Title

Disparate Pathways for Extrachromosomal DNA Biogenesis and Genomic DNA Repair.

Journal

Cancer Discovery, 15(1)

Authors

Rose, John

Belk, Julia

Wong, Ivy

et al.

Publication Date

2025-01-13

DOI

10.1158/2159-8290.CD-23-1117

Peer reviewed

RESEARCH BRIEF

Disparate Pathways for Extrachromosomal DNA Biogenesis and Genomic DNA Repair

John C. Rose[1], Julia A. Belk[1], Ivy Tsz-Lo Wong[2,3], Jens Luebeck[4], Hudson T. Horn[5], Bence Daniel[2,3], Matthew G. Jones[1], Kathryn E. Yost[1], King L. Hung[1], Kevin S. Kolahi[5], Ellis J. Curtis[2,3], Calvin J. Kuo[5], Vineet Bafna[4,6], Paul S. Mischel[2,3], and Howard Y. Chang[1,7]

ABSTRACT

Oncogene amplification on extrachromosomal DNA (ecDNA) is a pervasive driver event in cancer, yet our understanding of how ecDNA forms is limited. In this study, we couple a CRISPR-based method for ecDNA induction with extensive characterization of newly formed ecDNAs to examine their biogenesis. We find that DNA circularization is efficient, irrespective of 3D genome context, with the formation of 800 kb, 1 Mb, and 1.8 Mb ecDNAs reaching or exceeding 15%. We show nonhomologous end joining and microhomology-mediated end joining both contribute to ecDNA formation, whereas inhibition of DNA-PK catalytic subunit and ATM have opposing impacts on ecDNA formation. ecDNA and the corresponding chromosomal excision scar can form at significantly different rates and respond differently to DNA-PK catalytic subunit and ATM inhibition. Taken together, our results support a model of ecDNA formation in which double-strand break ends dissociate from their legitimate ligation partners prior to joining of illegitimate ends to form the ecDNA and excision scar.

SIGNIFICANCE: Our study harnesses a CRISPR-based method to examine ecDNA biogenesis, uncovering efficient circularization between double-strand breaks. ecDNAs and their corresponding chromosomal scars can form via nonhomologous end joining or microhomology-mediated end joining, but the ecDNA and scar formation processes are distinct. Based on our findings, we establish a mechanistic model of excisional ecDNA formation.

[1]Center for Personal Dynamic Regulomes, Stanford University, Stanford, California. [2]Sarafan ChEM-H, Stanford University, Stanford, California. [3]Department of Pathology, Stanford University School of Medicine, Stanford, California. [4]Department of Computer Science and Engineering, UC San Diego, La Jolla, California. [5]Division of Hematology, Department of Medicine, Stanford University School of Medicine, Stanford, California. [6]Halicioglu Data Science Institute, UC San Diego, La Jolla, California. [7]Howard Hughes Medical Institute, Stanford, California.

Corresponding Authors: Howard Y. Chang, Department of Dermatology, Stanford University, 269 Campus Drive, CCSR 2155, Stanford, CA 94305. E-mail: howchang@stanford.edu; and Paul S. Mischel, Department of Pathology, Stanford University, 300 Pasteur Drive, L235, Stanford, CA 94305. E-mail: pmischel@stanford.edu

Cancer Discov 2025;15:69–82

doi: 10.1158/2159-8290.CD-23-1117

INTRODUCTION

Extrachromosomal DNA (ecDNA) drives accelerated tumor evolution and is associated with worse outcomes for patients with a wide variety of cancers (1–4). Potent oncogenes are frequently encoded on ecDNA and can become massively amplified in copy number and upregulated in gene expression (2, 4). ecDNA is circular and acentric, ranging in size from 100 kb to several Mb (2, 5). Due to their lack of centromeres, ecDNA is randomly segregated at cell division and drives intratumoral heterogeneity, facilitating rapid genomic changes to adapt to treatment or metabolic stress (6). Despite these advances, the biogenesis of ecDNA remains poorly understood.

Multiple mechanisms of ecDNA formation have been proposed, including circularization of segments generated by two or more double-strand breaks (DSB) on the same chromosome arm ("excisional ecDNA formation"), or ligation of multiple segments resulting from chromothripsis. Mechanistic investigations of ecDNA formation have proved challenging, as current studies rely on patient samples, patient-derived cell lines, or cell lines subjected to drug selections (7–10). In these studies, the requirement for selection—either within the tumor microenvironment or *in vitro*—of rare cells possessing spontaneously generated ecDNA obscures the process of ecDNA formation. More recently, ecDNAs have been engineered in human cells and mice using Cre-lox, enabling examination of ecDNA without extended selection for ecDNA$^+$ cells (bioRxiv 2023.06.25.546239). However, due to the use of Cre-mediated recombination to form ecDNA, this method cannot be used to examine the process of ecDNA formation itself.

Mechanistic understanding of excisional ecDNA biogenesis has thus remained elusive. We consider two primary models for how excisional ecDNA formation may occur based on the literature: (i) the end-swapping model and (ii) the dissociation model. In the end-swapping model, the DSBs undergo synapsis, holding each end close to its legitimate partner. Subsequently, these two synapses come into physical proximity, repair foci merge, and the ends swapped and ligated, generating the ecDNA and chromosomal excision scar. The end-swapping model is analogous to the proposed mechanism of reciprocal translocations (11). In the dissociation model, a failure to form or maintain the DSB synapses is followed by dissociation of the legitimate DSB end partners from one another. Approximation and ligation of the ecDNA and excision scar thus proceed as relatively independent processes. These two models yield divergent predictions for (i) the impact of chromosome conformation on circularization efficiency, (ii) the relative rates of ecDNA versus scar formation, (iii) the similarity of repair signatures on ecDNA and scar junctions, and (iv) whether inhibition of DNA damage response (DDR) proteins would differentially impact ecDNA versus scar formation. To test these predictions and distinguish between these models, there is a need for model systems that enable us to study ecDNA molecules immediately after formation. Furthermore, such models would provide an opportunity to directly identify factors that may influence ecDNA formation.

We recently applied CRISPR-C, a method for generating circular DNA (12), to the study of early ecDNA evolutionary dynamics (6). CRISPR-C offers temporal control and flexibility in locus selection, while relying on endogenous DNA repair pathways for circularization. Here, we use CRISPR-C to examine the process of ecDNA biogenesis itself. Using droplet digital PCR (ddPCR), fluorescence microscopy, chromosome conformation capture (Hi-C), and targeted deep sequencing of ecDNA and scar junctions, we directly examine ecDNA formation following excision from chromosomes. We find that circularization of megabase-scale fragments between DSBs is efficient and can be mediated by nonhomologous end joining (NHEJ) or microhomology-mediated end joining (MMEJ). Inhibition of DNA-PK or ATM produce opposing effects, with DNA-PK inhibition promoting ecDNA formation at most sites. Collectively, these results support the dissociation model of excisional ecDNA formation and shed light on how these amplifications arise in cancers.

RESULTS

Generation of ecDNA with CRISPR-C

To validate CRISPR-C for the study of ecDNA (Fig. 1A), we investigated whether we could generate ecDNAs containing *MYC*, which is frequently amplified on ecDNA in human cancers (4). We designed single-guide RNAs (sgRNA) to excise a 2.8-Mb region that encompassed the chromosome 8 span of a *MYC*-containing ecDNA found in COLO320-DM (2), with two guides targeting the centromeric end and three guides targeting the telomeric end (Fig. 1A–C). We delivered Cas9 with all pairwise permutations of centromeric and telomeric sgRNAs into the near-haploid Hap1 cell line. At 24 hours, circularization was detected for all six combinations by inverse PCR and ddPCR (Supplementary Fig. S1A–S1C). We repeated this experiment with the most efficient sgRNA pair, resulting in a mean circularization frequency of 6.1% (SEM = 0.81%; $n = 3$); however, no circle junctions were detected when only a single guide was delivered (Fig. 1D). Generation of a 1.8-Mb *DHFR*-containing ecDNA via CRISPR-C was more efficient (16.1%; SEM = 1.37%; $n = 3$) despite the individual *MYC* sgRNAs producing more indels than the *DHFR* sgRNAs at 24 hours (Fig. 1E; Supplementary Fig. S1D). FISH on metaphase chromosome spreads confirmed the extrachromosomal localization of CRISPR-C–excised *DHFR* ecDNA (Fig. 1F; ref. 6).

Spatial Proximity Is Dispensable for Efficient Excisional ecDNA Generation

Translocations occur more frequently between spatially proximal loci (13–16). Excisional ecDNA generation is a form of intrachromosomal translocation, so it may be necessary to consider genome conformation when designing CRISPR-C sgRNAs. The end-swapping model predicts that ecDNA formation should be more efficient for spatially proximal loci, as only limited motion is required to merge the repair foci. Conversely, ecDNA formation would be expected to be less influenced by spatial proximity under the dissociation model.

ecDNAs range from approximately 0.1 to 5 Mb (2), so we focused on the impact of chromosome conformation on ecDNA generation between sites separated by ~1 Mb. Hi-C mapping

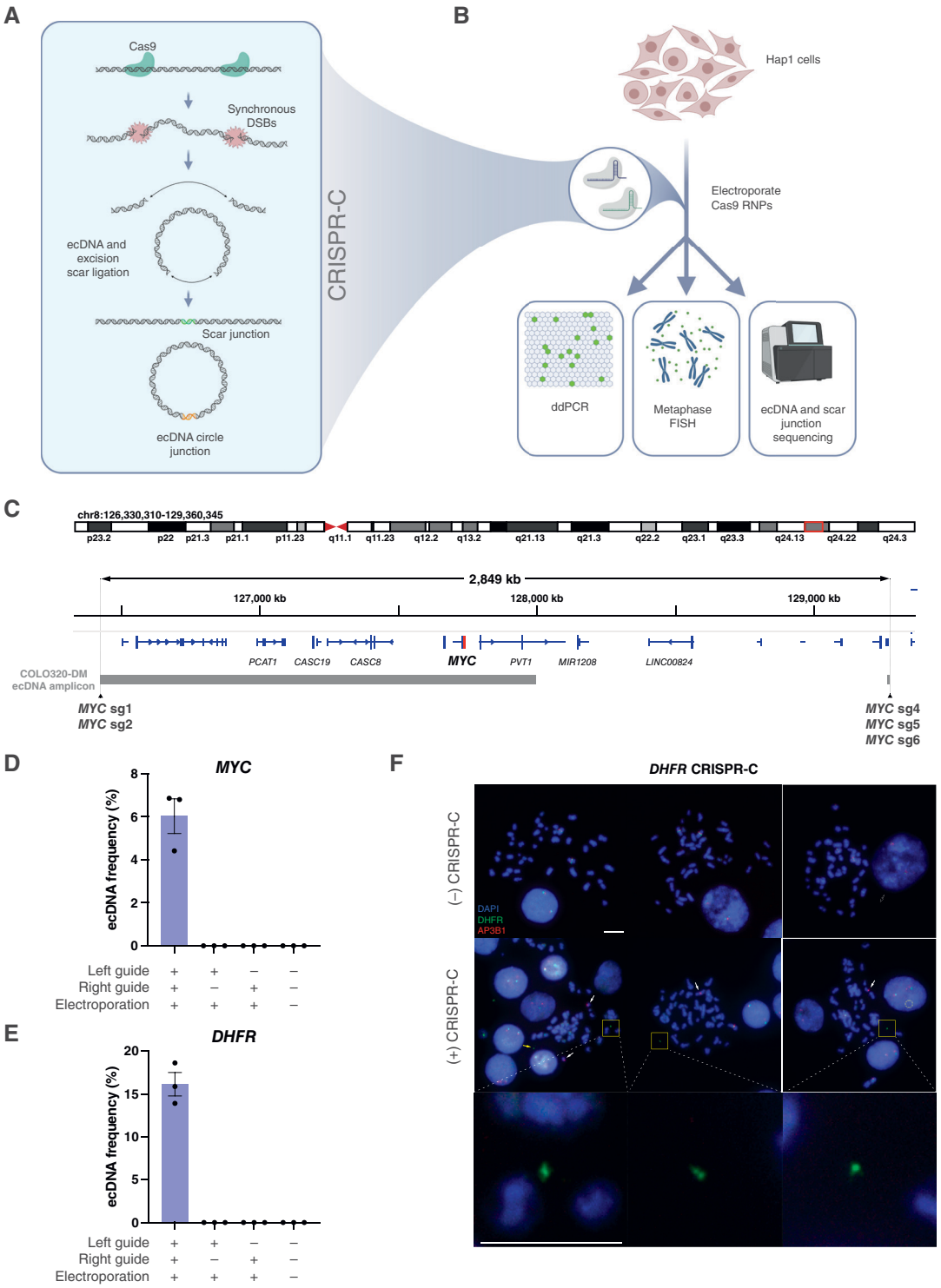

Figure 1. CRISPR-C can generate ecDNA at multiple chromosomal loci. **A,** Schematic depiction of ecDNA generation by CRISPR-C. **B,** Workflow for generating and analyzing induced ecDNA. **C,** Map of chromosome 8 locus indicating positions of guide RNAs used for CRISPR-C, the primary ecDNA amplicon from COLO320-DM, *MYC*, and other genes within the amplicon. **D,** *MYC* ecDNA frequency at 24 hours after electroporation of Cas9•RNPs complexed with both the left (MYC sg2) and right (MYC sg4) guides, or each Cas9•RNP individually, as determined by ddPCR. Bars depict means ($n = 3$ biological replicates; error bars = SEM). **E,** *DHFR* ecDNA frequency at 24 hours after electroporation of both Cas9•RNPs complexed with the left (DHFR gL) and right (DHFR gR) guides, or each Cas9•RNP individually, as determined by ddPCR. Bars depict means ($n = 3$ biological replicates; error bars = SEM). **F,** Representative FISH images of metaphase spreads of 24 hours after CRISPR-C, under non-electroporated control (top) and *DHFR* CRISPR-C (middle). Zoomed in images of the extrachromosomal *DHFR* signal (bottom). *AP3B1*, a *DHFR* proximal gene on chromosome 5, was stained by FISH to locate the native chromosome 5 locus. Yellow arrows, extrachromosomal *DHFR* signal; white arrows, location of chromosome 5 lacking *DHFR* signal. Scale bar, 10 μm. (**B,** Created with BioRender.com.)

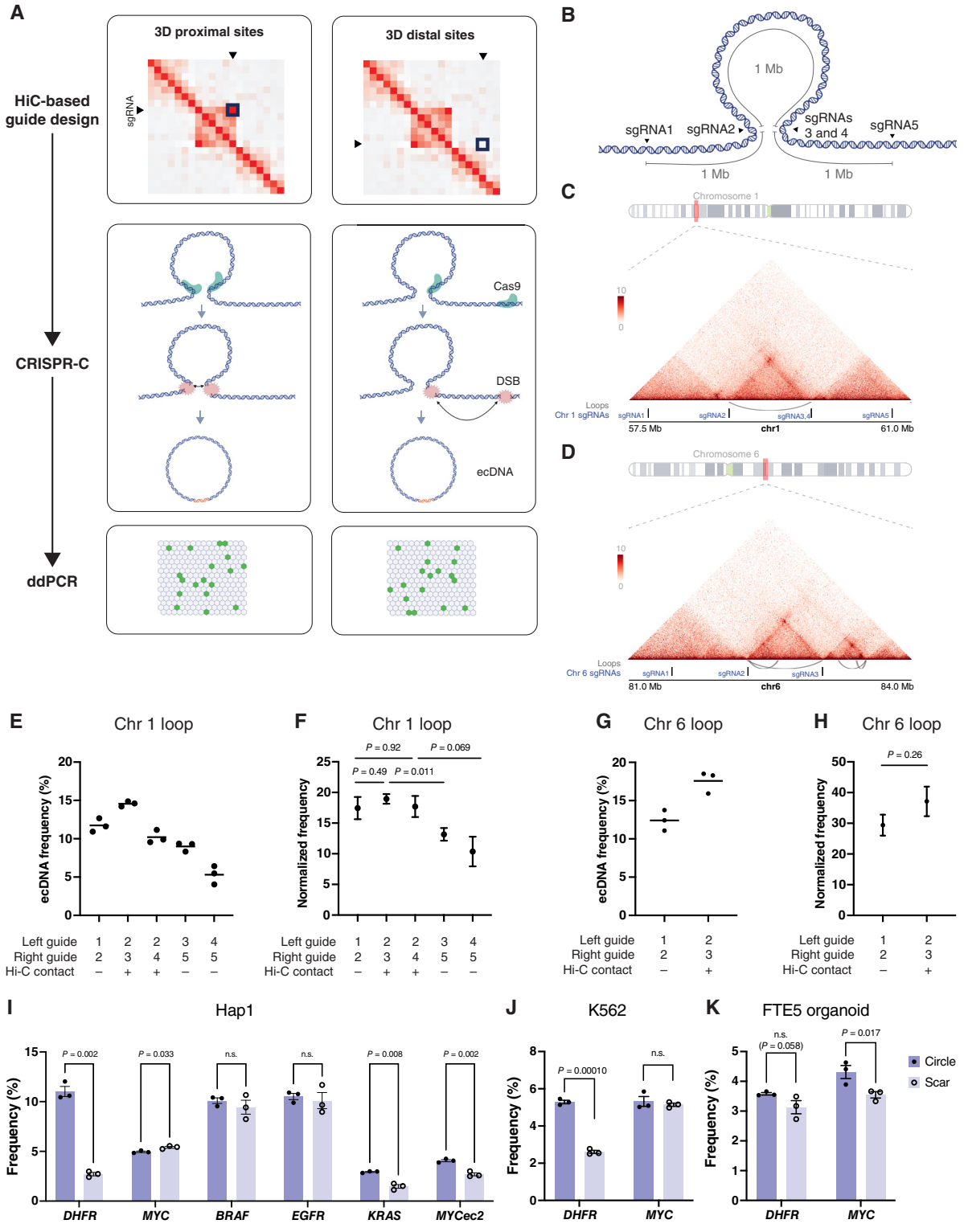

Figure 2. The impact of 3D chromosome conformation on ecDNA formation. Schematic depictions of (**A**) the workflow for examining the impact of chromosome conformation on circularization efficiency and (**B**) the chromosome 1 loop, with the approximate positions of the CRISPR-C guide RNAs used. Hap1 Hi-C interaction maps of (**C**) an approximately 1-Mb loop on chromosome 1 and (**D**) an approximately 800-kb loop on chromosome 6. Gray arches, loops; black lines, guide locations. **E,** ecDNA frequency generated by indicated chr1 guide pair. Line, mean. **F,** ecDNA frequency after normalization for individual guide editing efficiency ($n = 3$; error bars = SEM). **G,** ecDNA frequency generated by indicated chr6 guide pair. Line, mean. **H,** ecDNA frequency after normalization for individual guide editing efficiency. Dot, mean ($n = 3$; error bars = SEM). *P* values in **F** and **H** were calculated with unpaired two-sided *t* tests. ecDNA and scar junction frequencies at 24 hours in (**I**) Hap1 cells, (**J**) K562 cells, or (**K**) FTE5 ($TP53^{-/-}$) organoids. *P* values were calculated with paired two-sided *t* tests. Bars, means ($n = 3$ biological replicates; error bars = SEM). MYCe2, MYC ecDNA 2. (**B,** Created with BioRender.com.)

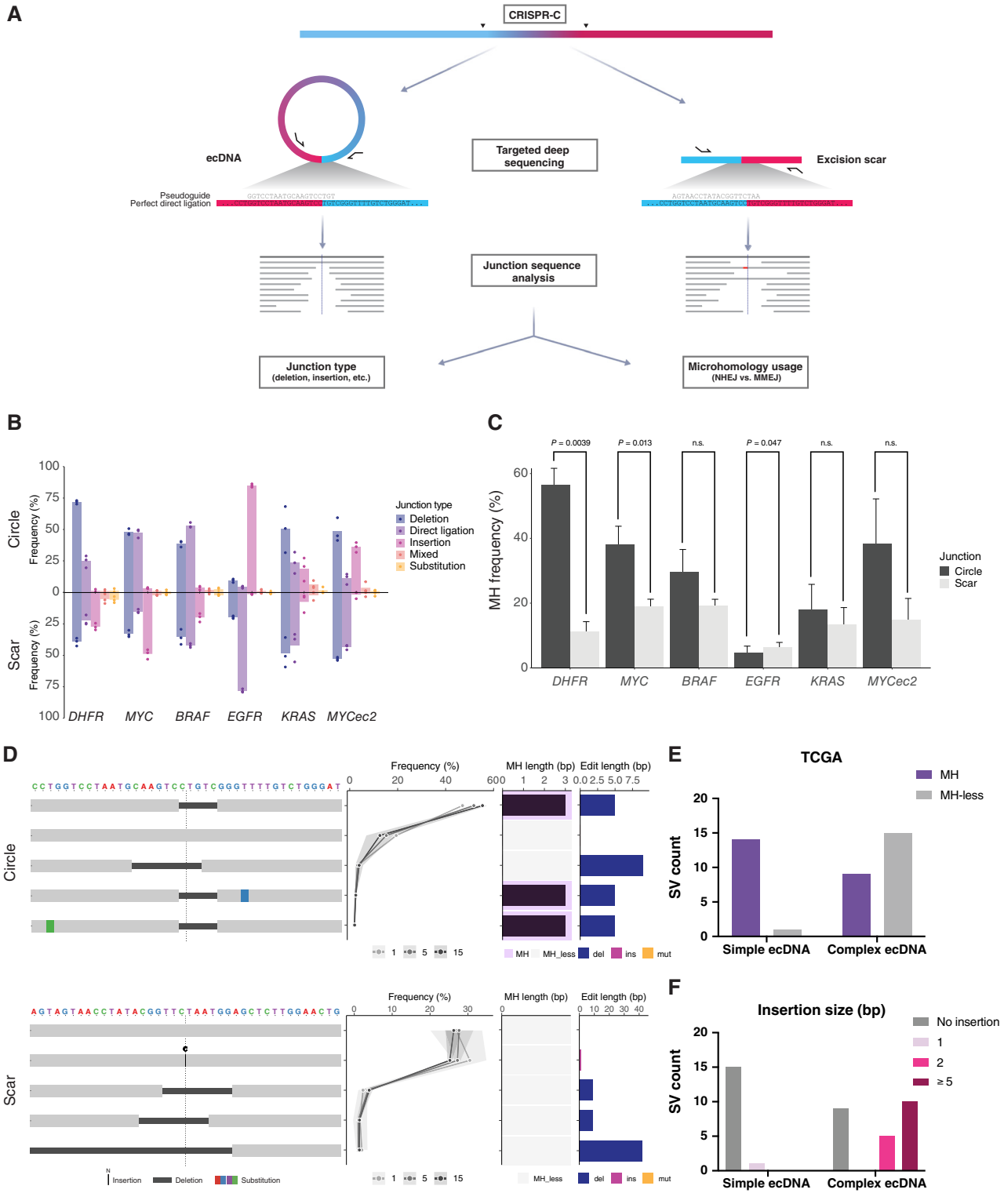

Figure 3. Analysis of junction sequences reveals roles for NHEJ and MMEJ in ecDNA and excision scar formation. **A,** Schematic of the junction sequencing strategy. **B,** Frequency of different junction types for circle and scar junctions for indicated loci. Mixed junctions contain more than one alteration (deletion, insertion, and/or substitution). Bars depict mean; $n = 3$ biological replicates. **C,** Microhomology frequency in circle and scar junctions. Bars depict means; error bars = SD; $n = 3$ biological replicates. *P* values calculated with paired two-sided *t* tests. **D,** Frequencies of the five most common circle and scar junction alleles for the DHFR neutral selection time course previously reported (6) on indicated days. Error bands = SD; $n = 3$ biological replicates. Microhomology length and the length of the indel or substitution for each junction allele are also indicated. Dashed line in left, junction site. Number of junctions on simple or complex ecDNA in TCGA samples that contain (**E**) MH or are MH-less or (**F**) insertions of the indicated lengths.

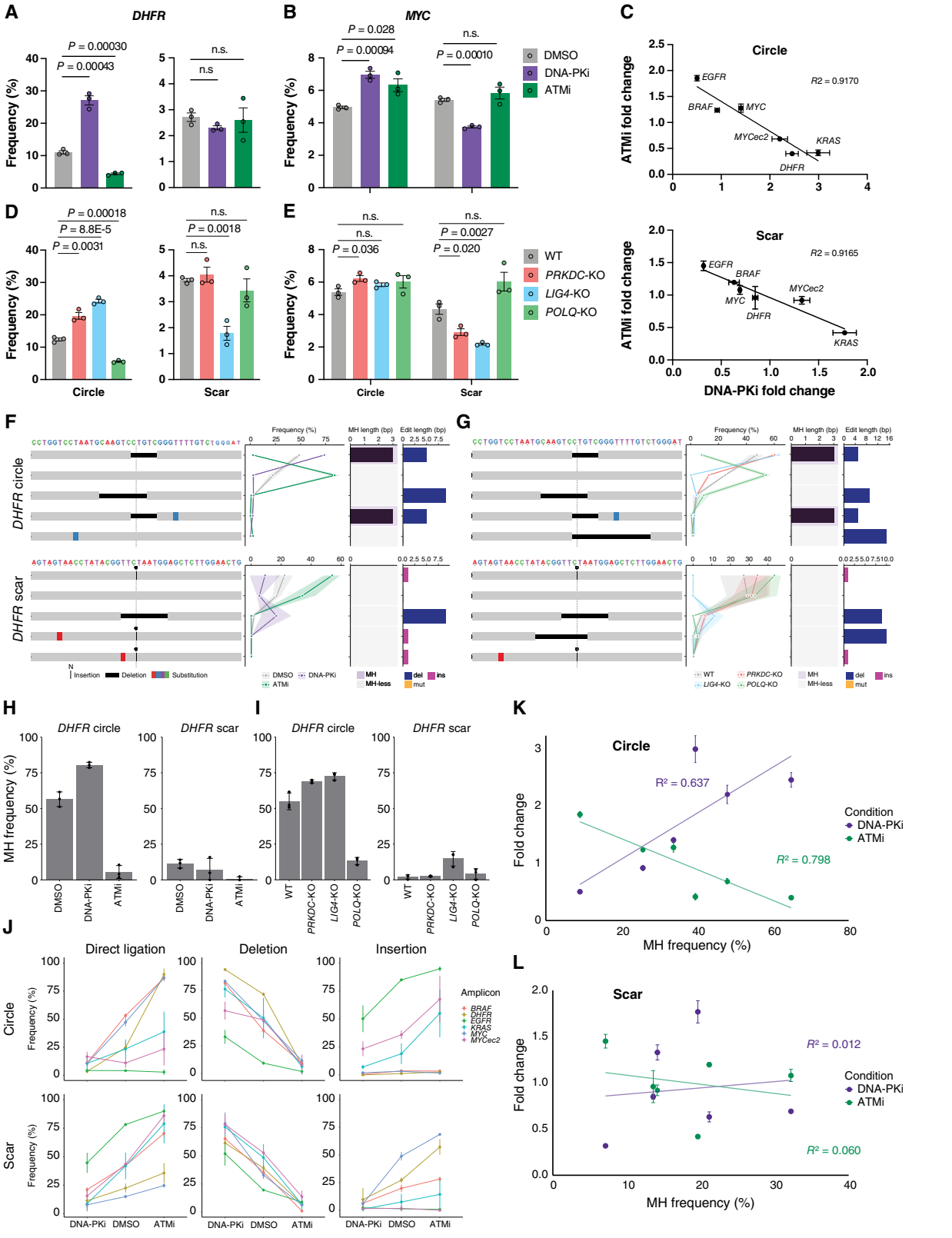

Figure 4. Pharmacologic and genetic modulation of ecDNA circularization. Impact of DNA-PKi or ATMi on CRISPR-C–mediated (**A**) *DHFR* or (**B**) *MYC* ecDNA and scar generation frequency. Bar indicates mean; error bars = SEM; $n = 3$ biological replicates. **C,** Fold change of ecDNA or scar frequency of DNA-PKi vs. ATMi relative to DMSO. Error bars = SEM; $n = 3$ biological replicates. Impact of DDR gene KO on CRISPR-C–mediated (**D**) *DHFR* or (**E**) *MYC* ecDNA and scar generation frequency. Bar indicates mean; error bars = SEM; $n = 3$ biological replicates. (*continued on following page*)

of intrachromosomal interactions in Hap1 cells identified an approximately 1-Mb loop on chromosome 1 and 800-kb loop on chromosome 6 that lacked the conformational complexity of the *MYC* and *DHFR* loci (Fig. 2A–D; Supplementary Fig. S2A and S2B). For chromosome 1, we designed guides targeting the loop ends, which exhibit high contact frequency with each other, or approximately 1 Mb outside of the loop, which exhibit minimal contact frequencies with the other target sites. CRISPR-C efficiently generated ecDNA for all sgRNA pairs tested, with moderately higher rates detected for guides targeting sites with high contact probabilities (Fig. 2E). After normalizing for individual guide activity, however, this increase in circularization frequency was not statistically significant for 3 of 4 pairwise comparisons (Fig. 2E and F; Supplementary Fig. S2C). Similarly, circularization was more efficient between spatially proximal sites at a loop on chromosome 6 in both Hap1 and the extensively characterized K562 cell line, but this effect was not significant after normalizing for the individual editing efficiencies of the sgRNAs (Fig. 2D, G, and H; Supplementary Fig. S2D–S2G). These data suggest that the spatial proximity of breakpoints may modestly promote circularization, but such proximity is not necessary for efficient excisional ecDNA formation, providing indirect evidence to support the dissociation model.

ecDNA and Excision Scar Formation Processes Exhibit Divergent, Site-Dependent Formation Rates

In some ecDNA$^+$ patient-derived samples and cell lines, a deletion exists in the corresponding native chromosomal locus—the excision scar (Fig. 1A; refs. 5, 6, 17). The end-swapping model predicts that the ecDNA and the scar should form at a 1:1 ratio. We quantified the efficiency of ecDNA and scar formation in six ecDNAs, including two distinct ecDNAs containing *MYC* (Fig. 2I). For four of six ecDNAs, the circle and scar exhibited unequal formation frequencies, in which circle formation was more efficient than scar formation for three ecDNAs, whereas scar formation was more efficient for the 2.8-Mb *MYC* ecDNA in Hap1 cells. Scar formation was more efficient for all sgRNA pairs initially screened to generate *MYC* ecDNA (Supplementary Fig. S3). Divergent ecDNA and scar formation rates were seen in the disparate model systems of K562 cells and FTE5, a nontransformed human *TP53*(−/−) fallopian epithelium organoid (Fig. 2J and K). High-grade serous ovarian cancers are frequently ecDNA positive, are universally *TP53* mutated, and arise from fallopian epithelium (4, 18). These results indicate that ecDNA and scar formation processes are, to some degree, independent processes, consistent with the dissociation model of ecDNA formation.

ecDNA and Excision Scar Formation Involves NHEJ and MMEJ

ecDNAs can form via several avenues (7), but the end-joining mechanisms to form a circle remain unclear. NHEJ is capable of generating translocations, but its rapidity opposes translocations (19). MMEJ can also mediate translocations (20), and microhomologies of 1 to 5 bp have been found in the ecDNA circle junctions of gliomas (21, 22). To examine the contributions of these pathways to ecDNA generation, we performed targeted deep sequencing of ecDNA circle and scar junctions after CRISPR-C (Fig. 3A).

Circle and scar junctions exhibited distinct sequence characteristics. Deletions and perfect direct ligations were prevalent in both types of junctions, but insertions were common either in the circle or scar junctions for each ecDNA, not both (Fig. 3B). For *DHFR* and *MYC* ecDNAs, the distributions of junction types were similar in K562 cells to Hap1 at both circle and scar junctions, whereas perfect direct ligations were more common in FTE5 (*TP53*$^{-/-}$) organoids (Supplementary Fig. S4A). We next classified junctions as containing microhomology (MH)—possessing ≥1 bp of microhomology—or microhomology-less (MH-less; refs. 23, 24). All insertions were classified as MH-less. For MH alleles, we also determined the MH length. Circle and scar junctions frequently exhibited MH, with MH more common in circle junctions for five of the six loci tested, albeit this difference was only statistically significant at two of those five sites (Fig. 3C). MH was less common at junctions in FTE5 (*TP53*$^{-/-}$) relative to Hap1 and K562 cells (Supplementary Figs. S4B and S5A). The different sequence characteristics of the ecDNA and scars suggest that distinct DNA repair processes are involved, supporting the dissociation model of excisional ecDNA formation.

We next examined the allele distributions (25) for the *DHFR* ecDNA and scar. The two most common circle junction sequences—a 5 bp deletion with 3 bp MH and perfect direct ligation—accounted for 66.40% ± 0.29% (SEM; $n = 3$) of circle junctions at 24 hours (Fig. 3D). Utilization of different repair pathways could impact or be reflective of cell fitness. We examined *DHFR* circle and scar junction sequence distributions over time under neutral selection, i.e., in which the ecDNA itself does not confer a fitness advantage or disadvantage (6). The sequence distributions of both the circle and scar remained stable over time, indicating that differential repair pathway use did not relate to or affect cell fitness.

Microhomology Is Common and Insertions Rare at ecDNA Junctions of Excisional ecDNA in Patient Samples

We next examined the junctions of ecDNA identified in The Cancer Genome Atlas (TCGA) samples. We limited our analysis to ecDNA for which reads spanning the junction exist to

Figure 4. ***(Continued)*** *P* values in **A**, **B**, **D**, and **E** calculated with unpaired two-sided *t* tests. Impact on the frequencies of the five most common DHFR circle and scar junction alleles for (**F**) pharmacologic or (**G**) genetic manipulation of DDR. Error bands = SD; $n = 3$ biological replicates. Microhomology length and the length of the indel or substitution for each junction allele are also indicated. Dashed line in left, junction site. Impact of (**H**) pharmacologic or (**I**) genetic manipulation of the DDR on MH frequency at the DHFR circle and scar junctions. Bars depict means; error bars = SD; $n = 3$ biological replicates. **J,** Impact of DNA-PKi and ATMi on the frequencies of direct ligations, deletions, or insertions, compared with DMSO control. Dots depict means; error bars = SD; $n = 3$ biological replicates. Fold change relative to DMSO of DNA-PKi or ATMi on (**K**) ecDNA frequency as a product of circle junction MH frequency or (**L**) scar frequency as a product of scar MH frequency. MH frequencies are the mean value for the DMSO condition. Dots indicate mean fold-change value. Error bars = SEM; $n = 3$ biological replicates. WT, wild type.

unambiguously identify the junction sequence. These ecDNAs were categorized as simple or complex: simple ecDNAs are predicted to have been derived via excisional formation; complex ecDNAs are predicted to have been formed via chromothripsis (bioRxiv 2024.05.06.592768). We hypothesized that complex ecDNA junctions would exhibit infrequent MH, as ligation of chromothripsis-derived DNA fragments is mediated via NHEJ (bioRxiv 2023.08.10.552800). Conversely, we predicted that MH would be more frequent in excision-derived simple ecDNAs, as their formation more closely mirrors our CRISPR-C–based model. Indeed, MH was significantly more common at junctions in simple ecDNA compared with complex ecDNA (Fig. 3E; P = 0.0007; Fisher exact test). Multibase insertions were common in complex ecDNA but not detected at simple ecDNA junctions (Fig. 3F; Supplementary Fig. S5B). Sequencing depth limitations of these samples precluded analyses of scar junctions.

DNA-PK and ATM Inhibition Produces Opposing Effects on ecDNA Formation

The DDR in humans is governed by three related kinases: DNA-PK, ATM, and ATR (26). DNA-PK is a regulator of canonical NHEJ, which repairs most DSBs in human cells. Rapid synapse formation by NHEJ machinery suppresses translocations (19), whereas DSBs that persist have more time to pair with other DSBs and generate translocations (20). Inhibition of DNA-PK catalytic subunit (DNA-PKcs) leads to its retention on DSB ends, greater DSB persistence, and an increase in the frequency of translocations (11, 20, 27–30). We reasoned that DNA-PKcs inhibition would similarly favor excisional ecDNA formation. ATM has been called the "master regulator of cellular responses" to DSBs, but most DSBs outside the S-phase can be repaired by NHEJ in an ATM-independent manner (26). ATM has been implicated in DSB mobility and the clustering of a subset of DSBs (31), whereas deficiency of ATM impairs MMEJ (32). We predicted that ATM inhibition would suppress ecDNA formation by favoring rapid NHEJ and limiting the clustering of DSBs. ATR is primarily involved in the response to DNA replication stress, so we reasoned that ATR inhibition would have limited impact (26).

To test these hypotheses, we performed CRISPR-C to generate *DHFR* ecDNA in Hap1 cells in the presence or absence of a DNA-PKcs inhibitor (DNA-PKi; AZD7648), ATM inhibitor (ATMi; KU-55933), or ATR inhibitor (berzosertib). We found that DNA-PKcs inhibition increased *DHFR* ecDNA formation by 2.46-fold (27.07% ± 1.41% vs. 11.02% ± 0.51% in DMSO control; SEM; $n = 3$; $P = 0.00043$), whereas ATM inhibition suppressed ecDNA formation by 2.5-fold (Fig. 4A; 4.40% ± 0.24% vs. 11.02% ± 0.51% in DMSO control; SEM; $n = 3$; $P = 3.0 \times 10^{-4}$). As expected, ATR inhibition had no impact on *DHFR* ecDNA formation (Supplementary Fig. S6). None of the drugs significantly affected the rate of *DHFR* excision scar formation.

Next, we examined the impact of DNA-PKi and ATMi on ecDNA and scar formation at additional sites. Including *DHFR*, DNA-PKi increased ecDNA formation at four of six sites, but decreased *EGFR* ecDNA formation, and had no significant impact on *BRAF* ecDNA formation (Fig. 4A–C; Supplementary Figs. S7 and S8). Scar formation was impaired at three sites and promoted at two sites by DNA-PKi. ATMi decreased ecDNA formation at three of six sites, but increased ecDNA formation at the other three sites, and had no significant effect on scar formation rates at most sites. Overall, the effects of DNA-PKi and ATMi on both ecDNA formation and scar formation were anticorrelated—stronger promotion of ecDNA formation by DNA-PKi was associated with stronger inhibition of that same ecDNA by ATMi and vice versa (Fig. 4C). Genetic knockout (KO) of the NHEJ components *LIG4* and *PRKDC* (encoding DNA-PKcs) and the MMEJ component *POLQ* validated the preceding pharmacologic experiments (Fig. 4D and E). We found, however, that the formation of the *DHFR* excision scar was sensitive to loss of *LIG4*, but not inhibition or KO of *PRKDC* (DNA-PKcs), suggesting that this scar can form via a DNA-PK–independent variant of the NHEJ pathway.

DNA-PKi favors MMEJ at CRISPR-induced DSBs over NHEJ (33), whereas ATMi has the opposite effect (32). DNA-PKi or KO of *PRKDC* or *LIG4* decreased the frequency of perfect direct ligation at the *DHFR* ecDNA circle junction, with a corresponding increase in the frequency of the primary MMEJ-mediated junction (Fig. 4F and G). Across all loci, DNA-PKi promoted MMEJ at the expense of NHEJ (Supplementary Figs. S9 and S10). Compared with *PRKDC*-KO, DNA-PKi elicited more pronounced effects on both ecDNA formation rate and junction sequences, likely a consequence of DNA-PKi trapping DNA-PKcs on DSB ends and preventing their further processing (Fig. 4A, B, and D–J; Supplementary Fig. S9; ref. 28). Intriguingly, *LIG4*-KO completely suppressed perfect direct ligations, whereas only partial loss was seen with inhibition or KO of DNA-PK, consistent with reports suggesting that DNA-PK is dispensable for direct ligation of blunt ends *in vitro* (34), and the LIG4/XRCC4 complex can be recruited to DSBs in the absence of DNA-PKcs (35). Inhibition of ATM or loss of *POLQ* biased repair toward NHEJ. ATMi, in particular, affected near complete suppression of MH, with corresponding increases in NHEJ-mediated insertions and perfect direct ligations, consistent with studies of Cas9-generated DSBs with individual guides (Fig. 4F–J; Supplementary Figs. S9 and S10; refs. 36, 37). Overall, suppression of NHEJ increased MH usage and deletions, whereas suppression of MMEJ decreased MH and increased insertions and perfect direct ligations. These shifts in pathway usage further validate the involvement of both NHEJ and MMEJ in ecDNA formation, and the ability of these two pathways to compensate for each other in this process.

We next considered whether MH usage at ecDNA or scar junctions in the absence of pharmacologic or genetic manipulation was associated with how the site responded to DNA-PKi and ATMi. MH frequency was positively associated with the effect of DNA-PKi and negatively associated with the effect of ATMi on ecDNA formation (Fig. 4K). However, the frequency of MH at the scar junction displayed no relationship with the impact of these two drugs on scar formation (Fig. 4L). The differential relationships of ecDNA circle and excision scar with repair outcomes and repair pathway modulation demonstrate they are separate processes, in agreement with the dissociation model.

DISCUSSION

In this report, we harnessed CRISPR-C to examine excisional ecDNA formation in human cells. Contrary to expectations (7), we found circularization of DNA to be remarkably efficient. Further work will be needed to determine the frequency of proximal (within 0.1–5 Mb), contemporaneous DSBs forming through endogenous processes or environmental insults (radiation or chemotherapeutics) and how different DSB etiologies affect ecDNA formation. The efficiency of circularization suggests that excisional ecDNA formation is likely more common than previously appreciated and that the circularization step is likely not the primary barrier to ecDNAs arising in healthy and diseased tissues.

It is well established that spatial proximity and Hi-C contact frequency specifically are associated with higher translocation rates (13–16). Here, greater Hi-C contact probability correlated with moderately, but not significantly, elevated CRISPR-C efficiency. More than 90% of intrachromosomal rearrangement breakpoints in breast cancer are within 2 Mb of each other (38), whereas genome-wide profiling of translocations revealed a similar preference for proximal end-joining (15). Within the typical ecDNA size range (0.1–5 Mb), circularization is likely efficient enough to overcome variations in 3D conformation. Thus, CRISPR-C can be used to efficiently induce ecDNA without first conducting expensive and laborious 3D genome conformation profiling, but chromosome conformation is likely not the primary determinant of what genome regions are amplified on ecDNA.

The DDR has been implicated in ecDNA generation for decades (39). Roles have been proposed for NHEJ and MMEJ (21, 22, 40, 41), yet the processes responsible for end-joining to form the circle have not been defined. In human cancers and cell lines, translocations are reported to occur primarily via NHEJ (42), yet we found that both NHEJ and MMEJ are proficient at generating ecDNA and excision scars. This finding was validated by the detection of MH in the junctions of numerous excision-derived ecDNAs in patient samples. The relative contribution of NHEJ and MMEJ varied between sites, but MH was generally lower at scar junctions, suggesting a greater reliance of scar end-joining on NHEJ. Future studies will be needed to determine how repair pathway contributions are influenced by DNA sequence, chromatin context, tissue type, and the transformation status of the cell.

At most sites, suppression of NHEJ promoted ecDNA formation, in line with the observation that inhibition or loss of DNA-PKcs increases DSB persistence and promotes translocations within and between chromosomes (43). Increased DSB persistence likely promotes ecDNA formation by (i) increasing the temporal overlap of DSBs; (ii) enabling dissociation of DSB ends from their legitimate partners; and (iii) favoring the intramolecular end-joining event (Supplementary Fig. S11A). The importance of increased DSB persistence is corroborated by the observation that DNA-PKi—which traps DNA-PKcs on DSB-ends—elicits a greater effect on ecDNA formation than *PRKDC*-KO. Recent work has suggested that DNA-PKi impedes the formation of ecDNA from chromothripsis (9, 44). The apparent discrepancy between these reports and ours may be explained by a greater reliance on NHEJ to join fragments generated by chromothripsis (bioRxiv 2023.08.10.552800), which aligns with our observation of lower MH frequency in chromothripsis-derived ecDNAs.

ATMi effectively suppressed MH and non-MH deletions at junctions, likely a consequence of ATM's role in activating end resection factors and other components of the DDR (45). Despite its impact on repair pathway usage, ATMi yielded only incomplete suppression of ecDNA formation at half of the sites, while promoting ecDNA formation at the others. We establish that ATMi and DNA-PKi have opposing effects on ecDNA formation from a given locus, but it remains unclear why DNA-PKi promotes ecDNA formation at some loci and ATMi promotes it at others. ATMi can limit the mobility and clustering of DSBs, whereas DNA-PKi can promote DSB clustering, but these effects are locus dependent (31, 46, 47). Mobile, clustered DSBs would be expected to more readily form ecDNA (Supplementary Fig. S11A). Our data also indicate that the balance of NHEJ versus MMEJ in circle formation predicts the response of the loci to DNA-PKi/ATMi (Fig. 4K and L), suggesting that inhibiting the less prominent pathway may shunt repair outcomes toward ecDNA formation. Notably, no such relationship was seen between scar junction MH and the response to inhibitors. Further elucidation of the mechanisms of ecDNA biogenesis will be no doubt aided by novel pharmacologic tools, such as the inhibitor of Polθ, ART558 (48).

Taken together, our work shows that excisional ecDNA formation occurs primarily via the dissociation model rather than the end-swapping model (Supplementary Fig. S11B). The end-swapping model predicts that (i) circularization of spatially proximal DSBs will be substantially more efficient than those at between sites with low Hi-C contact frequencies; (ii) ecDNA and excision scar formation rates will be equivalent; (iii) the repair signatures at circle junctions will be similar; and (iv) that the impact of DDR inhibitors on ecDNA and scar formation would be similar. All four tests proved contrary to these predictions, which positively supports the dissociation model.

In this study, we have demonstrated the power of ecDNA generation with CRISPR-C to yield meaningful insights into ecDNA biogenesis, establishing the dissociation model of excisional ecDNA formation and revealing mechanistic divergence in the formation of ecDNA and their chromosomal scars. Future therapeutic strategies to forestall the emergence of ecDNAs may target not only the mechanism of ecDNA formation but also the distinct processes governing repair of the linear locus from which they were excised. Aided by its portability, we anticipate this CRISPR-based approach will permit distinct and powerful investigations of the entire life cycle of ecDNAs in various contexts.

METHODS

Cell Culture

Hap1 (C631) and Hap1-KO cell lines (PRKDC-KO: HZGHC-024034c011; LIG4-KO: HZGHC000759c014; POLQ-KO: HZGHC-000647c003, Horizon Discovery) were maintained in Iscove's modified Dulbecco's medium (IMDM) supplemented with GlutaMAX and 10% FCS (Gibco). K562 cells (ATCC) were maintained in RPMI supplemented with 10% FCS. Cell lines were used as received without authentication. Cells regularly tested negative for *Mycoplasma* contamination.

Generation of Normal Fallopian Tube Organoids

The FTE5 organoid line was generated from phenotypically normal fresh surgical discards of a patient undergoing salpingectomy at Stanford University Hospital via Stanford University Hospital's Tissue Procurement Shared Resource Facility. Isolation of fallopian tube epithelial cells followed previously established protocols (49). Isolated fallopian tube cells were resuspended in Matrigel (R&D Systems, Basement Membrane Extract type 2 #3533-005-02) and plated in a 24-well plate as individual domes. Following solidification, organoid media supplemented with 10 μmol/L Y-27632 (PeproTech, #1293823) were added to each well. After 2 days, Y-27632 was removed from the media and changed every 3 to 4 days. Organoids were passaged every 14 days by dissolving Matrigel and dissociating fallopian tube organoids to single cells using TrypLE (Invitrogen, #12604-012) at 37°C for 15 minutes inverting 2 to 3 times every 5 minutes. TrypLE was quenched with FBS and removed following centrifugation for 3 minutes at 600 *g*. The single cells were resuspended in Matrigel and plated in a 24-well plate.

TP53 was knocked out of FTE5 by dissociating organoids to a single-cell suspension as above and electroporating in a PX330 plasmid transiently expressing both Cas9 and a gRNA targeting exon 4 of *TP53* (Addgene, #121917) using a Lonza 4D-Nucleofector X (#AAF-1003X) with protocol Primary Cell P3 code CA137. FTE5 ($TP53^{-/-}$) organoids were treated with 10 μmol/L nutlin-3a (Selleck Chemicals, #S1061) to select against remaining wild-type organoids and passaged multiple times before ecDNA generation.

Organoid Media

Advanced DMEM/F-12 (Thermo Fisher Scientific, #12634028) with 5% penicillin/streptomycin/glutamine (Thermo Fisher Scientific, #10378016), 5% FBS, 1 mmol/L HEPES (N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid, Thermo Fisher Scientific, #15630080), 1 mmol/L N-acetylcysteine (MilliporeSigma, #A9165), 1× B-27 supplement (Thermo Fisher Scientific, #12587001), 1× GlutaMax (Thermo Fisher Scientific, #35050061), 10 μmol/L SB-202190 (BioGems, #1523072), 10 mmol/L nicotinamide (MilliporeSigma, N0636), 50 ng/mL EGF (PeproTech, #AF-100-15), 100 μg/mL Normocin (InvivoGen, #ant-nr-1), and 50% Wnt-3A/R-spondin/Noggin conditioned media.

Guide RNA Design for CRISPR-C

sgRNAs were designed to target regions corresponding to the span of the canonical *MYC*-containing ecDNA from the COLO320-DM cell line, as shown in Fig. 1B. Guides for generating *DHFR*-containing ecDNA were previously described (6). To examine the role of chromosome conformation in CRISPR-C, regions of interest were identified by examining the HiC contact matrix (Fig. 2B). Guides for other sites were designed to mimic ecDNAs previously detected in cell lines (Supplementary Data Set S1). For all guides, regions of interest were examined within the UCSC Genome Browser, and RepeatMasker and GeneHancer tracks were used to avoid targeting repetitive regions and regulatory elements, respectively. Guides were designed using Integrated DNA Technologies (IDT) Custom Alt-R CRISPR-Cas9 guide RNA software (https://www.idtdna.com/site/order/designtool/index/CRISPR_CUSTOM). These sequences were ordered as Alt-R sgRNAs (IDT).

ecDNA Induction by CRISPR-C

Hap1 cells were trypsinized, quenched with IMDM (GlutaMAX and 10% FCS), counted, and centrifuged at 300 *g* for 5 minutes. Cells were washed once with PBS before resuspension in Neon Resuspension Buffer to 1.1×10^7 cells per mL for Hap1 cells and 2.67×10^7 cells per mL for K562 cells. Ribonucleoprotein (RNP) complexes were formed as follows: Cas9 (IDT) was diluted to 36 μmol/L in Neon Resuspension Buffer. Equal volumes of diluted Cas9 and sgRNA (44 μmol/L in TE (10 mmol/L Tris, 0.1 mol/L EDTA), pH 8.0) were mixed and incubated at room temperature for 10 to 20 minutes. Left and right sgRNA RNPs were assembled separately. Then 5.5 μL of each RNP, 5.5 22 μL of electroporation enhancer (10.8 μmol/L; IDT), and 99 μL of cells were mixed and electroporated according to the manufacturer's instructions using a 100-μL Neon pipette tip and electroporated using the Neon Transfection System (Thermo Fisher Scientific) using the following settings: Hap1 (1,575 V, 10 ms pulse width, and 3 pulses); K562 (1,700 V, 20 ms, and 1 pulse). Single-guide controls were prepared as above except 11 μL of the appropriate sgRNA was used. Volumes were scaled down for 10 μL electroporations. Cells were collected at 24 hours after electroporation as follows: Cells were washed with 1 mL per well prewarmed PBS (Gibco), followed by the addition of 100 μL TrypLE Express (Thermo Fisher Scientific) and incubation at 37°C for 5 to 10 minutes. TrypLE was quenched with 800 μL IMDM (GlutaMAX and 10% FCS), and the cell suspension was pelleted at 300 *g* for 5 minutes at 4°C. The supernatant was discarded, and the cell pellets were stored at −80°C.

For experiments including inhibition of DDR proteins, after electroporation, cells were plated in media containing either AZD7648 (1 μmol/L), SCR7 (1 μmol/L), berzosertib (80 nmol/L), KU-55933 (10 μmol/L), talazoparib (10 nmol/L), or DMSO. Samples were harvested 24 hours after electroporation. All compounds were obtained from Selleck Chemicals. The final DMSO concentration in all wells was 0.2%. Final drug concentrations were selected based on effective doses for cell culture reported in the literature (50–54).

For ecDNA generation via CRISPR-C in FTE5 organoids, organoid media were supplemented with Y-27632 1 day prior to electroporation. RNP complexes were made by incubating (per reaction) 10 μL Opti-MEM (Gibco, #31985062), 0.6 μl 10 μg/mL Cas9 (IDT, #1081059), and 44 μmol sgRNA for each sgRNA and incubated at room temperature for 15 minutes before combining. FTE5 organoids were dissociated as above, and 2×10^5 cells were resuspended in 10 μL Opti-MEM per reaction. Cells and RNP complexes were combined, and 20 μL was added per well of a Nucleocuvette strip (Lonza, #NC0828118) and electroporated using a Lonza 4D-Nucleofector X unit with protocol "Primary Cell P3 code CA137." Cells were collected from the cassette and resuspended in Matrigel before plating in a 24-well plate and given organoid media containing Y-27632 for 3 days before removing. Three reactions were performed per replicate. Twenty-four hours after electroporation, organoids were dissociated as before, and DNA was harvested from the organoids using a Qiagen DNeasy Blood and Tissue kit (Qiagen, #69506).

ddPCR to Determine ecDNA or Chromosomal Scar Frequency

Genomic DNA (gDNA) was isolated using DNeasy columns (Qiagen) according to the manufacturer's instructions, including incubation for 10 minutes at 56°C during the proteinase K digestion step; DNA was eluted with 100 μL EB (Elution Buffer). Amplicons for the ecDNA junction, chromosomal scar junction, and GAPDH were designed using IDT PrimerQuest software (https://www.idtdna.com/PrimerQuest/Home/Index). Dual-quenched probes (IDT) were used: FAM-labeled probes were used for both the ecDNA and chromosomal scar junction amplicons to facilitate multiplexing with the GAPDH amplicon utilizing a HEX-labeled probe. All probe and primer sequences are available in Supplementary Data Set S1. Droplets were created using droplet-generating oil for probes, DG8 cartridges, DG8 gaskets, and the QX200 Droplet generator (Bio-Rad Laboratories); amplification was performed using the ddPCR supermix for Probes (Bio-Rad Laboratories). The ddPCR supermix amplification reactions were set up according to the manufacturer's specifications (Bio-Rad Laboratories). Approximately 60 ng of gDNA was used in a 20 μL reaction with a final primer concentration of

900 nmol/L (225 nmol/L for each primer), 125 nmol/L FAM probe, and 125 nmol/L HEX probe. The reaction was partitioned into droplets for amplification according to the manufacturer's protocol (Bio-Rad Laboratories). Droplets were transferred to a 96-well PCR plate and heat-sealed using the PX1 PCR plate sealer (Bio-Rad Laboratories). Droplets were amplified using the following cycling conditions: 95°C for 10 minutes, 40 cycles (94°C for 30 seconds; 56.1°C for 60 seconds), and 98°C for 10 minutes. After thermal cycling, droplets were scanned individually using the QX200 Droplet Digital PCR system (Bio-Rad Laboratories). Positive and negative droplets in each fluorescent channel (HEX and FAM) were distinguished on the basis of fluorescence amplitude using a global threshold set by the minimal intrinsic fluorescence signal resulting from imperfect quenching of the fluorogenic probes (negative droplets) compared with the strong fluorescence signal from cleaved probes in droplets with amplified template(s). The frequency of ecDNA or chromosomal scar was calculated by dividing their measured concentration by the concentration of the *GAPDH* amplicon. Normalized circularization frequency (Fig. 2D) was calculated as follows:

$$\frac{\text{ecDNA frequency}}{(\text{Left guide indel frequency}) \times (\text{Right guide indel frequency})} \times 100$$

The indel frequencies were determined for each guide individually as described in ***CRISPR Editing Quantification***.

CRISPR Editing Quantification

Quantification of editing by CRISPR with individual sgRNAs was performed using the ICE analysis tool (Synthego, v3.0). Briefly, the target locus was amplified using Platinum SuperFi polymerase master mix (Thermo Fisher Scientific; for primers, see Supplementary Data Set S1). PCR purified and Sanger sequencing were performed by Elim Biopharmaceuticals, Inc. Sequence traces and sgRNA spacer sequences were then uploaded to the ICE webtool for analysis.

Metaphase FISH

Cells were arrested in mitosis with KaryoMAX Colcemid Solution (Gibco, #15122012) for 4 hours. The cells were then collected and washed once in 1× PBS. The cell pellet was resuspended in 50 μL 1× PBS and treated with 0.075 mol/L KCl buffer for 20 minutes at 37°C. The cells were then fixed in fresh Carnoy's fixative (3:1 methanol:glacial acetic acid) followed by three additional washes with the fixative. Cells were then dropped to a humidified slide and air-dried at room temperature. The slide was briefly equilibrated in 2× saline sodium citrate (SSC) buffer and subjected to dehydration with ascending series of ethanol (70%, 85%, and 100%) for 2 minutes each. The slides were air-dried completely, and FISH probes (Empire Genomics, *DHFR* gene: DHFR-20-GR; *AP3B1* gene: AP3B1-20-RE) diluted in a 1:6 ratio with hybridization buffer were applied to the sample. A coverslip was applied, and the sample was subjected to heat denaturation at 75°C for 3 minutes, followed by hybridization overnight at 37°C. The coverslip was removed, and the slide was washed in 0.4× SSC buffer and then in 2× SSC (0.1% Tween) for 2 minutes each. DAPI (4', 6-diamidino-2-phenylindole) stain (50 ng/mL) was used to stain for nuclei for 2 minutes and the slides were washed once briefly in ddH_2O. The slides were mounted with ProLong Diamond Antifade (Invitrogen, #P36961) and air-dried overnight. Images were acquired on a Leica DMi8 THUNDER imager using a 63× oil objective.

Targeted Deep Sequencing of ecDNA and Excision Scar Junctions

Twenty cycles of primary PCR to amplify the region of interest was performed using ~200 ng μL of gDNA in a 10 μL Platinum SuperFi II polymerase reaction (Thermo Fisher Scientific; for primers, see Supplementary Data Set S1). Illumina adapters and indexing sequences were added via 15 cycles of secondary PCR with 1 μL of primary PCR product in a 10 μL Platinum SuperFi II polymerase reaction. The final amplicons were run on a TAE (Tris-acetate-EDTA)-agarose gel (0.7%), and the product band was excised and extracted using the Freeze 'N Squeeze Kit (Bio-Rad) according to the manufacturer's instructions. Gel-purified amplicons were quantified with a double-stranded DNA HS Assay kit on a Qubit fluorometer (Thermo Fisher Scientific). Then amplicons were pooled and sequenced on the NextSeq 550 platform (NextSeq 500/550 Mid Output Kit v2.5, Illumina; for primers, see Supplementary Data Set S1). *DHFR* CRISPR-C time course samples in Fig. 3D were reported as previously described (6) and subsequently sequenced for this study.

Sequencing Analysis

Analysis of all junction sequences except the *DHFR* scar junction was performed with CRISPResso2 (55) in batch mode utilizing paired-end reads with the following command:

CRISPRessoBatch—batch_settings <batch_settings_file > -q 30 -n <output_name>

The *DHFR* scar junction was analyzed using CRISPResso2 using only read 2 because of the large size of the amplicon (read 1 did not overlap the cut site and was therefore not informative).

For each junction, the expected product of perfect direct ligation, in which the Cas9 cut sites of each guide are directly ligated without additional loss or insertion of a sequence, was used as the reference sequence, and a "pseudoguide" sequence was used in which the cut site was aligned to the junction (Fig. 3A). Downstream analysis and visualizations were based on the 40-bp trimmed and merged allele frequency table output by CRISPResso2. For large deletions, in which the deletion extended past the 40-bp trimmed sequence, the full sequence was used to extract the bases upstream and downstream of the deletion, for use in microhomology calculations (described in the following paragraph). For each sample, the top 2,000 repair outcomes were considered, which typically encompassed >99% of the repair outcomes by frequency.

The length of microhomology was quantified for each junction sequence as follows: The last n nucleotides of the upstream sequence were compared with the last n nucleotides of the deleted sequence. The largest n for which there was an exact sequence match was defined as the upstream microhomology *mh1*. This process was repeated to compare the first n nucleotides of the deletion with the first n nucleotides of the downstream process to calculate the downstream microhomology *mh2*. The microhomology length was identified as the greater of *mh1* versus *mh2*. Junctions with microhomology length ≥1 were defined as exhibiting MH, whereas deletions with no microhomology and insertions were defined as MH-less (23, 24). A small percentage (<0.2%) of repair outcomes contained two or more closely spaced deletions (e.g., CCTGCTCTC—CAG—CTGGGGA). These sequences were excluded from the MH calculations.

Hi-C

Hi-C was performed by following the reported HiChIP protocol and omitting the protein immunoprecipitation step (56). Briefly, 10^6 cells were cross-linked with 1% formaldehyde for 10 minutes. Formaldehyde was quenched by the addition of glycine. Nuclei were isolated in lysis buffer (1% Triton X-100, 0.1% SDS, 150 mmol/L NaCl, 1 mmol/L EDTA, and 20 mmol/L Tris, pH 8.0) by rotating the cells for 1 hour at 4°C. The nuclear pellet was resuspended in 0.5% SDS solution and incubated at 62°C for 10 minutes. SDS was quenched by the addition of 10% Triton-X. Chromatin was digested with 8 U of MboI enzyme overnight at 37°C. Biotin fill-in was performed for 1 hour at 37°C, followed by proximity ligation for 6 hours at room temperature. After proximity ligation, nuclei were subjected to proteinase K digestion and decrosslinking overnight at 68°C. DNA was column-purified (MinElute, Qiagen) and sonicated to a size distribution of

200 to 500 bp. DNA was subjected to streptavidin bead binding to pull down biotinylated DNA that represents ligation junctions. Tn5 was used to create libraries by on-bead tagmentation with sequencing adapters, followed by PCR. Libraries were quantified using Qubit, and size distribution was assessed using a bioanalyzer. Libraries were sequenced on NovaSeq 6000 with paired-end 150 bp configuration.

A K562 Hi-C dataset (57) was downloaded from the ENCODE consortium's portal, https://www.encodeproject.org/, via the accession code ENCFF616PUW.

Hi-C Analysis

Paired-end reads were aligned to the reference genome (hg38) using the HiC-Pro pipeline (version 2.11.0). Default settings were used to remove duplicate reads, assign reads to MboI restriction fragments, filter for valid interactions, and generate binned interaction matrices. The Juicer pipeline's HiCCUPS tool was used to identify loops. Filtered read pairs from the HiC-Pro pipeline were converted into .hic format files and input into HiCCUPS using default settings. Contact matrices were visualized using plotgardener (58) using KR normalization for Hap1 data, and VC normalization K562 data.

ecDNA Analysis in TCGA Samples

We collected focal amplifications calls from AmpliconRepository.org (https://ampliconrepository.org/project/655bddb5bba7c92509525039) for TCGA samples analyzed using AmpliconSuite-pipeline. We re-applied the latest version of AmpliconClassifier (AC, v1.0.0) to these samples to identify ecDNA-related structural variants (SV). We utilized the AmpliconClassifier annotations about the ecDNA genome background, to identify two classes of ecDNA:

Simple ecDNA (aka simple circular): derived from an excisional model, unenriched for chromothripsis.
Complex ecDNA: those from a background with heavy genomic rearrangements and multiple copy-number oscillations, suggestive of a chromothriptic origin (bioRxiv 2024.05.06.592768).

To summarize the classification of ecDNA into these states, the heavily rearranged background ecDNA contained multiple copy-number states in the ecDNA region (>6), as well as more than four SVs in which their break ends crossed each other in the genome (termed "cross-edges"). The number of changes in copy-number state between connected genome segments was also required to be at least four times as many as the number of distinct states in the ecDNA (many oscillations). By contrast, ecDNA from a simple excisional model required the existence of an SV which joined the lowest and highest genome coordinates in a head-to-tail fashion. The ecDNA also needed to show three or fewer cross-edges, up to six distinct copy-number states as well as a low fraction of copy-number transitions to copy-number states (ratio of less than 4:1).

Of the 761 ecDNAs discovered in the TCGA dataset, 417 were reported as "simple circular" ecDNAs. This category is likely the most enriched for ecDNA not derived from chromothripsis and for which we could identify the initial head-to-tail junction (the "founder" SV). Of the 417 simple circular ecDNAs, 330 had at least one SV with a duplication-like orientation (head-to-tail on the same chromosome), having at least one endpoint of the SV within 150 bp of the ecDNA intervals and which had a distance between SV endpoints that was at least 50 kbp long as the length of the ecDNA intervals (spanning or nearly spanning the whole ecDNA). If multiple SVs were found that matched this, the largest valid SV was used to represent the founder SV.

In short, AmpliconSuite-pipeline identifies microhomology by examining reads having split-alignments covering the two SV break ends and then compares the coordinates of the break ends with the reference genome to identify sequence homology in each direction. For simple circular ecDNA (those unenriched for chromothripsis), we extracted the SV microhomology of the "founder SVs" from AC's SV summary files and counted the frequencies of each length. For ecDNA from a complex background, we utilized all SVs in the ecDNA region when extracting microhomology from the AC's SV summary files.

Statistics and Reproducibility

Unpaired two-sided *t* tests were performed using the SciPy stats package v1.10.1 and GraphPad Prism v9.5.1. Schematics in Figs. 1A and B; Fig. 2A and B and Supplementary Fig. S11 were created with BioRender. The Fisher exact test was performed for analysis of Fig. 3E using GraphPad Prism v9.5.1.

Data Availability

Sequencing data were deposited at Gene Expression Omnibus under accession number GSE273081. All other data that support the findings of this study are available from the corresponding authors upon request.

Authors' Disclosures

J.A. Belk reports grants from Howard Hughes Medical Institute outside the submitted work. J. Luebeck reports a patent for US20220364182A1 pending. K.E. Yost reports personal fees from Cartography Biosciences outside the submitted work. K.L. Hung reports a patent for Methods for Targeted Purification and Profiling of Human Extrachromosomal DNA pending and a patent for DNA Element Responsive to Extrachromosomal DNA in Cancer Cells pending. C.J. Kuo reports personal fees from Surrozen, Inc., Mozart Therapeutics, and NextVivo, Inc. outside the submitted work. V. Bafna reports grants from Cancer Research UK, NCI, and National Institute of General Medical Sciences during the conduct of the study, as well as holding stock in Boundless Bio, Inc. and equity in Abterra, Inc. and serving on the scientific advisory board of both companies. P.S. Mischel reports being a cofounder of Boundless Bio, Inc., as well as holding equity in the company and serving on the scientific advisory board, for which he receives compensation. H.Y. Chang reports grants from NCI, Cancer Research UK, and Howard Hughes Medical Institute during the conduct of the study, as well as personal fees from Accent Therapeutics, Boundless Bio, Cartography Bio, Orbital Therapeutics, 10x Genomics, Arsenal Bio, Chroma Medicine, and Spring Discovery outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

J.C. Rose: Conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing–original draft, writing–review and editing. **J.A. Belk**: Data curation, software, formal analysis, writing–review and editing. **I.T-L. Wong:** Investigation. **J. Luebeck:** Formal analysis, software, writing–review and editing. **H.T. Horn:** Investigation. **B. Daniel:** Investigation. **M.G. Jones:** Methodology. **K.E. Yost:** Methodology, resources. **K.L. Hung:** Methodology. **K.S. Kolahi:** Resources. **E.J. Curtis:** Resources. **C.J. Kuo:** Resources, supervision. **V. Bafna:** Resources, supervision. **P.S. Mischel:** Funding acquisition, supervision, writing–review and editing. **H.Y. Chang:** Funding acquisition, supervision, writing–review and editing.

Acknowledgments

We thank members of the Chang and Mischel labs for discussion. This project was supported by Cancer Grand Challenges CGCSDF-2021\100007 with support from Cancer Research UK and

the NCI (H.Y. Chang and P.S. Mischel). H.Y. Chang is an investigator of the Howard Hughes Medical Institute. J.C. Rose is supported by NIH K99-CA279512 and the A.P. Giannini Foundation. J.A. Belk is supported by a Hanna Gray Fellowship from the Howard Hughes Medical Institute.

Note

Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

Received September 28, 2023; revised May 16, 2024; accepted August 5, 2024; published first August 7, 2024.

REFERENCES

1. Turner KM, Deshpande V, Beyter D, Koga T, Rusert J, Lee C, et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature 2017;543:122–5.
2. Wu S, Turner KM, Nguyen N, Raviram R, Erb M, Santini J, et al. Circular ecDNA promotes accessible chromatin and high oncogene expression. Nature 2019;575:699–703.
3. Nathanson DA, Gini B, Mottahedeh J, Visnyei K, Koga T, Gomez G, et al. Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. Science 2014;343:72–6.
4. Kim H, Nguyen N-P, Turner K, Wu S, Gujar AD, Luebeck J, et al. Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. Nat Genet 2020;52:891–7.
5. Hung KL, Luebeck J, Dehkordi SR, Colón CI, Li R, Wong IT-L, et al. Targeted profiling of human extrachromosomal DNA by CRISPR-CATCH. Nat Genet 2022;54:1746–54.
6. Lange JT, Rose JC, Chen CY, Pichugin Y, Xie L, Tang J, et al. The evolutionary dynamics of extrachromosomal DNA in human cancers. Nat Genet 2022;54:1527–33.
7. Bafna V, Mischel PS. Extrachromosomal DNA in cancer. Annu Rev Genomics Hum Genet 2022;23:29–52.
8. Alt FW, Kellems RE, Bertino JR, Schimke RT. Selective multiplication of dihydrofolate reductase genes in methotrexate-resistant variants of cultured murine cells. J Biol Chem 1978;253:1357–70.
9. Shoshani O, Brunner SF, Yaeger R, Ly P, Nechemia-Arbely Y, Kim DH, et al. Chromothripsis drives the evolution of gene amplification in cancer. Nature 2021;591:137–41.
10. Song K, Minami JK, Huang A, Dehkordi SR, Lomeli SH, Luebeck J, et al. Plasticity of extrachromosomal and intrachromosomal BRAF amplifications in overcoming targeted therapy dosage challenges. Cancer Discov 2021;12:1046–69.
11. Roukos V, Voss TC, Schmidt CK, Lee S, Wangsa D, Misteli T. Spatial dynamics of chromosome translocations in living cells. Science 2013;341:660–4.
12. Møller HD, Lin L, Xiang X, Petersen TS, Huang J, Yang L, et al. CRISPR-C: circularization of genes and chromosome by CRISPR in human cells. Nucleic Acids Res 2018;46:e131.
13. Zhang Y, McCord RP, Ho Y-J, Lajoie BR, Hildebrand DG, Simon AC, et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell 2012;148:908–21.
14. Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho Y-J, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell 2011;147:107–19.
15. Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell 2011;147:95–106.
16. Rocha PP, Micsinai M, Kim JR, Hewitt SL, Souza PP, Trimarchi T, et al. Close proximity to Igh is a contributing factor to AID-mediated translocations. Mol Cell 2012;47:873–85.
17. Storlazzi CT, Fioretos T, Surace C, Lonoce A, Mastrorilli A, Strömbeck B, et al. 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 2006;15:933–42.
18. Shih I-M, Wang Y, Wang T-L. The origin of ovarian cancer species and precancerous landscape. Am J Pathol 2021;191:26–39.
19. Stinson BM, Loparo JJ. Repair of DNA double-strand breaks by the nonhomologous end joining pathway. Annu Rev Biochem 2021;90:137–64.
20. Roukos V, Misteli T. The biogenesis of chromosome translocations. Nat Cell Biol 2014;16:293–300.
21. Vogt N, Lefèvre S-H, Apiou F, Dutrillaux A-M, Cör A, Leuraud P, et al. Molecular structure of double-minute chromosomes bearing amplified copies of the epidermal growth factor receptor gene in gliomas. Proc Natl Acad Sci U S A 2004;101:11368–73.
22. Gibaud A, Vogt N, Hadj-Hamou N-S, Meyniel J-P, Hupé P, Debatisse M, et al. Extrachromosomal amplification mechanisms in a glioma with amplified sequences from multiple chromosome loci. Hum Mol Genet 2010;19:1276–85.
23. Shen MW, Arbab M, Hsu JY, Worstell D, Culbertson SJ, Krabbe O, et al. Predictable and precise template-free CRISPR editing of pathogenic variants. Nature 2018;563:646–51.
24. Chen W, McKenna A, Schreiber J, Haeussler M, Yin Y, Agarwal V, et al. Massively parallel profiling and predictive modeling of the outcomes of CRISPR/Cas9-mediated double-strand break repair. Nucleic Acids Res 2019;47:7989–8003.
25. Hussmann JA, Ling J, Ravisankar P, Yan J, Cirincione A, Xu A, et al. Mapping the genetic landscape of DNA double-strand break repair. Cell 2021;184:5653–69.e25.
26. Blackford AN, Jackson SP. ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. Mol Cell 2017;66:801–17.
27. Dobbs TA, Tainer JA, Lees-Miller SP. A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. DNA Repair (Amst) 2010;9:1307–14.
28. Uematsu N, Weterings E, Yano K, Morotomi-Yano K, Jakob B, Taucher-Scholz G, et al. Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. J Cell Biol 2007;177:219–29.
29. Liu Y, Zou RS, He S, Nihongaki Y, Li X, Razavi S, et al. Very fast CRISPR on demand. Science 2020;368:1265–9.
30. Callén E, Jankovic M, Wong N, Zha S, Chen H-T, Difilippantonio S, et al. Essential role for DNA-PKcs in DNA double strand break repair and apoptosis in ATM deficient lymphocytes. Mol Cell 2009;34:285–97.
31. Arnould C, Rocher V, Saur F, Bader AS, Muzzopappa F, Collins S, et al. Chromatin compartmentalization regulates the response to DNA damage. Nature 2023;623:183–92.
32. Bhargava R, Carson CR, Lee G, Stark JM. Contribution of canonical nonhomologous end joining to chromosomal rearrangements is enhanced by ATM kinase deficiency. Proc Natl Acad Sci U S A 2017;114:728–33.
33. Schep R, Brinkman EK, Leemans C, Vergara X, van der Weide RH, Morris B, et al. Impact of chromatin context on Cas9-induced DNA double-strand break repair pathway balance. Mol Cell 2021;81:2216–30.e10.
34. Chang HHY, Watanabe G, Gerodimos CA, Ochi T, Blundell TL, Jackson SP, et al. Different DNA end configurations dictate which NHEJ components are most important for joining efficiency. J Biol Chem 2016;291:24377–89.
35. Mari P-O, Florea BI, Persengiev SP, Verkaik NS, Brüggenwirth HT, Modesti M, et al. Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4. Proc Natl Acad Sci U S A 2006;103:18597–602.
36. Bermudez-Cabrera HC, Culbertson S, Barkal S, Holmes B, Shen MW, Zhang S, et al. Small molecule inhibition of ATM kinase increases CRISPR-Cas9 1-bp insertion frequency. Nat Commun 2021;12:5111.
37. Bhargava R, Lopezcolorado FW, Tsai LJ, Stark JM. The canonical non-homologous end joining factor XLF promotes chromosomal deletion rearrangements in human cells. J Biol Chem 2020;295:125–37.
38. Stephens PJ, McBride DJ, Lin M-L, Varela I, Pleasance ED, Simpson JT, et al. Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature 2009;462:1005–10.

39. Windle B, Draper BW, Yin YX, O'Gorman S, Wahl GM. A central role for chromosome breakage in gene amplification, deletion formation, and amplicon integration. Genes Dev 1991;5:160–74.
40. Storlazzi CT, Lonoce A, Guastadisegni MC, Trombetta D, D'Addabbo P, Daniele G, et al. Gene amplification as double minutes or homogeneously staining regions in solid tumors: origin and structure. Genome Res 2010;20:1198–206.
41. Ly P, Teitz LS, Kim DH, Shoshani O, Skaletsky H, Fachinetti D, et al. Selective Y centromere inactivation triggers chromosome shattering in micronuclei and repair by non-homologous end joining. Nat Cell Biol 2017;19:68–75.
42. Ghezraoui H, Piganeau M, Renouf B, Renaud J-B, Sallmyr A, Ruis B, et al. Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. Mol Cell 2014;55:829–42.
43. Wang J, Sadeghi CA, Frock RL. DNA-PKcs suppresses illegitimate chromosome rearrangements. Nucleic Acids Res 2024;52:5048–66.
44. Dharanipragada P, Zhang X, Liu S, Lomeli SH, Hong A, Wang Y, et al. Blocking genomic instability prevents acquired resistance to MAPK inhibitor therapy in melanoma. Cancer Discov 2023;13:880–909.
45. Cejka P, Symington LS. DNA end resection: mechanism and control. Annu Rev Genet 2021;55:285–307.
46. Becker A, Durante M, Taucher-Scholz G, Jakob B. ATM alters the otherwise robust chromatin mobility at sites of DNA double-strand breaks (DSBs) in human cells. PLoS One 2014;9:e92640.
47. Caron P, Choudjaye J, Clouaire T, Bugler B, Daburon V, Aguirrebengoa M, et al. Non-redundant functions of ATM and DNA-PKcs in response to DNA double-strand breaks. Cell Rep 2015;13:1598–609.
48. Zatreanu D, Robinson HMR, Alkhatib O, Boursier M, Finch H, Geo L, et al. Polθ inhibitors elicit BRCA-gene synthetic lethality and target PARP inhibitor resistance. Nat Commun 2021;12:3636.
49. Kessler M, Hoffmann K, Brinkmann V, Thieck O, Jackisch S, Toelle B, et al. The Notch and Wnt pathways regulate stemness and differentiation in human fallopian tube organoids. Nat Commun 2015;6:8989.
50. Fok JHL, Ramos-Montoya A, Vazquez-Chantada M, Wijnhoven PWG, Follia V, James N, et al. AZD7648 is a potent and selective DNA-PK inhibitor that enhances radiation, chemotherapy and olaparib activity. Nat Commun 2019;10:5065.
51. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 2015;33:543–8.
52. Fokas E, Prevo R, Pollard JR, Reaper PM, Charlton PA, Cornelissen B, et al. Targeting ATR in vivo using the novel inhibitor VE-822 results in selective sensitization of pancreatic tumors to radiation. Cell Death Dis 2012;3:e441.
53. Hickson I, Zhao Y, Richardson CJ, Green SJ, Martin NMB, Orr AI, et al. Identification and characterization of a novel and specific inhibitor of the Ataxia-telangiectasia mutated kinase ATM. Cancer Res 2004;64:9152–9.
54. Pantelidou C, Sonzogni O, De Oliveria Taveira M, Mehta AK, Kothari A, Wang D, et al. PARP inhibitor efficacy depends on $CD8^+$ T-cell recruitment via intratumoral STING pathway activation in BRCA-deficient models of triple-negative breast cancer. Cancer Discov 2019;9:722–37.
55. Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol 2019;37:224–6.
56. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods 2016;13:919–22.
57. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 2014;159:1665–80.
58. Kramer NE, Davis ES, Wenger CD, Deoudes EM, Parker SM, Love MI, et al. Plotgardener: cultivating precise multi-panel figures in R. Bioinformatics 2022;38:2042–5.