBS and mounted in ProLong Gold antifade (Life Technologies) on glass slides.

Imaging was performed on a spinning disk confocal microscope (Nikon Ti-E with Yokogawa CSU-22, described above). Z-stacks of nine images at 0.5- μ m spacing were collected using a CoolSnap HQ2 CCD camera (Photometrics) or a Prime BSI back-thinned sCMOS camera (Photometrics), with a 60 \times /1.40 NA or a 100 \times /1.45 NA Plan Apochromat oil immersion objective (Nikon).

Antibody	Manufacturer	Dilution
γ -H2AX	MilliporeSigma #05-636-I	1:500
RPA1	Cell Signaling Technology #2267S	1:100

LAP2	BD Biosciences #611000	1:750
SUN1	Abcam #ab124770	1:200
SUN2	Abcam #ab124916	1:200
pSer19-Myosin light chain	Cell Signaling Technology #3675S	1:200
Myosin heavy chain	Biomedical Technologies Inc. #BT-567	1:100
Paxillin	Abcam #ab32084	1:250

Same cell correlative live- and fixed-imaging

Cells were seeded in 35-mm ibiTreat Grid-500 dishes (ibidi) with a gridded imaging surface. For live/fixed imaging of chromosome bridges, asynchronous cells were treated with 100 ng/ml doxycycline for 14-16 hours, followed by washout for 16-20 hours to induce TRF2-DN and generate chromosome bridges. Cells were treated with 9 μ M RO-3306, and immediately mounted on the microscope for live-cell imaging throughout the 18-hour arrest in G2. During a gap between image acquisition intervals, samples were released into mitosis by washing seven times. After 45 minutes, cells were fixed with 4% paraformaldehyde and indirect immunofluorescence was performed as described above. Based on live-cell imaging videos and using the gridded dish surface for cell location, cells with intact chromosome bridges at the start of the experiment were analyzed by spinning disk confocal microscopy.

For live/fixed imaging of intact micronuclei, asynchronous cells were treated with 0.8 μ M NMS-P715 (MPS1i) for 16-20 hours to generate micronuclei and synchronized in G2 with 9 μ M RO-3306 for 18 hours. Live-cell imaging was started in the last 1-2 hours of G2 arrest. During a gap between image acquisition intervals, samples were released into mitosis by washing seven times. After 45 minutes, extraction was performed for 1 minute in CSK buffer (10 mM PIPES pH 6.8, 100 mM NaCl, 1 mM MgCl₂, 1 mM EGTA, 0.5% Triton X-100) prior to fixation with 4% paraformaldehyde. Extraction was used before fixation to remove RFP-NLS signal, enabling staining in the near-red fluorescence channel; indirect immunofluorescence was performed as described above. Based on live-cell imaging videos and using the gridded dish surface for cell location, cells with micronuclei that remained intact until mitotic entry (judged by RFP-NLS signal) were analyzed by spinning disk confocal microscopy.

Fluorescence in situ hybridization

Cells were seeded on #1.5 glass coverslips and synchronized by serum starvation: samples were exchanged to serum-free DMEM/F12 for 24 hours, followed by release into medium with 10% FBS and 1 μ g/ml doxycycline to induce CRISPR-mediated Chr4 bridge formation at the next mitosis. Doxycycline was washed out after 14-16 hours, and samples were collected 24 and 48 hours after washout to analyze 1st generation and 2nd generation cells as follows. Cells were swelled in hypo-osmotic solution (75 mM KCl) for 15-30 minutes before fixation by dropwise addition of a one-half volume of -20°C Carnoy's solution (3:1 mixture of methanol:acetic acid). After 5 minutes, samples were exchanged to fresh -20°C Carnoy's solution twice more, incubating at room temperature for 20 minutes each time. Coverslips were removed from the fixative and left to air-dry overnight at room temperature. Chr4 centromere and whole-chromosome 4 "paint" probes (Leica Biosystems) were prepared in hybridization buffer according to the

manufacturer's instructions and applied to samples, sealed onto glass slides with rubber cement, and denatured at 75°C for 2-5 minutes before hybridization at 37°C in a humidified chamber for 1-2 days. Samples were removed from slides, washed twice in 50% formamide/2X SSC at 42°C for 10 minutes each, washed twice in 2X SSC at room temperature for 5 minutes each, and mounted on glass slides in ProLong Gold antifade with DAPI (Life Technologies).

Image analysis

Live-cell imaging videos were analyzed using NIS-Elements (Nikon) or Metamorph (Molecular Devices). In all experiments, bridge lifetime was measured as the time from cytokinetic furrow ingression to bridge breakage, scored visually based on GFP-BAF. Cell cycle progression was scored visually in cells expressing an mCherry-tagged fragment of Geminin (83). Timing of S phase onset was measured as the time interval from completion of mitosis to the first appearance of mCherry-Geminin signal. Total cell cycle duration was measured as the time interval from completion of the first mitosis until nuclear envelope breakdown in the second mitosis. Nuclear envelope (NE) rupture was scored visually in cells expressing RFP-NLS, imaged at 5-minute intervals, and was defined as the time from loss of nuclear accumulation until the first evidence of re-accumulation. For cell motility measurements, cell tracking was performed with the ImageJ/FIJI plugin, TrackMate, applied to detect primary nuclei based on GFP-BAF signal. Cell velocity was measured as the average root-mean-square distance traveled over 10-minute intervals, as previously described (84). Only cells that were tracked over at least five imaging intervals were analyzed.

Quantitative analysis of fixed-cell images was performed using ImageJ/FIJI. Image segmentation was performed using maximum intensity projections from Z-stacks. For interphase cells, images were segmented based on Hoechst staining (primary nuclei and micronuclei) or GFP-BAF signal (bridges). For mitotic cells, the main chromosome mass was segmented with Hoechst and the micronuclear/bridge chromosome was segmented based on RPA1 staining. Live-cell imaging showed that nearly all broken chromosome bridges or intact micronuclei become strongly positive for RPA1 upon mitotic entry, validating RPA as a reliable marker for bridge or micronuclear chromosomes in mitosis (Fig. 7E and Fig. S13A). The resulting image masks were used to measure γ -H2AX and/or EdU levels in the primary nucleus and micronucleus/bridge from a sum intensity projection.

Single-cell isolation, genome amplification, and DNA sequencing

Long-term live-imaging and correlative single-cell whole-genome sequencing ("Look-Seq") was performed as described (15) or with modifications detailed below. Flow sorting was used to deposit single cells into the wells of a 384-well μ Clear plate (Greiner). Widefield fluorescence imaging was performed at 15-20 minute intervals, and cells were treated to form bridges in the next mitosis. Following bridge breakage, the two daughter cells were trypsinized and isolated by limiting dilution into a new 384-well μ Clear plate, where they were allowed to attach for 3 hours before lysis.

During the course of this study, we developed a new method to isolate cells directly from imaging dishes, and this was employed for many Look-Seq experiments. Specifically, cells were plated in a 35-mm gridded ibiTreat dish (ibidi) and imaged at 10-

minute intervals. After sufficient time for bridge formation and breakage, the plate was transferred to another Nikon inverted microscope equipped with a CellEctor single-cell isolation system (Molecular Machines & Industries). The sample was exchanged into PBS-based, non-enzymatic dissociation reagent (Sigma) to loosen cell attachment for isolation as follows. Cells of interest were identified based on video recordings, using the gridded dish surface for cell location. Using a glass capillary with an inner diameter of 40 μm , single cells were directly picked from the imaging dish in a volume of 80 nL, and transferred into a $\sim 5\text{-}\mu\text{L}$ droplet of PBS in a PCR tube cap. Cell picking from a single imaging dish was performed within 30 minutes of applying cell dissociation reagent, and cells were kept on ice until lysis (within 5-20 minutes after isolation).

In our Look-Seq experiments, we collected cells on average 8.8 hr after bridge breakage. This likely provides ample time for DNA repair by non-homologous end joining (NHEJ), which occurs on a timescale of <2 hours (85).

Cell lysis and whole-genome amplification was then performed using the REPLI-g Single Cell kit (Qiagen), and the amplified DNA was purified, sheared to ~ 500 bp size, and processed with a Library Preparation Kit (KAPA) for multiplexed next-generation sequencing, as previously described (15).

Chromium single-cell CNV library construction

Hundreds of single cells were individually barcoded for whole-genome sequencing using the Chromium Single Cell DNA kit (10X Genomics). Libraries were successfully generated for nine primary clones (Fig. S18), while libraries for the remaining three primary clones failed to capture sufficient numbers of single cells (<100 unique cells identified, see details below).

Quality assessment of sequencing libraries by low-pass whole-genome sequencing

Sequencing libraries of single cells and bulk populations were subjected to low-pass whole-genome sequencing ($\sim 0.1\times$ mean coverage) on the MiSeq (Illumina) platform to assess library quality based on library complexity and chimeric frequency. Low-pass whole-genome sequencing further enabled us to (i) assess the uniformity of single-cell whole-genome amplification, (ii) estimate haplotype DNA copy number to identify bridge chromosomes or micronucleated chromosomes, and (iii) estimate the number of cells captured in the Chromium single-cell CNV libraries. Sequencing libraries passing quality control were sequenced to a greater depth on the HiSeq 2500 (Illumina) or the NovaSeq (Illumina) platforms. Single-cell samples isolated after bridge resolution were sequenced to $\sim 25\times$ mean coverage (range 20-34 \times), except for those for which the bridge was broken mechanically ($\sim 5\times$ mean coverage). Single-cell samples with intact micronuclei (either before or after cell division/reincorporation) were sequenced to 5-10 \times mean coverage. Single-cell derived bulk populations were each sequenced to $\sim 10\times$ mean coverage. Single-cell derived subclones were each sequenced to $\sim 1\times$ mean coverage. Each of the Chromium single-cell CNV libraries contained ~ 500 -800 cells and were sequenced in aggregate to $\sim 60\times$ mean coverage, yielding $\sim 0.1\times$ coverage per cell.

Alignment and preprocessing of reads from whole-genome sequencing

All sequencing reads were aligned to the human genome reference (GRCh38 primary assembly) with additional “sponge reference” compiled from repeat-rich

sequences (86) using **bwa** (<http://bio-bwa.sourceforge.net/>) v.0.7.12 in the paired-end mode by “bwa mem”. For read pairs with multiple alignments, the choice of primary alignment is based on the following criteria in order of priority: proper alignment of both mates (i.e., following the forward-reverse orientation and having a fragment size within the inferred insert size distribution), accurate mapping (highest combined mapping qualities of alignments for one or both mates), most aligned bases (largest combined blocks aligned to the reference); the other alignments are marked as supplementary alignments. Read pairs with identical alignment positions are marked as PCR duplicates by **MarkDuplicates** from the Picard software suite (<http://picard.sourceforge.net/>). For read pairs with multiple alignments, PCR duplicate status was determined based on primary alignment positions but marked for both primary and supplementary alignments.

Single-cell variant calling and detection of kataegis

We performed short variant detection using GATK HaplotypeCaller v4.1.2.0 running jointly on all single-cell bridge samples sequenced to $\sim 25\times$ mean depth. In addition to the default read filters, we further excluded reads of low mapping quality (--minimum-mapping-quality 10), excessive clipping (--filter-too-short 50), or having discordant alignment positions (NonChimericOriginalAlignmentReadFilter, MateOnSameContigOrNoMappedMateReadFilter). These filters served to eliminate potential errors due to alignment inaccuracy, including both misalignment of the entire read and incorrect placement of mismatched bases in split or discordant reads. We note that the exclusion of split reads and discordant reads may reduce evidential support for variants near rearrangement breakpoints. The rationale for imposing a more stringent filter was that the majority of split or discordant reads are not related to true structural variants but result from single-cell amplification artifacts. To further exclude variants caused by amplification artifacts that do not generate chimeric DNA, we removed all variants with fewer than four supporting reads.

To identify de novo single-nucleotide variants (SNVs) in each single-cell sample, we excluded variants detected in the bulk RPE-1 data and those with supporting reads in more than one single-cell sample. We further imposed the requirement that the reference base at a de novo variant site be observed in at least 50% of all samples in order to exclude poorly mappable regions of the genome, which are particularly prone to alignment artifacts. Finally, we also excluded variants present in the dbSNP database (build 146) to exclude both common variants not detected in the bulk data and frequent false variants due to recurrent mapping artifacts.

Even with this stringent filtering, the list of de novo variants was dominated by a high frequency of C>T SNVs (~ 1 per 100 kb). We believe most of these variants reflect cytosine deamination due to heating during library preparation, as previously reported (87). However, in a few samples we observed clusters of mutations (“kataegis”) where the tight localization of strand-coordinated C>T variants could not be explained by the background rate of deamination. To identify bona fide kataegis events due to AID/APOBEC activity, we searched for strand-coordinated clusters of C>T, C>G, and C>A mutations in the TpC context as described previously (43). We further required that kataegis clusters contain at least five strand-coordinated mutations in the TpC context with a minimum inter-mutation distance of 2 kb (28). Based on these criteria, kataegis was either only observed on the bridge chromosome near the site of breakage (T-2 and 4-

8), or in multiple loci across the genome where they were not associated with detectable copy number alterations (daughter pairs 4-4, 4-5, and 4-6).

Calculation of total sequence coverage

Raw sequence coverage. We first counted the number of reads in each 10-kb bins using the GATK (v.4.0.12.0-6) CollectReadCounts module with the following read filters to remove non-properly aligned read fragments:

```
--read-filter FragmentLengthReadFilter --max-fragment-length 1000  
--read-filter MateOnSameContigOrNoMappedMateReadFilter  
--read-filter MateDifferentStrandReadFilter
```

Normalization of GC-dependent bias. We then normalized read counts in 10-kb bins based on the local average GC content. For bulk libraries, the local average GC content was calculated directly from the reference sequence in each bin. For single-cell libraries, the local average GC content was calculated from the sequence in a 50-kb flanking region of each bin as follows. The average GC content of the i th bin was calculated as the arithmetic mean of GC content in bins $i-2$ to $i+2$, excluding any bin with 5,000 or more non-N bases. For bins at chromosome ends, the GC-averaging interval was truncated at the ends. Each 10-kb bin was first grouped into a GC stratum based on the local average GC content. Normalization of sequencing coverage was then performed independently for each (single-cell or bulk) sequencing library. First, the raw read counts were converted to log-read counts and centered by the genome-wide median of log-counts estimated from all bins having $\geq 9,000$ non-N bases and average GC percentage in the range of [0.32,0.61] (roughly corresponding to 0.1 to 99.9 percentile of average GC content of all bins). Second, for bins within each GC stratum, we calculated the median of normalized log-read counts, again excluding those with $\geq 9,000$ non-N bases or those having zero coverage (these bins reflect incompleteness in the human genome reference or sequence variation between the RPE-1 genome and the human genome reference). For GC strata with ≤ 100 bins, the estimates of median coverage were inaccurate, and we excluded all the bins in these strata from downstream calculations. Third, we determined the multiplicative GC correction factor for each GC stratum by normalizing the median of log-scale coverage of that stratum against the mean of medians across all strata. This normalization ensured uniform median log-scale coverage in all GC strata. The multiplicative GC correction factors were then applied to each bin by their GC strata to generate GC-normalized log-scale coverage. Finally, we centered the GC-normalized log-scale coverage by the median in all 10-kb bins (excluding those with zero coverage) and then converted the centered, normalized log-scale coverage to linear-scale normalized coverage. For bulk libraries, we further calculated bin-level normalized sequence coverage (30 kb, 90 kb, 250 kb) as the mean of 10-kb normalized sequence coverage in each bin. For single-cell libraries, we performed additional normalization of recurrent amplification bias.

Estimation and normalization of recurrent bias. After correcting for GC-dependent bias, we still observed significant variation that was often recurrent among independently amplified libraries. The source of this recurrent bias is unclear but may be related to priming during multiple-displacement amplification, which depends on local sequence composition and chromatin accessibility. To correct for this locus-specific bias, we took advantage of our large number of independently amplified single RPE-1 libraries and determined the median of normalized sequencing coverage in each 10-kb bin across

all samples after GC-correction. Variation in the median coverage across bins (assuming no copy-number changes) reflects recurrent amplification bias; we normalized this variation by dividing GC-normalized sequencing coverage in each bin in each sample by the median across all samples, with the exception of de novo or recurrent copy number variants which were identified as follows.

To exclude regions with copy-number variation in the calculation of median coverage, we first calculated arm-level coverage in each sample using the GC-normalized coverage and excluded any arm with average coverage deviating from the median (across all samples) by 25% or more. The median values determined in this way were robust to sporadic segmental or arm-level copy-number changes in one or a few samples. Special treatment was given to regions with altered DNA copy number in a large number of cells. For RPE-1 cells these regions were Chr10q (clonal single-copy gain from 61.8 Mb to q-ter as determined from bulk DNA-Seq) and Chr12 (frequent 2-copy gain of 12p from iso-Chr12p or loss-of-heterozygosity in 12p). For Chr12p, we excluded any sample with an average sequencing depth above 1.45 in the calculation of expected median coverage; this excludes samples with iso-Chr12p. For Chr10q, we first calculated the median coverage for every bin in the gained region and then multiplied the median by the ratio between the average median coverage in 10p relative to the average median coverage in the 10q gain; this ensures a constant median coverage across Chr10.

Putting all of this information together, we determined the expected sequencing coverage for every 10-kb bin in disomic regions reflecting recurrent amplification bias. For each single-cell sample, we divided the GC-normalized coverage by the mean coverage to normalize away recurrent amplification bias. While the correction for GC content resulted in a modest improvement in coverage uniformity, the correction for recurrent bias at the bin level resulted in a far greater improvement. With these two corrections, the normalized sequencing coverage better reflected the relative abundance of DNA sequence in each bin. We further calculated the bin-level normalized sequence coverage (50kb, 250kb, 1Mb) as the mean of 10-kb normalized sequence coverage in each bin.

Calculation of the haplotype fraction

Determination of complete RPE-1 haplotypes. Whole-chromosome haplotypes of the RPE-1 genome were determined from a combination of linked-reads and Hi-C bulk sequencing of RPE-1 cells using a recently developed computational method (88) and validated using our own sequencing data of monosomic cells. Details of variant calling, filtering, haplotype determination and validation can be found in our bioRxiv preprint (88).

Calculation of haplotype copy ratio from allelic coverage. We first counted the number of reads showing either the reference or the alternate genotype at each heterozygous site (only single-nucleotide variants were included) using an in-house version of the GATK (v.4.0.12.0-6) ASEReadCounter module that was modified to output counts at every site (including those with zero coverage). We then converted allelic coverage to haplotype coverage using the haplotype phase at each variant site, generating read counts supporting either parental haplotype (**A** or **B**) at all phased variant sites. Finally, we calculated the haplotype fraction at each variant site, defined as the fraction of reads originating from either haplotype. We further calculated the haplotype

fraction in 50-kb, 250-kb, and 1-Mb bins as the average haplotype fraction over all variants in each bin. For bulk libraries, the haplotype fraction was calculated for 30-kb, 90-kb, and 250-kb bins.

Calculation of haplotype DNA copy number

The haplotype coverage in each bin (50 kb, 250 kb, or 1 Mb for single-cell libraries, 30 kb, 90 kb, and 250 kb for bulk libraries) was calculated as

$$C_{A,B}^{(i)} = D^{(i)} \cdot \overline{R_{A,B}^{(i)}}$$

where $D^{(i)}$ is the normalized total sequence coverage and $\overline{R_{A,B}^{(i)}}$ is the mean haplotype fraction in the i th bin. The haplotype copy number was determined by normalizing $C_{A,B}^{(i)}$ by the median of $C_{A,B}^{(i)}$ across the genome (of both haplotypes). In a diploid genome, this median value corresponds to a single copy and haplotype copy number should take integer values (0, 1, 2, ...); in tetraploid or close-to-tetraploid genomes, this corresponds to two copies and haplotype copy number can take half-integer values (0, 1/2, 1, 1 1/2, 2, ...). Determining the mean coverage for a single homolog can be challenging if half of the chromosomes have a different copy number (e.g., in a perfect triploid genome, half of all chromosomes have one copy and the other half have two); such cases are rare (not encountered in the current study) but can be easily identified and corrected by manual review.

Segmentation of haplotype DNA copy number

Haplotype DNA copy number segmentation was done using a similar iterative strategy as described before (15) and performed on the 250-kb bin-level haplotype copy number data. During each round of iteration, the copy number in each bin was first calculated as the local average of 5 consecutive bins (1.25 Mb window) and then rounded to the nearest integer. This procedure was repeated until the copy number of all bins converged to a constant value. The choice of window size (1.25 Mb) determines the minimum length of copy-number alteration (CNA) that is retained after iterative local averaging. For single-copy changes (i.e., one copy loss or gain), the minimum CNA length is 1.25 Mb; for a half-copy change (e.g., in a G2 cell), the minimum CNA would be 2.5 Mb; for 2-copy changes (rare), the minimum CNA length is ~0.6 Mb. We further coalesced segmental breakpoints within 1.5 Mb to eliminate short CNA segments. The segmental haplotype DNA copy number was then calculated as the mean haplotype copy number across all 250-kb bins within each segment. The threshold of minimum CNA length was conservative by choice to accommodate varying amplification variability in a large number of samples. For samples with better amplification uniformity, or for events with better detection power (in particular, the detection of loss or retention), the threshold could be significantly smaller (~ 100 kb).

Fine-scale segmentation of haplotype retention and loss

Our haplotype phasing information enabled us to detect smaller (> 100 kb) single-copy deletions in our samples, evident from a near-complete loss of SNPs supporting one of the two haplotypes. For each single-cell sample sequenced to ~30× mean coverage, we divided the genome into 10-kb bins and defined haplotype retention or loss within each

bin as the majority coverage status of all SNPs in that bin (0: majority absent, 1: majority covered by at least one read). Bins not containing a phased SNP site were excluded from subsequent analysis.

To identify haplotype copy number change-points, we modeled allelic retention or loss using a Hidden Markov Model (HMM) with two “hidden” states (“RET” and “DEL”), and two observed states (0 and 1) denoting bin-level haplotype coverage as defined above. Using this framework, the emission probabilities reflected the frequency of two types of errors: erroneous coverage of a haplotype that was lost due to sequencing or phasing errors ($\beta_1 = p(1 | DEL)$), and coverage dropout in a region of haplotype retention ($\beta_2 = p(0 | RET)$).

We determined the emission and transition probabilities empirically from a sample with average amplification uniformity (T-6, daughter a). This sample exhibited an arm-level loss of Chr19p that allowed for estimation of $\beta_1 = 0.03$. Calculating the rate of haplotype coverage dropout on disomic chromosomes yielded $\beta_2 = 0.15$. We assumed equal transition probability (α) between states (RET→DEL and DEL→RET) and set its value such that the detection limit of CNAs would be 100 kb (or 10 bins) by the following argument. The probability that the HMM is in the DEL state and emits ten “1” observations is equal to β_2^{10} . Alternatively, the probability that the HMM switches from DEL to RET, emits ten “1” observations, and then switches back is $\alpha^2(1 - \beta_1)^{10}$. Setting these two expressions equal to each other and solving for α yielded $\alpha = 5.5 \times 10^{-8}$. With these choices of emission and transition probabilities, the minimum detectable length of loss flanked by large segmental retention was ~200 kb by a similar mathematical argument. The better detection of segmental retention than segmental loss by our model is a consequence of the rate of sequencing errors resulting in erroneous haplotype retention calls (β_1) being lower than the rate of bin-level coverage dropouts (β_2). We used the Viterbi algorithm to solve for the most likely sequence of hidden (RET or DEL) states given the observed data.

For samples sequenced to lower depth (~5×), we modified our approach to account for the higher rate of haplotype coverage dropouts. To accomplish this, we increased the bin size from 10 kb to 20 kb. Additionally, we determined the haplotype coverage status in each bin based on the presence of any read support for the observed haplotype (0: no supporting reads in bin, 1: at least one supporting read). We then calculated $\beta_1 = 0.06$ and $\beta_2 = 0.15$ using the same procedure as described above. We set the transition probability $\alpha = 1.75 \times 10^{-6}$ to set the minimum size of CNAs to 10 bins (> 200 kb genomic distance).

Our choices of parameters ensured that when fine-scale segmentation was performed on the whole chromosome, copy number breakpoint calls were restricted to within a few Mb of the main copy number transition. This indicates that the segments identified here are of biological origin and not due to dropout from uneven coverage or sequencing artifacts. Validating this general approach, we observed a striking concordance between breakpoints detected from this segmentation analysis and, independently, structural variants detected in one or both kindred cells (Fig. 4).

Detection of chromosomal rearrangements

Chromosomal rearrangements were detected from discordantly mapped read pairs (including split reads) using a previously described algorithm (15). For 25× sequencing

data, we required at least 3 discordant reads within the range of the average insert size (300 bp) on both breakpoints to be considered as a rearrangement event. We estimate that this minimal threshold allows detection of 78% of genetic variants on each homolog (median across all samples, 53-87%). To exclude amplification artifacts that can appear as false variants with low allele fractions, we used more stringent thresholds for structural variant detection. For genome-wide structural variant detection, we set a conservative threshold by requiring at least 10 variant-supporting reads (including one split read) for each structural variant; with this stringent threshold, the median detection sensitivity is estimated to be 45% (29-58%). For detection of structural variants on the bridge chromosome, we relaxed the threshold to 5 or more variant-supporting reads to increase the median sensitivity to 66% (43-79%).

For 5-10 \times sequencing data, we required at least 2 discordant reads within the range of the average insert size (300 bp) on both breakpoints to be considered as a rearrangement event. We removed intra-chromosomal rearrangements with distance between breakpoints <150 kb based on the analysis of false positive rearrangement detection due to MDA chimeras (15). The only exception to the exclusion of short-range rearrangements was when such events were supported by additional evidence including copy-number alterations (fine-scale retention and loss) or chained templates (see below).

Assembly of rearrangement chains (TST jumps)

We assembled complex chains of short templates in three steps. We first performed read-based assembly around each breakpoint detected by discordant reads using SvABA (<https://github.com/walaj/svaba>) (65). SvABA can both identify the exact breakpoint location at base-pair level resolution and assemble longer contigs from multiple reads spanning the same breakpoint. Second, we classified regions containing more than 3 breakpoints within 10 kb as rearrangement hotspots and determined all the rearrangement junctions in these regions. Some of these junctions were already assembled by SvABA. Those missed by SvABA were usually located at or near regions of interspersed repeats (Alu or L1) and/or had very few supporting split reads (1-2); these events were manually assembled using both split reads and discordant mates. For most breakpoints in hotspots, we were able to assemble short contigs consisting of more than one rearrangement junction. In the last step, we merged these short contigs into rearrangement chains using three types of long-range linkage from discordant read pairs. The first type (most frequent) is junction-junction linkage provided by discordant pairs with each pairmate is mapped to a different rearrangement junction (i.e., a split read spanning both breakpoints of a rearrangement junction) in different contigs. The second type is junction-segment linkage provided by discordant pairs with one pairmate mapped to a rearrangement junction but the other mapped to a segment in a different contig. The third type (less frequent) is segment-segment linkage provided by discordant pairs with pairmates mapped to segments in different contigs. Across each contig, every rearrangement junction was supported by at least one read providing junction-junction type linkage or junction-segment linkage to a non-adjacent segment; segment-segment linkage was used for additional support.

Haplotype copy number variation from massively parallel single-cell sequencing

Cell barcode extraction and filtration. We first extracted the cell barcodes from the fastq files using the Cell Ranger DNA software (10X Genomics). After barcode extraction, all the sequencing reads were processed using the same workflow as other data. To identify cell barcodes associated with real DNA sequence from single cells, we generated a histogram of read counts for each cell barcode. The distribution of read counts was bimodal for all libraries and we excluded barcodes with lower read counts.

Haplotype coverage calculation. We calculated the haplotype coverage at each phased variant site from sequence reads associated with each eligible cell barcode using the same workflow as described above. As the sequence yield is very low for each cell, most variant sites only show coverage of one haplotype, resulting in haplotype fractions at most sites to be either 0 or 1. We therefore used a different approach to estimate haplotype copy number. First, we calculated the fraction of haplotype coverage as the percentage of phased variants showing coverage for either haplotype in each 1 Mb bin. Second, we estimated haplotype copy number from haplotype coverage f using

$$c = \ln(1 - f) / \ln(1 - f_0)$$

The rationale behind this calculation is as follows. Assume that the fraction of coverage of a single chromatid to be f_0 . As f_0 is a constant in each library, it can be estimated from the haplotype coverage across the genome when most chromatids have only one copy. When a homolog is present in only one copy, the expected fraction of non-coverage is $1 - f_0$. When a homolog has n independently amplified copies, the expected fraction of this homolog that is not covered by any of the n copies is given by $(1 - f_0)^n$.

Therefore, the DNA copy number can be estimated from

$$n = \ln(1 - f) / \ln(1 - f_0)$$

In practice, it is convenient to use $\ln(1 - f)$ as the signal of haplotype coverage after normalization by its median across the genome for both haplotypes. This normalization will produce haplotype copy number that is centered at 1 copy. For purposes of visualization (Fig. S20), we next applied a rolling median over the 1 Mb haplotype copy number calls and subsequently rounded to the nearest half-integer state. For whole-genome plots, the median filter window size was 21 bins (10 flanking bins on each side) and for Chr4 plots the window size was 5 bins. We finally performed hierarchical clustering using Ward's method to group together single cells with similar copy number states.

Sequencing of the sample from a patient with renal cell carcinoma

We identified a previously published clear cell renal cell carcinoma (89) where chromothripsis was associated with a chromosome fusion event that joined Chrs 3p and 5q. Structural variants were called from the short read sequencing data using the BRASS (breakpoint via assembly) algorithm (90). Additional DNA from the same tumor from fresh frozen tissue (DIAMOND study; Evaluation of biomarkers in urological disease - NHS National Research Ethics Service reference 03/018) was extracted and eluted into nuclease-free water. DNA (4.6 mg) was prepared using library construction kit SQK-LSK109 (Oxford Nanopore Technologies) in accordance with the manufacturer's protocol. The library was sequenced using the PromethION device (flow cell version FLO-PRO002 using MinKNOW software version 1.14.2, Oxford Nanopore Technologies) with standard run time. Sequencing reads were then aligned to GRCh37d5 using Minimap2 (91), allowing the reconstruction of split reads traversing the known

rearrangement hotspots. The average sequencing depth for covered regions was 15×. Data are shown for one long read (~40 kb) spanning a complex structural variant exhibiting the TST jump signature. Other reads across this region (not shown) independently supported the presence of the TST jump signature in this sample.

Supplementary Text

Mechanisms of chromosome bridge breakage

Our findings point to a mechanism of bridge breakage in human cells that requires mechanical forces from interphase actomyosin contractility. The simplest explanation of our findings is that bridge breakage either occurs directly as a result of these mechanical forces, or that force is required in conjunction with additional factors, such as nucleases or topoisomerases. It is also possible that some DNA breakage could occur if the nuclear envelope surrounding the bridge itself is fragile and prone to rupture, like that of micronuclei. This possibility is supported by the fact that during live-imaging, we often observe accumulations of GFP-BAF along the bridge, which visually resemble the focal accumulations of BAF that accompany primary nuclear rupture events (92) and the marked BAF accumulations on ruptured micronuclei (17). However, it is important to note that our data indicates that NE rupture alone is not sufficient for bridge breakage (Fig. S2D and Fig. S4). To definitively determine whether bridges exhibit nuclear envelope fragility will require correlative live-cell imaging and electron microscopy.

Chromosome bridges have long been studied in other organisms as well, including in plants and fungi, where bridge breakage appears to be more closely associated with the completion of mitosis (18, 19, 93-95). One major difference between mammalian cells and plants or fungi is the presence of cell wall: in plants and fungi, mitotic exit additionally involves the deposition of cell wall between the newly forming daughter cells (“septation” in fungi). Studies in fungi led to proposals that bridge breakage might be mediated by contraction of the actomyosin cytokinetic ring or during ESCRT-III-mediated abscission, and some (but not all) studies have suggested that these processes may be monitored by checkpoints to ensure clearance of DNA from the cleavage furrow before the complete separation of daughter cells (26, 96). However, recent studies have indicated that the breakage of dicentric chromosomes in yeast is primarily mediated by forces generated during cell wall deposition (94, 95). Therefore, bridge breakage in plants and fungi may differ in some respects from the mechanism we describe here in mammalian cells. However, we note that the genomic consequences may nevertheless be similar, at least in plants. Micronuclei are common in plants undergoing genome elimination after interspecies hybridization, which leads to chromothripsis (97) (by contrast, fungi undergo closed mitosis and thus rarely generate micronuclei-like structures). Given the similarities between micronuclei and chromosome bridges, we speculate that plants may also undergo chromothripsis-like phenomena after bridge formation.

Insights into patterns of structural variation in cancer genomes from in vitro evolution experiments

The original BFB model predicts that a series of foldback inversions and large copy number step transitions (i.e., large palindromes) are generated over multiple generations by cycles of chromosome bridge formation and breakage (18, 19). Our study

expands this model and identifies additional mutational mechanisms associated with bridge formation and breakage that can generate more complex chromosomal alterations, including chromothripsis. These findings explain why the simple palindromic pattern predicted by the classical BFB model is not commonly observed in isolation in cancer genomes (1, 11, 98).

The combination of live-cell imaging and single-cell sequencing of sibling cells enabled us to definitively distinguish rearrangement mechanisms involving DNA fragmentation [termed “cut-and-paste” in Li et al., the Pan Cancer Analysis of Whole Genomes or PCAWG study (42)] from DNA replication-based mechanisms (termed “copy-and-paste” in Li et al.). This information complements and clarifies the interpretation of chromosomal rearrangement patterns observed in cancer genomes.

Segmental copy-number alterations larger than 1 Mb. The most obvious mutational outcome of chromosome bridge breakage is the gain or loss of large terminal chromosome segments [“telomere-bounded” copy-number alterations as defined by Zack et al. (99)]. Breakage is expected to occur at any location within the bridge, giving rise to large segmental CNAs (often 10 Mb or above), as observed in our single-cell sequencing experiments and consistent with the size of telomere-bounded CNAs in cancer genomes [median size 20 Mb (99)]. Our findings additionally revealed a mechanism to explain large CNAs affecting internal chromosome segments (Fig. 2C), which are also observed in cancer genomes (99) but whose origin has remained unclear. These internal CNAs can be explained by independent breaks on two separate dicentric chromosomes after the formation of a chromosome bridge. Additionally, bridge breakage may provide an explanation for at least some unbalanced translocations observed in cancer. This can occur from breakage of bridges containing multiple chromosomes, whose subsequent re-ligation can produce inter-chromosomal rearrangements (Fig. 3C).

Medium-sized (~0.1-1 Mb) segmental copy-number alterations and local jumps. We found that bridge breakage can result in local chromosome fragmentation. The reciprocal distribution of these fragments gives rise to a pattern of oscillating copy number in both daughter cells (between 0 and 1 in one daughter, and between 1 and 2 in the other daughter). The number of fragments varies, but is generally in the range of 1-10. Rejoining of these fragments results in chromosome rearrangement patterns resembling “local n -jumps” or “local-distant” rearrangements (hereafter “local jumps”) described in the PCAWG study (42). The frequent association of local jumps with copy-number gains was invoked to support the inference that local jumps have a replicative origin (“copy-and-paste”). However, our analysis of sibling cells demonstrates that bridge fragmentation and re-ligation (“cut-and-paste”) can also generate these rearrangements with segmental gain. Therefore, we note that although bridge fragmentation often results in fewer copy-number changes than is operationally used to define chromothripsis (usually $\geq \sim 10$ copy-number oscillations), at least some local jumps and chromothripsis may share a common underlying mechanism.

The TST jump signature. Although our evidence suggests that many of the larger-scale chromosomal rearrangements we identified after bridge breakage do not have a replicative origin, we did identify a rearrangement signature, the TST jumps, that appears to be generated by template switching replication errors. We observed chains of insertions that originated from 1-10kb hotspots, distributed across the segment(s) of the chromosome(s) inferred to be in the bridge. Both the strong breakpoint clustering and the

sharply restricted size distribution of the insertions are consistent with an origin from DNA replication errors. We suggest that TST jumps may be generated in a variety of aberrant nuclear structures, as we previously identified features of the TST jump signature in the progeny of a micronucleated cell (15).

Tandem chains of rearrangements have been identified in a variety of contexts, including in cancer genomes and in non-homologous end joining-deficient cells after telomere crisis (42, 65, 66). Although it is appealing to speculate that these chains have a similar underlying mechanism to the TST jumps, many of the described insertions originate from multiple different chromosomes, rather than the predominantly local origin we observed. The size distribution of insertions reported in the PCAWG study (42) is also broader than we observed and appears to be multimodal. Currently, we favor the interpretation that TST jumps reflect a specific mutational process corresponding to a subset of the chained insertions described in cancer genomes.

We also note that the identification of template chains requires local sequence assembly using both split and unmapped reads, which has not routinely been done in cancer genome analysis. We also note that the size of shotgun sequencing fragments (200-500 bps) imposes a limit on the size of short templates that can be identified. These considerations suggest that the frequency of chained insertions may be underestimated in cancer genomes, and they motivate further studies employing long-read sequencing, as we have done here for a renal cell carcinoma genome.

In summary, results from our in vitro experiments not only recapitulate multiple patterns of chromosomal rearrangements seen in cancer genomes, but add mechanistic insight that explains, refines, and clarifies the interpretation of mutational patterns identified by data-driven cancer genome analysis.

Supplemental Figure 1

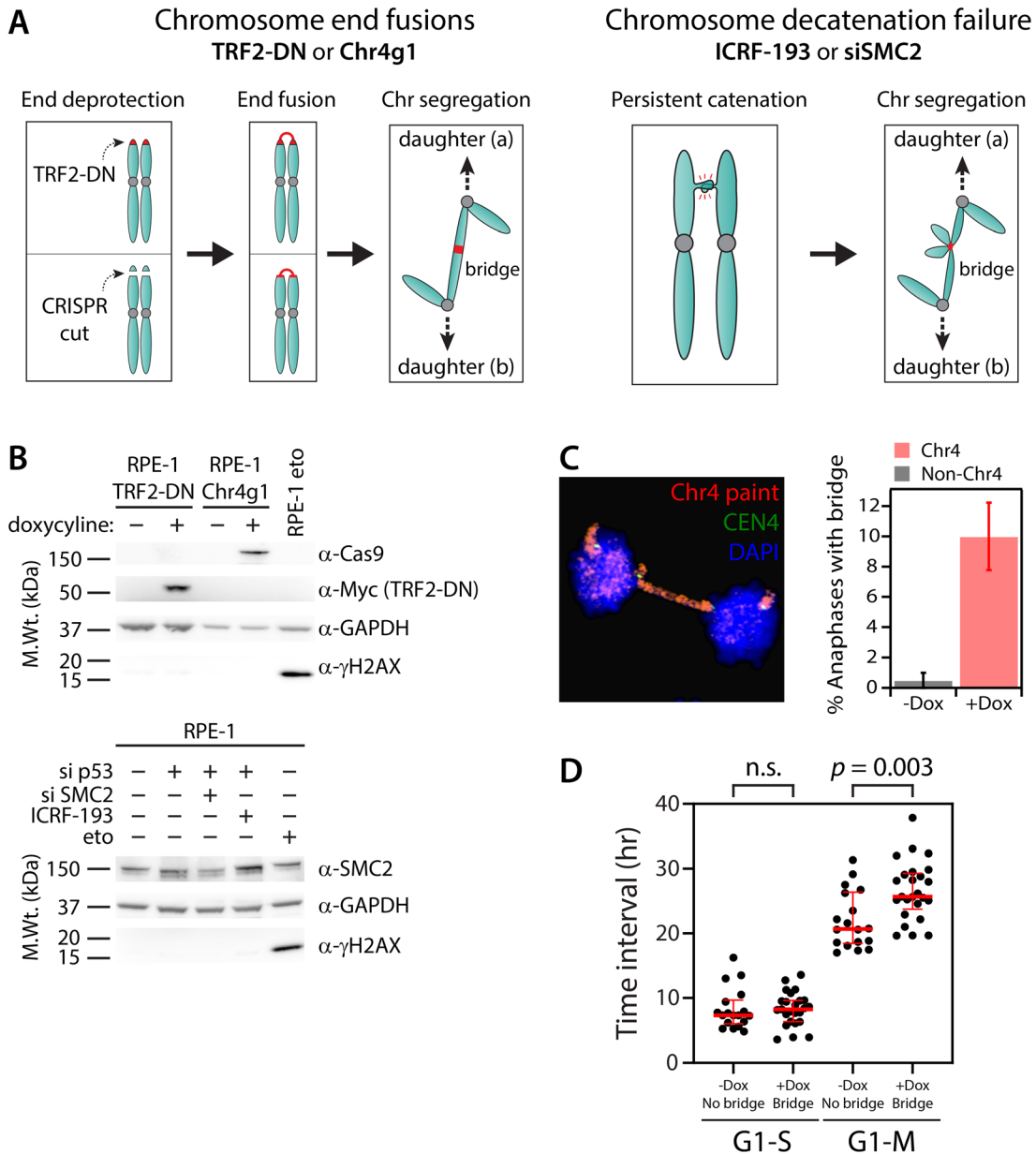


Fig. S1. Four methods to induce the formation of chromosome bridges.

(A) Schematic of experimental methods to generate chromosome bridges. **Left:** TRF2-DN or Chr4 CRISPR methods produce uncapped chromosome ends, which can be fused by the DNA repair machinery to generate dicentric fusions. Note that Chr4 CRISPR can cut both ends of both homologs of Chr4, so like TRF2-DN, chromosome fusions can involve either or both termini of the chromosome. **Right:** low-dose topoisomerase II inhibition (ICRF-193) or partial knockdown of condensin subunit SMC2 interferes with chromosome decatenation during mitosis.

- (B) Control western blots for experimental methods used to induce chromosome bridges. **Top:** Western blot of RPE-1 cell lines showing doxycycline-inducible expression of Myc-tagged TRF2-DN (α -Myc), or Cas9 (α -Cas9). Note that there is no detectable DNA damage in either case (blot for γ -H2AX) ~40 hr post-induction. **Bottom:** Western blot showing that partial knockdown of condensin (α -SMC2, depleted to ~70% of control levels), or low-dose topoisomerase II inhibition (ICRF-193) also did not elicit detectable DNA damage. In both panels, α -GAPDH serves as a loading control, and RPE-1 cells treated with etoposide (eto) are a positive control for γ -H2AX.
- (C) Terminal Chr4 CRISPR-mediated DNA breaks generate chromosome bridges containing Chr4. **Left:** Representative image showing that Cas9 induction produces Chr4 bridges as observed by fluorescence in situ hybridization with chromosome 4 “paint” (red) and chromosome 4-specific centromere (CEN4, green) probes, counterstained with DAPI (blue). **Right:** quantification of images as shown at left. Bars show the frequency of anaphase bridges that were positive (red shading) or negative (gray) for Chr4 paint. Error bars indicate uncertainty based on counting statistics; $n = 200$ anaphase figures examined per condition.
- (D) Optimized conditions for transient expression of TRF2-DN and live-cell imaging minimize previously reported effects on cell cycle progression. Cells were untreated (-Dox, No bridge), or treated with doxycycline for 14 hours, followed by washout, to induce TRF2-DN expression and chromosome bridge formation (+Dox, Bridge). Live-cell imaging was performed using GFP-BAF to visualize cell nuclei and chromosome bridges, and an mCherry-tagged fragment of Geminin was used to monitor S phase onset (83). Cells were imaged by widefield fluorescence microscopy using a 20 \times objective, with four z-planes acquired every 10 min. Plot shows time from the beginning of G1 until S phase initiation, and from G1 until the second mitosis (each dot represents one cell). No significant difference was observed in the timing of S phase initiation, and there was a minor ~1.2-fold delay in total cell cycle duration (p -values calculated by Mann-Whitney test).

Note: these cell cycle transit times in our imaging conditions contrast sharply with those reported in a prior study that utilized the same cells (28). In this prior study (48 hr TRF2-DN induction; 60 \times objective confocal imaging with 9-35 z-planes acquired every 10 min), only 20% of cells entered S phase within 20 hr after mitosis. Our use of the GFP-BAF reporter rather than GFP-H2B, which is lost from bridges under mechanical force, enabled us to employ lower light exposure conditions.

Supplemental Figure 2

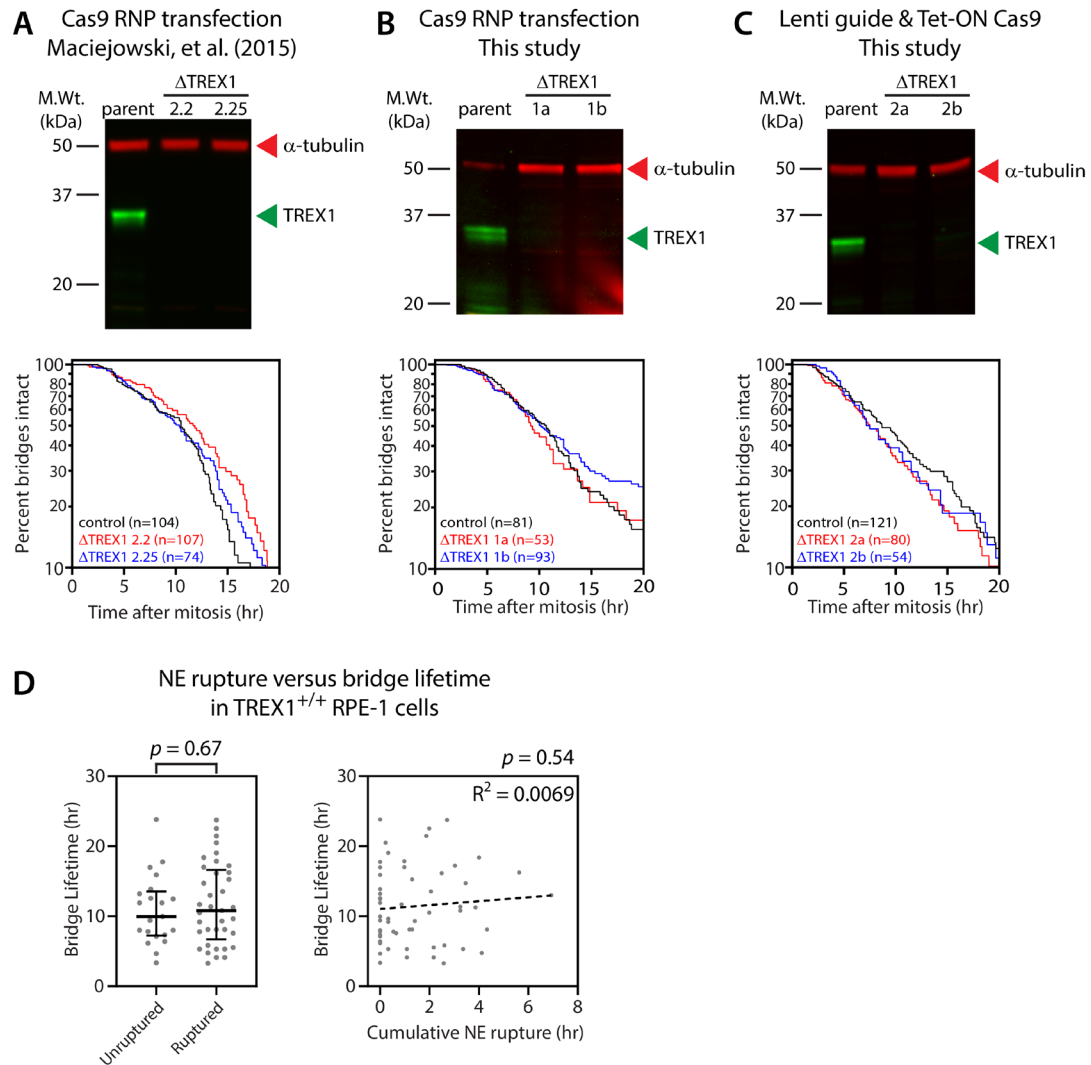


Fig. S2. Neither TREX1 deletion nor nuclear envelope disruption significantly affects the timing of chromosome bridge breakage.

- (A) **Top:** Western blot showing TREX1 knockout (α -TREX1, green) in published clones generated by transfection of Cas9/sgRNA ribonucleoprotein complex (28). α -tubulin (red) is used as a loading control. **Bottom:** Bridge lifetime curves for control and two TREX1 knockout clones are shown (control versus clone 2.2, $p = 0.62$; control versus clone 2.25, $p = 0.098$; p -values from Mann-Whitney test). For all lifetime experiments, widefield imaging was performed with a 20 \times objective at 10 min intervals, acquiring a z-stack of four images at 2- μ m spacing.
- (B) As in (A), Western blot showing TREX1 knockout clones generated in this study by Cas9/sgRNA transfection. Plot shows bridge lifetime curves for control and TREX1 knockout clones (control versus clone 1a, $p = 0.59$; control versus clone 1b, $p = 0.90$).
- (C) A second method to generate TREX knockout clones in RPE-1 cells expressing inducible Cas9. As in (A), Western blot showing TREX1 knockout clones

generated in this study by constitutive sgRNA expression (lentiviral delivery) with transient, doxycycline-inducible expression of Cas9 (79). Rather than by transient TRF2-DN expression (A-B), chromosome bridges were induced with ICRF-193. Bridge lifetime curves for control and two TREX1 knockout clones are shown (control versus 2a, $p = 0.38$; control versus 2b, $p = 0.75$).

(D) Nuclear envelope rupture is neither required for nor accelerates chromosome bridge breakage. **Left:** Comparison of bridge lifetime for cells that experienced nuclear envelope disruption (Ruptured), versus cells where no nuclear envelope disruption was detected (Unruptured) prior to bridge breakage ($p = 0.67$, Mann-Whitney test). Middle line denotes median bridge lifetime and whiskers indicate interquartile range. **Right:** Plot shows no significant correlation between bridge lifetime and cumulative nuclear envelope (NE) rupture duration ($p = 0.54$). Bridges were induced by transient TRF2-DN expression and visualized by GFP-BAF; nuclear envelope integrity was monitored using RFP-NLS.

Supplemental Figure 3

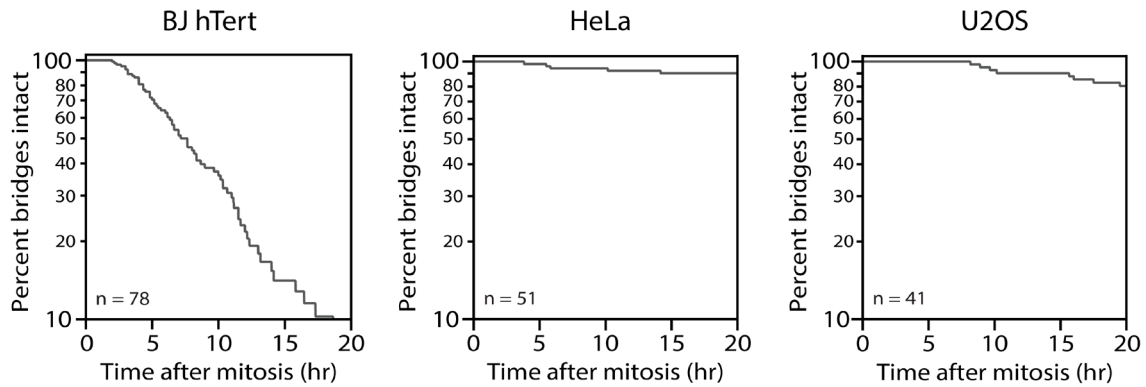


Fig. S3. Cell motility and bridge extension correlates with bridge breakage.

Plots show chromosome bridge lifetimes in BJ hTert foreskin fibroblasts (Left), HeLa cells (Middle), and U2OS cells (Right). BJ cells exhibit high motility and break chromosome bridges in interphase, whereas HeLa and U2OS cells show low motility and rarely break bridges during interphase. Bridges were induced in BJ hTert cells with low-dose topoisomerase II inhibition (ICRF-193); HeLa and U2OS cells were untreated, and spontaneously formed bridges were analyzed.

Supplemental Figure 4

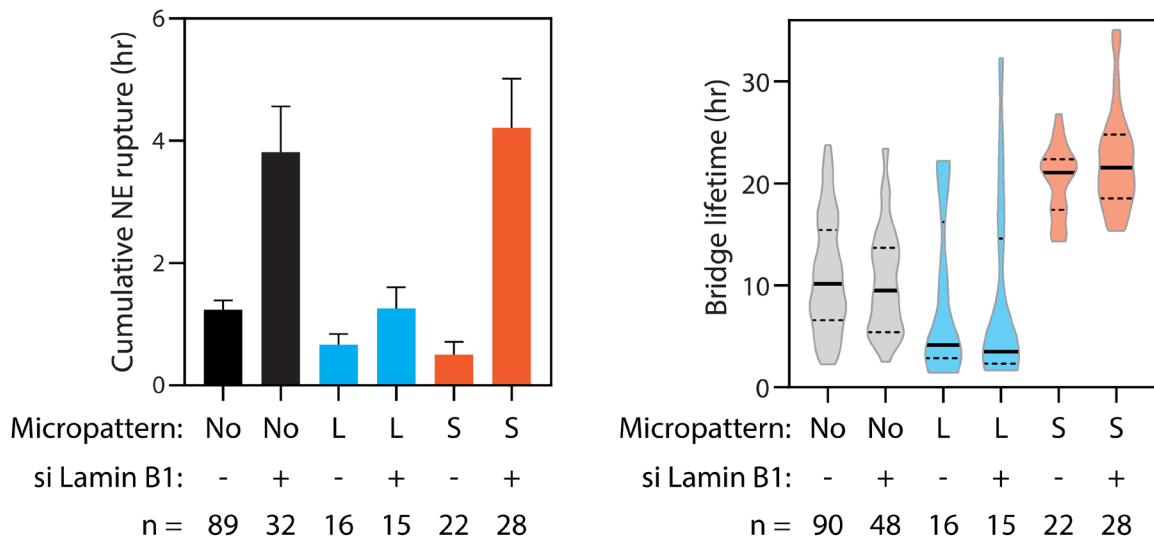


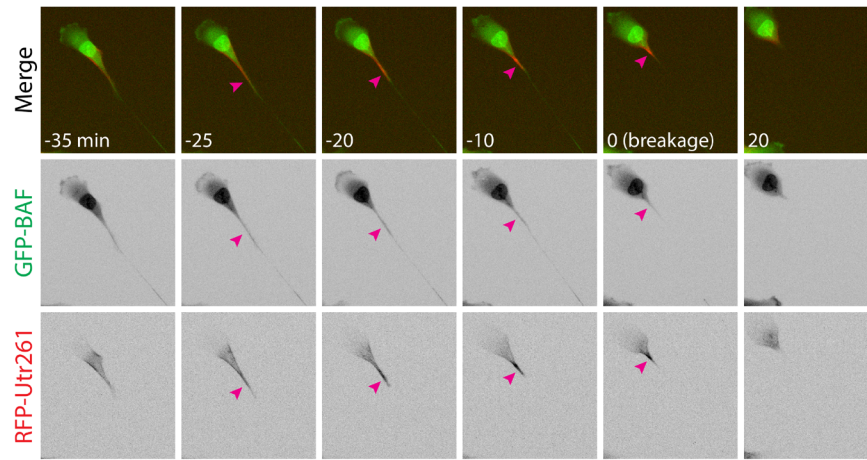
Fig. S4. Increased NE rupture time does not shorten bridge lifetime.

Left: Plot shows cumulative NE rupture time (assessed by RFP-NLS), comparing Lamin B1 knockdown cells and controls. “No” micropattern indicates unrestricted cells migrating freely in 2D culture; “L” or “S” refer to cells plated on long or short micropatterns, respectively. N-values are number of cells analyzed per condition. Bars represent the mean \pm standard error of the mean. No micropattern \pm siLB1, $p < 0.0001$; Long micropattern \pm siLB1, $p = 0.22$; Short micropattern \pm siLB1, $p < 0.0001$.

Right: Violin plot showing distribution of bridge lifetimes (time from mitosis until bridge breakage or entry into the next mitosis), with or without Lamin B1 knockdown, in the conditions described above. Middle solid line shows median bridge lifetime, dashed lines show interquartile range. N-values are number of cells analyzed per condition. No micropattern \pm siLB1, $p = 0.69$; Long micropattern \pm siLB1, $p = 0.65$; Short micropattern \pm siLB1, $p = 0.22$. Note that the observed decrease in bridge lifetime on long micropatterns relative to unrestricted cells migrating freely in 2D culture is due to the effects of fibronectin on increasing cell contractility (see Fig. 1F).

Supplemental Figure 5

A RPE-1 cells with chromosome bridges induced by TRF2-DN



BJ hTert cells with chromosome bridges induced by ICRF-193

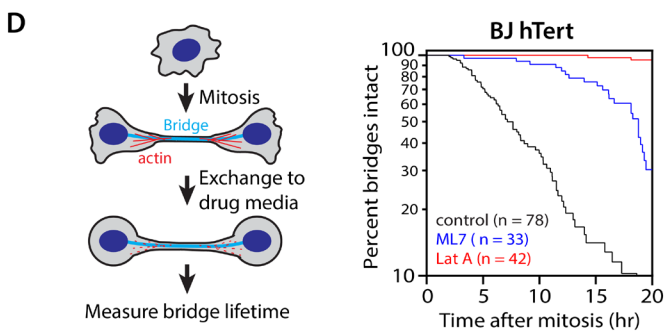
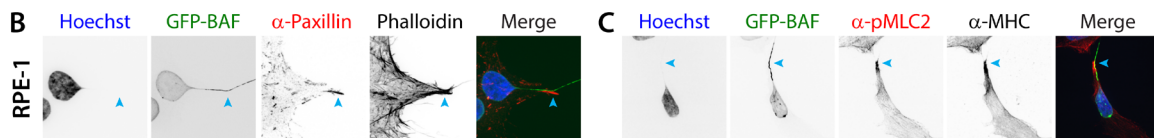
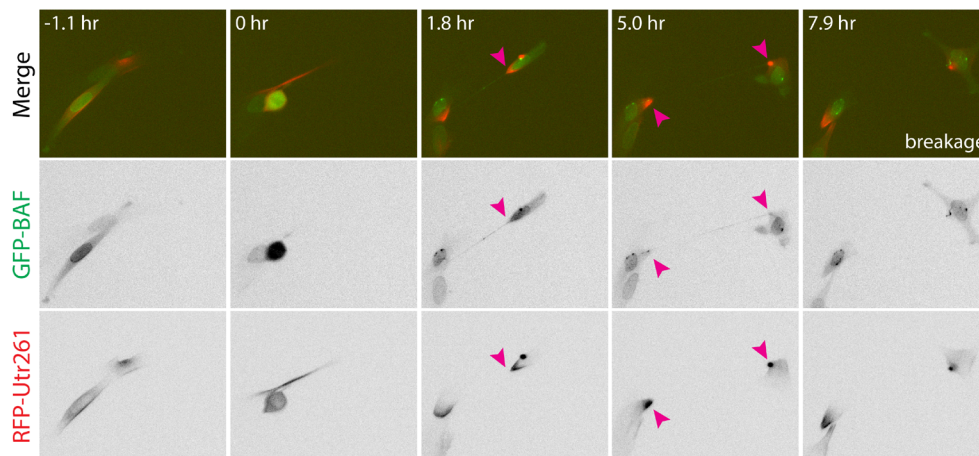


Fig. S5. Like RPE-1 cells, bridge breakage in BJ hTert cells requires actin and myosin II-dependent contractility.

- (A) Robust actin concentration at the base of a chromosome bridge prior to breakage. **Top:** Time-lapse images show actin dynamics (RFP-Utr261, red) during chromosome bridge breakage (GFP-BAF, green) in RPE-1. Actin accumulated and appeared to contract (-25 to -10 min, magenta arrowheads) leading to bridge breakage (0 min). **Bottom:** As above, for BJ hTert. A cell divided to form a bridge between the two daughters (-1.1 to 1.8 hours). Large actin accumulations (magenta arrowheads) formed at the “base” of the bridge in each daughter cell (1.8 to 5.0 hours), whose apparent contraction preceded bridge breakage (7.9 hours).
- (B) Representative images (RPE-1 cells) show large focal adhesions (α -Paxillin) and actin fibers (phalloidin) at the bent region of a chromosome bridge (GFP-BAF), indicated by cyan arrowheads.
- (C) As in (B), representative accumulation of contractile myosin II (α -myosin heavy chain, MHC; α -phospho-myosin light chain 2, pMLC2) at the transition between taut and slack segments of a chromosome bridge.
- (D) **Left:** schematic of actin and myosin-II inhibition experiments. During live-cell imaging, cells were allowed to divide and form bridges prior to exchange into drug medium. Bridges that were intact at the time of drug addition were analyzed. **Right:** plot of bridge lifetime for BJ hTert cells treated with DMSO (control), Latrunculin A, or ML7, after bridge formation and extension, as in Fig. 1E. Bridges were induced by low-dose ICRF-193. Control data are from Fig. S3.

Supplemental Figure 6

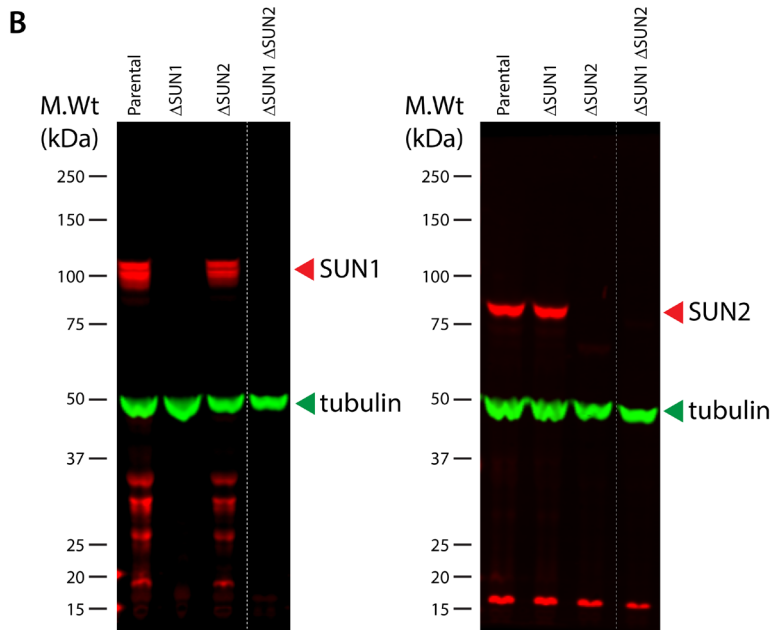
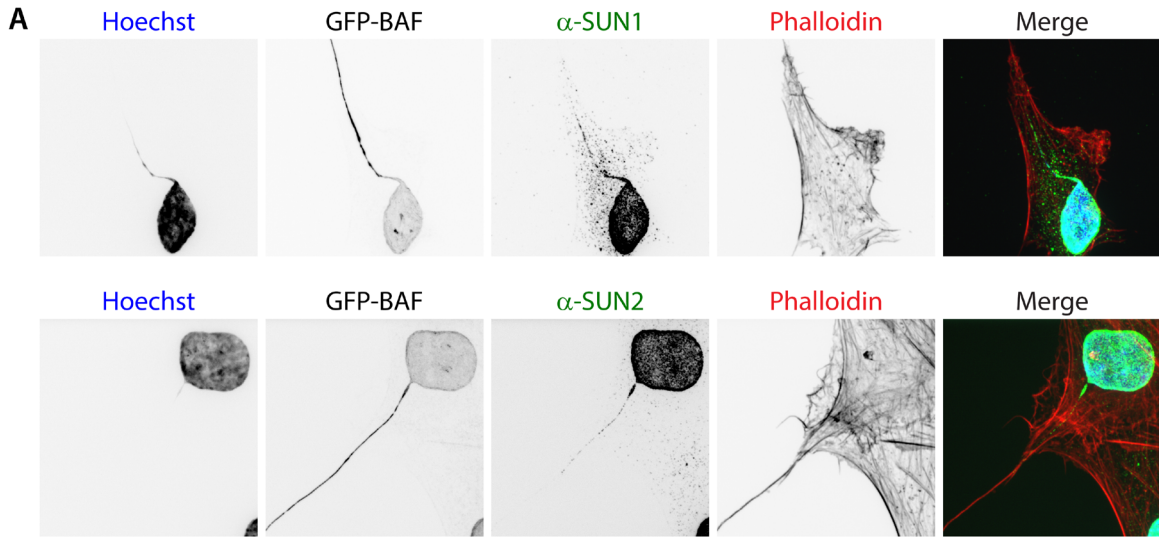
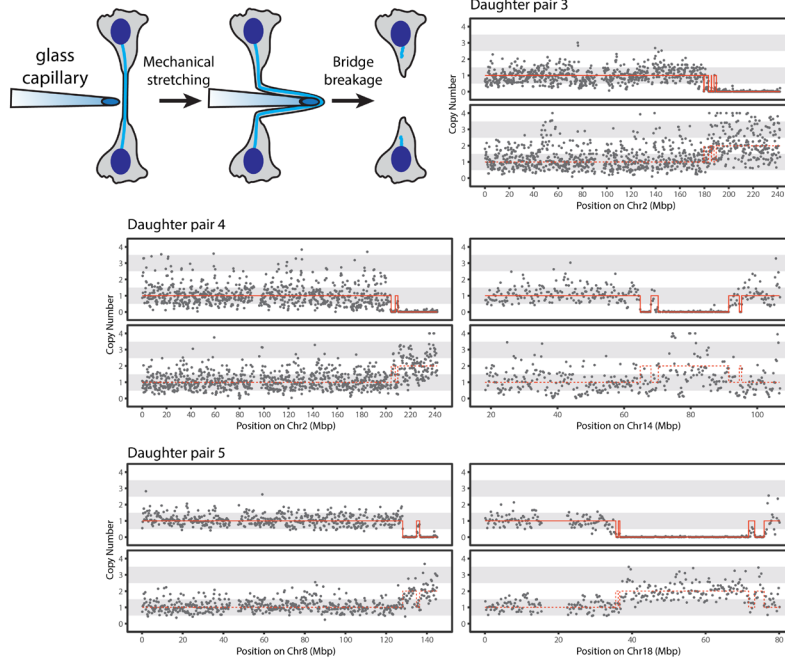


Fig. S6. LINC complex components SUN1 and SUN2 localize to chromosome bridges.

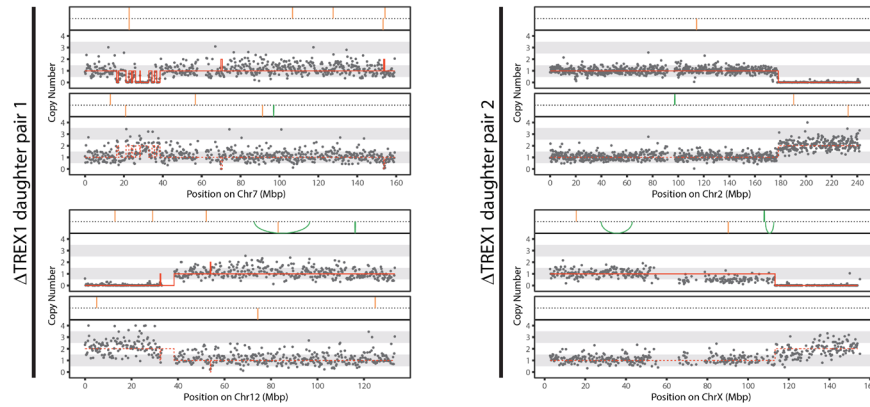
- (A) **Top:** Immunofluorescence images showing localization of SUN1 (green) to the base of chromosome bridges near cytoplasmic actin filaments (phalloidin, red). **Bottom:** Immunofluorescence images showing similar localization pattern for SUN2. For both SUN1 and SUN2, the staining patterns were validated to be specific because the signal was absent from knockout clones.
- (B) Western blots showing CRISPR knockout of SUN1 (left, red) and/or SUN2 (right, red) in RPE-1 cells; α -tubulin serves as a loading control (green). Dotted white lines indicate digital merging of non-adjacent lanes from the same blot.

Supplemental Figure 8

A Mechanical bridge breakage



B Spontaneous bridge breakage in Δ TREX1 cells



Supplemental Figure 8

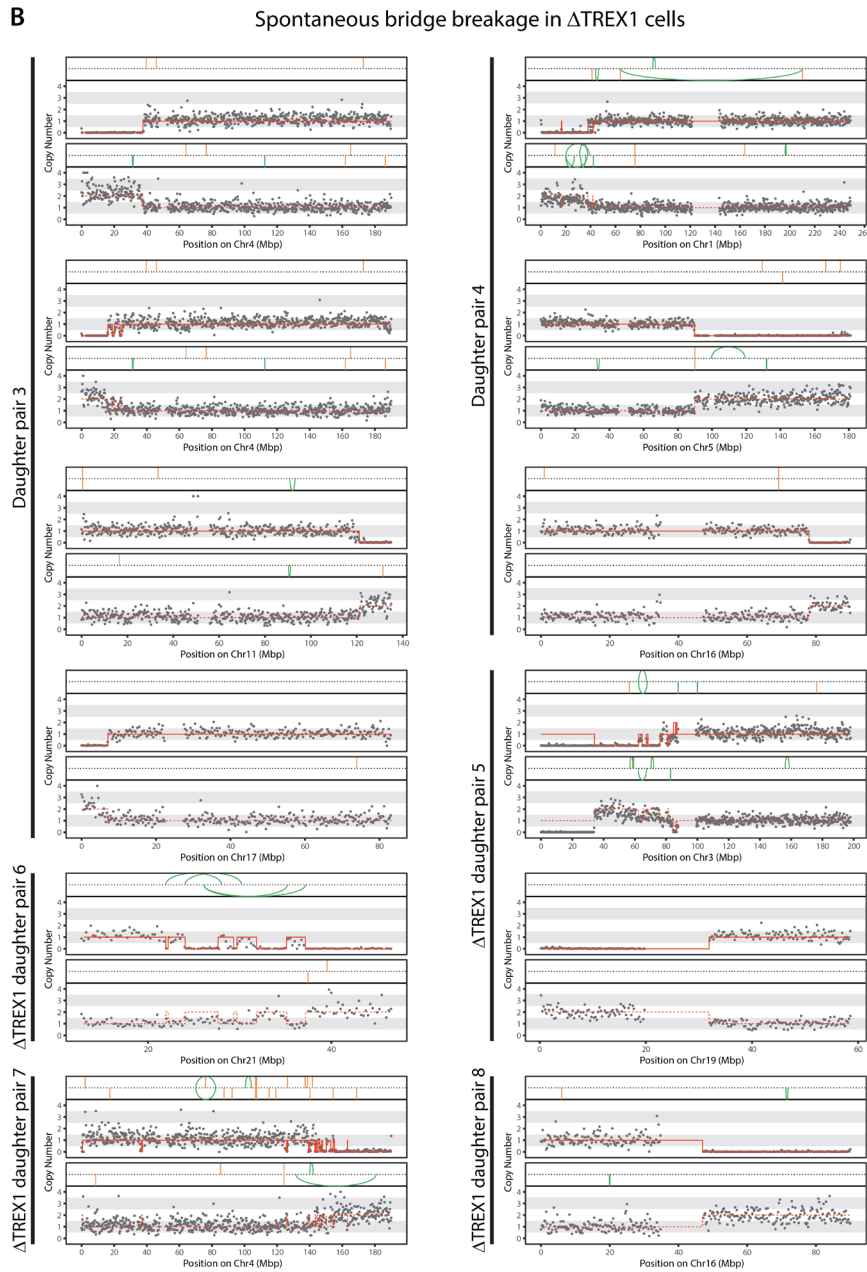


Fig. S8. Mechanical bridge breakage generates local fragmentation of bridge chromosomes (additional examples).

(A) Mechanical bridge breakage using a glass capillary produces copy number alterations consistent with simple breakage and local fragmentation. Schematic shows the experimental design. As in Fig. 4A, plots show copy number (gray

- dots, 1-Mb bins) and copy number segments (solid and dashed red lines) for mechanically broken bridge chromosomes. Note that, in contrast to the samples with spontaneous bridge breakage, in this experiment the cells were isolated immediately after bridge breakage, without time for DNA ligation to generate rearrangements.
- (B) TREX1 knockout cells undergo bridge breakage with localized fragmentation, as observed in control cells. Bridges were induced with transient TRF2-DN expression. Whole-chromosome CN/SV plots (as in Fig. 4B) are shown for each chromosome in the TREX1 knockout bridge pair samples with reciprocal copy number alterations. Note that Chr17 from Δ TREX1 daughter pair 3, and Chr21 from Δ TREX1 daughter pair 6 are also shown in Fig. 4B.

Supplementary Figure 9

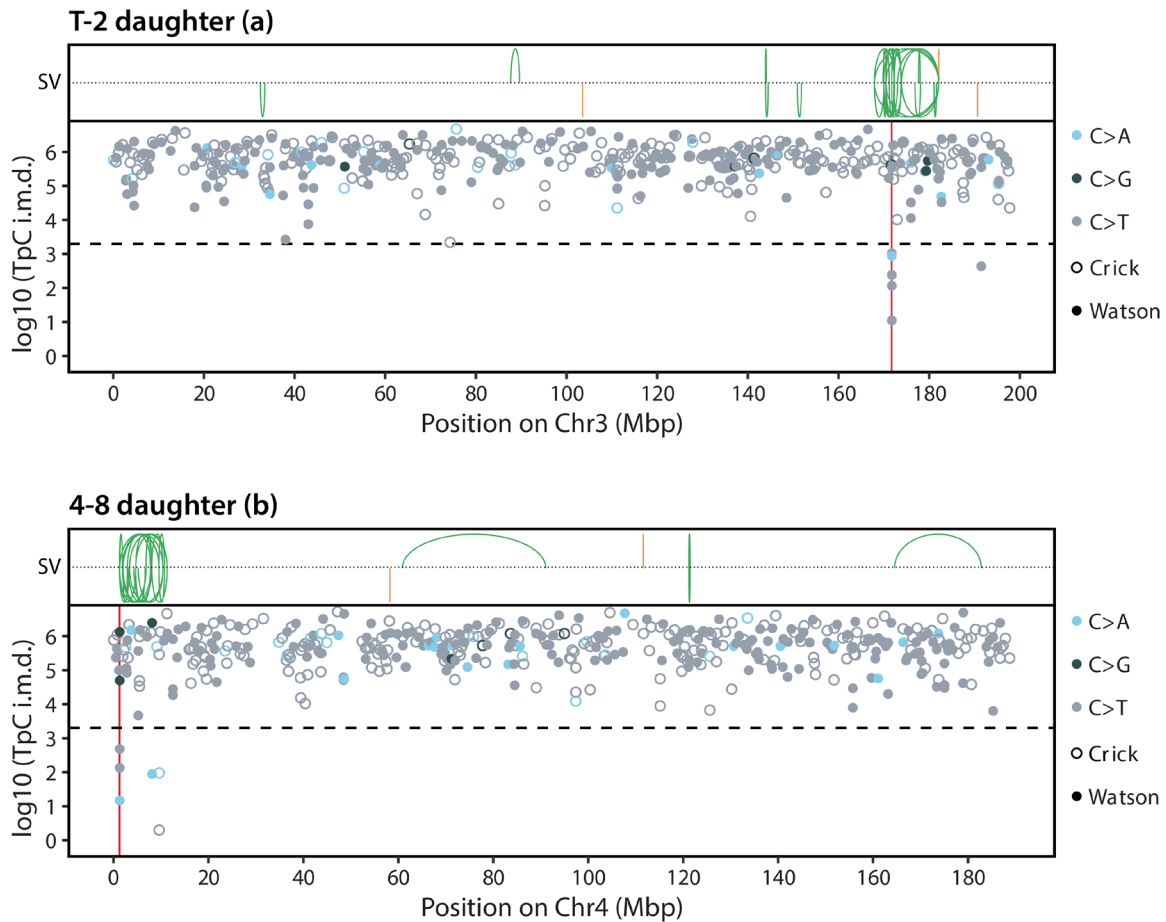


Fig. S9. Kataegis in single cells after chromosome bridge breakage.

Clusters of strand-coordinated mutations in the TpC dinucleotide context (“kataegis”) detected by single-cell sequencing. “Rainfall” plots representing inter-mutation distance (i.m.d.) between adjacent TpC *de novo* single nucleotide variants (SNVs) are shown for T-2 daughter (a) and 4-8 daughter (b). Each SNV is shown by a point, colored to indicate the class of mutation (C>A, C>G, C>T), assuming the alteration occurred on the C of the C:G base pair. In order to show the relative strand coordination of the kataegis clusters, points are filled according to whether the mutated C is present on the top (“Watson”) or bottom (“Crick”) strand of the reference genome. Clusters of at least 4 SNVs with i.m.d. ≤ 2 kb (dotted line) are highlighted with red vertical bars. Above each rainfall plot, chromosome rearrangements (SV) are shown as in Fig. 3.

Supplemental Figure 10

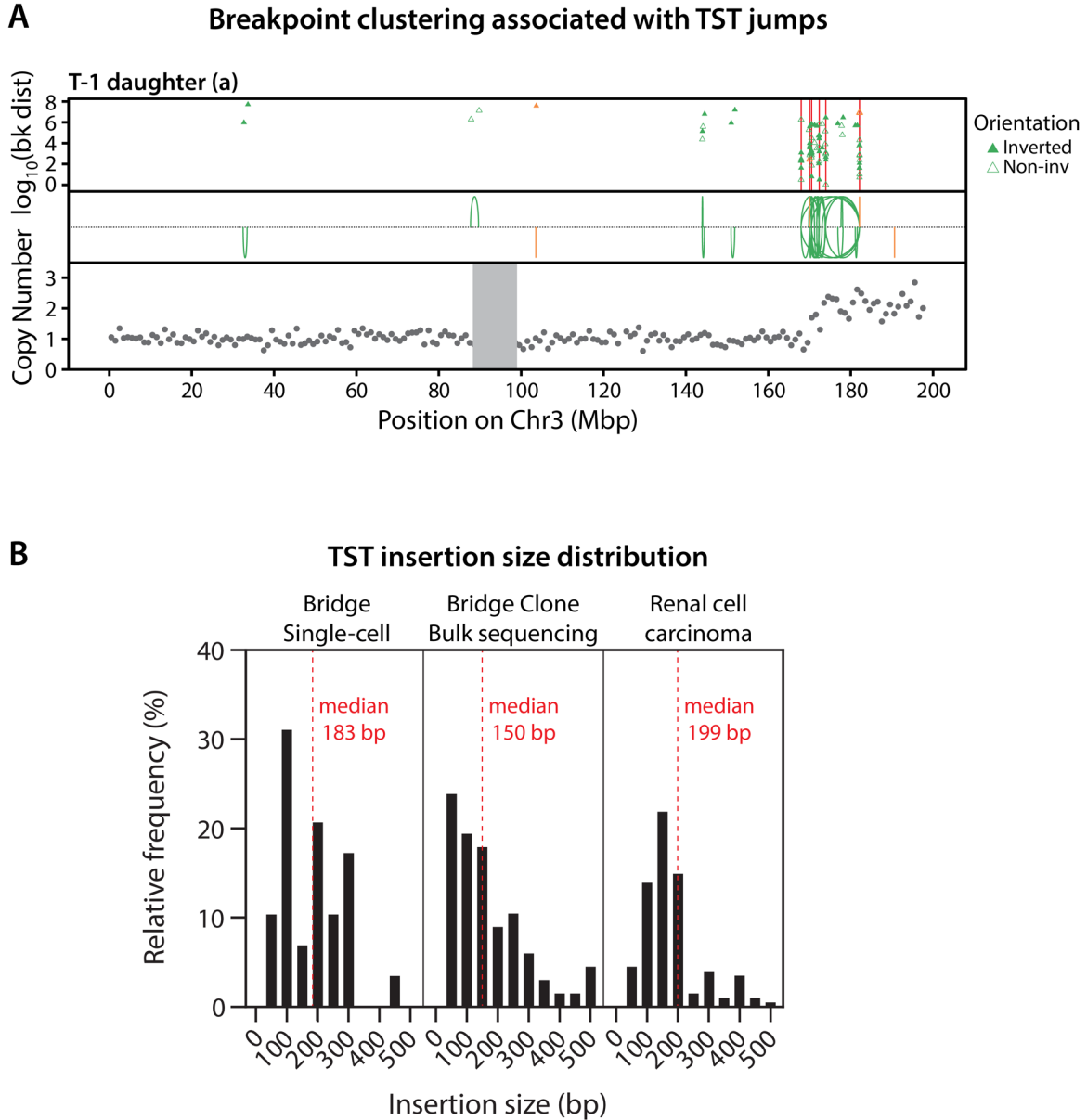


Fig. S10. Features of the Tandem Short Template jump signature.

- (A) Extreme breakpoint clustering near the site of chromosome bridge breakage. Whole-chromosome view of the sample shown in Fig. 5A: copy number (1-Mb bins, gray dots) and rearrangements (green curves and orange lines) are shown as in Fig. 3. Uppermost (“rainfall”) plot shows the distance between adjacent breakpoints (\log_{10} scale); clustering of breakpoints is indicated from marked drops in inter-breakpoint distance (red lines correspond to rearrangement hotspots shown in Fig. 5A).
- (B) Chains of short insertions identified in bridged cells and in a primary tumor sample exhibit a similar fragment size distribution. Histograms show the size

distribution for TST insertions from single-cell sequencing of a daughter cell after bridge breakage (left; data from Fig. 5A), from bulk sequencing of progeny derived from a single cell after bridge breakage (center; data from Fig. 5B), and from long-read sequencing data from the renal cell carcinoma sample (right; data from Fig. 5C). Red dashed lines show median fragment size.

Supplemental Figure 11

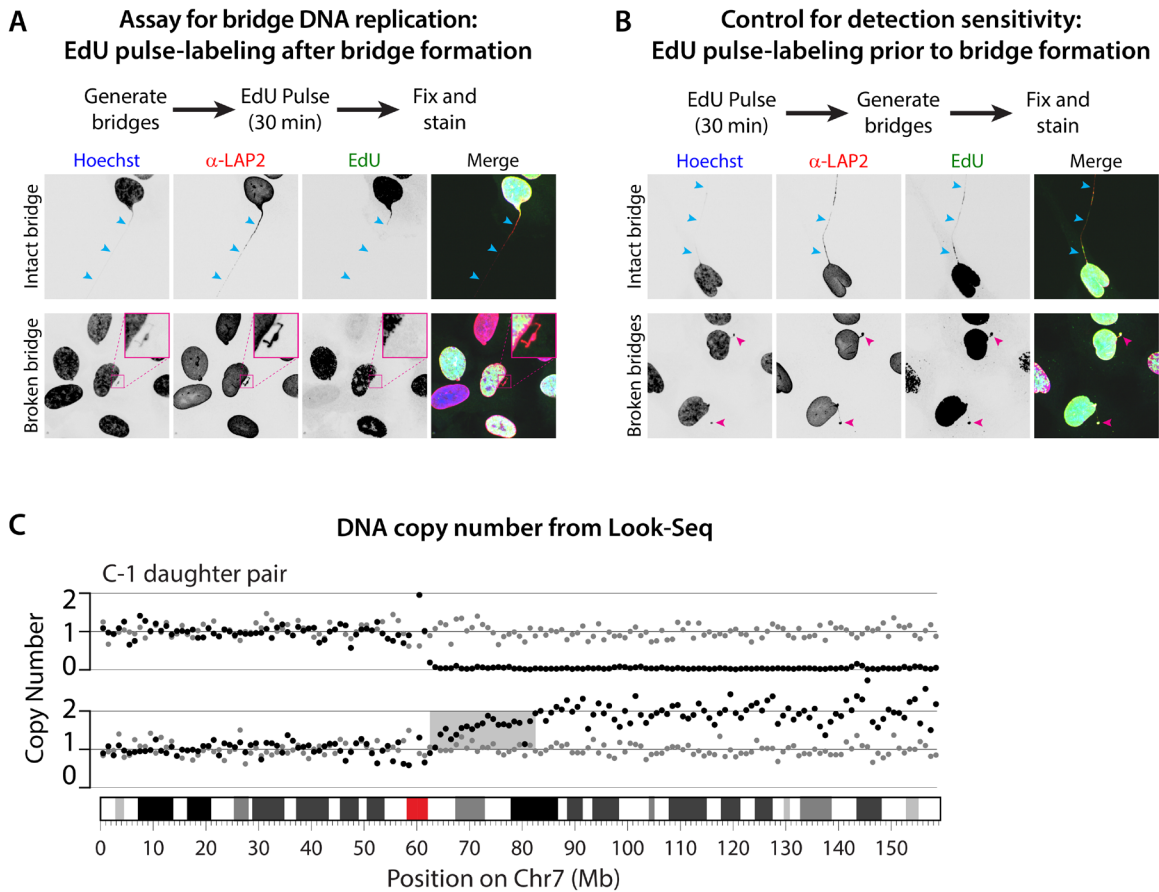


Fig. S11. Defective DNA replication within chromosome bridges.

- (A) Cytological observation of under-replication of chromatin in bridges. **Top:** scheme of the experiment. Bridges were generated (TRF2-DN) and EdU pulse-labeling was performed in S phase. **Bottom:** Representative images of cells with intact bridges (cyan arrowheads) or broken bridges, visualized by staining for LAP2. Insets (magenta boxes) show the broken bridge stub.
- (B) Control experiment demonstrating that EdU can be detected within chromosome bridges despite their small amount of DNA. **Top:** scheme of the control experiment. EdU pulse-labeling was performed in S phase prior to bridge induction (TRF2-DN), so that all chromosomes were labeled within a normal primary nucleus. Pre-labeled bridges then formed in the next mitosis and were imaged in the following interphase. **Bottom:** images as in (A).
- (C) Example of interphase under-replication of DNA in bridges detected by single-cell sequencing. Shown are copy number plots for the bridge haplotype (black dots) and the control, non-bridge haplotype (gray dots). Gray shading: region of under-replication of bridge haplotype. The mean copy number in this 20-Mb region for the bridge haplotype, $CN=1.56$, is lower than the expected gain ($CN=2$) for this region. Partial retention of that haplotype in the sister cell (median CN of bridge haplotype=0.05) cannot explain the extent of “missing” DNA.

Supplemental Figure 12

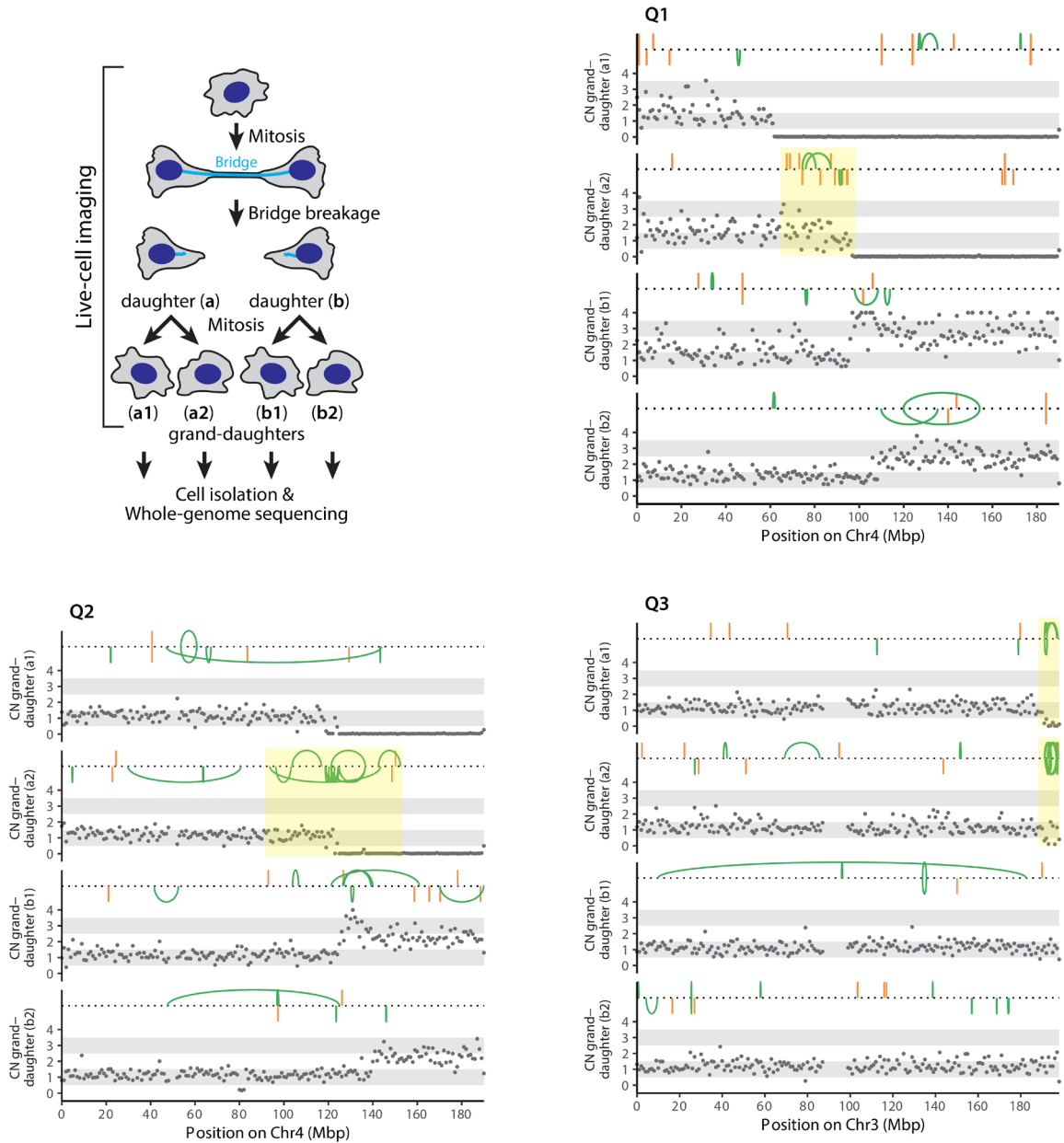


Fig. S12. Frequent chromothripsis after cells with broken bridges pass through mitosis.

Apparent high frequency of complex rearrangement affecting bridge chromosomes after passing through the next mitosis. **Top Left:** Schematic of the two-generation Look-Seq experiment. A cell divides to form a bridge between daughter cells in the first generation. After the bridge breaks, the daughter cells with broken bridge stubs divide, generating four “granddaughters” in the second generation that were isolated for whole-genome sequencing. **Top Right and Bottom panels:** DNA copy number plots (gray dots: 1-Mb bins) with chromosome rearrangements, as in Fig. 3, for sets of four grand-daughter cells.

In each case, yellow shading indicates the region of the bridge chromosome with chromothripsis.

Supplemental Figure 13

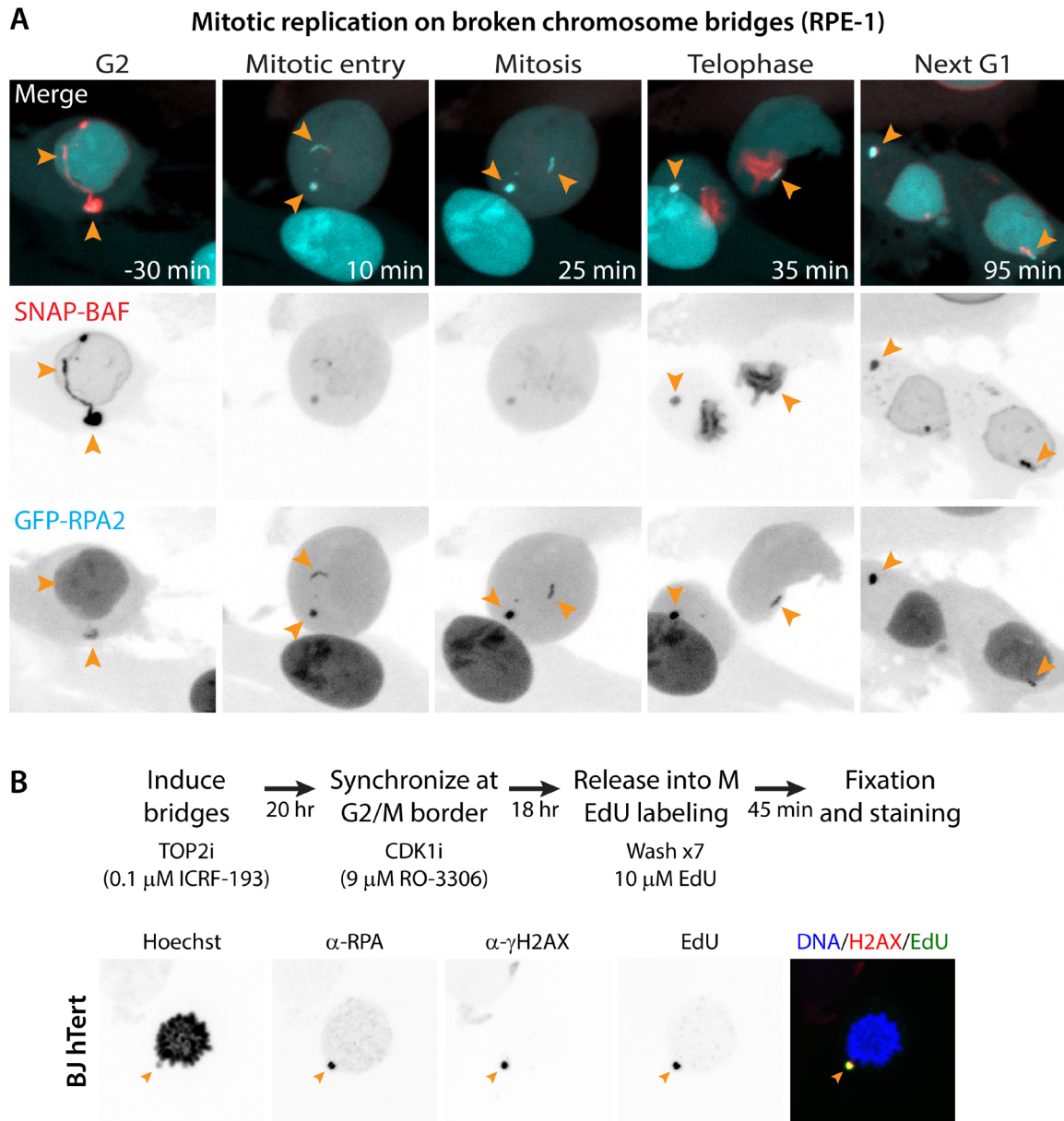


Fig. S13. DNA damage coincident with a burst of DNA replication on broken bridge chromosomes upon entry into mitosis.

(A) Representative time-lapse images show a burst of mitotic DNA replication specifically on a chromosome from a broken bridge (RPE-1 cells). Mitotic replication was visualized by GFP-RPA2; bridge with SNAP-BAF. During mitosis, high activity of Vaccinia Related Kinase (VRK) inactivates DNA binding by BAF (*100*) (10 to 25 min). Orange arrowheads indicate the broken bridge chromosome. Note that an unrelated interphase cell migrates through the bottom

of the field of view in several frames (10 to 35 min). Confocal imaging was performed with a 40× objective, 7 z-slices at 1-μm spacing, acquired every 5 min.

(B) Mitotic replication on chromosomes from bridges in BJ hTert cells. **Top:** scheme of the experiment. Cells were treated with ICRF-193 to generate bridges, followed by G2 synchronization with RO-3306. Cells were then released into mitosis in the presence of EdU, fixed, and then stained to detect EdU, γ -H2AX, and RPA1. **Bottom:** Representative images showing mitotic replication (EdU and RPA1) and associated massive DNA damage (γ -H2AX) on the bridge chromosome.

Supplemental Figure 14

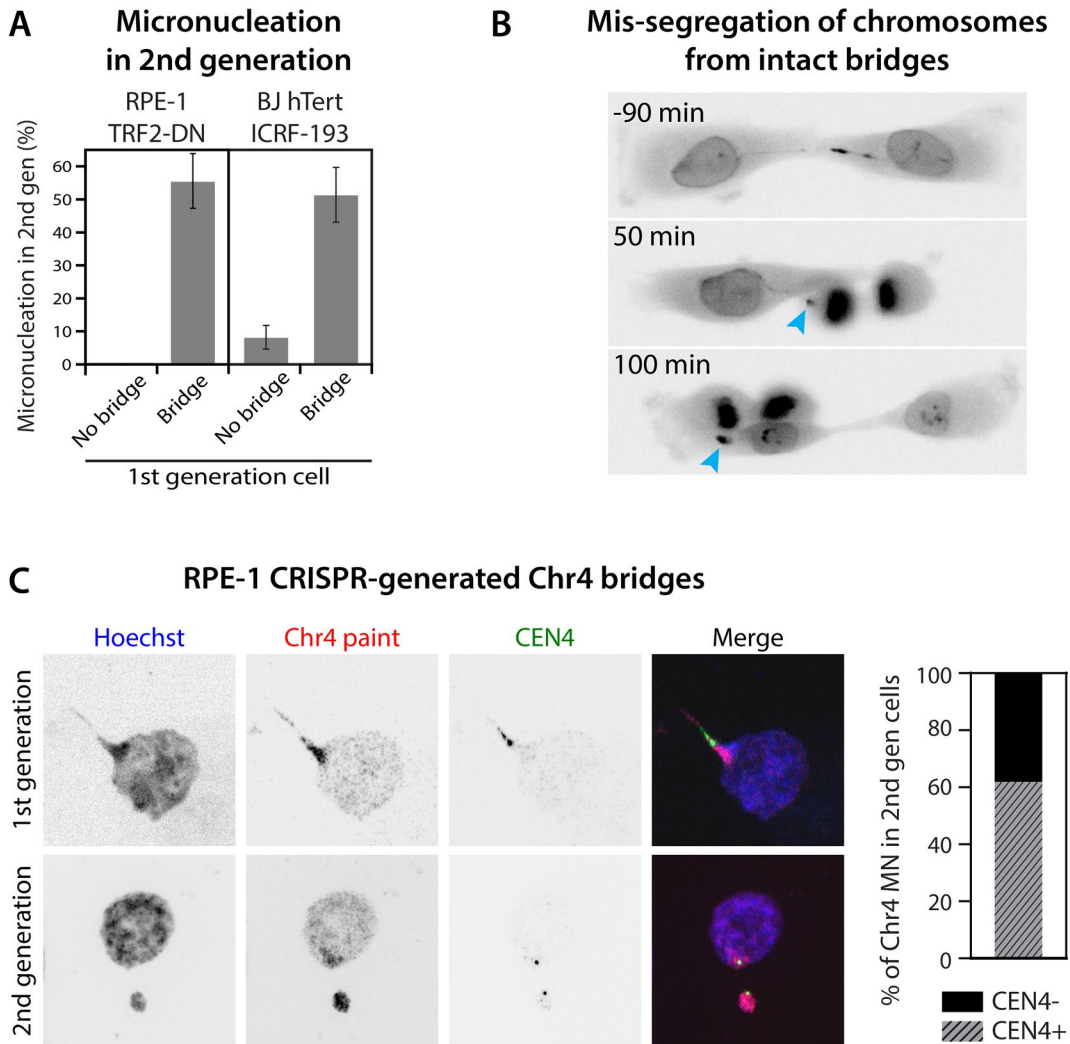


Fig. S14. Chromosome bridge formation predisposes to a high rate of micronucleation in the subsequent cell division.

- (A) A high rate of micronucleation is common in the second generation after chromosome bridge formation. Scheme of the live-cell imaging experiment is depicted in Fig. 6D. Bar graph shows the results from additional experiments in RPE-1 and BJ hTert cells. In RPE-1 cells, bridges were generated by transient TRF2-DN expression; control cells lacking chromosome bridges ($n = 85$) and cells with broken bridges ($n = 87$) were present in the same imaging dish and treated identically. In BJ hTert cells, bridges were generated with ICRF-193; as above, both control cells lacking chromosome bridges ($n = 74$) and cells with broken bridges ($n = 62$) were analyzed from the same imaging experiments. Error bars indicate uncertainty based on counting statistics.
- (B) Chromosome bridges that remain intact throughout interphase also frequently mis-segregate in the next mitosis. Time-lapse images show RPE-1 cells with a

bridge (visualized by GFP-BAF); cells were plated on a short micropattern to prevent bridge breakage during the first interphase. Daughter cells are shown in G2 (-90 min) and at the completion of their 2nd mitoses (50 and 100 min), during which the bridge chromosome is mis-segregated into a micronucleus-like structure (cyan arrowheads). Timestamps are relative to onset of the 2nd mitosis in the daughter cell on the right.

(C) Micronuclei derived from a bridge chromosome usually retain their centromeres.

Left: Images show first- and second-generation cells after CRISPR-mediated Chr4 bridge formation (see Fig. 6D for schematic). Fluorescence in situ hybridization (FISH) was used to detect whole-chromosome 4 (Chr4 paint, red) and Chr4 centromere (CEN4, green). **Right:** quantification of CEN4 status for Chr4-containing micronuclei in second-generation cells.

Supplemental Figure 15

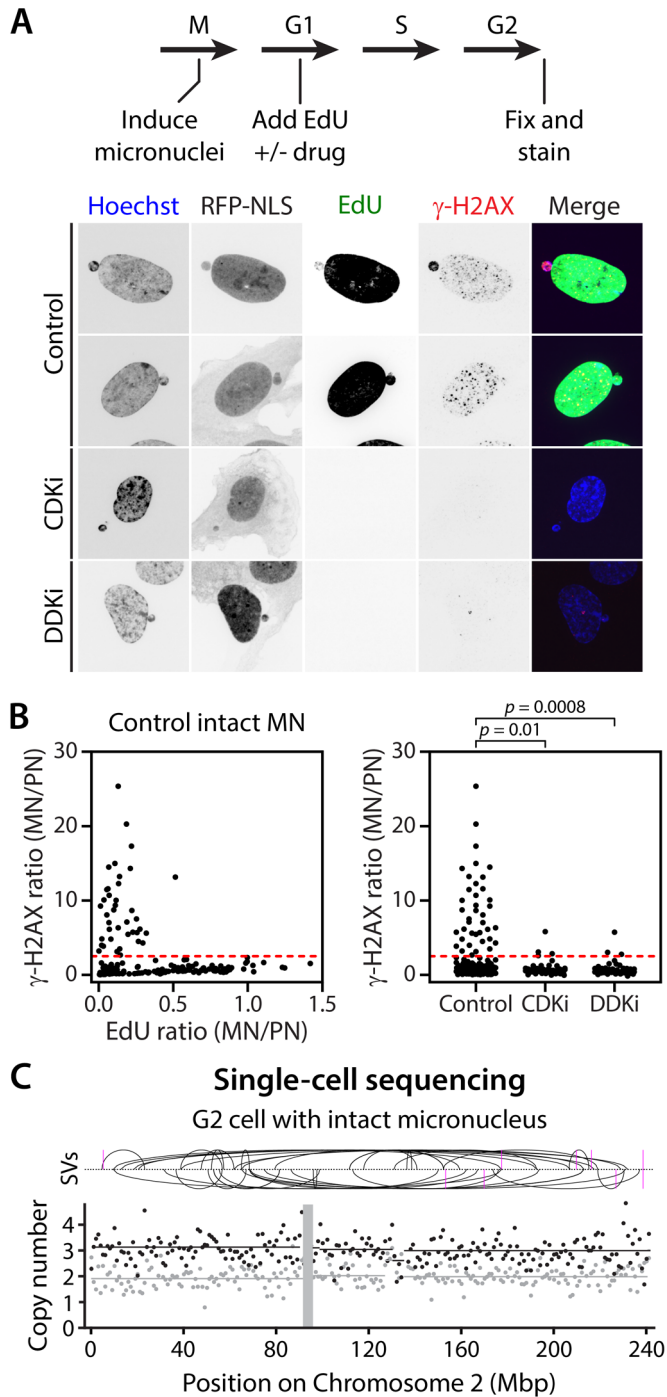


Fig. S15. Low frequency of chromothripsis associated with defective interphase DNA replication in intact micronuclei.

(A) Modest DNA damage is associated with defective replication in intact micronuclei of G2 cells. **Top:** schematic of the experiment. Micronuclei were

- induced by a nocodazole washout procedure (15), and EdU was added in G1 to visualize all DNA replication during the following S phase. Cells were then fixed in G2 (22 hours after mitosis). Where indicated, small-molecule inhibitors of Dbf4-dependent kinase (PHA-767491) or cyclin-dependent kinase (flavopiridol) were also added in G1 to block the initiation of DNA replication. **Bottom:** example images show intact micronuclei (assessed by RFP-NLS), with counterstaining to relate the extent of DNA replication (EdU) to the amount of DNA damage (γ -H2AX). Robust γ -H2AX signal was correlated with diminished EdU signal and was blocked by the DDK or CDK inhibitors.
- (B) Quantification of (A) showing DNA damage in micronuclei with poor DNA replication. **Left:** DNA damage in intact micronuclei (ratio of γ -H2AX intensity in the micronucleus relative to the primary nucleus) relative to replication proficiency (EdU intensity ratio). Dashed red line indicates the threshold (three standard deviations above the mean intensity for primary nuclei) above which micronuclei were scored as positive for DNA damage. **Right:** compared to control, CDKi ($p = 0.01$) or DDKi ($p = 0.0008$) prevents DNA damage in intact micronuclei; p -values determined by Mann-Whitney test.
- (C) Complex rearrangement of a chromosome from an intact micronucleus in a G2 cell. The chromosome from the micronucleus (Chr2) is under-replicated and was identified by its odd-numbered copy number state. The mis-segregation generating this micronucleus resulted in a diploid cell with an extra copy of Chr2 from the micronucleus ($2N+1$) (15). Black dots: 1-Mb bins for the haplotype of the micronuclear chromosome, which together with the fully replicated copy of this Chr2 haplotype in the primary nucleus, leads to black copy number of ~ 3 . Gray dots: the other Chr2 haplotype, which was fully replicated in the primary nucleus and present at a copy number of 2. $N=1$ of 10 cells examined; the remaining 9 G2 micronucleated cells did not exhibit rearrangement of the micronucleated chromosome.

Supplemental Figure 16

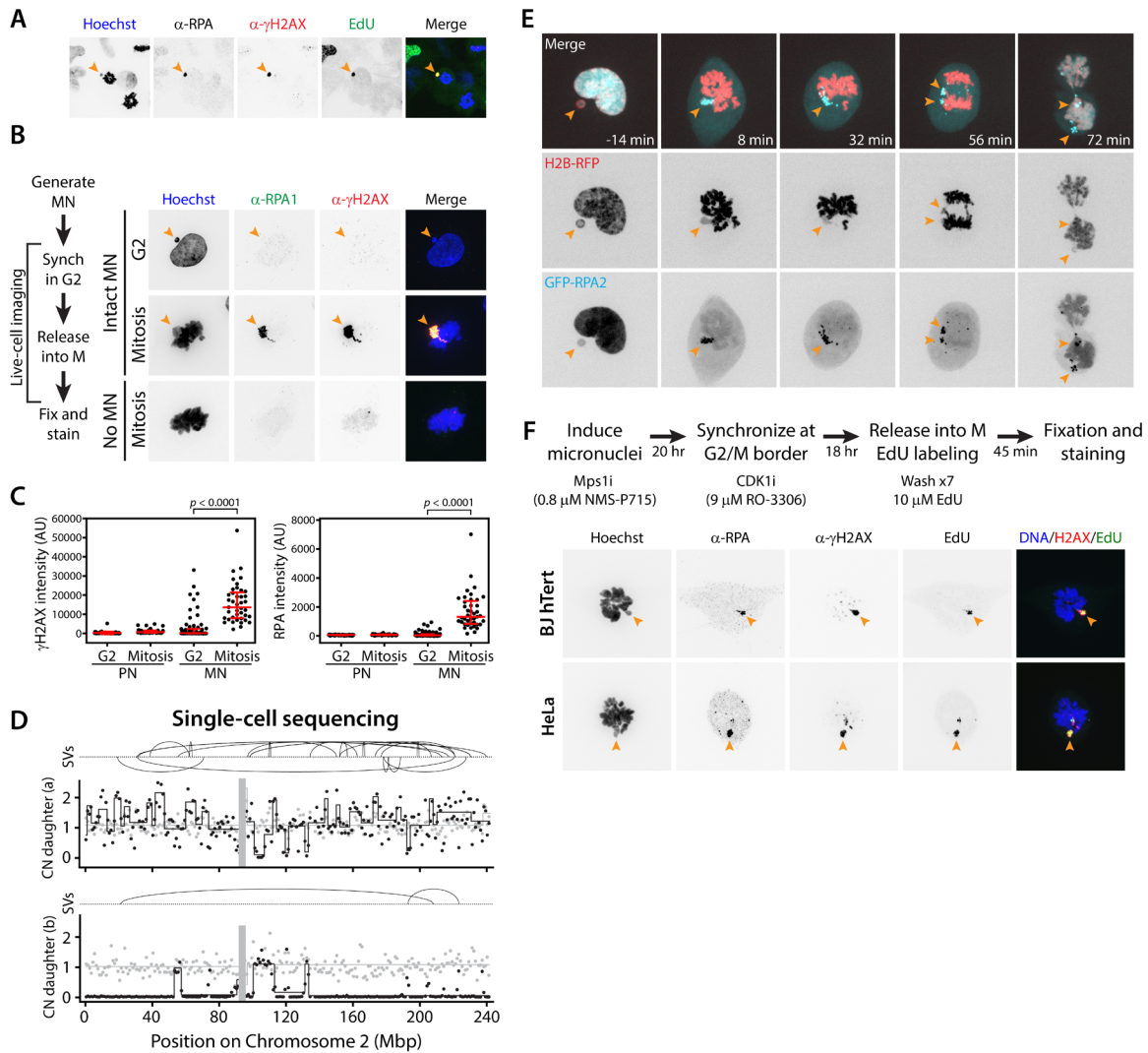


Fig. S16. Micronuclei develop extensive DNA damage associated with a burst of mitotic DNA synthesis, which promotes chromothripsis.

- (A) Mitotic DNA replication and DNA damage (synchronized fixed cells), similar to Fig. 6C, for cells induced to form micronuclei by nocodazole washout. Orange arrowheads: micronucleated chromosome.
- (B) Parallel experiments as in Fig. 6A-B, for cells with intact micronuclei that were released into mitosis. To avoid confounding DNA damage from interphase nuclear envelope rupture, only cells with intact micronuclei (RFP-NLS) were analyzed. Orange arrowheads: micronucleated chromosome.
- (C) Quantification from (B). Levels of DNA damage on the micronucleated chromosome increased ~10-fold in mitotic cells compared to G2 cells ($p < 0.0001$, left plot), concomitant with abrupt initiation of mitotic DNA synthesis as indicated by RPA1 accumulation ($p < 0.0001$, right plot); p -values calculated by Mann-Whitney test.

- (D) Complex rearrangement of a chromosome from an intact micronucleus after passing through mitosis. Copy number and rearrangements are shown for the micronucleated chromosome (one haplotype of Chr2), identified by its odd copy number as described in Fig. S15C. N=8 of 9 daughter pairs examined evidenced complex rearrangements on the micronucleated chromosome.
- (E) Representative time-lapse images show a burst of mitotic DNA replication specifically on a chromosome from an intact micronucleus. Micronuclei were induced by nocodazole washout. Mitotic replication was visualized by GFP-RPA2; chromosomes with H2B-RFP. Confocal imaging was performed with a 100× objective, 7 z-slices at 1- μ m spacing, acquired every 2 min. Orange arrowheads: micronuclear chromosome.
- (F) Mitotic replication on chromosomes from micronuclei in BJ hTert and HeLa cells. **Top:** scheme of the experiment. Micronuclei were induced by MPS1 inhibition, followed by G2 synchronization with RO-3306. Cells were then released into mitosis in the presence of EdU, fixed, and then stained to detect EdU, γ -H2AX, and RPA1. **Bottom:** Representative images showing mitotic replication (EdU and RPA1) and associated massive DNA damage (γ -H2AX) on the micronuclear chromosome (orange arrowheads).

Supplemental Figure 17

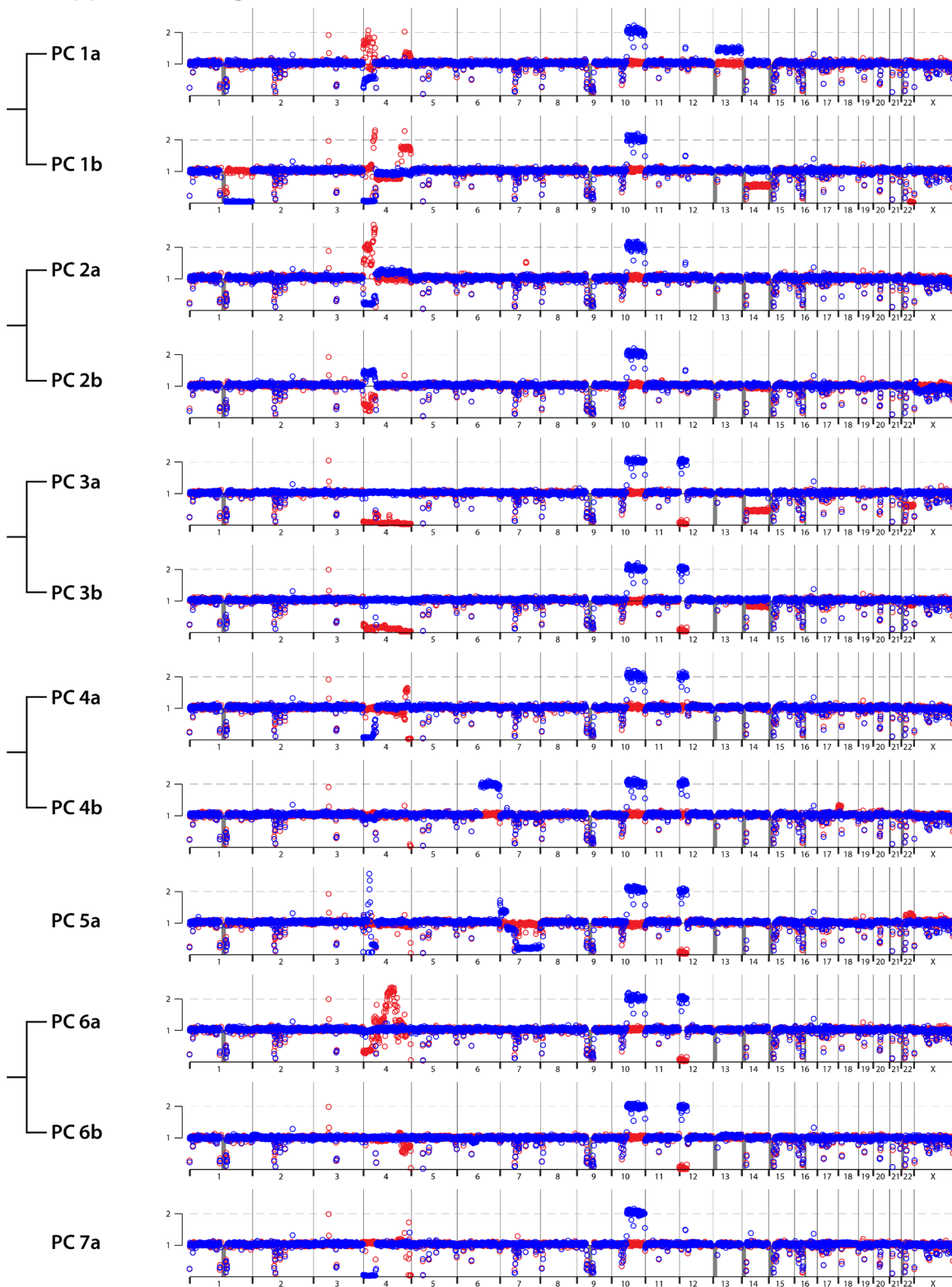


Fig. S17. Daughter cells derived from the breakage of a CRISPR-generated Chr4 bridge produced primary clones with extensive Chr4 copy number alterations (bulk sequencing of all 12 primary clones).

After bridge breakage, daughter cells were separated, grown into primary clones, and then analyzed by whole-genome sequencing. Plots show copy number data for the two homologous copies of each chromosome (red and blue circles for the respective haplotypes). In most cases, both daughter cells that were isolated after breakage grew into primary clones (PC); pairing information for these populations is indicated by cladograms at left. For the PC 5a and PC 7a samples, the corresponding “b” sister cell failed to grow into a clone. Gain of 10q blue haplotype is clonal in RPE-1 cells, and iso(12p) blue haplotype is a subclonal event in RPE-1 cells (based on copy number and karyotype data).

Supplemental Figure 18

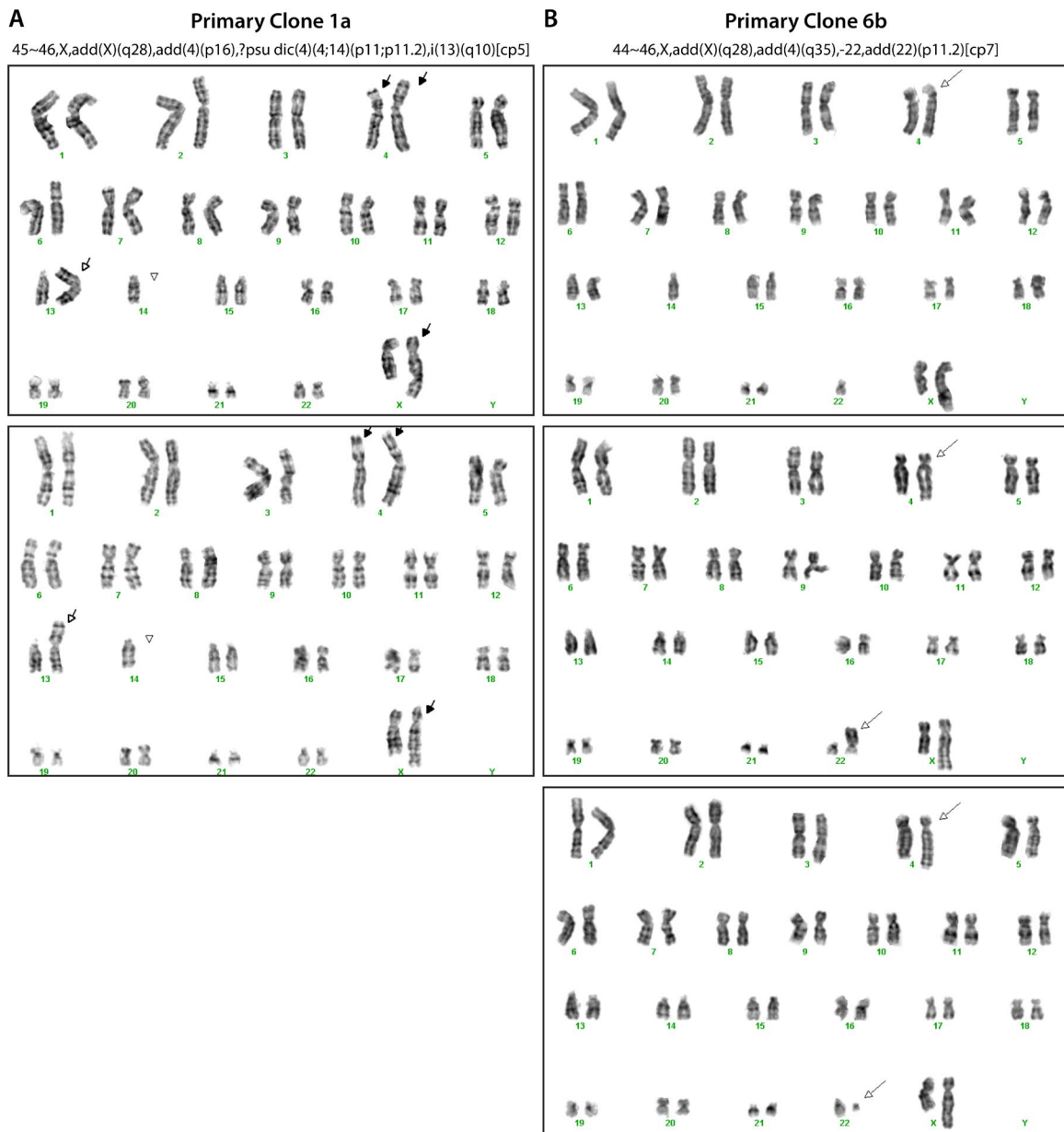


Fig. S18. Subclonal karyotype aberrations in primary clones derived from single cells after the breakage of CRISPR-generated Chr4 bridges.

Left: Karyotypes were determined by Giemsa staining (G-banding) of metaphase spreads from Primary Clone 1a. Solid arrows indicate near-clonal aberrations of Chr4, such as dicentric fusion of chromosomes 4 and 14, and inverted duplications of 4p. Hollow arrows and arrowheads indicate subclonal aberrations: isochromosome 13q (upper panel), and dicentric fusion of chromosomes 13 and 14 (lower panel).

Right: Karyotypes for Primary Clone 6b. Subclonal aberrations (hollow arrows) of Chr4, including dicentric fusion with Chr22 (upper panel), as well as possible inverted duplications of Chr4q (middle and lower panels) are indicated. Additional subclonal

aberrations were observed on Chr22: fusion of unidentified material to the p-arm (middle panel), or large deletions of material from both p- and q-arms (lower panel). Note that translocations involving Chr4 were observed as fusions with acrocentric chromosomes more frequently than expected by random chance ($p = 0.024$, Fisher's exact test). Other translocations, such as the 13:14 fusion in Primary Clone 1a may be derived from an original Chr4:14 fusion, although we cannot exclude that some events not involving Chr4 could be due to CRISPR off-target DNA cleavage. It was previously reported that TRF2-DN induces a high frequency of acrocentric fusions in RPE-1 cells (52), supporting the idea that fusion to acrocentric chromosomes is a general mechanism for the resolution of DNA breaks resulting from chromosome bridge breakage, irrespective of how the bridge is formed.

Supplemental Figure 19

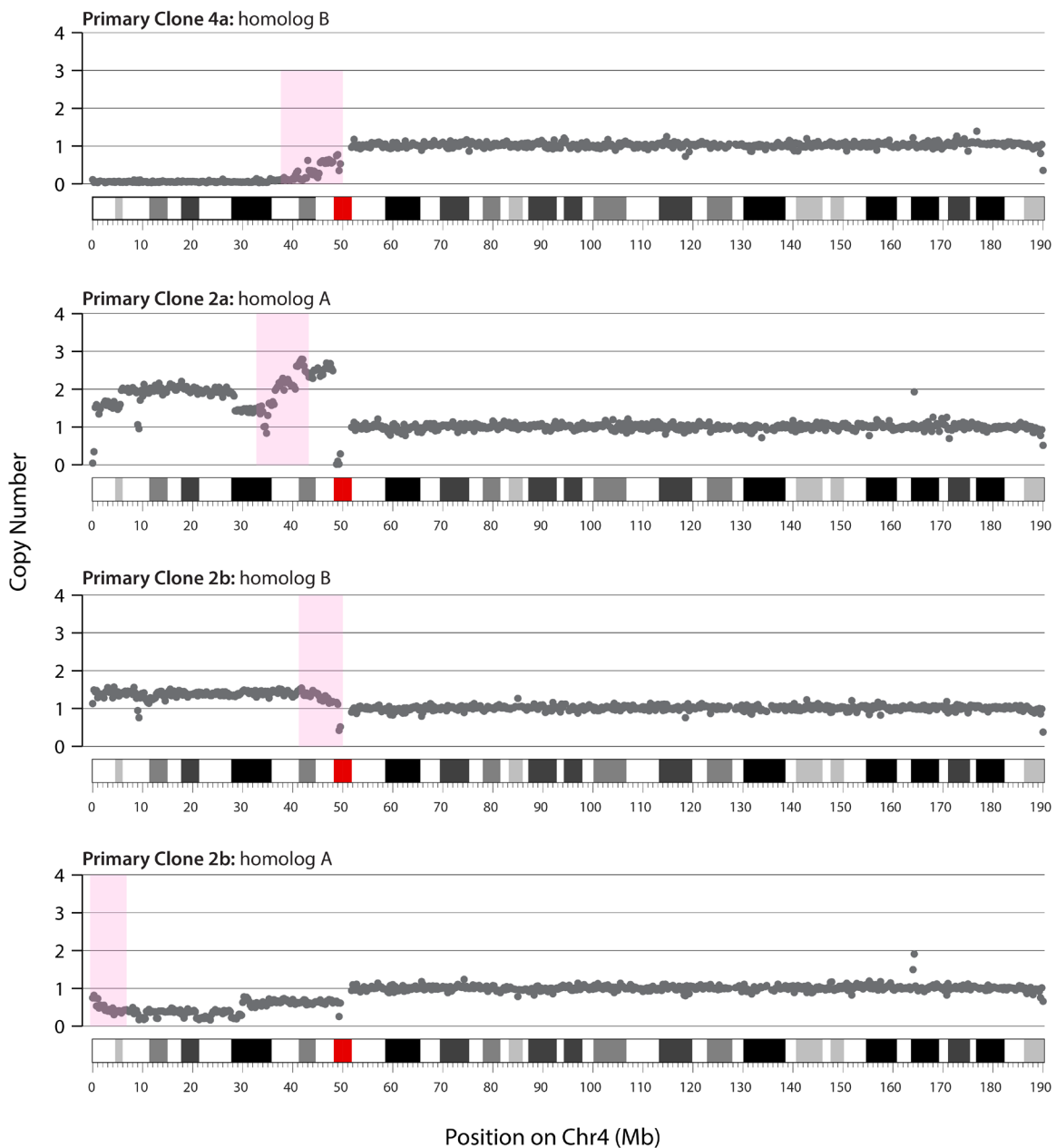


Fig. S19. Gradual copy number transitions indicate ongoing genome instability within populations of cells.

Heterogeneity in the size of segmental copy number alterations among single cells within a population results in a gradual, sloping copy number transition when the population is analyzed by bulk DNA sequencing (see illustration in Fig. 7D). Shown are four examples of gradual sloping copy number transitions (pink shading) from bulk sequencing data of primary clones derived from single cells after the breakage of CRISPR-mediated Chr4 bridges.

Supplemental Figure 20

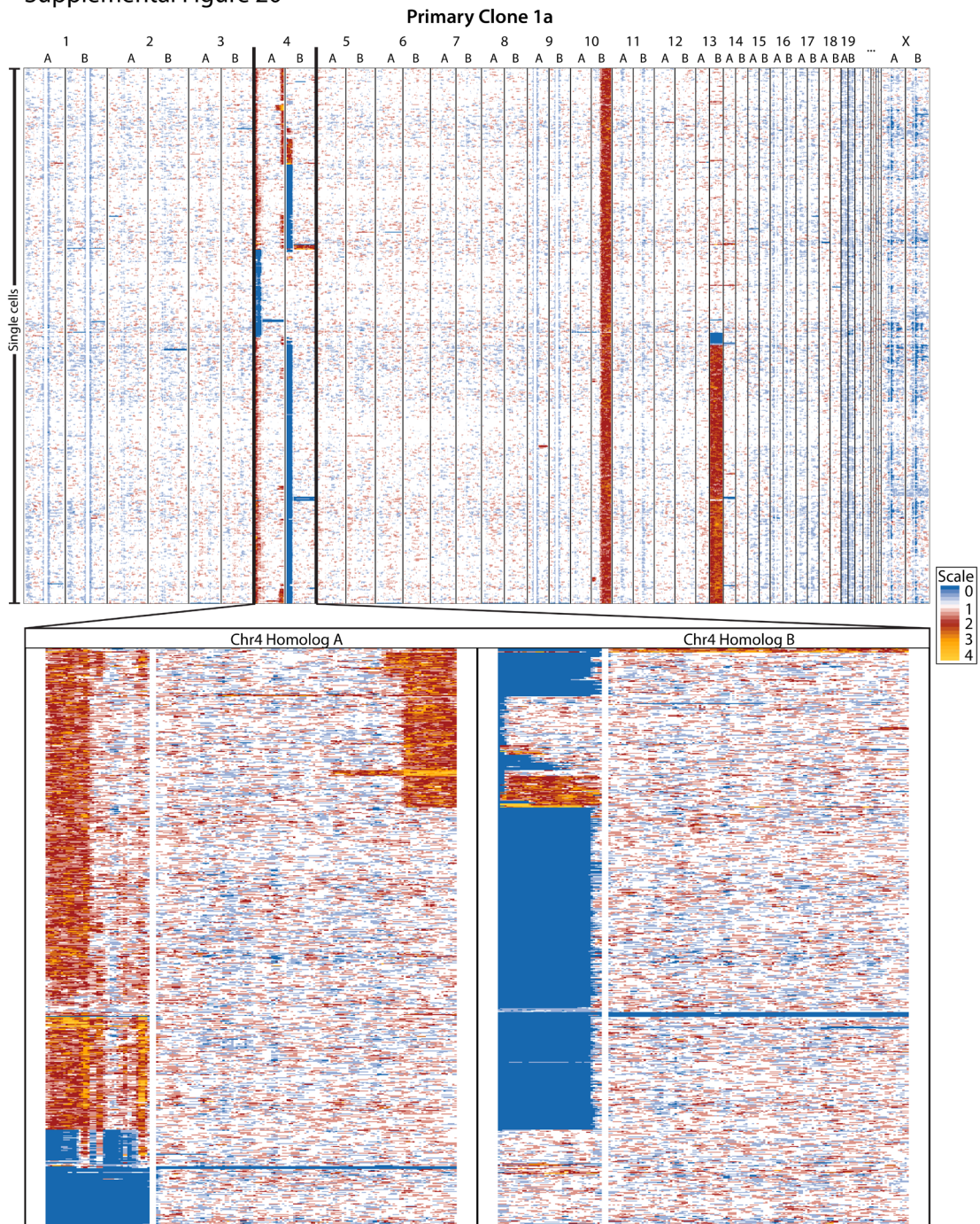


Fig. S20. Copy number heterogeneity among single cells within primary clones derived from cells after the breakage of CRISPR-generated Chr4 bridges.

Plots showing genome-wide (top) and Chr4 (bottom) haplotype copy number profiles obtained from low-pass sequencing of ~500-800 single cells from the primary clones shown in Fig. 7. Haplotype copy number estimates in 1Mb bins (rounded to the nearest

half-integer) are shown as a heatmap, where each row of the grid represents a single cell from the population. Rows are ordered according to hierarchical clustering based on the haplotype copy number profiles (Methods). Note: clustering was performed independently for the whole-genome (top) and Chr4-only (bottom) plots.

Supplemental Figure 21

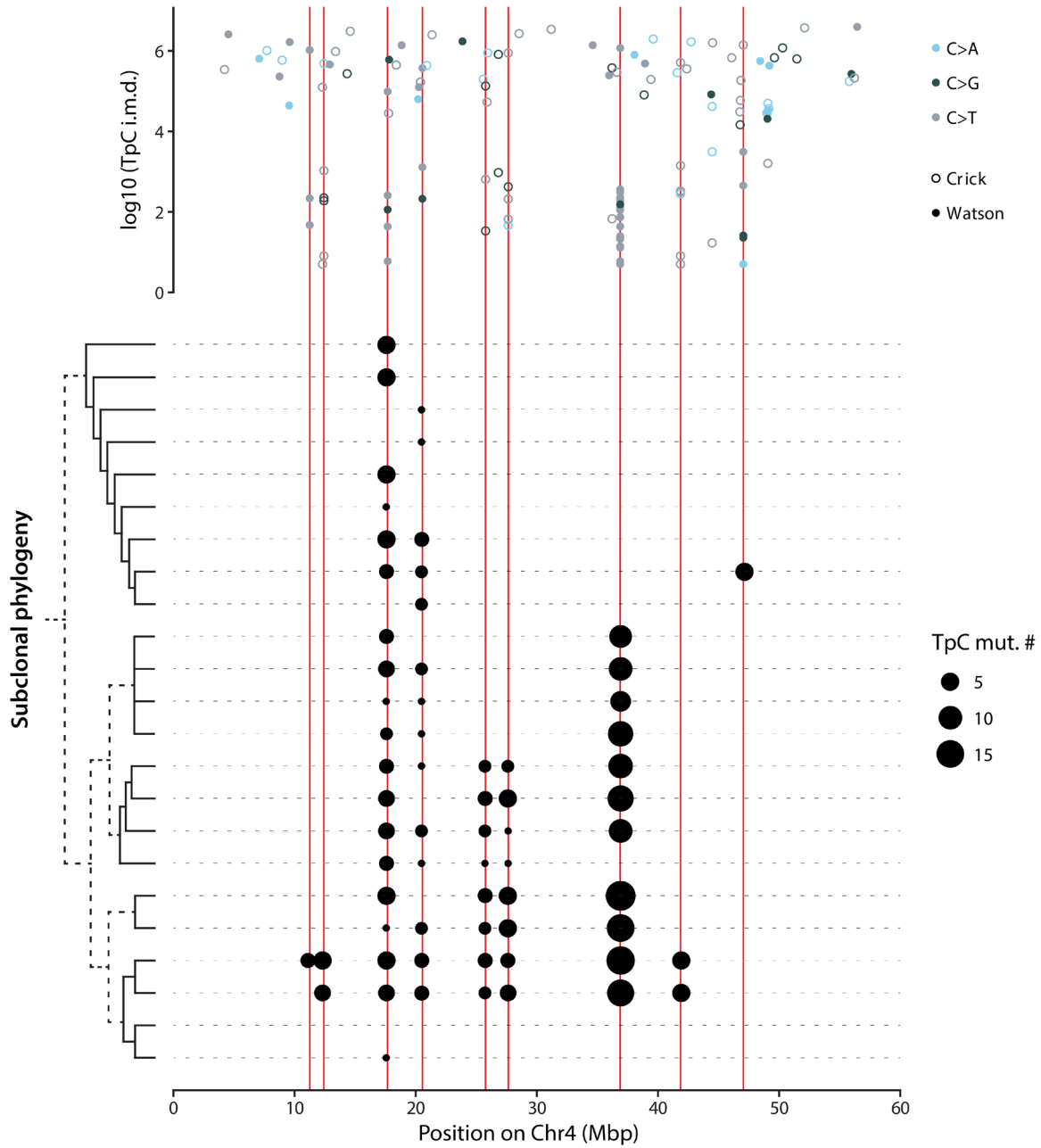


Fig. S21. Kataegis in single-cell subclones derived from Primary Clone 1a.

Multiple subclonal kataegis clusters in Primary Clone 1a. **Top:** as in Fig. S9, rainfall plot of TpC SNVs on Chr4p (supported by at least two sequencing reads in any subclone). Clusters of at >2 SNVs are highlighted (red lines). **Bottom:** dot plot shows the detection of kataegis clusters in each subclone, where dot size represents the number of TpC SNVs detected. **Left:** Dendrogram shows the phylogeny of subclones inferred from shared copy number breakpoints. Groups of subclones with similar copy number patterns were readily

apparent (solid lines) whereas the hierarchical relationship between more divergent subclones could not be unambiguously resolved from copy number alone (dashed lines).

Supplemental Figure 22

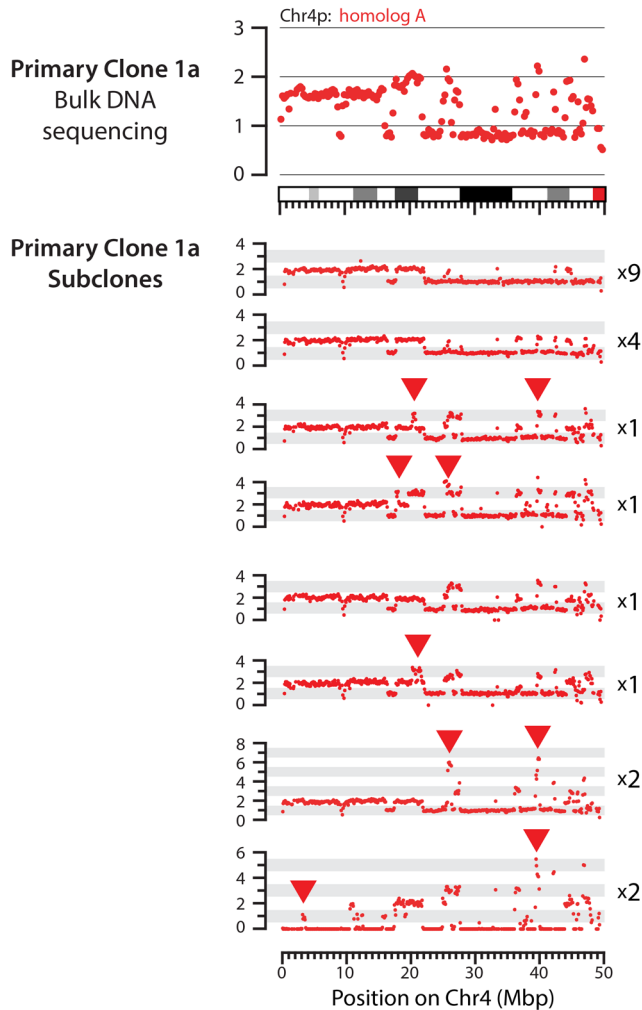


Fig. S22. Focal copy number variations in subclones derived from the same primary clone.

Focal copy number variations are observed downstream of bridge breakage and contribute to subclonal heterogeneity. **Top:** Copy number for Chr4p homolog A from bulk sequencing of primary clone 1a (homolog B is shown in Fig. 7D). **Bottom:** eight unique copy number profiles for Chr4 homolog A identified from 21 single-cell derived subclones obtained from the primary clone. The number of individual subclones represented in each profile is listed next to each plot. Red arrowheads indicate selected regions that exhibit focal copy number variations among the subclones.

Supplemental Figure 23

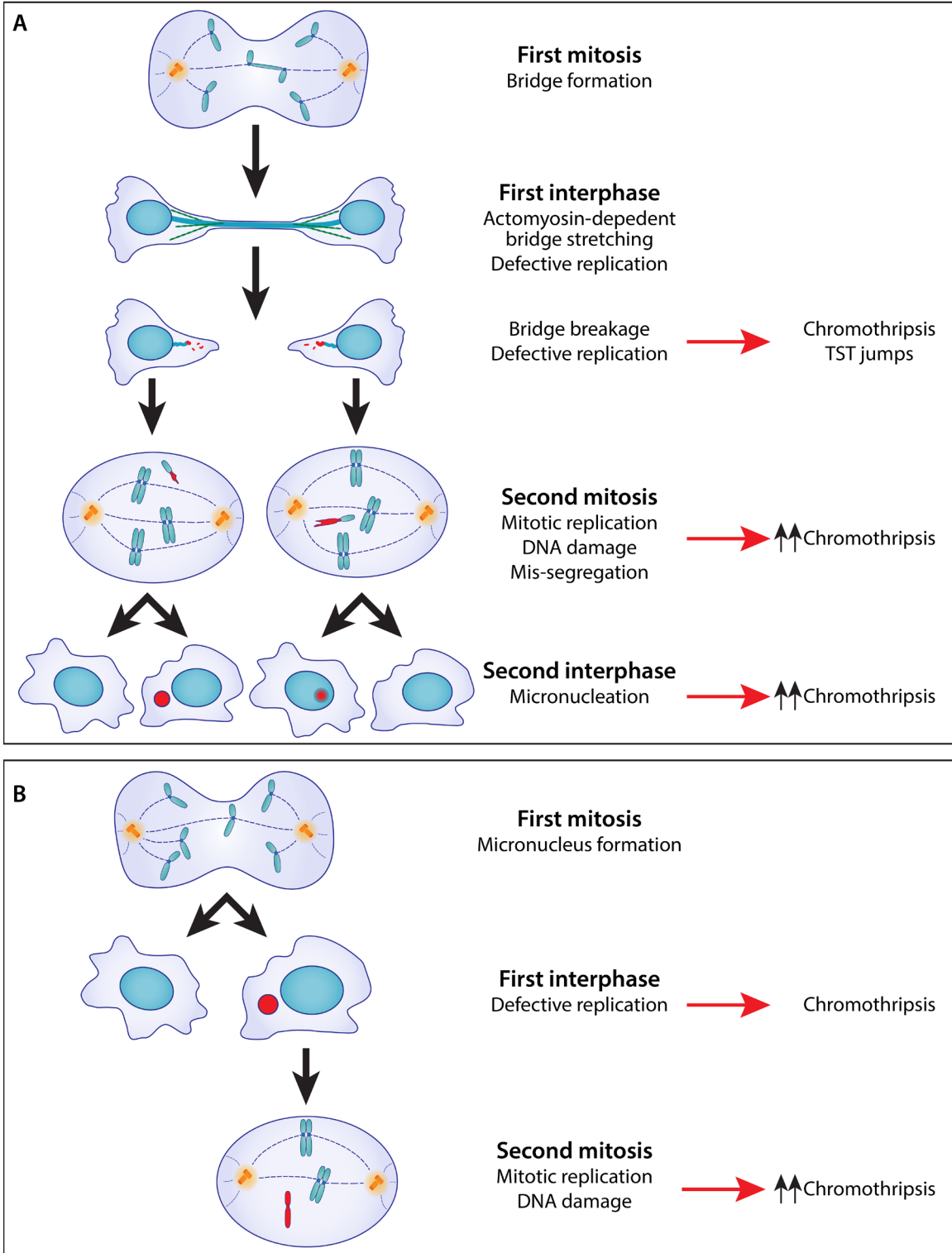


Fig. S23. Model for the generation of cycles of chromothripsis from either chromosome bridges or micronuclei.

- (A) Chromothripsis from chromosome bridges. **First mitosis:** Chromosome bridges result from the segregation of dicentric or catenated chromosomes. Nuclear envelope (NE) assembly on bridge DNA is abnormal, such that only one subset of NE proteins (“core” proteins) assemble on the bridge, while the remaining NE proteins, including nuclear pore complexes (“non-core”) do not assemble (not shown; (17)). **First interphase:** Bridges extend during interphase because of cell migration, and experience stretching forces from actomyosin contractility. The dearth of nuclear pore complexes, combined with the fact that the long, thin geometry of the bridge is predicted to act as a barrier to diffusion (55), likely results in a local nucleoplasm deficient for DNA replication and repair proteins (12, 15, 17). Bridge DNA is therefore poorly replicated. Furthermore, during extension and stretching of the bridge, local contractile forces may eject histones (28, 74), including the centromeric histone, CENP-A. Next, the bridge breaks in a manner that requires actomyosin mechanical forces. This event generates simple DNA breakage or local DNA fragmentation, which can lead to “local jumps” and chromothripsis-like rearrangements by DNA end-joining. In addition, these broken DNA ends may undergo error prone replicative repair (e.g., by MMBIR), generating the TST jump signature and chromothripsis. Overall, chromothripsis occurs at a relatively low frequency in the first interphase. **Second mitosis.** Upon entry into the second mitosis, under-replicated DNA from the broken chromosome bridge stub undergoes DNA replication coupled to heavy DNA damage, which increases the frequency of chromothripsis. Moreover, chromosomes from broken bridges frequently mis-segregate during the second mitosis (as shown in the progeny of the left daughter cell), presumably due to defects in centromere/kinetochore function. In some cases, the bridge chromosome may be reincorporated into the primary nucleus of one or both grand-daughter cells (progeny of the right daughter cell). **Second interphase.** Mis-segregation of the broken bridge chromosome generates micronuclei, which then amplifies the frequency and extent of chromothripsis (12, 15). DNA damage acquired during the initial bridge breakage event, during mitotic replication, and/or within subsequently forming micronuclei, can all lead to more chromosome bridges (BFB cycles) and more micronuclei during subsequent cell divisions. Altogether, these events generate ongoing complex genome evolution from a single cell division error—the formation of a chromosome bridge.
- (B) Parallels between chromothripsis from micronuclei and from chromosome bridges. **First mitosis.** Lagging chromosomes result from mitotic errors or the generation of acentric chromosome fragments that do not normally segregate. **First interphase.** Nuclear envelope (NE) assembly on the lagging chromosome is abnormal, as described above for chromosome bridges (not shown; (17)). Accordingly, there is poor DNA replication and an initial low frequency of chromothripsis. Due to defects in the import of proteins required to maintain nuclear envelope stability, such as B-type lamins, the micronuclear envelope can also undergo spontaneous rupture (not shown; (16)), leading to further DNA damage by an unknown mechanism. Note: the nuclear envelope surrounding

chromosome bridges may also undergo rupture, but the geometry of the bridge precludes cytological assessment by monitoring for loss of nuclear integrity. **Second mitosis.** Like chromosome bridges, under-replicated chromosomes from micronuclei undergo a burst of aberrant DNA replication, leading to additional DNA damage and further promoting chromothripsis.

Table S1.

List of bridge induction methods used in all experiments. Italicized text indicates methods that were employed in the same series of experiments, but not shown in the cited figure panel.

MAIN FIGURES		SUPPLEMENTAL FIGURES (continued)	
Figure Panel	Bridge induction method(s)		
1A	TRF2-DN, <i>siSMC2</i> , ICRF-193, Chr4 CRISPR	S5B-C	TRF2-DN
1B-H	TRF2-DN	S5D	ICRF-193
2	TRF2-DN, <i>siSMC2</i> , Chr4 CRISPR	S6A	TRF2-DN
3	<i>TRF2-DN</i> , <i>siSMC2</i> , Chr4 CRISPR	S7	TRF2-DN, <i>siSMC2</i> , Chr4 CRISPR
4A	<i>TRF2-DN</i> , Chr4 CRISPR	S8A	TRF2-DN, <i>Chr4 CRISPR</i>
4B	TRF2-DN	S8B	TRF2-DN
5A	TRF2-DN, <i>siSMC2</i> , <i>Chr4 CRISPR</i>	S9	TRF2-DN, <i>siSMC2</i> , Chr4 CRISPR
5B	Chr4 CRISPR	S10A	TRF2-DN, <i>siSMC2</i> , <i>Chr4 CRISPR</i>
6A-C	TRF2-DN	S10B	TRF2-DN (single-cell), Chr4 CRISPR (bulk)
6D	Chr4 CRISPR	S11A-B	TRF2-DN
7	Chr4 CRISPR	S11C	<i>TRF2-DN</i> , <i>siSMC2</i> , <i>Chr4 CRISPR</i>
		S12	TRF2-DN (Q2, Q3), <i>siSMC2</i> (Q1)
		S13A	TRF2-DN
		S13B	ICRF-193
		S14A	TRF2-DN (RPE-1), ICRF-193 (BJ)
		S14B	TRF2-DN
		S14C	Chr4 CRISPR
		S17	Chr4 CRISPR
		S18	Chr4 CRISPR
		S19	Chr4 CRISPR
		S20	Chr4 CRISPR
		S21	Chr4 CRISPR
		S22	Chr4 CRISPR
SUPPLEMENTAL FIGURES			
Figure Panel	Bridge induction method(s)		
S1B	TRF2-DN, <i>siSMC2</i> , ICRF-193, Chr4 CRISPR		
S1C	Chr4 CRISPR		
S1D	TRF2-DN		
S2A-B	TRF2-DN		
S2C	ICRF-193		
S3	ICRF-193 (BJ), uninduced (Hela/U2OS)		
S4	TRF2-DN		
S5A	TRF2-DN (RPE-1), ICRF-193 (BJ)		

Table S2.

Summary of cytogenetic analysis for the uninduced parental cell line for CRISPR-mediated Chr4 bridge formation, and for primary clones derived from single cells after breakage of induced Chr4 bridges.

Sample	Karyotype	Metaphases counted	Karyotyped	Aberrations of Chromosome 4	Aberrations involving all other chromosomes
Parental cell line	46,X,add(X)(q28),dup(12)(q24.1q15)[5]/46,X,add(X)(q28)[15]	80	20	None	Loss of one copy of Chr16 in 1/20 karyotypes Loss of one copy each of Chr3 and Chr22 in 1/20 karyotypes
Primary Clone 1a	45~46,X,add(X)(q28),add(4)(p16),?psu dic(4)(4;14)(p11;p11.2),i(13)(q10)[cp5]	10	5	Homolog 1: add(4)(p16) in 5/5 karyotypes, with subclonal variation in the amount of additional material fused to 4p Homolog 2: ?psu dic(4)(4;14)(p11;p11.2) in 3/5 karyotypes; non-clonal aberrations in remaining 2/5 karyotypes	?psu dic(4)(4;14)(p11;p11.2) in 3/5 karyotypes i(13)(q10) in 3/5 karyotypes ?psu dic(4)(4;13) together with der(14) in 1/5 karyotypes
Primary Clone 1b	79~85,XX,add(X)(q28)x2,add(1)(q23.1)x2,-4,-4,?psu dic(4)(4;14)(p11;p11.2)x2,i(4)(q10),-14[cp4]	10	4	Homolog 1: i(4)(q10) in 4/4 karyotypes Homolog 2: ?psu dic(4)(4;14)(p11;p11.2) in 3/4 karyotypes None of the aberrant Chr4s were seen in duplicate in any karyotype	Loss of one copy of Chr14 in 4/4 karyotypes ?psu dic(4)(4;14)(p11;p11.2) in 3/4 karyotypes Two copies of add(1)(q23.1) in 4/4 karyotypes
Primary Clone 2a	44~45,X,add(X)(q28),add(4)(p12),?psu dic(4)(4;14)(p11;p11.2)[cp5]	10	5	Homolog 1: ?psu dic(4)(4;14)(p11;p11.2) in 5/5 karyotypes Homolog 2: add(4)(p12) in 3/5 karyotypes	?psu dic(4)(4;14)(p11;p11.2) in 5/5 karyotypes add(22)(p) in 4/5 karyotypes
Primary Clone 2b	44~46,X,add(X)(q28),?psu dic(4)(4;14)(p11;p11.2)[cp8]	10	8	Homolog 1: ?psu dic(4)(4;14)(p11;p11.2) in 8/8 karyotypes, but with subclonal variation in the amount of additional material fused to 4p Homolog 2: add(4)(p16) in 2/8 karyotypes	?psu dic(4)(4;14)(p11;p11.2) in 8/8 karyotypes
Primary Clone 3a	43~45,X,add(X)(q28),-4,der(14;22)(q10;q10)[cp5]	11	5	Homolog 1: Loss of Chr4 in 5/5 karyotypes Homolog 2: Apparently normal, except for unusual banding pattern in pericentromere region	der(14;22)(q10;q10) with aberrant banding near the centromere in 5/5 karyotypes
Primary Clone 3b	43~44,X,add(X)(q28),-4,-10,der(14;22)(q10;q10)[cp4]	10	4	Homolog 1: Loss of Chr4 in 4/4 karyotypes Homolog 2: Apparently normal	der(14;22)(q10;q10) with aberrant banding on 14q (possibly Chr4 material) in 4/4 karyotypes
Primary Clone 4a	45~46,X,add(X)(q28),-4,?psu dic(4)(4;14)(p11;p11.2),add(12)(q15),+r[cp4]	10	4	Homolog 1: ?psu dic(4)(4;14)(p11;p11.2) in 3/4 karyotypes Homolog 2: ring chromosome containing 4q in 4/4 karyotypes	?psu dic(4)(4;14)(p11;p11.2) in 3/4 karyotypes add(12)(q15) in 4/4 karyotypes
Primary Clone 4b	43~45,X,add(X)(q28),der(4)add(4)(p16)add(4)(q35),add(12)(q15),der(14;22)(q10;q10),+19,-21[cp7]	10	7	Homolog 1: der(4)add(4)(p16)add(4)(q35) in most karyotypes with subclonal variation in the amount of additional material fused to 4p Homolog 2: ?psu dic(4)(4;14)(p11;p11.2) in 1/5 karyotypes	der(14;22)(q10;q10) in 3/7 karyotypes add(12)(q15) in 7/7 karyotypes, which in 1 case contains additional material compared to the others

Primary Clone 5a	42~46,X,add(X)(q28),-4,add(4)(p14),-7,-22,+r[cp6]	10	6	Homolog 1: add(4)(p14) in 6/6 karyotypes Homolog 2: ring chromosome containing 4q in 6/6 karyotypes	Loss of one copy of Chr7 Loss of one copy of Chr22 Material from Chrs 7 and/or 22 could be present in the ring chromosome
Primary Clone 6a	43~45,X,add(X)(q28),-4,add(4)(p16),-22,add(22)(p11.2),+r[cp4]	10	4	Homolog 1: add(4)(p16) in 4/4 karyotypes, with subclonal variation in the amount of additional material fused to 4p Homolog 2: ring chromosome containing 4q in 2/4 karyotypes	Loss of one copy of Chr22 in 4/4 karyotypes add(22)(p11.2) in 2/4 karyotypes
Primary Clone 6b	44~46,X,add(X)(q28),add(4)(q35),-22,add(22)(p11.2)[cp7]	10	7	Homolog 1: add(4)(q35) in most karyotypes	add(22)(p11.2) in 2/7 karyotypes Loss of one copy of Chr22 in 2/7 karyotypes del(22) in 2/7 karyotypes Putative dic(4)(4;22) in 1/7 karyotypes

Table S3.

Description of cell lines used in this study.

RPE-1 hTert cells and derivatives

		CELL LINE ID	REFERENCE
	RPE-1		ATCC
	RPE-1 sh p21/Rb (neo)	PM554	
	RPE-1 H2B-RFP (puro); GFP-RPA2	PM539	
	RPE-1 H2B-GFP (puro); tdRFP-NLS		Zhang, et al. 2015
	RPE-1 CENP-A-Halo (endogenous)		Swartz, et al. 2019
	RPE-1 CENP-A-Halo (endogenous); sh p21/Rb (neo); rtTA (blast); dox(mycTRF2-DN); GFP-BAF	PM685	
siSMC2, ICRF-193	RPE-1 GFP-BAF (neo)	PM621	
	RPE-1 GFP-BAF (neo); sh p21/Rb (blast)	PM686	
TRF2-DN	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]		Maciejowski, et al., 2015
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF	PM533	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF; mCherry-Geminin	PM638	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF; tdRFP-NLS	PM553	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF; TagRFP-T-Utr261	PM602	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; SNAP-BAF	PM687	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; SNAP-BAF; GFP-RPA2	PM688	

Chr4g1	RPE-1 rtTA-T2A-neo; dox(Cas9)		McKinley & Cheeseman, 2017
	RPE-1 rtTA-T2A-neo; dox(Cas9); sg Chr4g1 (puro); sh p21/Rb (blast); GFP-BAF	PM600	
TREX1 ^{-/-}	RPE-1 H2B-mCherry; mTurquoise2-RPA70; sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; TREX1 ^{-/-} clone 2.2		Maciejowski, et al., 2015
	RPE-1 H2B-mCherry; mTurquoise2-RPA70; sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; TREX1 ^{-/-} clone 2.25		Maciejowski, et al., 2015
	RPE-1 SNAP-BAF; H2B-mCherry; mTurquoise2-RPA70; sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; TREX1 ^{-/-} clone 2.2	PM689	
	RPE-1 SNAP-BAF; H2B-mCherry; mTurquoise2-RPA70; sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; TREX1 ^{-/-} clone 2.25	PM690	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF; TREX1 ^{-/-} clone 1a	PM691	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF; TREX1 ^{-/-} clone 1b	PM662	
	RPE-1 rtTA-T2A-neo; dox(Cas9); GFP-BAF (neo); TREX1 ^{-/-} clone 2a	PM692	
	RPE-1 rtTA-T2A-neo; dox(Cas9); GFP-BAF (neo); TREX1 ^{-/-} clone 2b	PM693	
LINC knockout	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF; SUN1 ^{-/-}	PM694	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF; SUN2 ^{-/-}	PM695	
	RPE-1 sh p21/Rb (puro); rtTA (blast); dox[mycTRF2-DN(IRESneo)]; GFP-BAF; SUN1 ^{-/-} SUN2 ^{-/-}	PM696	

BJ hTert cells and derivatives

BJ	ATCC
BJ sh p21/Rb (blast); GFP-BAF (neo)	PM615
BJ sh p21/Rb (blast); GFP-BAF (neo); TagRFP-T-Utr261	PM641

U2OS cells and derivatives

U2OS	PM437
U2OS GFP-BAF (neo)	PM599

HeLa cells and derivatives

HeLa	PM324
HeLa GFP-BAF (neo)	PM697

Movie S1

A pair of RPE-1 daughter cells with a chromosome bridge (GFP-BAF is grayscale) on a long fibronectin micropattern. Broken bridge ends are indicated by blue arrowheads. Timestamp shows relative time in hours. Widefield imaging was performed with a 20x objective. Related to Fig. 1B, left.

Movie S2

A pair of RPE-1 daughter cells with a chromosome bridge (GFP-BAF is grayscale) on a short fibronectin micropattern. Timestamp shows relative time in hours. Widefield imaging was performed with a 20x objective. Related to Fig. 1B, right.

Movie S3

Non-uniform stretching of a chromosome bridge and subsequent breakage in the taut region (RPE-1 cells; GFP-BAF is grayscale). The time interval between frames is 10 minutes; total video duration is 3.7 hours. Imaging was performed on a spinning disk confocal microscope with a 40x objective. Related to Fig. 1D.

Movie S4

Actin dynamics associated with chromosome bridge breakage in RPE-1 cells. Left panel: Merge overlay of GFP-BAF (green) with the actin reporter, RFP-Utr261 (red). Middle panel: GFP-BAF (monochrome). Right panel: RFP-Utr261 (monochrome). Contraction of the actin-rich structure (red arrowheads, 10 to 35 min) immediately precedes bridge breakage (35 min), and then the actin structure rapidly disassembles (40 to 55 min). Cyan arrowheads indicate broken bridge ends. Timestamp shows relative time in minutes. Widefield imaging was performed with a 20x objective. Related to Fig. S5A, top.

Movie S5

Representative video showing that disruption of the actin cytoskeleton (Latrunculin A) blocks chromosome bridge breakage (GFP-BAF is grayscale). Red arrowheads mark the daughter cells of interest. At the start of the video, cells connected by a long bridge are in normal growth medium (“No drug”) and are later exchanged into drug medium (“Latrunculin A”). Near the end of the video, the upper cell passes through the next mitosis as indicated (but fails cytokinesis due to actin inhibition). The time interval between frames is 10 minutes; total video duration is 11 hours. Widefield imaging was performed with a 20x objective. Related to Fig. 1E.

Movie S6

Video illustrating mechanical breakage of chromosome bridges with a glass capillary. Initial frames: differential interference contrast image showing a pair of daughter cells connected by a bridge (open arrow) and the glass capillary (filled arrow). The following frames show timelapse imaging of GFP-BAF (inverted greyscale). Open cyan arrows: broken bridge ends. Widefield imaging was performed with a 20x objective. Related to Fig. 4A.

Movie S7

Sudden onset of DNA replication on a broken bridge stub after entry into mitosis. Left panel: Merge overlay, GFP-RPA2 (green) is a marker for mitotic replication; SNAP-BAF (red) labels the broken bridge stub. Middle panel: GFP-RPA2 (monochrome). Right panel: SNAP-BAF (monochrome). Imaging was performed on a spinning disk confocal microscope with a 40x objective. Timestamp shows relative time in minutes. Related to Fig. S13A.

Movie S8

Sudden onset of DNA replication (GFP-RPA2, green) on a micronucleated chromosome (H2B-RFP, red) after entry into mitosis. Imaging was performed on a spinning disk confocal microscope with a 100x objective. Time interval between frames is two minutes; total video duration is 92 minutes. Related to Fig. S16E.