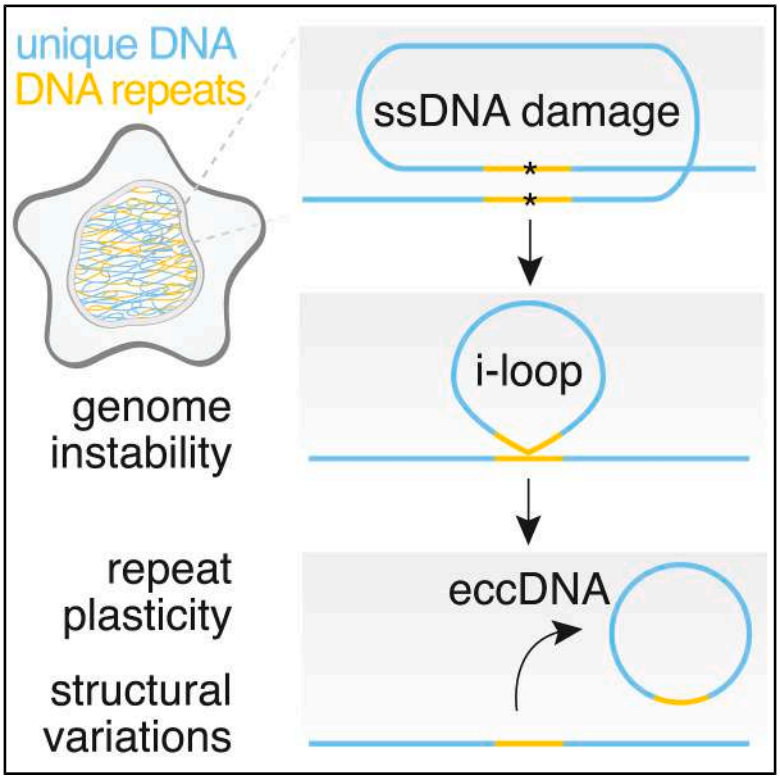
Molecular Cell

Article

Damage-induced i-loops generate eccDNA from repetitive elements

Graphical abstract



Authors

Elia Zanella, Michele Giannattasio, Sara Bisi, Francesco Nicassio, Ylli Doksani

Correspondence

ylli.doksani@ifom.eu

In brief

Zanella et al. reveal a mechanism of eccDNA formation. Nicks or gaps at repetitive sequences promote i-loop formation that can be excised as eccDNA. Given the abundance of repetitive elements and the high endogenous rate of single-strand lesions, this mechanism may substantially contribute to genome plasticity and structural variation.

Highlights

- Nicks and gaps at repeats induce internal loops that can be excised as eccDNA
- Telomeres, alpha satellite, and rDNA are hotspots of i-loop formation
- More rare, damage-induced interactions of distant repeats can generate larger eccDNA
- During cell death, repeat-derived eccDNA is generated via i-loop excision

Zanella et al., 2026, Molecular Cell *86*, 1856–1869
May 21, 2026
https://doi.org/10.1016/j.molcel.2026.04.003

CellPress

CellPress

Molecular Cell

Article

Damage-induced i-loops generate eccDNA from repetitive elements

Elia Zanella,[1] Michele Giannattasio,[1,2] Sara Bisi,[3] Francesco Nicassio,[3] and Ylli Doksani[1,4,*]
[1]IFOM ETS - The AIRC Institute of Molecular Oncology, Via Adamello 16, 20139 Milano, Italy
[2]Dipartimento di Oncologia ed Emato-Oncologia, Università degli Studi di Milano, Via Festa del Perdono 7, 20122 Milan, Italy
[3]Center for Genomic Science of IIT@SEMM, Fondazione Istituto Italiano di Tecnologia, Via Adamello 16, 20139 Milano, Italy
[4]Lead contact
*Correspondence: ylli.doksani@ifom.eu
https://doi.org/10.1016/j.molcel.2026.04.003

SUMMARY

Extrachromosomal circular DNA (eccDNA) drives genome instability and tumorigenesis, warranting thorough investigations of its biogenesis. Here, we report a mechanism of eccDNA formation in human cells, distinct from the one mediated by joining of broken DNA ends. We show that repeats such as telomeres and centromeric alpha satellite form internal loops (i-loops) as a consequence of single-stranded DNA (ssDNA) breaks or gaps rather than double-stranded breaks (DSBs). I-loops are precursors for the excision of eccDNA, visible by electron microscopy (EM) and detectable via rolling-circle amplification. Apoptosis triggers the formation of i-loops and eccDNA at telomeric and alpha satellite repeats. Nanopore sequencing revealed other repetitive elements, including rDNA and retrotransposons, as sources of eccDNA in apoptosis. Based on the prevalence of SSBs over DSBs and the abundance of repeats in the human genome, we propose that the i-loop mechanism contributes substantially to all forms of eccDNA, with implications for tumor biology and genome evolution.

INTRODUCTION

Extrachromosomal circular DNA (eccDNA) is a ubiquitous byproduct of DNA metabolism that varies widely in size, ranging from a hundred nucleotides to Mb-sized circles. Smaller eccDNA forms are present both in normal and cancer cells and can contribute to genome plasticity, structural variations, and complex genomic rearrangements.[1–4] Larger eccDNA forms containing replication origins and oncogenes are frequently observed in cancer cells, likely as a result of positive selection. They were initially identified in metaphase spreads as double minute chromosomes and are now more generally referred to as ecDNA.[5–10] eccDNA is a newly appreciated hallmark of cancer, which can promote rapid gene amplification through asymmetric segregation and drive tumor heterogeneity and evolution.[11–15]

The simplest mechanism of eccDNA generation is circularization of broken DNA fragments, which can result from DNA damage, telomere dysfunction, or active cleavage by endonucleases involved in chromothripsis or programmed cell death.[16–22] A substantial fraction of eccDNA derives from repetitive DNA[23–25] and many classes of tandem repeats, including telomeric, centromeric, and rDNA, have long been known for their tendency to generate eccDNA.[26–28] This increased propensity of repeats to form circular DNA points to a recombination-based process underlying eccDNA generation. However, the molecular mechanism of eccDNA formation through repeat recombination remains unclear. Given that ~50% of the human genome is composed of repetitive elements, understanding how repeat sequences contribute to eccDNA formation is important.

Here, we document a general mechanism for eccDNA formation involving repetitive sequences. We previously showed that, in the presence of DNA damage, telomeric repeats form internal loops (i-loops) that can be excised, leading to the generation of telomeric circles.[29] We now show that the process of i-loop formation and excision is not unique to telomeres but represents a more general mechanism of eccDNA formation at repetitive elements. Similar to telomeres, pericentric alpha satellite repeats[30] accumulate i-loop structures in response to DNA damage, which can be excised to generate eccDNA. We show that i-loop formation is not efficiently induced by double-stranded breaks (DSBs) but is strongly promoted by single-stranded breaks, which are at least 1000-fold more prevalent than DSBs.[31–34] In addition, the early stages of apoptosis, a condition reported to generate large amounts of eccDNA,[22] are associated with a strong accumulation of i-loops and eccDNA generation from telomeric and alpha satellite repeats. Long-read sequencing of eccDNA products from apoptotic cells revealed a strong contribution of rDNA, retrotransposons, satellite DNA, and other repetitive elements to eccDNA formation. Our results reveal a second mechanism of eccDNA biogenesis, initiated by homologous interactions at SSBs in two repetitive elements that are in close three-dimensional proximity. This leads to the formation of i-loops, a subset

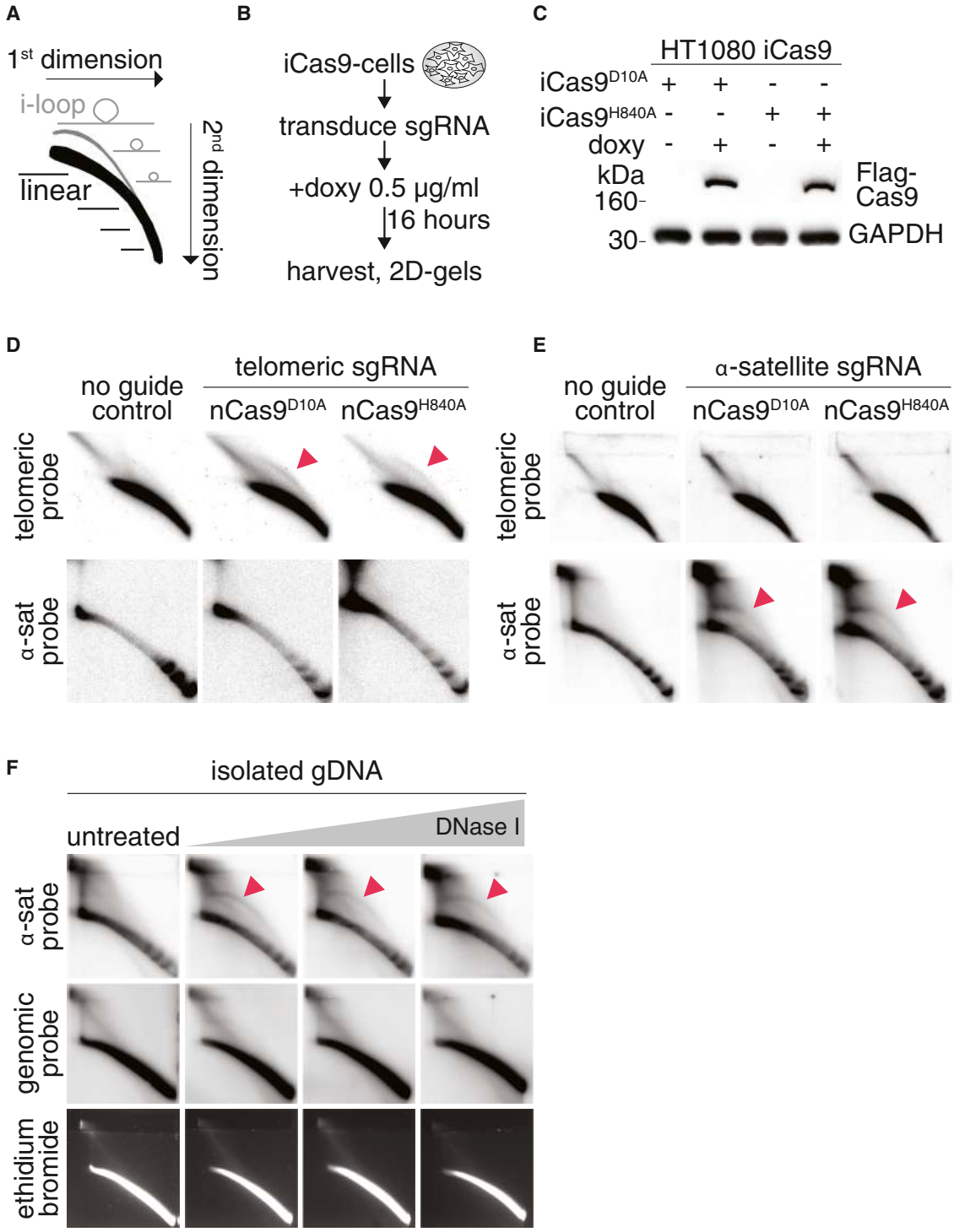


Figure 1. Accumulation of i-loop structures at telomeric and alpha satellite repeats upon single-stranded DNA damage

(A) Schematic representation of the main arcs detected by 2D-gels.

(B) Schematic representation of the experimental procedure.

(C) Western blot showing inducible expression of the Cas9 nickase variants. HT1080 cells containing doxy-inducible, Flag-tagged Cas9^{D10A} or Cas9^{H840A}, were transduced with a lentiviral construct expressing a telomere-specific sgRNA, and then Cas9 expression was induced by addition of 0.5 µM doxycycline for 16 h.

(D) 2D-gel analysis of telomeric alpha satellite repeats following the experimental conditions described in (B), with a telomeric sgRNA. Total DNA was digested with HinfI and RsaI, for telomeric repeats, or with MspI for alpha satellite repeats, and fragments above 3 kb were separated in 2D-gels. The DNA was then blotted

(legend continued on next page)

of which are excised as eccDNA, resulting in the deletion of the chromosomal sequence between the initial damage sites. This mechanism provides a molecular framework to investigate the mutants and cellular conditions associated with eccDNA accumulation. Given the abundance of repetitive elements in the genome, the i-loop excision mechanism may contribute to the generation of ecDNA species that confer selective advantages to cancer cells.

RESULTS

I-loop formation at alpha satellite repeats following ss DNA damage

We have previously shown that damaged telomeric repeats generate i-loop structures that can be visualized by electron microscopy (EM) and by two-dimensional agarose gel electrophoresis (2D-gels) as a slow-migrating arc (Figure 1A).[29] This arc has been previously attributed to circular DNA forms.[35–37] However, we showed that the arc can also be detected in the absence of circular telomeric DNA and is populated by molecules with telomeric i-loops, which are precursors for formation of circular DNAs.[29] These observations raised the question of whether other repetitive elements can form i-loops. We focused on alpha satellite DNA, which consists of 171 bp units repeated in tandem and constituting around 2.8%–3% of the genome and is enriched in peri-centromeric regions.[38] To monitor i-loop formation upon single-stranded DNA (ssDNA) damage, we induced SSBs in human HT1080 cells by transient expression of doxy-inducible Cas9 nickase variants ($Cas9^{D10A}$ or $Cas9^{H840A}$) along with a specific single-guide RNA (sgRNA) targeting alpha satellite repeats[39] (Figures 1B and 1C). In parallel, telomeric repeats were targeted with a specific sgRNA. The formation of DNA structures at these repetitive elements was monitored by 2D-gels. In control samples, both telomeric and alpha satellite repeats migrate as linear forms of heterogeneous length, with a more regular banding pattern generated by the alpha satellite repeats (Figures 1D and 1E). In the presence of the Cas9-induced SSBs, the slower-migrating arc, indicative of i-loop formation, became visible in 2D-gels, both at telomeres and alpha satellite repeats (Figures 1D and 1E). I-loop formation was site-specific with no cross-induction observed at the non-targeted repeats, consistent with i-loops forming in situ, as a direct consequence of DNA damage. Alkaline gels revealed that only a fraction of telomeres contained nicks, while DSBs (that would manifest as a low molecular weight (MW) signal on both telomeric strands) were not detected (Figure S1A), indicating that i-loops formation is induced by ssDNA damage and not by their conversion into DSBs. Consistently, DSBs generated at mouse telomeres with a TRF1-FokI fusion protein[40] or at alpha satellite with the wild-type Cas9 did not result in i-loop induction, unlike nCas9-induced SSBs (Figures S1B–S1D).

I-loop formation could initiate directly at SSBs or in a subpopulation of nicks that are processed into small ssDNA gaps. In the first case, i-loop formation would rely on strand-exchange activities in vivo, whereas in the second case, it would occur spontaneously, via annealing of exposed ssDNA gaps at damaged repeats. Indeed, we have shown that inductions of nicks and short gaps on isolated nuclei or genomic DNA (gDNA) with a low dose of DNase I treatment[41] can lead to a robust accumulation of i-loop structures at telomeric repeats.[29] In order to address whether the same type of transition occurs also at alpha satellite repeats, we performed the DNase I treatment on isolated HT1080 gDNA and monitored alpha satellite repeats structures by 2D-gels. This analysis revealed the presence of a damage-dependent i-loop arc (Figure 1F), indicating that alpha satellite repeats share with telomeric repeats a similar propensity to form i-loop structures spontaneously upon ssDNA damage. Importantly, the i-loop arc is not detectable with a pan-genomic probe, indicating that it is repeat-specific and not a general feature of DNase-I-treated DNA (Figure 1F).

EM analysis of alpha satellite repeats reveals damage-induced i-loops

To corroborate the presence of the i-loop structures in damaged alpha satellite repeats, we aimed to visualize i-loops by EM. To this end, we employed CenRICH, a recently developed procedure that consists of selective digestion of gDNA with restriction enzymes that cut rarely in the alpha satellite DNA, followed by size fractionation on sucrose gradients, which eliminates low MW non-centromeric DNA and yields preparations highly enriched in alpha satellite repeats.[42] I-loop formation was induced by mild DNase I treatment of isolated nuclei, as described above. The DNA was then psoralen crosslinked in nuclei to prevent branch migration during extraction and subjected to CenRICH (Figure S2A). This procedure led to 20–40-fold enrichment of alpha satellite repeats (Figures S2B and S2C). Under these conditions, we estimate that around 60% of the DNA molecules contain alpha satellite repeats.[42] In non-enriched samples, the background level of molecules with i-loops detected by EM was 4%, in line with previous experiments in human and mouse cells.[29] EM analysis of the damaged CenRICH preparations revealed a ~5-fold increase in the fraction of molecules containing i-loops, compared with the untreated enriched sample or DNase-I-treated bulk DNA (Figures 2A and 2B). Higher magnification of the EM images of i-loop junctions revealed the presence of nicks and ssDNA, similarly to what was observed for telomeric i-loops[29] (Figure S3). I-loop size in the alpha-satellite-enriched fraction ranged from 200 bp to 35 kb, with the median size of 3.11 kb (Figure 2C).

onto a membrane and hybridized with a probe recognizing telomeric, or alpha satellite repeats, respectively. Red triangles indicate the i-loop signal induced by the Cas9 nickase.

(E) 2D-gel analysis performed as described in (C), except that in this case, cells were transduced with a lentiviral construct expressing an alpha satellite-specific sgRNA.

(F) 2D-gel analysis showing the spontaneous formation of i-loop structures at alpha satellite repeats upon induction of nicks and single-stranded gaps on isolated DNA. gDNA from HT1080 cells was incubated with 0, 0.01, 0.02, and 0.03 ng/μl of DNase I for 8 min at RT. The reaction was stopped, and the samples were then processed for 2D-gels as described in (D) and hybridized with a probe recognizing alpha satellite repeats (top). After exposure, the probe was stripped, and the membrane was hybridized with a ^{32}P-labeled pan-genomic probe (middle). Ethidium bromide staining of the second dimension is also shown (bottom).

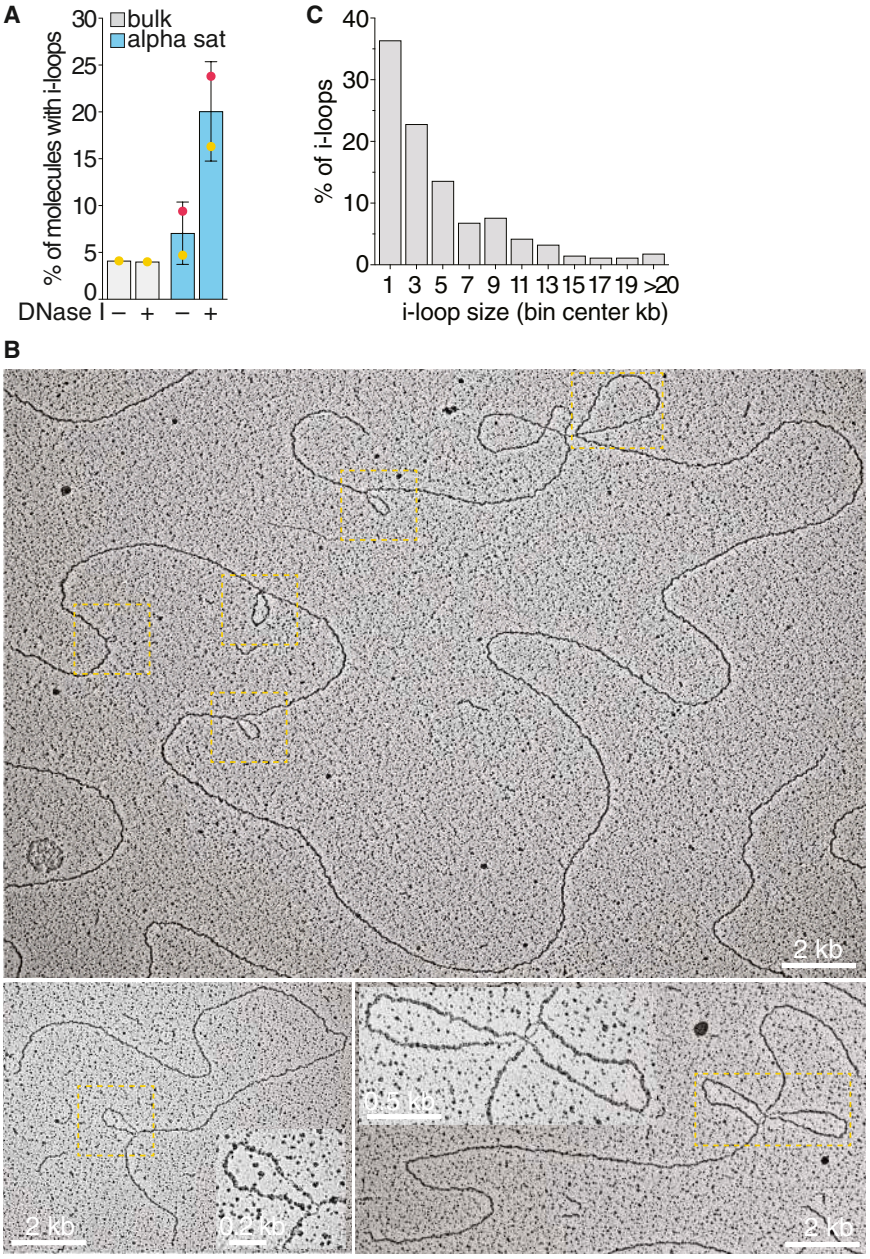


Figure 2. EM analysis of i-loops formed at damaged alpha satellite repeats

(A) Quantification of molecules with i-loops visualized by EM. Nuclei, prepared from HT1080 cells, were incubated with 0 and 6.5 ng/μl DNase I for 8 min at RT. Then, the DNA was psoralen cross-linked to prevent branch migration, before nuclear lysis. Total DNA was extracted and subjected to the CenRICH procedure (Figure S2). The DNA was then prepared for EM analysis by BAC-formamide spreading and acute-angle rotary shadowing with platinum. Bulk DNA was digested with KpnI before spreading. The graph reports the percentage of molecules containing at least one i-loop over all molecules. For the bulk sample, 1,357 and 995 molecules were analyzed from one experiment for NT and DNase-I-treated, respectively. Bars represent the mean with SD from two independent experiments for the enriched samples with 979; 646 molecules analyzed for the NT sample, and 1,152; 425 molecules for the DNase-I-treated sample.

(B) Examples of molecules with i-loops observed at damaged alpha satellite repeats, from the experiments described in (A). Yellow rectangles indicate i-loop structures. Insets in the bottom images represent 4× magnifications of the area inside the yellow rectangles.

(C) Size distribution of i-loops visualized at alpha satellite repeats in EM images from the experiments described above. n = 619 i-loops. Median i-loop size is 3.11 kb.

This value is about twice the median size of the i-loops observed at telomeres (1.6 kb),[29] perhaps due to the larger alpha satellite repeat length (171 bp versus 6 bp), which may result in fewer small (<400 bp) loops in the alpha satellite repeats. Finally, the frequency of spontaneous i-loops at alpha satellite repeats (detected in the untreated, enriched sample), as well as the absolute level of i-loops induced by DNA damage (Figure 2A), was lower than what was previously observed at telomeres,[29] indicating a higher propensity of telomeric repeats to form i-loops compared with alpha satellite repeats. This is consistent with the probability of exposure of homology at two sites of damage, and therefore the probability of i-loop formation, being inversely correlated with the length of the repeated unit.

I-loops are a precursor of the alpha satellite repeat eccDNA

Apart from the accumulation of i-loops, EM analysis of damaged alpha satellite repeats revealed a substantial presence of circular DNA. These forms reached up to ∼1.8 % of all molecules in the DNase-I-treated CenRICH samples, while being only occasionally detected in the control samples (Figures 3A and 3B). This result suggests that some of the i-loops may be excised and released as eccDNA by Holliday junction resolvases or other endonucleases active in the nuclear preparation. In both cases, the excised circular forms would contain at least one nick on each strand.[43] Indeed, rolling-circle amplification (RCA) reaction did not detect closed eccDNA forms of telomeric or alpha satellite origin in these samples (Figures 3C and 3D). However, complementing the nuclear preparation with ATP, to allow nick sealing, led to accumulation of closed circular species detectable by RCA at both telomeric and alpha satellite repeats (Figures 3C and 3D). Alpha satellite i-loop excision was confirmed with a biochemical assay developed at telomeric repeats,[29] in which i-loops are first induced by mild DNase I digestion of genomic DNA, then briefly incubated with a cellular extract to allow excision and detection of eccDNA by RCA (Figures S4A and S4B). Collectively, these

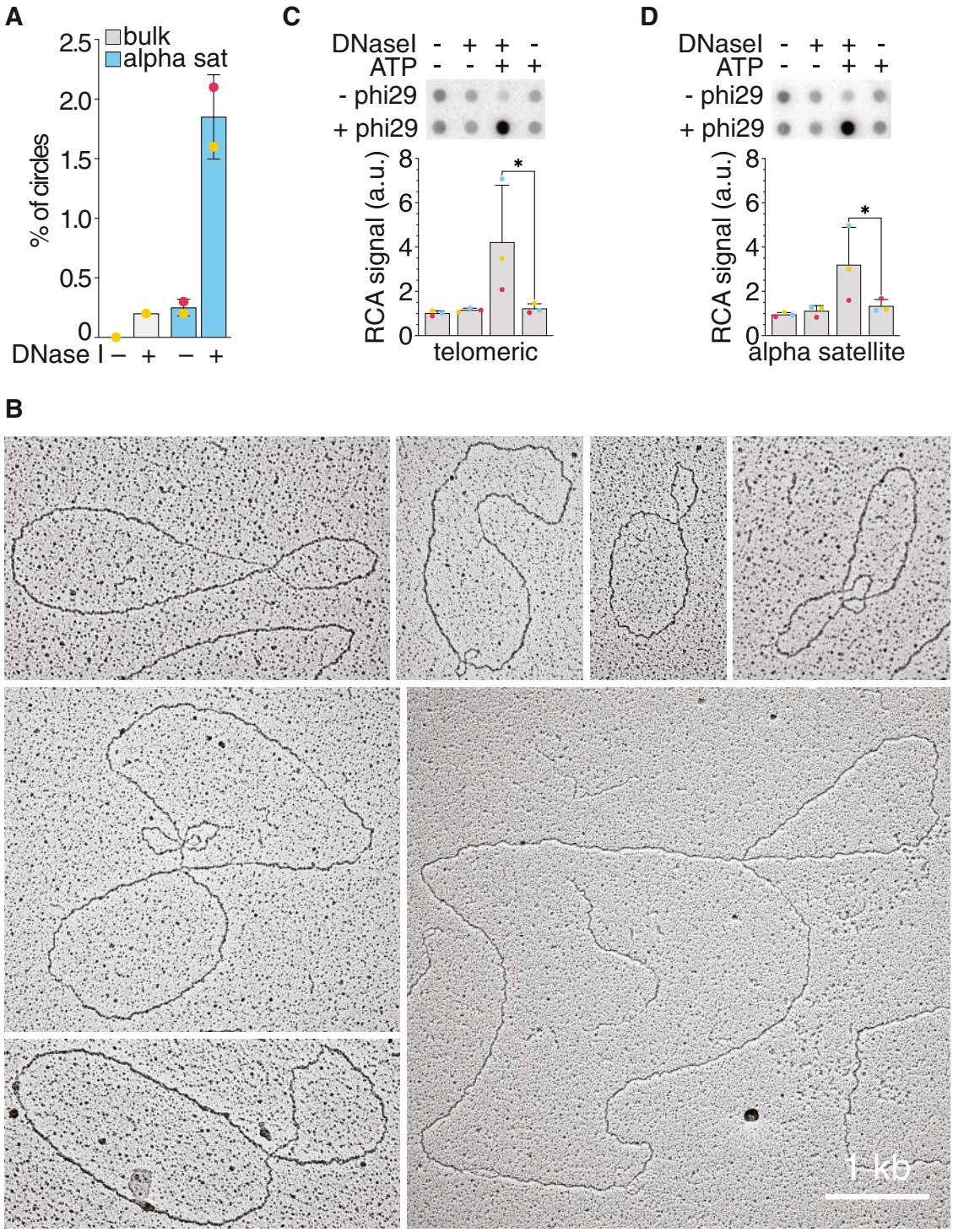


Figure 3. Formation of eccDNA at damaged alpha satellite repeats detected by EM and RCA

(A) Quantification of the circular molecules visualized by EM from the same experiments described in Figure 2A. Bars represent the mean with SD from two independent experiments for the enriched samples. NT n = 979 and 646 molecules, DNase-I-treated n = 1,152; 425 molecules.

(B) Examples of the circular molecules visualized by EM in the alpha satellite-enriched samples exposed to DNA damage.

(C) RCA of eccDNA induced by DNA damage at telomeric repeats. Isolated HT1080 nuclei were incubated with 0 and 6.5 ng/μl DNase I for 8 min at RT. In parallel, one preparation was supplemented with 2 mM ATP. The DNA was then isolated, subjected to a RCA with the phi29 polymerase, in the absence of oligonucleotides, allowing replication from pre-existing nicks or gaps. Top: Dot-blot analysis of RCA products. The reaction was blotted on a membrane and hybridized with the Sty11 probe recognizing telomeric repeats. Bottom: Quantification of the dot-blot signal from 3 independent experiments. The values were reported over

(legend continued on next page)

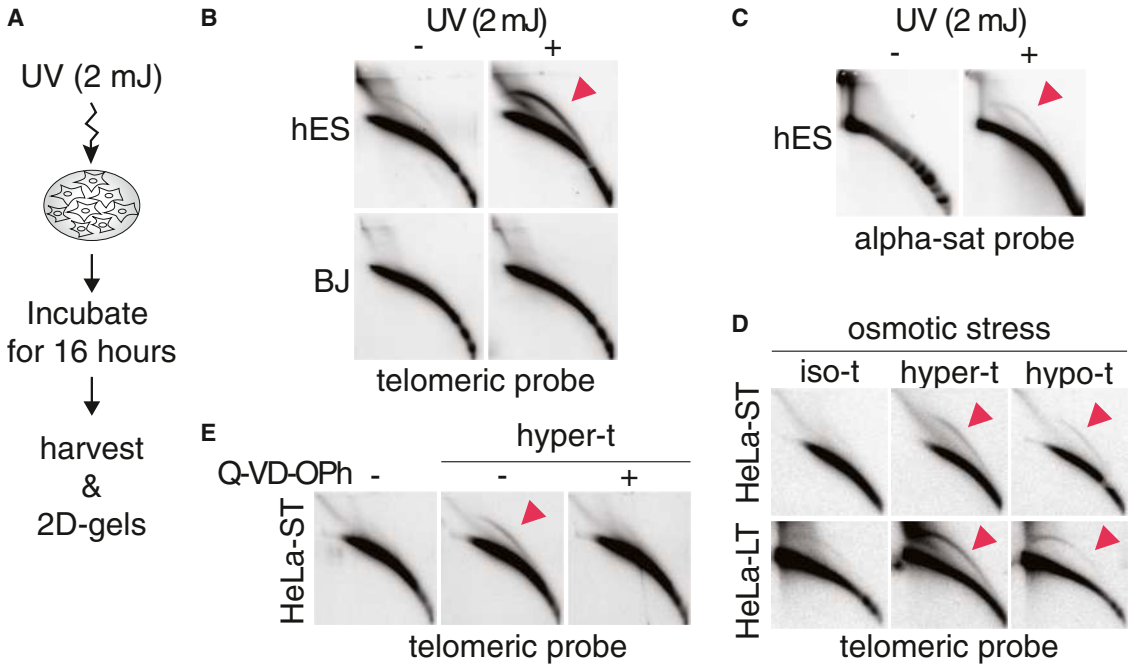


Figure 4. Accumulation of i-loops at telomeric and alpha satellite repeats during apoptosis

(A) Schematic representation of the experimental procedure. The indicated cell lines were exposed to a dose of 2 mJ UVC and collected 16 h later. Under these conditions, the hES cells undergo apoptosis, while the differentiated BJ fibroblasts do not (Figure S5A).

(B) 2D-gel analysis showing a strong accumulation of i-loops at telomeres in apoptotic cells, from an experiment as the one described in (A). Total DNA was collected, digested with RsaI and HinfI, and fragments above 3 kb were separated by 2D-gels. The DNA was then blotted onto a membrane and hybridized with the Sty11 probe, recognizing telomeric repeats. The red triangle indicates the i-loop arc.

(C) 2D-gel analysis showing a strong accumulation of i-loops at alpha satellite repeats in apoptotic cells. The same experimental procedure as in (A) was performed, except that the DNA was digested with MspI and the membrane was hybridized with a probe recognizing the alpha satellite repeats. The red triangle indicates the i-loop arc.

(D) 2D-gels showing that induction of apoptosis by osmotic stress leads to the accumulation of i-loops. HeLa cells with long (~20 kb) and short (3–10 kb) telomeres were incubated with 1M sorbitol (hypertonic) for 1 h or with H2O (hypotonic) for 30 min. Cells were then incubated with media for another 3 h before harvesting and extraction of total DNA. The DNA was then separated in 2D-gels as described in (A).

(E) 2D-gels showing that accumulation of i-loops in apoptotic cells is dependent on caspase activity. HeLa cells with short telomeres were exposed to hypertonic stress as described in (C). In one sample, the pan-caspase inhibitor Q-VD-OPh was added at a 10 μM final concentration. The DNA was then collected and analyzed by 2D-gels with the procedure described in (C).

results show that, similar to telomeres, centromeric alpha satellite repeats, experiencing ssDNA damage, may undergo the i-loop formation and excision mechanism leading to eccDNA generation.

I-loop formation and eccDNA generation at repetitive elements during programmed cell death

A recent study has reported the generation of abundant eccDNA forms during apoptosis, the most common form of programmed cell death.[22] In these conditions, eccDNA could form as a result of circularization of DNA fragments generated by caspase-activated DNases (CAD).[22] Given the involvement of DNases in programmed cell death, we wondered whether ssDNA damage, accumulating during the early stages of the apoptotic process, would also promote the formation of i-loops at repetitive elements and their excision as eccDNA.[44–48] To test this hypothesis, we exposed cells to a 2 mJ dose of UV, which leads to a robust activation of apoptosis in human embryonic stem (hES) cells but not in BJ fibroblasts (Figures 4A and S5A). This treatment was associated with a robust accumulation of the i-loop arc at telomeric repeats in hES cells but not in BJ cells (Figures 4B and S5B). Induction of i-loop structures was also clearly visible at alpha satellite repeats, albeit at a lower magnitude than telomeres (Figures 4C and S5C). Formation of i-loops was not dependent on the type of stimulus that induced cell death. Indeed, cell death induced by exposure to osmotic stress or over-confluence led to a clear accumulation of telomeric i-loops in HeLa cells with either short or long

each of the -phi29 controls. Bars represent the mean with SD. *p* values were obtained by one-way ANOVA followed by Sidák's post-hoc correction for multiple comparisons.

(D) RCA of eccDNA induced by DNA damage at alpha satellite repeats, following the same procedure described in (C), except that the membrane was hybridized with a probe recognizing the alpha satellite repeats. Bars represent the mean with SD. *p*-values were obtained by one-way ANOVA followed by Sidák's post-hoc correction for multiple comparisons.

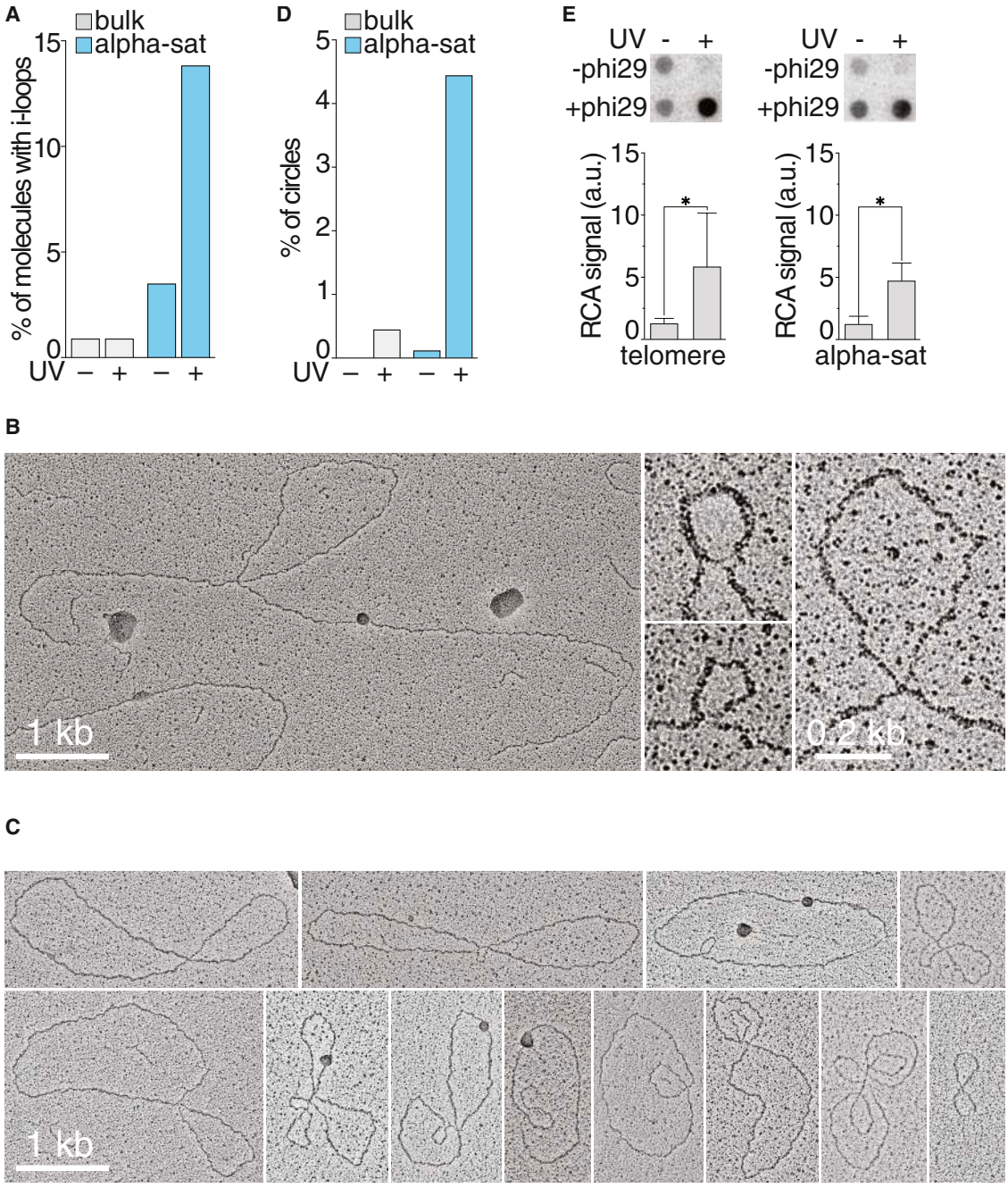


Figure 5. EM and RCA analysis of i-loops and eccDNA at alpha satellite repeats in apoptotic cells

(A) Quantification of molecules with i-loops visualized by EM. Apoptosis was induced in hES cells by UV treatment as described in Figure 4A. Then, the DNA was psoralen crosslinked to prevent branch migration, before nuclear lysis. Total DNA was extracted and subjected to the CenRICH procedure (Figure S6). The DNA was then prepared for EM analysis by BAC-formamide spreading and acute-angle rotary shadowing with platinum. Bulk DNA was digested with KpnI before spreading. The graph reports the percentage of molecules containing at least one i-loop over all molecules. Bulk NT, n = 449; bulk apoptotic, n = 448; alpha sat NT, n = 854; alpha sat apoptotic, n = 962 molecules.

(B) Examples of molecules with i-loops from the alpha sat apoptotic sample from the experiment described in (A).

(C) Examples of circular molecules from the alpha sat apoptotic sample from the experiment described in (A).

(D) Quantification of circular forms as a percentage of all molecules, from the experiment described in (A).

(E) RCA of eccDNA generated from telomeres and alpha satellite repeats during apoptosis. Total DNA was prepared from an experiment similar to (A), except that the psoralen crosslinking step was omitted. The DNA was subjected to a RCA with the phi29 polymerase, in the absence of

(legend continued on next page)

telomeres (Figures 4D and S5D). Finally, accumulation of the i-loop arc was abolished by the treatment with the pan-caspase inhibitor Q-VD-OPh,[49] consistent with its dependence on the CAD activity (Figures 4E, S5E, and S5F).

In addition to 2D-gel analysis, we directly visualized the structure of alpha satellite repeats in hES cells undergoing apoptosis, induced by UV treatment. gDNA was psoralen crosslinked before cell lysis, to prevent branch migration. Alpha satellite repeats were then enriched using the CenRICH procedure explained above (Figure S6). EM analysis in these conditions revealed a ~3-fold increase in the fraction of molecules containing i-loops (Figures 5A and 5B), consistent with the results obtained by 2D-gels. Apart from i-loops, we also detected many eccDNA forms, reaching up to 4% of the DNA molecules in the centromere-enriched fraction from apoptotic cells (Figures 5C and 5D). EM analysis revealed an increase in circular forms also in the bulk sample from apoptotic cells, consistent with the reported genome-wide generation of eccDNA during apoptosis.[22] Finally, generation of eccDNA forms at telomeres and alpha satellite repeats was detected also by an RCA assay (Figure 5E).

Given the induction of i-loops and eccDNA from telomeres and centromeres during apoptosis, we determined the contribution of other repetitive elements to eccDNA generation during this process. To this end, we induced apoptosis in hES cells by exposure to 2 mJ UVC, as above, and performed long-read sequencing of eccDNA products with a procedure similar to the one adopted in Wang et al.,[22] with some modifications as follows. We anticipated that a substantial fraction of the eccDNA generated by i-loop excision might contain nicks and/or small gaps (Figure S3).[29,50] Therefore, instead of purifying eccDNA via alkaline lysis, total gDNA from apoptotic cells was isolated and subjected to RCA. The RCA reaction was performed in the absence of specific primers or random hexamers, which leads to the specific amplification of circular molecules containing pre-existing nicks or gaps.[51] The RCA product was subjected to long-read nanopore sequencing, alongside gDNA from untreated cells as a control. We obtained 275,973 reads with a mean size of 4.8 kb from the RCA sample. Using a modified version of the full-length eccDNA caller pipeline[22] (see STAR Methods), we identified 112,744 reads derived from RCA of circular molecules with at least two full passes (i.e., two complete cycles of amplification of a circular molecule, identified as reads that contain at least two mapped fragments that align to the same genomic location). In contrast, the same pipeline identified only one false-positive RCA product (i.e., a read that matched the two-pass criterion) in 128,853 reads of the gDNA control. Coverage analysis showed that these eccDNA forms were distributed throughout the genome (Figure 6A), which is consistent with them being generated via two independent mechanisms: re-ligation of broken DNA fragments[22] and excision of i-loops from repetitive elements distributed throughout the genome. The same analysis, however, revealed that rDNA loci are hotspots for eccDNA generation in apoptotic cells (Figure 6B), which would not be expected simply by random fragmentation and re-ligation, but is more consistent with a DNA-damage- and homology-driven i-loop formation and excision mechanism.

Repeat masker analysis was used on the RCA monomers identified above, in order to assess potential differences among different repetitive elements in the formation of eccDNA. All classes of repeats were detected in the eccDNA population, with the more abundant interspersed repeats being the most represented (Figure 6C). However, when coverage data were normalized for the relative abundance of each repetitive element, certain classes of repeats showed a significantly higher propensity to generate eccDNA in these conditions (Figure 6D). These classes included satellite DNA, retrotransposons, and rDNA. Among repeat families, the highest enrichment in the RCA sample compared with the genomic control was shown by the LSU-rRNA_Hsa (the large subunit ribosomal RNA); centromeric satellite repeats, consistent with the previous data; retrotransposon SVA (a small family of transposons, containing SINE, VNTR, and Alu elements[53]) and other satellite repeats (Figure S7).

These data suggest a second mechanism of eccDNA generation, acting in parallel to re-ligation of broken DNA fragments. The findings are consistent with the long-known propensity of repetitive elements to generate eccDNA. Based on the results shown here, we propose that ssDNA damage exposing homology at repetitive elements contributes to eccDNA generation, via i-loop formation and excision (Figure 7).

Involvement of repetitive elements in ecDNA formation in tumors

Based on the data above, we decided to investigate the involvement of repetitive elements in the generation of large eccDNA forms that accumulate in tumors. Because these forms are well above the mechanical fragmentation limit of DNA, they cannot be monitored directly by 2D-gels or EM. Therefore, we decided to look for the presence of repetitive elements in proximity to the ecDNA junctions in publicly available ecDNA data from the AmpliconRepository.[54] Specifically, we scored events in which repeat elements, belonging to the same family, occur on both sides of the amplicon, within defined windows across the junction breakpoints. These data were then compared with a randomized set of events that had the same length distribution and the same chromosomal origin as the real set. The results indicate a clear and significant enrichment of the occurrence of same-family repeats on both junction breakpoints of simple ecDNA events (the class of events compatible with the i-loop excision mechanism) (Figures 6E–6G). Importantly, the same level of enrichment was not found in amplifications with a foldback structure, compatible with breakage-fusion-bridge (BFB) cycles.[55] Depending on the window, these events reached up to 10% of the analyzed amplicons (Figure 6G). Considering the aggressive filtering against repetitive elements employed by the amplicon architect,[56] we

oligonucleotides, allowing replication from pre-existing nicks or gaps. Top: Dot-blot analysis of RCA products. The reaction was blotted on a membrane and hybridized either with the Sty11 probe recognizing telomeric repeats (left) or with a probe recognizing the alpha satellite repeats (right). Bottom: Quantification of the dot-blot signal from 3 independent experiments. Bars represent the mean with SD. *p* values were calculated using an exact one-tailed Mann-Whitney test.

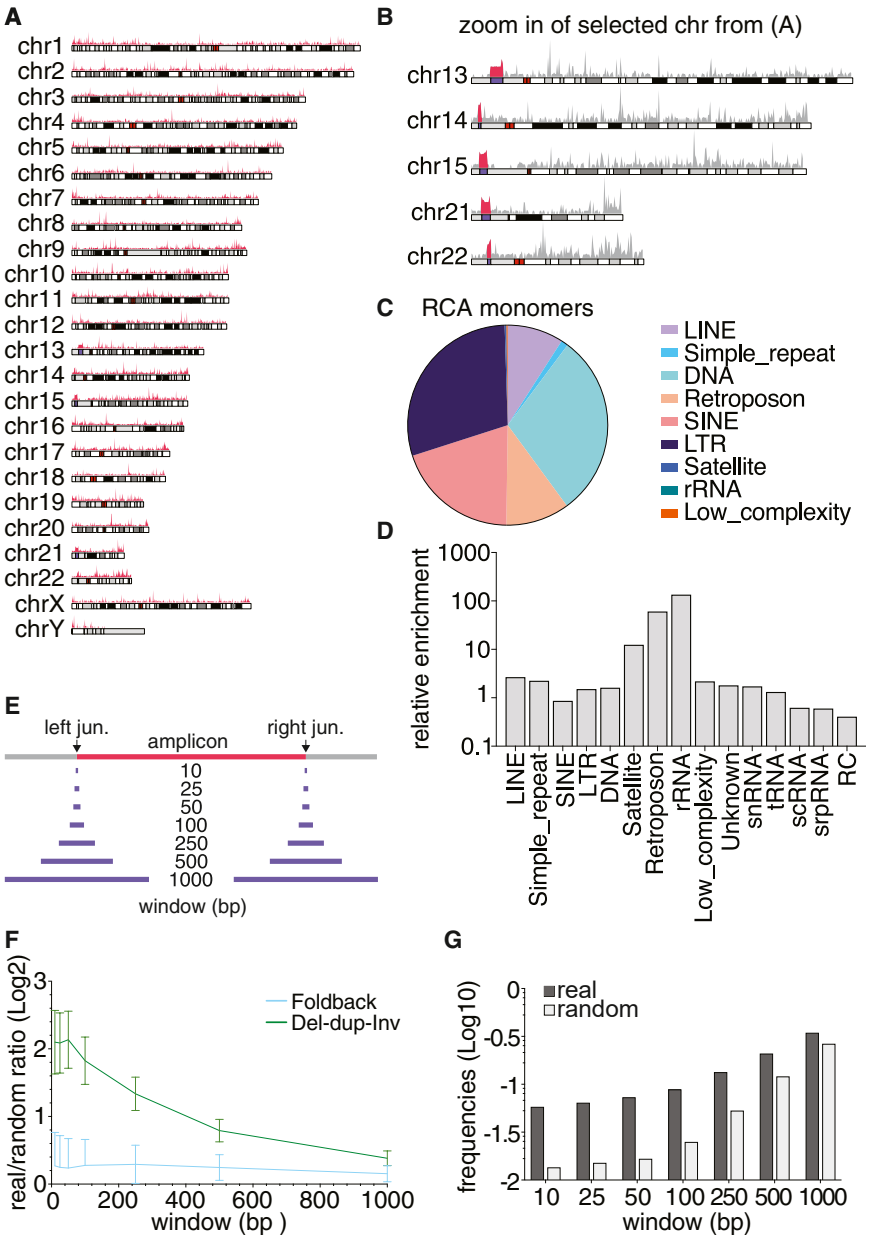


Figure 6. Genome-wide location of eccDNA forms generated during apoptosis and enrichment of repetitive elements at ecDNA junctions

(A) Coverage map of RCA monomers. Apoptosis was induced in hES cells with the same procedure described in Figure 4A. Total DNA was subjected to RCA in the absence of oligonucleotides, allowing replication from pre-existing nicks and gaps, and the reaction was subjected to long-read nanopore sequencing. RCA products and the monomer unit were identified with a pipeline similar to the one described in Wang et al.[22] (see STAR Methods). The coverage map shows the overall distribution of RCA monomers across the genome.

(B) Zoom in on the indicated chromosomes from the coverage graphs shown in (A). In order to highlight the coverage, the region corresponding to the rDNA loci is shown in orange.

(C) Pie chart showing the distribution of different repeat classes in the RCA monomers.

(D) Quantification of the relative abundance of different repeat classes in the RCA monomers, normalized to their abundance on the genome. The values represent the percentage of eccDNA molecules containing a certain repeat class divided by the relative fraction of that class among all genomic repeats.

(E) Schematic representation of the analysis of repetitive elements at the amplicon junctions. Amplicon intervals of simple ecDNA forms from the Hartwig Medical Foundation (HMF) project, processed by the Verhaak and Kim labs,[52] were first stratified in distinct classes (e.g., deletion-like, duplication-like, foldback, simple circular, etc). The subsequent analysis was performed independently for each class. Junction intervals were converted to left (L) and right (R) junctions. These junctions were symmetrically expanded by ±W base pairs to generate the indicated windows. The windows were then intersected with hg19 RepeatMasker annotations (grouped by family). A "match" was defined as a junction sharing at least one common repeat family between its L and R windows. A length and chromosome-preserving randomized control was generated for each class by relocating original intervals to new random positions within their respective chromosomes. These control sets were processed identically. Finally, match frequencies were compared between the real and random data for each window size.

(F) Ratio of real/random frequencies for the deletion-like, duplication-like, inversion (del-dup-inv), and foldback events. Error bars represent 95% confidence intervals computed using the Katz log method. (n = 3,152 for the del-dup-inv and 880 for the foldback.) The same type of graph for the PCAWG data is shown in Figure S6E.

(G) Absolute frequencies of real and random events from the data shown in (F). Significance was assessed per window, using a two-sided Fisher's exact test < 3.93×10^{-12}. The same type of graph for the PCAWG data is shown in Figure S6F.

suggests that this result indicates an involvement of repetitive elements in the generation of simple ecDNA forms during tumor progression.

DISCUSSION

The propensity of repetitive elements to form eccDNA has been reported in various organisms, including yeast, plants, humans, and others.[27,57–61] In addition, DNA damage promotes eccDNA formation.[29,35,62] Collectively, these studies point toward recombination-based mechanisms; however, the precise pathways underlying eccDNA generation from repetitive sequences remain unclear.

Recombination-based vs spontaneous i-loop formation

Multiple DSBs can generate eccDNA genome-wide through intramolecular end-joining of broken fragments. Here we show that eccDNA formation at repeats can occur efficiently via

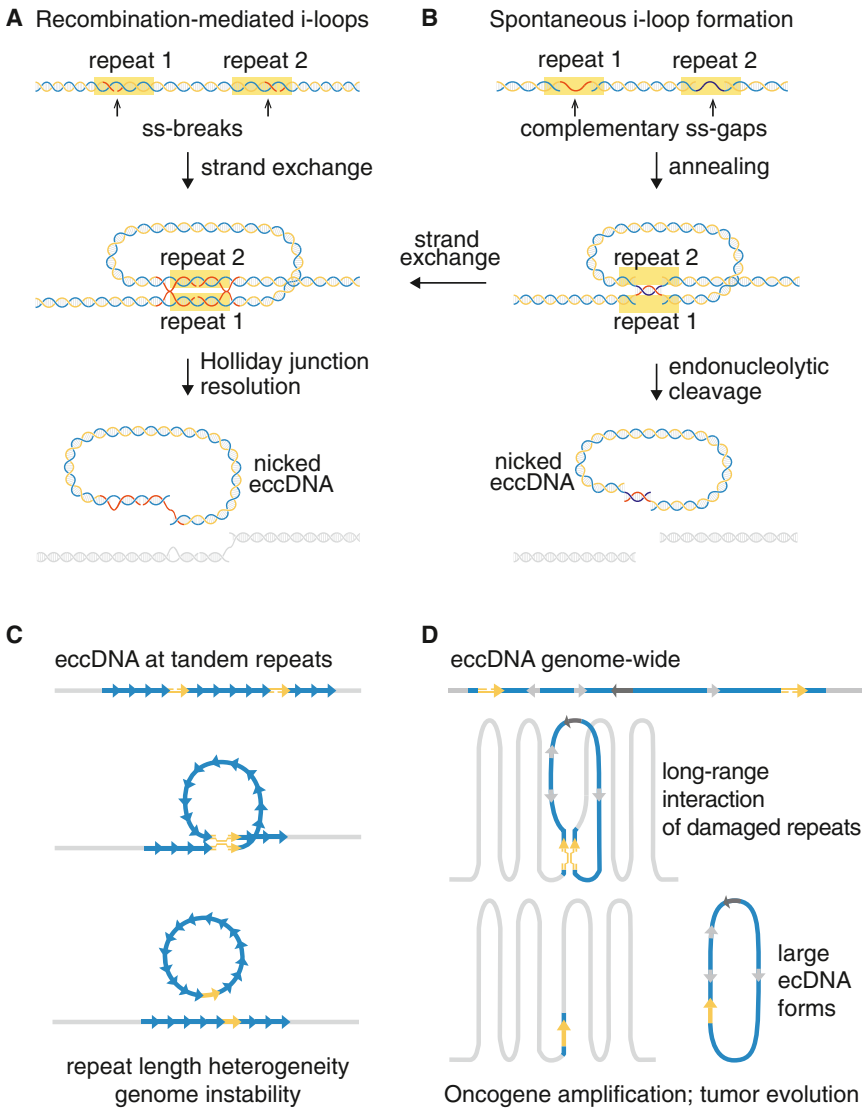


Figure 7. Schematic representation of possible models of i-loops and eccDNA generation from repetitive elements

(A) I-loop formation at two single-stranded breaks would be dependent on strand-exchange activities, normally involved in homologous recombination. In this case, branch migration would generate Holliday junctions at the base of the loop that, when resolved by endonucleolytic cleavage, could give rise to eccDNA excision.

(B) Spontaneous i-loop formation from complementary single-stranded gaps. In this scenario, an i-loop structure would be generated simply by annealing of the complementary strands, as proposed in.[29] Nucleolytic cleavage of the i-loop junction could lead to the formation of eccDNA form, held together by base pairing. In both scenarios, i-loop excision alone would give rise to nicked eccDNA, stabilized by base pairing. Formation of covalently closed circles would require nick sealing, typically done by Lig1 or Lig3 activities.

(C) The mechanism of i-loop formation and excision would be very efficient at tandem repeats, due to the local abundance of homology. In regions such as telomeres and alpha satellite repeats, this mechanism leads to abundant formation of i-loops ranging from a few hundred bp to 20–30 kb in size. These structures and the derived eccDNA forms are suitable for structural analysis by EM and 2D-gels.

(D) The formation and excision of i-loops might contribute to the generation of larger ecDNA forms. Interaction between two damaged repeats located at distal sites on the chromosome, but found in close spatial proximity by chromatin looping, could lead to the formation of large eccDNA molecules. Such events are expected to be far less frequent than recombination at tandem repeats, and the size of the resulting i-loops and eccDNA forms would exceed the size limit of detection by 2D-gel electrophoresis and EM. However, the abundance and distribution of repetitive elements across the genome might create numerous opportunities for these distal interactions. Such interactions would generate large, replicating eccDNA forms that, depending on their genetic content, may provide a growth advantage and therefore undergo amplification, potentially driving tumorigenesis.

excision of i-loop intermediates that are strongly induced by ssDNA damage, rather than DSBs. Indeed, in the presence of homology, RecA-like strand-exchange activities may facilitate intramolecular recombination at two sites of damage, even in the absence of a DSB intermediate.[63–66]

Importantly, our results indicate that in the presence of short ss-gaps within repetitive elements, i-loop formation may occur spontaneously through annealing of complementary sequences, independently of strand-exchange activities (Figures 7A and 7B). This finding is consistent with previous genetic studies in *E. coli* and *S. cerevisiae*, which demonstrated that recombination between repeats leading to fragment deletions can occur independently of RecA or Rad51.[67–69] Taken together, our results underscore the critical role of cellular mechanisms that suppress recombination between repetitive elements in preventing the generation of eccDNA.[70]

Small eccDNAs at tandem repeats versus long-range interactions

The mechanism of i-loop formation and excision is particularly prominent at tandem repeat regions such as telomeres, alpha satellite repeats, and rDNA, likely due to the abundance of local homology. This feature, combined with selective enrichment methods for telomeric or alpha satellite DNA, allowed us to directly visualize i-loop intermediates by using EM. Such direct structural characterization of i-loops and eccDNA, however, is

constrained to DNA molecules smaller than ~50 kb due to mechanical fragmentation. Nonetheless, the mechanism itself is not restricted to specific repeat types, but rather, it is initiated whenever homologous sequences become exposed. Given that our chromosomes are densely populated by repetitive elements constituting up to 50% of the human genome, even rare interactions between damaged repeats at distant genomic locations but positioned in close three-dimensional proximity could substantially contribute to the formation of larger ecDNA molecules (Figures 7C and 7D). This pathway likely acts alongside end ligation-mediated joining of DNA fragments generated by multiple double-strand breaks during chromothripsis, BFB cycles, or apoptosis.[17,21,22] The potential relevance of the i-loop excision pathway to ecDNA generation is further strengthened by the fact that it is triggered by single-stranded nicks and gaps, forms of DNA damage that are substantially more common and less lethal than DSBs.

In addition, our observation that early stages of apoptosis coincide with increased eccDNA generation from repetitive elements, including telomeric repeats, warrants a careful evaluation of the genetic conditions previously linked to telomeric circle accumulation[71] to exclude indirect effects arising from increased cell death. While i-loops can be strongly induced at different repeats genome-wide during apoptosis, our results with the Cas9 nickase targeted at telomeres and alpha satellite repeats clearly indicate that i-loops form also in non-apoptotic cells in a site-specific manner, as a direct consequence of ssDNA damage.

Our results elucidate mechanistic details and structural intermediates involved in eccDNA formation, offering insights into how repetitive sequences contribute to genome instability and eccDNA production in response to DNA damage.

Limitations of the study

The size of the i-loops and eccDNA that can be separated in 2D-gels and visualized by EM is ultimately limited to ~50–100 kb, due to mechanical fragmentation of the DNA and resolution of agarose electrophoresis. Therefore, although our data support i-loop formation and excision as a mechanism of large ecDNA generation, this study cannot directly visualize the larger ecDNA species that may be produced in the same conditions.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ylli Doksani (ylli.doksani@ifom.eu).

Materials availability

Plasmids and other reagents generated in this study are available from the lead contact upon reasonable request.

Data and code availability

- Nanopore sequencing data of RCA amplification products have been deposited at EBI-EMBL: https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-16581 and are publicly available as of the date of publication. Original images and source data have been deposited at Mendeley: 10.17632/hwf38mddyt.1 and are publicly available as of the date of publication. The data for ecDNA junction analysis were publicly available at Amplicon Repository: https://www.ampliconrepository.org/project/655c060abba7c925095555da and https://www.ampliconrepository.org/project/68e805aac9508d1776e523b5.
- All original code for analyzing eccDNA from Nanopore sequencing as well as for amplicon junction analysis has been deposited at Github: https://github.com/ifom/NERP/ and is publicly available at Zenodo: 10.5281/zenodo.19344238 as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We are grateful to Giulia Mazzucco for the assistance with the initial EM experiments, Elisa Allievi and Cinzia Cancellieri for assistance with the ES cultures, and Francesco Ferrari and Martin Schaefer for feedback and advice on the bioinformatic analysis. We are grateful to the IFOM Cell Biology Unit for their assistance with the handling and storing of cell lines and the IFOM Genomics Unit for support and advice for the sequencing experiments. We are grateful to Titia de Lange, Francisca Lottersberger, Giorgio Scita, and YD lab members for their feedback on the manuscript. Y.D.'s lab is supported by the Associazione Italiana per la Ricerca Sul Cancro (AIRC IG 28954), the Concern Foundation, and Worldwide Cancer Research (WWCR 24-0166). F.N.'s lab is supported by the Associazione Italiana per la Ricerca Sul Cancro (AIRC IG 22851 and 32172), Marie Skłodowska-Curie Actions (MSCA) LongTREC (ID: 101072892) and INT2ACT (ID: 101226670), and from the National Center for Gene Therapy and Drugs based on RNA Technology (CN00000041) supported by European Union - Next Generation EU, Mission 4, Component 2, CUP B93D21010860004.

AUTHOR CONTRIBUTIONS

E.Z. performed the experiments and the bioinformatic analysis and contributed to the experimental design. M.G. assisted with EM sample preparation and acquisition. S.B. and F.N. assisted with the nanopore sequencing experiment. Y.D. designed and supervised the research and prepared the manuscript with inputs from E.Z. and feedback from the other authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Bacterial strains and culture
  - Human embryos and/or stem cells
  - Cell lines and culture conditions
- METHOD DETAILS
  - Generation of inducible Cas9 nickase cell lines
  - Isopropanol precipitation of DNA
  - Standard PCR
  - Mutagenic PCR
  - List of primers use for mutagenic PCR
  - DNA fragment purification from gel
  - DNA ligation
  - Transformation of chemically competent E. coli cells
  - Miniprep DNA preparation
  - Maxiprep DNA preparation
  - Cell lysates preparation and immunoblotting
  - DNA damage induction in isolated nuclei and genomic DNA
  - Apoptosis induction

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2026.04.003.

Received: July 28, 2025
Revised: March 19, 2026
Accepted: April 3, 2026
Published: April 30, 2026

REFERENCES

1. Shibata, Y., Kumar, P., Layer, R., Willcox, S., Gagan, J.R., Griffith, J.D., and Dutta, A. (2012). Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. Science *336*, 82–86. https://doi.org/10.1126/science.1213307.
2. Møller, H.D., Mohiyuddin, M., Prada-Luengo, I., Sailani, M.R., Halling, J.F., Plomgaard, P., Maretty, L., Hansen, A.J., Snyder, M.P., Pilegaard, H., et al. (2018). Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. Nat. Commun. *9*, 1069. https://doi.org/10.1038/s41467-018-03369-8.
3. Koche, R.P., Rodriguez-Fos, E., Helmsauer, K., Burkert, M., MacArthur, I.C., Maag, J., Chamorro, R., Munoz-Perez, N., Puiggròs, M., Dorado Garcia, H., et al. (2020). Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. Nat Genet *52*, 29–34. https://doi.org/10.1038/s41588-019-0547-z.
4. Pascarella, G., Hon, C.C., Hashimoto, K., Busch, A., Luginbühl, J., Parr, C., Hin Yip, W., Abe, K., Kratz, A., Bonetti, A., et al. (2022). Recombination of repeat elements generates somatic complexity in human genomes. Cell *185*, 3025–3040.e6. https://doi.org/10.1016/j.cell.2022.06.032.
5. Cox, D., Yuncken, C., and Spriggs, A.I. (1965). Minute chromatin bodies in malignant tumours of childhood. Lancet *1*, 55–58. https://doi.org/10.1016/S0140-6736(65)90131-5.
6. Lubs, H.A., and Salmon, J.H. (1965). THE CHROMOSOMAL COMPLEMENT OF HUMAN SOLID TUMORS. II. KARYOTYPES OF GLIAL TUMORS. J. Neuro.Surg. *22*, 160–168. https://doi.org/10.3171/jns.1965.22.2.0160.
7. Schwab, M., Alitalo, K., Klempnauer, K.H., Varmus, H.E., Bishop, J.M., Gilbert, F., Brodeur, G., Goldstein, M., and Trent, J. (1983). Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. Nature *305*, 245–248. https://doi.org/10.1038/305245a0.
8. Kohl, N.E., Kanda, N., Schreck, R.R., Bruns, G., Latt, S.A., Gilbert, F., and Alt, F.W. (1983). Transposition and amplification of oncogene-related sequences in human neuroblastomas. Cell *35*, 359–367. https://doi.org/10.1016/0092-8674(83)90169-1.
9. Storlazzi, C.T., Fioretos, T., Surace, C., Lonoce, A., Mastrorilli, A., Strömbeck, B., D'Addabbo, P., Iacovelli, F., Minervini, C., Aventin, A., et al. (2006). MYC-containing double minutes in hematologic malignancies: evidence in favor of the episome model and exclusion of MYC as the target gene. Hum. Mol. Genet. *15*, 933–942. https://doi.org/10.1093/hmg/ddl010.
10. Bafna, V., and Mischel, P.S. (2022). Extrachromosomal DNA in Cancer. Annu. Rev. Genomics Hum. Genet. *23*, 29–52. https://doi.org/10.1146/annurev-genom-120821-100535.
11. Turner, K.M., Deshpande, V., Beyter, D., Koga, T., Rusert, J., Lee, C., Li, B., Arden, K., Ren, B., Nathanson, D.A., et al. (2017). Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature *543*, 122–125. https://doi.org/10.1038/nature21356.
12. Bailey, C., Pich, O., Thol, K., Watkins, T.B.K., Luebeck, J., Rowan, A., Stavrou, G., Weiser, N.E., Dameracharla, B., Bentham, R., et al. (2024). Origins and impact of extrachromosomal DNA. Nature *635*, 193–200. https://doi.org/10.1038/s41586-024-08107-3.
13. Hung, K.L., Jones, M.G., Wong, I.T.L., Curtis, E.J., Lange, J.T., He, B.J., Luebeck, J., Schmargon, R., Scanu, E., Brückner, L., et al. (2024). Coordinated inheritance of extrachromosomal DNAs in cancer cells. Nature *635*, 201–209. https://doi.org/10.1038/s41586-024-07861-8.
14. Pradella, D., Zhang, M., Gao, R., Yao, M.A., Gluchowska, K.M., Cendon-Florez, Y., Mishra, T., La Rocca, G., Weigl, M., Jiao, Z., et al. (2025). Engineered extrachromosomal oncogene amplifications promote tumorigenesis. Nature *637*, 955–964. https://doi.org/10.1038/s41586-024-08318-8.
15. Fiorini, E., Malinova, A., Schreyer, D., Pasini, D., Bevere, M., Alessio, G., Rosa, D., D'Agosto, S., Azzolin, L., Milite, S., et al. (2025). MYC ecDNA promotes intratumour heterogeneity and plasticity in PDAC. Nature *640*, 811–820. https://doi.org/10.1038/s41586-025-08721-9.
16. Stephens, P.J., Greenman, C.D., Fu, B., Yang, F., Bignell, G.R., Mudie, L.J., Pleasance, E.D., Lau, K.W., Beare, D., Stebbings, L.A., et al. (2011). Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell *144*, 27–40. https://doi.org/10.1016/j.cell.2010.11.055.
17. Zhang, C.Z., Spektor, A., Cornils, H., Francis, J.M., Jackson, E.K., Liu, S., Meyerson, M., and Pellman, D. (2015). Chromothripsis from DNA damage in micronuclei. Nature *522*, 179–184. https://doi.org/10.1038/nature14493.
18. Maciejowski, J., Li, Y., Bosco, N., Campbell, P.J., and de Lange, T. (2015). Chromothripsis and Kataegis Induced by Telomere Crisis. Cell *163*, 1641–1654. https://doi.org/10.1016/j.cell.2015.11.054.
19. Ly, P., Teitz, L.S., Kim, D.H., Shoshani, O., Skaletsky, H., Fachinetti, D., Page, D.C., and Cleveland, D.W. (2017). Selective Y centromere inactivation triggers chromosome shattering in micronuclei and repair by non-homologous end joining. Nat. Cell Biol. *19*, 68–75. https://doi.org/10.1038/ncb3450.
20. Umbreit, N.T., Zhang, C.Z., Lynch, L.D., Blaine, L.J., Cheng, A.M., Tourdot, R., Sun, L., Almubarak, H.F., Judge, K., Mitchell, T.J., et al. (2020). Mechanisms generating cancer genome complexity from a single

cell division error. Science *368*, eaba0712. https://doi.org/10.1126/science.aba0712.

21. Shoshani, O., Brunner, S.F., Yaeger, R., Ly, P., Nechemia-Arbely, Y., Kim, D.H., Fang, R., Castillon, G.A., Yu, M., Li, J.S.Z., et al. (2021). Chromothripsis drives the evolution of gene amplification in cancer. Nature *591*, 137–141. https://doi.org/10.1038/s41586-020-03064-z.
22. Wang, Y., Wang, M., Djekidel, M.N., Chen, H., Liu, D., Alt, F.W., and Zhang, Y. (2021). eccDNAs are apoptotic products with high innate immunostimulatory activity. Nature *599*, 308–314. https://doi.org/10.1038/s41586-021-04009-w.
23. Paulsen, T., Kumar, P., Koseoglu, M.M., and Dutta, A. (2018). Discoveries of Extrachromosomal Circles of DNA in Normal and Tumor Cells. Trends Genet. *34*, 270–278. https://doi.org/10.1016/j.tig.2017.12.010.
24. Krolewski, J.J., Schindler, C.W., and Rush, M.G. (1984). Structure of extrachromosomal circular DNAs containing both the Alu family of dispersed repetitive sequences and other regions of chromosomal DNA. J. Mol. Biol. *174*, 41–54. https://doi.org/10.1016/0022-2836(84)90364-4.
25. Mouakkad-Montoya, L., Murata, M.M., Sulovari, A., Suzuki, R., Osia, B., Malkova, A., Katsumata, M., Giuliano, A.E., Eichler, E.E., and Tanaka, H. (2021). Quantitative assessment reveals the dominance of duplicated sequences in germline-derived extrachromosomal circular DNA. Proc. Natl. Acad. Sci. USA *118*, e2102842118. https://doi.org/10.1073/pnas.2102842118.
26. Jones, R.S., and Potter, S.S. (1985). Characterization of cloned human alphoid satellite with an unusual monomeric construction: evidence for enrichment in HeLa small polydisperse circular DNA. Nucleic Acids Res. *13*, 1027–1042. https://doi.org/10.1093/nar/13.3.1027.
27. Cohen, S., Agmon, N., Sobol, O., and Segal, D. (2010). Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. Mobile DNA *1*, 11. https://doi.org/10.1186/1759-8753-1-11.
28. Regev, A., Cohen, S., Cohen, E., Bar-Am, I., and Lavi, S. (1998). Telomeric repeats on small polydisperse circular DNA (spcDNA) and genomic instability. Oncogene *17*, 3455–3461. https://doi.org/10.1038/sj.onc.1202250.
29. Mazzucco, G., Huda, A., Galli, M., Piccini, D., Giannattasio, M., Pessina, F., and Doksani, Y. (2020). Telomere damage induces internal loops that generate telomeric circles. Nat. Commun. *11*, 5297. https://doi.org/10.1038/s41467-020-19139-4.
30. McNulty, S.M., and Sullivan, B.A. (2018). Alpha satellite DNA biology: finding function in the recesses of the genome. Chromosome Res. *26*, 115–138. https://doi.org/10.1007/s10577-018-9582-3.
31. Lindahl, T., and Nyberg, B. (1972). Rate of depurination of native deoxyribonucleic acid. Biochemistry *11*, 3610–3618. https://doi.org/10.1021/bi00769a018.
32. Lindahl, T., and Barnes, D.E. (2000). Repair of endogenous DNA damage. Cold Spring Harb. Symp. Quant. Biol. *65*, 127–133. https://doi.org/10.1101/sqb.2000.65.127.
33. Vilenchik, M.M., and Knudson, A.G. (2003). Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. Proc. Natl. Acad. Sci. USA *100*, 12871–12876. https://doi.org/10.1073/pnas.2135498100.
34. Lieber, M.R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu. Rev. Biochem. *79*, 181–211. https://doi.org/10.1146/annurev.biochem.052308.093131.
35. Cohen, S., and Lavi, S. (1996). Induction of circles of heterogeneous sizes in carcinogen-treated cells: two-dimensional gel analysis of circular DNA molecules. Mol. Cell. Biol. *16*, 2002–2014. https://doi.org/10.1128/MCB.16.5.2002.
36. Wang, R.C., Smogorzewska, A., and de Lange, T. (2004). Homologous recombination generates T-loop-sized deletions at human telomeres. Cell *119*, 355–368. https://doi.org/10.1016/j.cell.2004.10.011.
37. Cesare, A.J., and Griffith, J.D. (2004). Telomeric DNA in ALT Cells Is Characterized by Free Telomeric Circles and Heterogeneous t-Loops. Mol. Cell. Biol. *24*, 9948–9957. https://doi.org/10.1128/MCB.24.22.9948-9957.2004.
38. Altemose, N., Logsdon, G.A., Bzikadze, A.V., Sidhwani, P., Langley, S.A., Caldas, G.V., Hoyt, S.J., Uralsky, L., Ryabov, F.D., Shew, C.J., et al. (2022). Complete genomic and epigenetic maps of human centromeres. Science *376*, eabl4178. https://doi.org/10.1126/science.abl4178.
39. Gao, X.D., Tu, L.C., Mir, A., Rodriguez, T., Ding, Y., Leszyk, J., Dekker, J., Shaffer, S.A., Zhu, L.J., Wolfe, S.A., and Sontheimer, E.J. (2018). C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. Nat. Methods *15*, 433–436. https://doi.org/10.1038/s41592-018-0006-2.
40. Doksani, Y., and de Lange, T. (2016). Telomere-Internal Double-Strand Breaks Are Repaired by Homologous Recombination and PARP1/Lig3-Dependent End-Joining *Cell Rep*. Cell Rep. *17*, 1646–1656. https://doi.org/10.1016/j.celrep.2016.10.008.
41. Riley, D.E. (1980). Deoxyribonuclease I generates single-stranded gaps in chromatin deoxyribonucleic acid. Biochemistry *19*, 2977–2992. https://doi.org/10.1021/bi00554a024.
42. Gamba, R., Mazzucco, G., Wilhelm, T., Velikovsky, L., Salinas-Luypaert, C., Chardon, F., Picotto, J., Bohec, M., Baulande, S., Doksani, Y., and Fachinetti, D. (2022). Enrichment of centromeric DNA from human cells. PLoS Genet. *18*, e1010306. https://doi.org/10.1371/journal.pgen.1010306.
43. Ip, S.C.Y., Rass, U., Blanco, M.G., Flynn, H.R., Skehel, J.M., and West, S.C. (2008). Identification of Holliday junction resolvases from humans and yeast. Nature *456*, 357–361. https://doi.org/10.1038/nature07470.
44. Walker, P.R., LeBlanc, J., and Sikorska, M. (1997). Evidence that DNA fragmentation in apoptosis is initiated and propagated by single-strand breaks. Cell Death Differ. *4*, 506–515. https://doi.org/10.1038/sj.cdd.4400273.
45. Iglesias-Guimarais, V., Gil-Guiñon, E., Sánchez-Osuna, M., Casanelles, E., García-Belinchón, M., Comella, J.X., and Yuste, V.J. (2013). Chromatin collapse during caspase-dependent apoptotic cell death requires DNA fragmentation factor, 40-kDa subunit-/caspase-activated deoxyribonuclease-mediated 3 -OH single-strand DNA breaks. J. Biol. Chem. *288*, 9200–9215. https://doi.org/10.1074/jbc.M112.411371.
46. Larsen, B.D., Benada, J., Yung, P.Y.K., Bell, R.A.V., Pappas, G., Urban, V., Ahlskog, J.K., Kuo, T.T., Janscak, P., Megeney, L.A., et al. (2022). Cancer cells use self-inflicted DNA breaks to evade growth limits imposed by genotoxic stress. Science *376*, 476–483. https://doi.org/10.1126/science.abi6378.
47. Liu, Q.Y., Ribecco-Lutkiewicz, M., Carson, C., Testolin, L., Bergeron, D., Kohwi-Shigematsu, T., Walker, P.R., and Sikorska, M. (2003). Mapping the initial DNA breaks in apoptotic Jurkat cells using ligation-mediated PCR. Cell Death Differ. *10*, 278–289. https://doi.org/10.1038/sj.cdd.4401146.
48. Arnoult, D., Gaume, B., Karbowski, M., Sharpe, J.C., Cecconi, F., and Youle, R.J. (2003). Mitochondrial release of AIF and EndoG requires caspase activation downstream of Bax/Bak-mediated permeabilization. EMBO J. *22*, 4385–4399. https://doi.org/10.1093/emboj/cdg423.
49. Caserta, T.M., Smith, A.N., Gultice, A.D., Reedy, M.A., and Brown, T.L. (2003). Q-VD-OPh, a broad spectrum caspase inhibitor with potent anti-apoptotic properties. Apoptosis *8*, 345–352. https://doi.org/10.1023/A:1024116916932.
50. Cohen, S., and Segal, D. (2009). Extrachromosomal circular DNA in eukaryotes: possible involvement in the plasticity of tandem repeats. Cytogenet. Genome Res. *124*, 327–338. https://doi.org/10.1159/000218136.
51. Henson, J.D., Cao, Y., Huschtscha, L.I., Chang, A.C., Au, A.Y.M., Pickett, H.A., and Reddel, R.R. (2009). DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. Nat. Biotechnol. *27*, 1181–1185. https://doi.org/10.1038/nbt.1587.

52. Kim, H., Kim, S., Wade, T., Yeo, E., Lipsa, A., Golebiewska, A., Johnson, K.C., An, S., Ko, J., Nam, Y., et al. (2024). Mapping extrachromosomal DNA amplifications during cancer progression. Nat. Genet. *56*, 2447–2454. https://doi.org/10.1038/s41588-024-01949-7.

53. Wang, H., Xing, J., Grover, D., Hedges, D.J., Han, K., Walker, J.A., and Batzer, M.A. (2005). SVA elements: a hominid-specific retroposon family. J. Mol. Biol. *354*, 994–1007. https://doi.org/10.1016/j.jmb.2005.09.085.

54. Luebeck, J., Huang, E., Kim, F., Liefeld, T., Dameracharla, B., Ahuja, R., Schreyer, D., Prasad, G., Adamaszek, M., Kenkre, R., et al. (2024). AmpliconSuite: an end-to-end workflow for analyzing focal amplifications in cancer genomes. Preprint at bioRxiv. https://doi.org/10.1101/2024.05.06.592768.

55. Raeisi Dehkordi, S., Wong, I.T.L., Ni, J., Luebeck, J., Zhu, K., Prasad, G., Krockenberger, L., Xu, G., Chowdhury, B., Rajkumar, U., et al. (2025). Breakage fusion bridge cycles drive high oncogene number with moderate intratumoural heterogeneity. Nat. Commun. *16*, 1497. https://doi.org/10.1038/s41467-025-56670-8.

56. Deshpande, V., Luebeck, J., Nguyen, N.D., Bakhtiari, M., Turner, K.M., Schwab, R., Carter, H., Mischel, P.S., and Bafna, V. (2019). Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. Nat. Commun. *10*, 392. https://doi.org/10.1038/s41467-018-08200-y.

57. Sinclair, D.A., and Guarente, L. (1997). Extrachromosomal rDNA circles–a cause of aging in yeast. Cell *91*, 1033–1042. https://doi.org/10.1016/S0092-8674(00)80493-6.

58. Møller, H.D., Parsons, L., Jørgensen, T.S., Botstein, D., and Regenberg, B. (2015). Extrachromosomal circular DNA is common in yeast. Proc. Natl. Acad. Sci. USA *112*, E3114–E3122. https://doi.org/10.1073/pnas.1508825112.

59. Navrátilová, A., Koblízková, A., and Macas, J. (2008). Survey of extrachromosomal circular DNA derived from plant satellite repeats. BMC Plant Biol. *8*, 90. https://doi.org/10.1186/1471-2229-8-90.

60. Zellinger, B., Akimcheva, S., Puizina, J., Schirato, M., and Riha, K. (2007). Ku suppresses formation of telomeric circles and alternative telomere lengthening in Arabidopsis. Mol. Cell *27*, 163–169. https://doi.org/10.1016/j.molcel.2007.05.025.

61. Cohen, S., Yacobi, K., and Segal, D. (2003). Extrachromosomal circular DNA of tandemly repeated genomic sequences in Drosophila. Genome Res. *13*, 1133–1145. https://doi.org/10.1101/gr.907603.

62. Paulsen, T., Malapati, P., Shibata, Y., Wilson, B., Eki, R., Benamar, M., Abbas, T., and Dutta, A. (2021). MicroDNA levels are dependent on MMEJ, repressed by c-NHEJ pathway, and stimulated by DNA damage. Nucleic Acids Res. *49*, 11787–11799. https://doi.org/10.1093/nar/gkab984.

63. Mozlin, A.M., Fung, C.W., and Symington, L.S. (2008). Role of the Saccharomyces cerevisiae Rad51 paralogs in sister chromatid recombination. Genetics *178*, 113–126. https://doi.org/10.1534/genetics.107.082677.

64. Ma, W., Westmoreland, J.W., and Resnick, M.A. (2013). Homologous recombination rescues ssDNA gaps generated by nucleotide excision repair and reduced translesion DNA synthesis in yeast G2 cells. Proc. Natl. Acad. Sci. USA *110*, E2895–E2904. https://doi.org/10.1073/pnas.1301676110.

65. Adar, S., Izhar, L., Hendel, A., Geacintov, N., and Livneh, Z. (2009). Repair of gaps opposite lesions by homologous recombination in mammalian cells. Nucleic Acids Res. *37*, 5737–5748. https://doi.org/10.1093/nar/gkp632.

66. Vriend, L.E.M., and Krawczyk, P.M. (2017). Nick-initiated homologous recombination: Protecting the genome, one strand at a time. DNA Repair (Amst) *50*, 1–13. https://doi.org/10.1016/j.dnarep.2016.12.005.

67. Bi, X., and Liu, L.F. (1994). recA-independent and recA-dependent intramolecular plasmid recombination. Differential homology requirement and distance effect. J. Mol. Biol. *235*, 414–423. https://doi.org/10.1006/jmbi.1994.1002.

68. Lovett, S.T., Drapkin, P.T., Sutera, V.A., and Gluckman-Peskind, T.J. (1993). A sister-strand exchange mechanism for recA-independent deletion of repeated DNA sequences in Escherichia coli. Genetics *135*, 631–642. https://doi.org/10.1093/genetics/135.3.631.

69. Klein, H.L. (1995). Genetic control of intrachromosomal recombination. BioEssays *17*, 147–159. https://doi.org/10.1002/bies.950170210.

70. George, C.M., and Alani, E. (2012). Multiple cellular mechanisms prevent chromosomal rearrangements involving repetitive DNA. Crit. Rev. Biochem. Mol. Biol. *47*, 297–313. https://doi.org/10.3109/10409238.2012.675644.

71. Doksani, Y. (2019). The Response to DNA Damage at Telomeric Repeats and Its Consequences for Telomere Function. Genes (Basel) *10*, 318. https://doi.org/10.3390/genes10040318.

72. Choo, K.H., Vissel, B., Nagy, A., Earle, E., and Kalitsis, P. (1991). A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. Nucleic Acids Res *19*, 1179–1182. https://doi.org/10.1093/nar/19.6.1179.

73. Catacchio, C.R., Ragone, R., Chiatante, G., and Ventura, M. (2015). Organization and evolution of Gorilla centromeric DNA from old strategies to new approaches. Sci Rep *5*, 14189. https://doi.org/10.1038/srep14189.

74. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

75. Smit, A.F.A., Hubley, R., and Green, P. (2013). RepeatMasker Open-4.0. http://www.repeatmasker.org.

76. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

77. Gel, B., and Serra, E. (2017). karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics *33*, 3088–3090. https://doi.org/10.1093/bioinformatics/btx346.

78. Storer, J., Hubley, R., Rosen, J., Wheeler, T.J., and Smit, A.F. (2021). The Dfam community resource of transposable element families, sequence models, and genome annotations. Mobile DNA *12*, 2. https://doi.org/10.1186/s13100-020-00230-y.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-FLAG (mouse monoclonal)	Millipore	Cat#F1804; RRID:AB_262044
Anti-GAPDH (mouse monoclonal)	Sigma-Aldrich	Cat#G8795; RRID: AB_1078991
Anti-mouse IgG, HRP-linked (horse polyclonal)	Cell Signaling	Cat#7076; RRID: AB_330924
Bacterial and virus strains		
Lentiviral particles: Lenti-iCas9-Neo	Addgene	Cat#85400; RRID:Addgene_85400
Lentiviral particles: Lentiguide-Puro	Addgene	Cat#52963; RRID:Addgene_52963
Retroviral particles: FokI-TRF1	Doksani and de Lange[40]	N/A
One shot top10 chemically competent cells	Invitrogen	Cat# C4040-03
One shot Stbl3 chemically competent cells	Invitrogen	Cat# C737303
Chemicals, peptides, and recombinant proteins		
4,5',8-trimethylpsoralen	Sigma-Aldrich	Cat#T6137; CAS:3902-71-4
4-hydroxytamoxifen	Sigma-Aldrich	Cat# H7904
Accutase	Euroclone	Cat#ECB3056D
Agarose	Euroclone	Cat# EMRO10500
Agarose Low EEO	US Biological	Cat#A1015
Doxycycline hydrochloride	Sigma-Aldrich	Cat#D9891; CAS:24390-14-5
D-sorbitol	Carlo Erba	Cat#70-70-4; CAS:50-70-4
DNase I	Worthington	Cat#LS006333
Ethidium Bromide	Sigma-Aldrich	Cat#E1510
Geneticin G418 sulphate crystalline	Life Technologies	Cat#1181031
L-glutamine 200mM	Euroclone	Cat#LOBE1760F
Polybrene (Hexadimethrine bromide)	Sigma-Aldrich	Cat#H9268
Polyetylenimine Max 40000 (PEI 25000)	Histo-Line Laboratories	Cat#24765-1
Proteinase K	Roche	Cat#3115887001
Protease inhibitor cocktail	Roche	Cat#11836170001
Puromycin	Adipogen	Cat#AGCN20078M100
Q-VD-OPh (pan-caspase inhibitor)	MedChemExpress	Cat#HY-12305; CAS:1135695-98-5
SYBR safe gel stain	Thermo Scientific	Cat#S33102
RNase A	Sigma-Aldrich	Cat#R5000
Phi29 DNA polymerase	NEB	Cat#M0269L
Q5 High-fidelity polymerase	NEB	Cat# M0491S
T4 Gene 32 Protein (T4GP32)	NEB	Cat#M0300S
T4 DNA Ligase	NEB	Cat# M0202S
Phenol:chloroform:isoamyl alcohol 25:24:1	Sigma-Aldrich	Cat#P2069
ATP	NEB	Cat#P0756
Restriction enzymes: HinfI	NEB	Cat#R0155
Restriction enzymes: RsaI	NEB	Cat#R0167
Restriction enzymes: MspI	NEB	Cat#R0106S
Restriction enzymes: MboI	NEB	Cat#R0147
Restriction enzymes: AluI	NEB	Cat#R0137
Restriction enzymes: KpnI	NEB	Cat#BA0142S
Restriction enzymes: ScrFI	NEB	Cat# R0110S
Restriction enzymes: Eco0109I	NEB	Cat# R0503S

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Restriction enzymes: NlaIV	NEB	Cat# R0126S
Restriction enzymes: DpnI	NEB	Cat# R0176L
α-32P-dCTP 3000 Ci/mmol	PerkinElmer	Cat#NEG513Z
γ-32P-ATP 6000 Ci/mmol	PerkinElmer	Cat#NEG5502Z
Critical commercial assays		
Prime-a-gene Labeling System	Promega	Cat#U1100
MycoAlert Detection Kit	Lonza	Cat#LT07-418
Oxford Nanopore Rapid Barcoding Kit	Oxford Nanopore Technologies	Cat#SQK-RBK114.24
Amicon Ultra-15 Centrifugal Filter Units, 30 kDa	Millipore	Cat#UFC903024
QubitTM dsDNA Assay Kit BR	Invitrogen	Cat#Q32850
QubitTM dsDNA Assay Kit HS	Invitrogen	Cat#Q33231
Super Signal West Pico/Dura chemiluminescent substrate	Thermo Scientific	Cat#34580
Deposited data		
RCA sequencing data (Nanopore)	This study	EBI-EMBL: https://ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-16581
Original images and source data	This study	Mendeley Data: 10.17632/hwf38mddyt.1
Experimental models: Cell lines		
Human: HT1080 (male)	ATCC	Cat#CCL-121; RRID:CVCL_0317
Human: HEK293-T (female)	ATCC	Cat#CRL-3216; RRID:CVCL_0063
Human: HeLa (female)	ATCC	Cat# CRM-CCL-2;RRID:CVCL_0030
Human: HeLa 1.3 (female)	Titia de Lange Lab	Cat#CRM-CCL-2; RRID:CVCL_B5DA
Human: BJ fibroblasts (male)	ATCC	Cat#CRL-2522; RRID:CVCL_3653
Human: hES cells (female)	WiCell Research Institute	RRID:CVCL_9773
Mouse: SV40-MEFs	Laboratory stock	RRID:CVCL_U630
Human: HT1080 Lenti-iCas9(D10A)-Neo	This study	N/A
Human: HT1080 Lenti-iCas9(H840A)-Neo	This study	N/A
Human: HT1080 Lenti-iCas9-Neo	This study	N/A
Oligonucleotides		
sgRNA telomeric: 5'-GTTAGGG TTAGGGTTAGGGTTA-3'	This study	N/A
sgRNA alpha satellite: 5'-GGAAT CTGCAAGTGGATATTG-3'	This study	N/A
TelC probe: $(CCCTAA)_4$	This study	N/A
TelG probe: $(TTAGGG)_4$	This study	N/A
Alpha satellite primer α27: 5'-CATCACA AAGAAGTTTCTGAGAATGCTTC-3'	Choo et al.[72] and Catacchio et al.[73]	N/A
Alpha satellite primer α30: 5'-TGCATTCA ACTCACAGAGTTGAACCTTCC-3'	Choo et al.[72] and Catacchio et al.[73]	N/A
Alt-R CRISPR-Cas9 crRNA (telomeres)	Integrated DNA Technologies	Custom
Alt-R CRISPR-Cas9 tracrRNA	Integrated DNA Technologies	Cat#1072532
Recombinant DNA		
Lenti-iCas9-Neo	Addgene	Cat#85400; RRID:Addgene_85400
Lenti-iCas9(D10A)-Neo	This study; Doksani lab plasmid #123	N/A
Lenti-iCas9(H840A)-Neo	This study; Doksani lab plasmid #124	N/A
Lentiguide-Puro	Addgene	Cat#52963; RRID:Addgene_52963

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Teloguide (telomere-specific sgRNA)	This study; Doksani lab plasmid #129	N/A
Alphaguide (alpha satellite-specific sgRNA)	This study; Doksani lab plasmid #138	N/A
pSty11 (telomeric probe)	De Lange Lab; Addgene	Cat#12401; RRID:Addgene_12401
α-satellite pSP73	This study; Doksani lab plasmid #139	N/A
Retroviral FokI-TRF1	Doksani and de Lange[40]	N/A
Retroviral FokI^{D450A}-TRF1	Doksani and de Lange[40]	N/A
Software and algorithms		
NERP pipeline	This study	Github: https://github.com/ifom/NERP/; Zenodo: https://doi.org/10.5281/zenodo.19344238
T2T-CHM13v2.0 reference genome	Genome Reference Consortium	https://github.com/marbl/CHM13
ImageJ/Fiji	Schneider et al.[74]	https://imagej.nih.gov/ij/; RRID: SCR_002285
Data annotation and storage macro (ImageJ)	Mazzucco et al.[29]	N/A
Digital Micrograph software	Gatan	https://www.gatan.com/; RRID: SCR_016434
RepeatMasker v4.1.2	Smit et al.[75]	http://www.repeatmasker.org; RRID: SCR_012954
bedtools	Quinlan and Hall[76]	https://bedtools.readthedocs.io/; RRID: SCR_006646
karyoploteR v1.18.0	Gel and Serra[77]	https://bioconductor.org/packages/karyoploteR/; RRID: SCR_021170
R v4.1.0	R Core Team	https://www.r-project.org/; RRID: SCR_001905
Bash	GNU Operating System	RRID: SCR_021268
DFAM	Storer et al.[78]	RRID: SCR_021168
01_read_mapping.sh	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
02_coverage_preprocessing.sh	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
03_circular_consensus.py	Modified from Wang et al.[22]	Zenodo: https://doi.org/10.5281/zenodo.19344238
04_hotspot_detection.sh	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
05_repeatmasker.sh	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
06_repeat_profiling.py	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
07_repeat_analysis.py	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
08_statistical_analysis.r	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
09_karyoplot_coverage.R	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
AA_tsv_analyzer.sh	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
generate_junctions_tsv.py	This study	Zenodo: https://doi.org/10.5281/zenodo.19344238
Typhoon scanner software	GE Healthcare	N/A
Python 3	Python Software Foundation	https://www.python.org/; RRID: SCR_008394
Other		
Matrigel hESC-qualified Matrix	Corning	Cat#354277
mTeSR Plus Medium	STEMCELL Technologies	Cat#100-0276
DMEM High Glucose	Euroclone	Cat#ECB7501L
TET-approved fetal bovine serum	Euroclone	Cat#ECS0182L
MEM media	Euroclone	Cat#ECB2071L
MEM non-essential amino acids	Microtech	Cat#X0557
Penicillin/Streptomycin	Euroclone	Cat#ECB3001L
GenePrint® 10 System	Promega	CAT#B9510
MinION flow cells R10.4.1	Oxford Nanopore Technologies	Cat#FLO-MIN114
NucleoBond Xtra Maxi Kit	Macherey-Nagel	Cat#740414.50
Wizard Plus SV Minipreps DNA	Promega	Cat# A1330

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Wizard SV gel and PCR clean-up system	Promega	Cat# A9281
Gilder Grids, 400 mesh, 3.05mm O.D., Center and Rim-Mark, Cu Coated Copper Grids	Ted Pella, INC	Cat#G400CU
Nitrocellulose membrane Protran 0.45 μm	GE Healthcare	Cat#10600002
Hybond-XL hybridization membraned	Cytiva	Cat#GERPN203S
Microspin G-50 Columns	GE Healthcare	Cat#27533001
Microspin G-25 Columns	GE Healthcare	Cat#27532501
PhosphorImager screen	GE Healthcare	Cat#28-9564-76
Amersham Typhoon 5 scanner	Cytiva	RRID: SCR_025702
HeLa cell extract	IPRACELL	Cat# CC012010
Gel Apparatus with recirculating valves	Thermo Scientific	Cat#A5
Gel Apparatus	Thermo Scientific	Cat#A2-BP
Hybridization Cassette	VWR	Cat#RPN11645
Stratalinker 1800	Stratagene	Cat#400072
Phosphor screen	GE Healthcare	Cat#GE28956476
Plug molds	BioRad	Cat#1703713
dNTP Set (100mM)	Invitrogen	Cat# 10297018
QuBit 4 fluorometer	Invitrogen	Cat#Q33226
Safe Imager2.0 Blue-Light Transilluminator	ThermoFisher Scientific	Cat#G6600
PFGE apparatus (CHEF-DRII-MAPPER)	Bio Rad	Cat# 170-3690 through 170-3703

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial strains and culture

The Escherichia coli (E. coli) Top10 and Stbl3 cells used for amplifying plasmids were cultured in Luria broth (LB) medium supplemented with antibiotics, at 37°C or 30°C for unstable constructs.

Human embryos and/or stem cells

Agreement 19-W0452 is the executed Memorandum of Understanding – Wisconsin Materials between IFOM and WiCell for the use of the Wisconsin Materials. Agreement 19-W0399 is your executed WiCell Simple Letter Agreement to use the WA09 human ES cells – Wisconsin Materials in your lab at IFOM. All the experiment were performed in compliance with ISSCR 2021 guidelines.

Cell lines and culture conditions

HT1080 (human fibrosarcoma, male origin), HEK293-T (human embryonic kidney, female origin), and HeLa 1.3 cells (human cervical adenocarcinoma, female origin) were grown in DMEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine and 100 U/ml penicillin, 0.1 μg/ml streptomycin at 37°C in 5% CO2 and 95% relative humidity. Human embryonic stem cells (hES, female origin) were grown feeder-free in mTeSR-plus medium on Matrigel-hESC-qualified coated plates at 37°C in 5% CO2. BJ cells (human fibroblasts, male origin) were grown in modified Eagle medium (MEM) supplemented with 10% fetal bovine serum, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 2 mM L-glutamine and 100 U/ml penicillin and 0.1 μg/ml streptomycin at 37°C in 5% CO2. SV40-LT-immortalized mouse embryonic fibroblasts (SV40-MEFs) were grown in DMEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin, 0.1 μg/ml streptomycin and 0.1 mM non-essential amino acids at 37°C in 5% CO2.

During transfection and transduction procedures, cells were grown in the absence of antibiotics. Cells infected with doxycycline-inducible constructs were maintained using TET-approved fetal bovine serum. All cell lines were tested for mycoplasma both upon arrival and after generation of new stocks using PCR analysis and biochemical testing (MycoAlert Detection Kit, Lonza Cat#: LT07-418); all tested negative for mycoplasma contamination. Cell line authentication was performed by the IFOM Cell Biology Unit with GenePrint® 10 System (10-Locus STR System for Cell Line Authentication) by Promega (CAT#B9510).

METHOD DETAILS

Generation of inducible Cas9 nickase cell lines

HT1080 cells were transduced with lentiviral constructs expressing doxycycline-inducible Flag-tagged $Cas9^{D10A}$ or $Cas9^{H840A}$ nickase variants. The $Cas9^{D10A}$ variant nicks the non-target strand, while $Cas9^{H840A}$ nicks the target strand. 48 hours post-infection, cells were selected in Neomycin (400 ug/ml) for 7 days and then amplified. Cells were subsequently transduced with lentiviral constructs expressing either telomere-specific sgRNA (5'-GTTAGGGTTAGGGTTAGGGTTA-3') or alpha satellite-specific sgRNA (5'-GGAATCTGCAAGTGGATATTG-3').Cells were harvested not earlier than 48 hours post-infection. Cas9 expression was induced by addition of 0.5 μM doxycycline for 16 hours before harvesting.

Isopropanol precipitation of DNA

DNA was precipitated by addition of 1/10 volume of Na-Acetate 3M pH 5.2 and 1 volume of isopropanol. The solution was thoroughly mixed by inversion and centrifuged at 16000 g at 4°C for 1 hour. The supernatant was removed before the addition of 1 volume of ethanol 70% without perturbing the pellet. The solution was centrifuged at 16000 g for 5 min, the supernatant was discarded, the empty vial was centrifuged again for 10 seconds, and residual ethanol was carefully removed with a pipette. The pellet was air-dried for 2 min and resuspended in Tris-HCl 10 mM pH 8.0 or TE (1X).

Standard PCR

PCR reaction mix was prepared as follows: polymerase buffer at 1X final concentration, dNTPs 200 μM each, 0.5 μM forward and 0.5 μM reverse primers, 100-200 ng of DNA template and 0.02 U/μl of Q5 polymerase (NEB). Thermocycling conditions were designed as follows: initial denaturation at 98°C for 1 minute, followed by 30-35 cycles of 10 sec at 98°C, 20 sec at the lower annealing temperature of primer set used, 30 sec/kb at 72°C, and a final extension of 7 min at 72°C. The PCR product was then purified by isopropanol precipitation or by using Wizard PCR clean-up system kit (Promega) and stored at -20°C.

Mutagenic PCR

Two reactions for each construct were set-up with 500 ng of template. For each reaction a standard PCR was carried out using just one of the mutagenic primers set. After the PCR reaction is complete, the two reactions are pooled together and annealed by heating at 98°C and let it cool down at room temperature. Once the annealing reaction is complete, it is put at 37°C and digested with 30 U of DpnI overnight. The reaction is directly used to transform Top10 competent bacteria cells.

List of primers use for mutagenic PCR

- Cas9D10A FW: 5'-cctggCcatcggcaccaactctgtgg-3'
- Cas9D10A RV: 5'-cgatgGccaggccgatgctgtacttc-3'
- Cas9H840A FW: 5'-gtggacGCtatcgtgcctcagagctttc-3'
- Cas9H840A RV: 5'-cacgataGCgtccacatcgtagtcggac-3'

DNA fragment purification from gel

DNA sample was loaded on a 0.8% agarose gel in TAE 1X supplemented with SYBR safe gel stain and run at 4.5 V/cm. At the end of the run, the gel was exposed at Safe Imager2.0 Blue-Light transilluminator (G6600) to avoid DNA nicking and the band of interest cut with a scalpel. The DNA was purified from the agarose slice following the Wizard DNA clean-up system protocol (Promega). Briefly: 10 μl of membrane binding solution was added for each 10 mg of gel slice. The slice was vortexed and incubated at 55°C until the gel is completely dissolved. An SV column was inserted into a collection tube, then the dissolved gel is transferred on the column and allowed to equilibrate for 1 min. The column was centrifuged at 16000 g for 1 min and the flow-through discarded. 700 μl of membrane wash solution was added and the column centrifuged at 16000 g for 1 min. After the flow-through was discarded, 500 μl of membrane wash solution was added and the column centrifuged at 16000 g for 5 min. The collection tube was emptied and the residual ethanol allowed to evaporate. The column is transferred to a clean tube, and the DNA was eluted with 20/30 μl of Tris HCl 10 mM pH 8.0 by centrifugation at 16000 g for 1 minute. DNA was stored at 4°C or -20°C.

DNA ligation

The ligation reaction was set up on ice as follows: T4 DNA Ligase Buffer (NEB) 1X final concentration, 400 U of T4 DNA Ligase (NEB), 50 ng of vector and the equivalent of three times molar excess of insert in 20-30 μl final volume. The reaction was gently mixed by pipetting and incubated overnight at 16°C in a PCR machine.

Transformation of chemically competent E. coli cells

Competent cells were thawed on ice and 25 μl of cell suspension were used for each transformation. 5 μl of DNA (10-100 pg) was added and incubated on ice for 30 min. Cells are heat-shocked at 42°C for 30 seconds (45 seconds for STBL3 cells) and then placed on ice for 2 min. 250 μl of pre-warmed S.O.C media (2% tryptone, 0.5% Yeast extract, 8.56mM NaCl, 2.5mM KCl, 10mM MgCl2,

10mM $MgSO_4$) was added to the cell mix, then tubes were incubated at 37°C for 1 hour in a thermomixer, shaking at 350 rpm. Next, tubes were centrifuged at 330 *g* for 3 min and all but 50 µl of the supernatant were discarded. The pellet was gently resuspended in the remaining media and plated on a pre-warmed LB agarose plate supplemented with 100 mg/ml ampicillin using glass beads and incubated overnight at 37°C.

Miniprep DNA preparation

Following Promega Wizard Plus SV Miniprep kit protocol: The bacterial culture was centrifuged for 5 min at 2700 g, then the supernatant was discarded. The pellet was resuspended in 250 µl of cell resuspension solution (Promega) then 250 µl of cell lysis solution (Promega) was added and mixed by inverting the tube four times. 10 µl of alkaline protease solution (Promega) was added, the tube was mixed by inverting four times and incubated for 5 min at RT. 350 µl of neutralizing solution was added and the tube was inverted four times to mix, then centrifuged at 16000 g for 10 min at RT. A spin column was inserted into a collection tube and the clear lysate was decanted into the spin column, then centrifuged at 16000 g for 1 min at RT. The flow-through was discarded and 750 µl of wash solution (Promega) was added, centrifuged at 16000 g for 1 min at RT and the flow-through was discarded. The previous step was repeated with 250 µl of wash solution and centrifugation at 16000 g for 2 min at RT. The spin column was transferred to a sterile tube, 80 µl of Tris-HCl 10 mM pH 8.0 was added to the column and eluted by centrifugation at 16000 g for 1 min at RT. The DNA was stored at -20°C.

Maxiprep DNA preparation

Following the NucleoBond Xtra Maxi protocol: 500 ml of an overnight bacterial culture was centrifuged at 6000 g for 15 min at 4°C. The supernatant was removed and the pellet was resuspended in 12 ml of resuspension buffer. 12 ml of lysis buffer was added, mixed by inversion 4-6 times and incubated for 5 min at RT. The column and the filter were equilibrated with 25 ml of equilibrating buffer. The lysis solution was neutralized with 12 ml of neutralization buffer and mixed by inverting until the solution became colorless, indicating complete neutralization. The lysate was loaded on the filter and after 15 min of decanting the filter was washed with 15 ml of equilibrating buffer. The filter was discarded and 25 ml of wash buffer added. Then elution was performed with 15 ml of elution buffer directly in a tube containing 10.5 ml of isopropanol and precipitated at 15000 g for 1 hour at 4°C. The supernatant was removed, carefully 4 ml of ethanol 70% was added, then centrifuged at 15000 g for 15 min at 4°C, the supernatant was removed and the pellet was air-dried to remove ethanol residues. DNA was resuspended in appropriate volume of Tris-HCl 10 mM pH 8.0.

Cell lysates preparation and immunoblotting

Cells were lysed in 2X Laemmli buffer (120 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 200 mM DTT, 0,02% bromophenol blue) at 10^7 cells/ml, denatured for 4 min at 95°C, and the genomic DNA was sheared by passing the lysate four times through a 28- gauge insulin needle. Lysates, prepared as above, were loaded at the equivalent of 1.2 x 105 cells per lane. Protein samples were separated by SDS-PAGE in an 8% polyacrylamide gel and transferred in cold room at 4°C overnight at 150 mA onto a 0,45 µm nitrocellulose membrane. Blotting was verified by Ponceau S staining, which was removed by washing with PBS-T. Membranes were blocked in 5% milk in PBS-T for 1 hour at RT and incubated with primary antibodies in 5% milk PBS-T 0.1% overnight at 4°C or for 2 hours at RT. Membranes were washed three times for 5 min in PBS-T, incubated with HRP-labeled secondary antibody in 5% milk PBS- T for 30-45 min at RT and washed three times with PBS-T at RT. Signal detection was performed by chemiluminescence using the SuperWest Substrate kit (Thermo Scientific) and acquired with Chemidoc Imaging System.

DNA damage induction in isolated nuclei and genomic DNA

For DNase I treatment of isolated nuclei cells are harvested by trypsinization, washed once with ice-cold PBS and pelleted at 330 g for 5 min at 4°C. The pellet is gently resuspended in ice-cold hypotonic buffer (Tris HCl pH 7.4 10 mM, MgCl2 1.5 mM, KCl 10 mM, sucrose 340 mM, supplemented right before use with the protease inhibitor cocktail (Roche, 11836170001) at 8 x 10^6 cells/ml. The cell suspension is pelleted again at 330 g for 5 min at 4°C. The supernatant is removed and cells are resuspended again at the same concentration in ice-cold hypotonic buffer and incubated on ice for 10 min. 1/30 v/v ratio of 10% triton X-100 is added to the cell suspension (0.33% v/v final concentration). The tube containing the cells is inverted twice and incubated on ice for 5 min. Nuclei are pelleted at 450 g for 5 min at 4°C. Nuclei are washed twice with 10 ml of NWB solution (Tris HCl pH 7.4 10 mM, KCl 60 mM, NaCl 15 mM, MgCl2 5mM, sucrose 300 mM), resuspended in NWB solution at 25 x 10^6 nuclei/ml and stored on ice.

Nuclei suspensions were treated with 0, 6.5 ng/µl DNase I for 8 min at room temperature (RT) in DNase I reaction buffer (NWB solution supplemented with $CaCl_2$ 2 mM, BSA 100 µg/ml). For ATP-supplemented conditions, 2 mM ATP was added to the reaction. Reactions were stopped by addition of stop buffer (50 mM EDTA, 10 mM EGTA). For genomic DNA treatments, purified gDNA (0.3 µg/µl) was incubated with 0, 0.01, 0.02, or 0.03 ng/µl DNase I for 8 min at RT, followed by immediate termination with stop buffer and genomic DNA extraction with phenol-chloroform.

Apoptosis induction

For UV-induced apoptosis in hES cells, cells at approximately 70% confluency were washed twice with PBS, leaving 1 ml PBS in the plate. Cells were irradiated in a UV Stratalinker 1800 equipped with UV-C bulbs at an energy setting of 20 (equivalent to 2 mJ/cm^2 or

20,000 μJ/ m^2). Immediately after irradiation, pre-warmed media was added, and cells were incubated for 16 hours before harvesting. Apoptosis was verified by morphological analysis and genomic DNA fragmentation on agarose gels.

For osmotic stress-induced apoptosis in HeLa cells, hyperosmotic stress was induced by replacing standard culture media with media supplemented with 1 M sorbitol for 1 hour. Cells were then washed twice with standard media and incubated for 3 hours before harvesting. Hypoosmotic stress was induced by replacing media with ddH2O for 30 min (HeLa) or 45 min, followed by two washes with standard media and 3 hours incubation. Where indicated, the pan-caspase inhibitor Q-VD-OPh was added at 10 μM final concentration. Cell death was assessed by morphological analysis and visualization of DNA fragmentation in agarose gels.

Psoralen crosslinking

Cells were harvested, washed, and resuspended in ice-cold PBS supplemented with MgCl2 and CaCl2 at $\leq 5 \times 10^7$ cells/ml. The suspension was incubated with 30 μg/ml 4,5',8-trimethylpsoralen for 5 min in the dark on ice while stirring, then irradiated for 8 min in a UV Stratalinker 1800 equipped with 365 nm UV bulbs. This cycle was repeated once with fresh psoralen. Cells were washed five times with ice-cold PBS containing MgCl2 and CaCl2 before genomic DNA extraction.

Genomic DNA extraction

Cell pellets were resuspended in TNE buffer (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM EDTA). Lysis was performed by adding one volume of TNES buffer (TNE with 0.5% SDS) containing RNase A (10 mg/ml) and incubating for 45 min at 37°C. Proteinase K (10 mg/ml) was added and the lysate incubated overnight at 37°C. Genomic DNA was purified by sequential extractions with phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform, followed by isopropanol precipitation.

2D-gel electrophoresis

20 μg of genomic DNA was digested overnight with restriction enzymes (HinfI and RsaI for human telomeres; MboI and AluI for mouse telomeres; MspI for alpha satellite repeats). Digested DNA was resolved on 0.35% agarose gels in 0.5× TBE, without ethidium bromide. The first dimension was run at room temperature (human samples: 0.5 V/cm for 15 h; mouse samples: 1.3 V/cm for 9.5 h). The first-dimension gels were stained for 45 min in TBE 0.5X with 0.3 μg/ml ethidium bromide and then, lanes containing fragments of interest were excised, at the desired size. For human telomeres and alpha satellite repeats, the first dimension slices were cut from the 3 kb band to the well. For mouse telomeres, the size cutoff was 4-6 kb. Slices, rotated by 90°, were embedded in 0.8% agarose gels containing 0.3 μg/ml ethidium bromide in 0.5× TBE. The second dimension was conducted at 4°C (human samples: 4.43 V/cm for 3 h; mouse samples: 2 V/cm for 16 h). Gels were imaged with Chemidoc Imaging System and processed for standard Southern blotting.

Southern blotting and probe hybridization

DNA was depurinated with 0.5 N HCl for 30 min, treated twice with denaturing solution (1.5 M NaCl, 0.5 N NaOH) for 30 min each, and twice with neutralizing solution (1.5 M NaCl, 0.5 M Tris-HCl pH 7.0) for 30 min each. DNA was transferred overnight onto nitrocellulose membranes by capillary action in SSC 10X (1.5 NaCl,0.15M sodium citrate) and UV-crosslinked at 1200 J with 254 nm light.

For telomeric strand-specific probes 2 μl of ddH2O, 1μl of 10X T4 PNK buffer (NEB), 1μl of T4 PNK 10U/μl (NEB),1μl of $(CCCTAA)_4$ or $(TTAGGG)_4$ oligo 50 ng/μl and 5 μl 10.0 mCi/ml γ-32P-ATP were mixed and incubated for 45 min at 37°C. 80 μl of TES buffer were added to stop the reaction. The probe-containing solution was loaded onto a G25 GE Healthcare column. The obtained purified probe was finally diluted in 25 ml Church mix.

For double-stranded telomeric probe preparation (Sty11), 200 ng of the 800 bp EcoRI fragment from pSty11 and 5 ng of (CCCTAA)4 oligonucleotide were heated at 95°C for 5 min, cooled on ice, and incubated with α-32P-dCTP and Klenow polymerase at RT for 1 h. For alpha satellite probe preparation, a 316 bp fragment was amplified from HT1080 gDNA using primers α27 (5'-CATCACAAAGAAGTTTCTGAGAATGCTTC-3') and α30 (5'-TGCATTCAACTCACAGAGTTGAACCTTCC-3') and labeled using the Prime-a-gene Labeling system (Promega).For pangenomic probe preparation 50 ng of HEK293 genomic DNA was vortexed for 15 seconds five times and used as template for the Prime-a-gene Labeling system (Promega). The probe-containing solution was loaded onto a G50 GE Healthcare column. The obtained purified probe was finally diluted in 25 ml Church mix.

Membranes were prehybridized in Church mix (0.5 M sodium phosphate pH 7.2, 1 mM EDTA pH 8.0, 7% SDS, 0.1% BSA) for 30 min at 65°C for double-stranded probes or 55°C for single-stranded probes. Labeled probes were added and hybridized for at least 4 hours. Membranes were washed three times for 15 min with Church mix (double-stranded probes) or SSC 4X (single-stranded probes), with 0.1% SDS added to the final wash. Signals were detected using a Typhoon scanner and quantified with Fiji software.

PFGE Analysis of Telomere Restriction Fragments

For each sample ~10^6 cells were resuspended in PBS, briefly heated to 50C and mixed with 2% agarose at 50C (1:1 ratio), and immediately casted in a plug mold. Plugs were digested overnight in proteinase K digestion buffer (10 mM Tris-HCl pH 8.0, 250 mM EDTA, 0.2% sodium deoxycholate, and 1% sodium lauryl sarcosine) at 55°C. After extensive washes with Tris EDTA (TE), plugs were incubated with 60 U each of AluI and MboI overnight at 37°C. Digested DNA was resolved on a 1% agarose/0.5X TBE gel using a CHEF-DRII PFGE apparatus (Bio-Rad) for 24 hr. The gels were then processed by standard southern blot and hybridization with the telomeric Sty11 probe.

Alkaline gel electrophoresis

Cells embedded in agarose plugs were washed twice in ddH2O and equilibrated in alkaline buffer (0.5 M NaOH, 0.1 M EDTA) for 2 h. Plugs were loaded onto 0.35% agarose gels prepared in alkaline buffer 1×. Gels were run at 70 V (2 V/cm) for 10 min, then a peristaltic pump set at 12 ml/min was turned on and gels run for an additional 16 h 20 min. Gels were neutralized in neutralizing solution for 45 min, stained with 33 μM ethidium bromide for 45 min, and processed for Southern blotting. Membranes were hybridized with end-labeled (CCCTAA)4 or (TTAGGG)4 oligonucleotides to detect C-strand or G-strand breaks, respectively.

Enrichment of centromeric alpha satellite DNA (CenRICH)

Approximately 50 × 10^6 cells were harvested and subjected to psoralen crosslinking as described above. Cells were lysed in TNES buffer with 50 μg/ml RNase A for 60 min at 37°C, then with 100 μg/ml proteinase K for 12 h at 37°C. DNA was extracted with phenol-chloroform-isoamyl alcohol followed by chloroform and precipitated with isopropanol. Approximately 200 μg of DNA was digested overnight with 32 units each of ScrFI, Eco0109I, and NlaIV. Digested DNA was loaded on a 10%-20%-30% sucrose gradient (3 ml each fraction in TNE buffer) and centrifuged in a SW41-Ti rotor at 30,100 rpm (111,265 g) for 16 hours at 4°C. High molecular weight fractions containing centromeric repeats (fractions 3-7) were collected, concentrated, and washed twice with 10 mM Tris pH 8.0 using Amicon Ultra-15 filters (30 kDa MWCO).

Electron microscopy sample preparation and analysis

For EM analysis of alpha satellite-enriched DNA, samples containing 5-20 ng of DNA were used. As a control, 30 ng of KpnI-digested genomic DNA was analyzed. Carbon-coated EM grids were treated with 33 μg/ml ethidium bromide in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) for 30-45 min at RT. DNA samples were mixed with formamide and 0.02% benzalkonium chloride (BAC) in TE buffer, immediately deposited onto a water surface, and a mica sheet was used to generate a monomolecular DNA layer, which was transferred to EM grids. Carbon grids with absorbed DNA molecules were immediately stained with a solution of uranyl acetate 0.2 μg/μl in ethanol and coated with 8 nm of Platinum using the MED020 evaporator modified with the low-angle grid shadowing kit (Leica 16770525) so that the sample holder was placed at an angle of 280.5 degrees and made an angle of around 3 degrees with the platinum gun fixed on the head of the instrument. For platinum e-beam evaporation, we utilized the parameters indicated in the MED020 instruction manual.

Grids were imaged using a FEI Tecnai12 Bio twin transmission electron microscope at 120 kV equipped with a GATAN SC-1000 Orius CCD camera. EM images were obtained by acquiring overlapping fields and stitching them using Digital Micrograph software. Images in.dm3 format were randomized in ImageJ using the Blind Analysis Tools plugin.[74] The number of molecules containing i-loops or circular DNA forms was annotated among all molecules in the acquired area using an ImageJ data annotation and storage macro. I-loop length measurements were performed in ImageJ. For BAC spreading preparations, 0.36 μm corresponded to 1 kb of double-stranded DNA.

Rolling circle amplification assay

For standard rolling circle amplification, 60 ng of isolated genomic DNA or DNA from enriched samples was mixed in a final volume of 20 μl of 1× NEB Phi29 Buffer, dNTPs (0.37 mM each), DTT (4 mM), and BSA (0.2 μg/ml). The reaction was initiated by adding 0.75 U of Phi29 DNA polymerase and incubated at 30°C for 12 hours. Reactions were heat-inactivated at 65°C for 20 min to terminate polymerase activity. RCA products were analyzed by dot blot hybridization or resolved on 0.8% agarose gels followed by Southern blotting and hybridization with sequence-specific radioactive probes.

For RCA sequencing experiments, eight parallel reactions were prepared to maximize sequence coverage of rare eccDNA molecules. Each reaction contained 30 ng input DNA, 10 units of Phi29 DNA polymerase, and 25 ng/μl T4 Gene 32 protein (T4GP32) in 20 μl final volume using 1× NEB Phi29 Buffer with dNTPs (0.37 mM each), DTT (4 mM), and BSA (0.2 μg/ml).The ssDNA-binding protein T4GP32 was included to stabilize the extended single-stranded regions of RCA products and facilitate sustained polymerase processivity during the extended incubation. Reactions were incubated at 30°C for 72 hours, allowing for multiple rounds of rolling circle amplification and preferential amplification of circular templates to high copy number. Following incubation, all eight reactions were pooled together into a single tube and the reaction stopped by heat-inactivation at 65 °C for 20 minutes. Pooled RCA products were extracted sequentially with phenol:chloroform:isoamyl alcohol (25:24:1), followed by an additional chloroform extraction, and finally concentrated by ethanol precipitation. The precipitate was resuspended in appropriate buffer and quantified using the Qubit ssDNA assay prior to library preparation. DNA quality was verified by agarose gel electrophoresis to confirm complete RCA product formation (expected as high molecular weight smear >10 kb).

Incubation of DNA with HeLa extracts

Both standardized commercially available, and home-made extracts have been used with the same results. All the steps for the nuclear extract preparation are performed at 4C and the lysis buffer contains: 20mM HEPES (pH 7.9 at 4C), 25% v/v glycerol 420 mM NaCl,1.5 mM MgCl2, 0.2 mM EDTA, 0.5 mM PMSF and 0.5 mM DTT. 1 μg of genomic DNA was incubated with 60 μg of HeLa nuclear extract (6 mg/ml) (IPRACELL, CC012010) in 50 mM Tris HCl pH8, 150 mM NaCl, 5 mM MgCl2, 2 mM ATP, 1 mM DTT for 35 min at 37 °C in 20 μl final volume. After incubation, the reaction was stopped with 0.1 volumes of ice-cold EDTA-EGTA 0.25 M each and extracted with 1 volume of phenol-chloroform isoamylalcohol, followed by precipitation in isopropanol.

Nanopore sequencing and library preparation

A total of 400 ng of RCA-amplified DNA was used for library preparation using the Oxford Nanopore Rapid Barcoding Kit (SQK-RBK114.24). Libraries were sequenced on R10.4.1 MinION flow cells using the Oxford Nanopore MinION sequencing machine.

Nanopore RCA bioinformatic analysis

A comprehensive, automated computational pipeline (Scripts 01-09) was developed to process Nanopore RCA sequencing data, identifying circular consensus sequences, enriched genomic hotspots, and analyzing repeat element composition.

Read mapping and coverage preprocessing

Raw FASTQ reads were mapped to the reference genome (T2T-CHM13v2.0) using Minimap2 with parameters optimized for Nanopore data (-x map-ont -c –secondary=no). Alignment files (PAF format) were processed to generate coverage tracks. For gDNA-NT control samples, alignment information (reference sequence, alignment start and end, read ID, mapping quality, strand) was extracted from PAF files and reformatted to BED format using awk. Files were sorted lexicographically by chromosome and numerically by start position. Genome coverage was calculated using bedtools genomecov with the T2T-CHM13v2.0 chromosome sizes file.[76]

For RCA samples, to mitigate bias from smaller eccDNAs exhibiting higher coverage due to more RCA passes, a specialized workflow was developed. Unique read coverage was calculated by ensuring each read contributed only once per genomic region. An awk script extracted read identifier, reference sequence, and alignment positions, constructing a unique key for each genomic region and outputting only the first occurrence of each read mapping to a specific region. A separate awk script quantified RCA passes for each read within each genomic region. Both unique read coverage and pass count files were converted to bedGraph format using bedtools genomecov. Files were combined and coverage values multiplied by pass counts to generate read-level normalized coverage, reducing bias from differential RCA amplification rates.

Circular DNA identification from RCA sequencing

Circular DNA molecules were identified using a modified version of the Python script from Wang et al.[22] (03_circular_consensus.py). The script processes PAF-format alignment files containing CIGAR strings generated by mapping nanopore reads to the T2T-CHM13v2.0 reference genome. The pipeline detects circular DNA structures by identifying reads with multiple passes over the same genomic region, indicative of rolling circle amplification products. For each read, fragments mapping to the reference genome were grouped based on genomic coordinates, with a maximum offset of 20 bp allowed between fragment boundaries. A minimum mapping quality threshold of 30 was enforced. The script tracks strand orientation patterns to identify strand-switching events during RCA. A consensus sequence was generated for each detected eccDNA molecule by combining information from multiple passes. Variants relative to the reference genome were called with parameters requiring minimum depth of 4 and minimum alternative allele frequency of 0.75.

Hotspot detection and enrichment analysis

Genomic regions enriched in eccDNA were identified by calculating fold-changes between RCA samples and non-amplified genomic DNA (gDNA) controls in 1 kb sliding windows (04_hotspot_detection.sh). A fold-change threshold of 2.0 was applied to define significant hotspots. Enriched regions were intersected with RepeatMasker annotations to quantify the overlap of specific repeat classes and families with eccDNA hotspots.

Repeat annotation and profiling

Consensus sequences derived from Script 03 were annotated de novo using RepeatMasker (05_repeatmasker.sh).[75,78] The resulting annotations were profiled (06_repeat_profiling.py) to determine the composition of repeat classes (e.g., LINE, SINE, DNA) and families (e.g., L1, Alu, hAT) within the unique eccDNA sequences. For each repeat element, the script calculated the percentage of total masked base pairs covered by each repeat class and family. Comparative analysis (07_repeat_analysis.py) was performed to calculate differential enrichment of repeat families between samples, utilizing log2 fold-change metrics. Repeat overlap analysis was performed by intersecting RepeatMasker annotations with enriched regions using bedtools.

Statistical analysis of repeat enrichment

A comprehensive statistical analysis was performed using R (08_statistical_analysis.r) to quantify repeat enrichment. Six key metrics were calculated for each repeat family: (1) Genome Abundance, (2) Percent Enriched (percentage of element type found in eccDNA), (3) Percent Coverage (percentage of eccDNA regions covered by the element), (4) Abundance-Normalized Enrichment (Percent Coverage normalized by Genome Abundance), (5) Enrichment Score (efficiency of enrichment), and (6) Composite Score (geometric mean of abundance-normalized enrichment and enrichment score). These metrics were used to analyze repeat element characterization quantifying association between eccDNA formation and genomic repetitive elements, calculating percent coverage and enrichment metrics for repeat classes and families; and chromosomal distribution analysis identifying chromosomal biases in eccDNA formation expressed as enrichment ratios relative to genomic background. Statistical summaries included mean, median, range parameters, and frequency distributions across predefined size categories.

Genomic coverage visualization

Genomic coverage visualization was performed using the karyoploteR package (v1.18.0) in R (v4.1.0).[77] RCA sequencing data was imported from bedgraph format with coverage values transformed using log2(x+1) function. Data was binned at 300 kb resolution. For each chromosome, coverage values were independently normalized to a [0,1] scale. Chromosome ideograms with cytobands from the T2T-CHM13v2.0 assembly (hs1) were generated and coverage data plotted as area graphs.

Viral transduction

Lentiviral particles were produced by transfecting HEK293-T cells with 45 μg of plasmid DNA (viral constructs and helper constructs) using 150 μg of polyethylenimine (PEI) Max 40000 per 15-cm plate. Viral supernatant was collected 36 hours post-transfection, filtered (0.45 μm), and supplemented with 8 μg/ml polybrene. Target cells were infected by replacing their growth medium with virus-containing medium. Stably transduced cells were selected using puromycin (2-3 days), hygromycin (4-5 days), or neomycin (7-10 days) until uninfected control cells were eliminated.

Telomeric DSB induction with FokI-TRF1

SV40 MEFs were transduced with retroviral pLPC-FokI-TRF1 construct (WT) or nuclease-dead FokI^{D450A}-TRF1 version (DA).[40] Cells were collected 96 hours later, DNA was prepared in agarose plugs, digested with AluI and MboI, and separated by pulse-field gel electrophoresis (PFGE). Gels were blotted and hybridized with the Sty11 probe. For 2D-gel analysis, DNA was digested with AluI and MboI, separated by 2D-gels, and hybridized with the Sty11 probe.

QUANTIFICATION AND STATISTICAL ANALYSIS

General statistical principles

For all the wet lab experiments statistical analyses was carried out using GraphPad Prism software with the one-way ANOVA followed by Šidák's correction for multiple comparisons and exact one-tailed Mann-Whitney U test as indicated in the figure legends. P-values are considered statistically significant and indicated by asterisks as follows: * P<0.05, ** P<0.01, ***P<0.001.

Electron microscopy quantification

Sample characterization

EM analysis (Figures 2, 3, and 5) was performed on two sample types: bulk genomic DNA (control) and enriched alpha satellite DNA (from CenRICH procedure; see STAR Methods section above). For each condition, samples were processed in parallel. Sample sizes were determined based on practical constraints of EM analysis and previous experience showing adequate statistical power with similar n values.

Blinding and randomization

All EM image files (.dm3 format) from a given experimental batch were randomized using the Blind Analysis Tools plugin in ImageJ/Fiji before any manual annotations were performed.

Molecule counting protocol

- Images from each sample were acquired in overlapping fields and computationally stitched using Digital Micrograph software to generate large montages areas.
- Within each montaged image, the investigator used an ImageJ macro to annotate and count:
  - Total number of DNA molecules (all structures observed)
  - Number of molecules containing at least one i-loop structure
  - Number of molecules containing circular forms
  - Structures that were ambiguous or partially degraded were excluded from analysis

Quantitative output: Results were expressed as the percentage of molecules containing i-loops = (number of molecules with $\geq$1 i-loop / total molecules analyzed) $\times$ 100, and similarly for circular forms.

Sample sizes and replication

- Figure 2A (DNase I-induced i-loop in enriched samples):
  - Bulk DNA control (NT): n=1357 molecules
  - Bulk DNA + DNase I: n=995 molecules
  - Enriched DNA (NT): n=979 and n=646 molecules analyzed from two independent experiments (total n=1625 molecules)
  - Enriched DNA + DNase I: n=1152 and n=425 molecules from two independent experiments (total n=1577 molecules)
  - Statistical presentation: Percentage values for enriched samples (NT and DNase I-treated) were calculated separately for each experiment, then reported as mean ± SD across the two experiments. Bulk samples, analyzed from only one experiment, are reported with no error bars.
- Figure 3A (quantification of circular eccDNA molecules):

- Circular form data are derived from the same EM fields analyzed for Figure 2A (same n values and experiment limitation).
- Figure 5A (apoptosis-induced i-loops and eccDNA in enriched samples):
  - Bulk NT: n=449 molecules
  - Bulk apoptotic: n=448 molecules
  - Enriched NT: n=854 molecules
  - Enriched apoptotic: n=962 molecules
- Figure 5D (quantification of circular eccDNA molecules):
  - Circular form data are derived from the same EM fields analyzed for Figure 5A (same n values and single experiment limitation).

I-loop size distribution

Measurement protocol: Individual i-loop structures were identified within EM images. For each i-loop, the contour length was measured from end to end of the loop using ImageJ measurement tools (Figure 2C).

Calibration: Measurements were calibrated using the standard conversion factor established from BAC-formamide spreading preparations: 0.36 μm of DNA on the EM grid corresponds to 1 kb of double-stranded DNA.

Sample analyzed: A total of 619 i-loop structures were measured from the enriched alpha satellite samples treated with DNase I (pooled across the two independent experiments shown in Figure 2A).

RCA monomer coverage and enrichment analysis

Sequencing data processing

Long-read Nanopore sequencing of RCA products yielded multiple "passes" over the same genomic regions. Circular DNA molecules were identified computationally using 03_circular_consensus.py (a modified version of the algorithm described in Wang et al.[22]), (Figure 6).

Coverage normalization

To account for the fact that eccDNA molecules of different sizes undergo different numbers of RCA amplification cycles (smaller molecules amplify to higher copy numbers), coverage was normalized by:

1. Unique read calculation: For each genomic region, the number of unique sequencing reads (identified by read ID) mapping to that region was counted, removing duplicate counts from individual reads that mapped multiple times to the same coordinates due to RCA.
2. Pass count determination: For each unique read-region pair, the number of times that specific read mapped to identical genomic coordinates was tabulated (representing the number of sequential passes the polymerase made through that region).
3. Normalized coverage: Normalized coverage values were calculated as: (unique read count) × (mean pass count for that region), which accounts for both the number of independent DNA templates (unique reads) and the amplification intensity per template (pass count).

This normalization approach reduces the bias whereby small, highly amplified eccDNAs appear over-represented in raw coverage maps.

Repeat element enrichment calculation

For each repeat class identified within eccDNA molecules (Figure 6D):

$$\text{Enrichment ratio} = (\%\text{ of eccDNA molecules containing repeat class X}) / (\%\text{ of genomic repeats that are class X})$$

Genomic repeat abundance reference: Reference frequencies for repeat classes were obtained from RepeatMasker annotations of the T2T-CHM13v2.0 human reference genome.

Statistical analysis of ecDNA junction repeat enrichment

Repeat enrichment at ecDNA junctions was evaluated by defining "matching family frequency," calculated as the proportion of junctions where the same repeat family flanked both the left and right breakpoints. This frequency was compared between real and randomized control datasets across multiple window sizes (10, 25, 50, 100, 250, 500, and 1000 bp) (Figures 6F, 6G, Figures S6E, and S6F). Statistical significance was assessed for each window size using Fisher's exact test (for odds ratios) and the Chi-squared test (for proportion differences), with p-values < 0.05 considered significant. All tests were two-sided and performed using R statistical packages invoked via the analysis pipeline.

ADDITIONAL RESOURCES

This study did not generate new websites, forums, or clinical trial data.