

BRCA1-A and LIG4 complexes mediate ecDNA biogenesis and cancer drug resistance

Oliver W. Chung,^{1,5} Shun Yao,^{1,5} Fu Yang,^{1,2,5} Ling Wang,¹ Christian Cerda-Smith,¹ Haley M. Hutchinson,¹ Kris C. Wood,¹ Weijia Su,¹ Mustafa Khasraw,³ Lee Zou,¹ Dale A. Ramsden,⁴ and ZZ Zhao Zhang^{1,6,*}

¹Department of Pharmacology & Cancer Biology, Duke University School of Medicine, Durham, USA.

²Present address: Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, TaiKang Center for Life and Medical Sciences, Wuhan University, Wuhan, China.

³The Preston Robert Tisch Brain Tumor Center at Duke, Department of Neurosurgery, Duke University Medical Center, Durham, USA.

⁴Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, USA.

⁵These authors contributed equally

⁶Lead contact

*Correspondence: z.z@duke.edu

SUMMARY

Extrachromosomal circular DNA (ecDNA) are commonly produced within the nucleus to drive genome dynamics and heterogeneity, enabling cancer cell evolution and adaptation. However, the mechanisms underlying ecDNA biogenesis remain poorly understood. Here using genome-wide CRISPR screening in human cells, we identified the BRCA1-A and the LIG4 complexes mediate ecDNA production. Following DNA fragmentation, the upstream BRCA1-A complex protects DNA ends from excessive resection, promoting end-joining for circularization. Conversely, the MRN complex, which mediates end resection and thus antagonizes the BRCA1-A complex, suppresses ecDNA formation. Downstream, LIG4 conservatively catalyzes ecDNA production in *Drosophila* and mammals, with patient tumor ecDNA harboring junctions marked by LIG4 activity. Notably, disrupting LIG4 or BRCA1-A in cancer cells impairs ecDNA-mediated adaptation, hindering resistance to both chemotherapy and targeted therapies. Together, our study reveals the roles of the LIG4 and BRCA1-A complexes in ecDNA biogenesis, and uncovers new therapeutic targets to block ecDNA-mediated adaptation for cancer treatment.

INTRODUCTION

Although originally reported six decades ago, extrachromosomal circular DNA (henceforth, abbreviated as “ecDNA” for all circles regardless of size) have recently emerged as a common DNA form broadly produced in eukaryotes¹⁻⁴. With the size ranging from hundreds to millions of bases, ecDNA often harbor genes or genetic regulatory elements³⁻¹⁰. Together with the feature of unrestricted copy numbers—often up to hundreds of copies per cell, ecDNA bring one important layer of genome dynamics that allows cells to fully exploit their genetic information for adaptation and evolution^{11,12}. Reflecting this fundamental function, ecDNA have been documented to play pivotal roles in diverse processes: magnifying histone genes in yeast to achieve dose compensation¹³, amplifying rDNA to drive frog egg development^{1,14}, and re-writing the cattle genome for coat color sidedness patterning¹².

Notably, forming ecDNA is a common mechanism for the amplification of oncogenes and genes that render cancer cells drug resistant^{3,4,15}. Strikingly, nearly one-in-three cancer patients harbor ecDNA in their cancer cells^{3,7,8,11,16-19}. This percentage increases even further in certain aggressive cancer types: to 60% in glioblastoma and ~40-50% in sarcoma and esophageal cancers^{8,19}. Given their high copy numbers and decompacted chromatin structure, ecDNA can massively amplify and express oncogenes residing on them^{3,7,8,11,16-18}. Meanwhile, ecDNA have shown strong structural and spatial dynamics, in which they can hop on and off the chromosomes during cancer evolution^{16,20}. As such, charactering the molecular mechanisms underlying ecDNA biology and correspondingly drugging these processes could provide new therapeutic strategies for cancer treatment.

Despite their essential functions, our understanding of the ecDNA formation process is limited. Our previous work elucidated the mechanism by which the

retrotransposons residing in the host genome form ecDNA during their life cycle²¹. These DNA circles represent replication products and typically only harbor retrotransposon sequences²¹. Given that retrotransposons are often silenced in the host genome²²⁻²⁴, retrotransposon-derived ecDNA are not frequently produced. A more common process of ecDNA biogenesis is directly circularizing one or more DNA fragments generated from chromosomal breaks^{3,4,15}. This process appears to be more prevalent in cancer cells, especially when these cells undergo genome rupture, such as chromothripsis^{16,17}. However, the molecular mechanisms that drive this process remain largely unexplored.

In this study we employed the genome-wide CRISPR screening to identify factors mediating ecDNA biogenesis. We found ecDNA formation is orchestrated by BRCA1-A and LIG4 complexes. BRCA1-A complex acts upstream by protecting DNA free ends from resection and then hands off these DNA molecules to LIG4 for end-to-end joining, resulting in ecDNA formation. Importantly, mutations in LIG4 or components of the BRCA1-A complex abrogate ecDNA-mediated adaptation to cancer therapeutic drugs. Our findings provide a crucial step in decoding the mechanism of ecDNA formation and offer promising drug targets for addressing therapeutic resistance in cancer.

RESULTS

CRISPR screens identify ecDNA biogenesis regulators

Current tools for studying and observing ecDNA are limited. To address this limitation, the CRISPR-C approach, using CRISPR/Cas9 cleavage to generate DNA fragments with the promoter located downstream of eGFP at the opposite end, was invented²⁵. Under this design, circularization of the fragments brings the promoter upstream of eGFP, thus giving rise to fluorescence that reflects ecDNA formation (Figure 1A)²⁵. Given that eGFP expression can be used to reflect ecDNA biogenesis, we sought to perform genome-wide

CRISPR screening to identify factors regulating this process. For the screen, we used HEK293T cells harboring the eGFP reporter to mutate one gene/cell for 18,761 genes, by transducing the genome-wide MinLib CRISPR lentiviral library²⁶. Upon inducing ecDNA production and eGFP expression, eGFP-positive and eGFP-negative cells were sorted to achieve 1,000-fold coverage of the sgRNA library. Under this design, sgRNAs targeting genes required for ecDNA biogenesis were enriched in the eGFP-negative population and excluded from the eGFP-positive population (Figure 1B). Conversely, sgRNAs targeting genes encoding factors that suppress ecDNA formation would show the reverse pattern (Figure 1B).

Given that ecDNA biogenesis is coupled with the DNA breakage, factors from DNA repair pathways were the most significantly enriched screen hits (Figure 1C and S1A). Meanwhile, we also detected factors from other processes as potential regulators for ecDNA biogenesis (Figure S1). For example, histone variant and factors related with rRNA processing or ribosome biogenesis may regulate ecDNA production under our screen design (Figure 1D and S1). While these findings point to directions for future investigations, here we primarily focus on the function of DNA repair factors in ecDNA biogenesis.

Among our top hits for driving ecDNA production are LIG4 and its physical partner XRCC4, and DNA-PKcs (encoded by the *PRKDC* gene) (Figure 1C-1E). Notably, DNA-PKcs has previously been implicated in mediating ecDNA production when cancer cells undergo chromothripsis^{16,27,28}, giving credence to our screening approach. Despite the LIG4/XRCC4 complex and DNA-PKcs being classified into the Non-homologous end joining (NHEJ) DNA repair process^{29,30}, the other factors from this pathway were not uncovered as significant hits from our screen (Figure 1E, see discussion). Meanwhile, although previous research has suggested the potential function of alternative end

joining (alt-EJ) in circular DNA formation under specific settings^{28,31,32}, our screening results showed that all factors from this DNA repair process are dispensable for ecDNA biogenesis from the designed reporter (Figure 1E).

Meanwhile, the other top hits for driving ecDNA biogenesis are factors can form a complex: BRCC36 encoded by gene *BRCC3*, BRCC45/BRE encoded by *BABAM2*, ABRA1/ABRAXAS encoded by *ABRAXAS1*, and RAP80 encoded by *UIMC1* (Figure 1C-1E). Together with MERIT40, these four factors physically interact with each other to build a complex that can sequester BRCA1 protein³³. This complex thus is named as BRCA1-A complex. It is worth noting that BRCA1 itself was not screened as a regulator, suggesting that its canonical function in promoting homologous recombination—executed by forming BRCA1-C complex—is unessential for ecDNA formation. Screening four different components from the same complex as strong hits indicates its pivotal role in mediating the circularization process. The BRCA1-A complex has been characterized for its function in preventing end-resection from the DNA break sites³⁴⁻³⁷. This suggests that preventing end-resection upon DNA segmentation promotes ecDNA production. Satisfyingly, our screen identified factors that drive end-resection as suppressors of ecDNA biogenesis (Figure 1C-1E and S1B). These factors are MRE11, NBN1, and RAD50 (Figure 1C-1E and S1B), which form the MRN complex that resects the DNA break ends and recruits ATM³⁸. Consistently, ATM is screened as a negative regulator of ecDNA production (Figure 1C and S1B). In summary, our screen results suggest that upon DNA fragmentation, BRCA1-A complex prevents end resection to drive the ecDNA production process that is potentially catalyzed by LIG4.

Validating BRCA1-A, MRN, and LIG4 complexes

To validate the screen hits that serve as drivers for ecDNA production, we used CRISPR/Cas9 on the eGFP reporter cells to generate mutant cell lines for the following seven genes: *LIG4*, *XRCC4*, *PRKDC*, *UIMC1*, *BRCC3*, *BABAM2*, and *ABRAXAS1* (Figure S2A and S2B). For each factor, we employed at least two different sgRNAs to generate distinct mutant clones (Figure S2A and S2B), minimizing the possibility of obtaining false-positive findings from the off-targeting mutation by an individual sgRNA. Using eGFP expression as the proxy for ecDNA production from the reporter, we obtained findings consistent with our screen results. Mutating any of the component from either *LIG4* or *BRCA1-A* complex showed no eGFP expression (Figure S3), suggesting that the cells without either complex lost the capability of ecDNA formation. Besides using eGFP to indicate ecDNA production, we also designed qPCR or digital droplet PCR (ddPCR) assay to precisely quantify the number of ecDNA produced from the reporter (Figure 2A-2C). Consistent with eGFP expression, mutating *LIG4*, *XRCC4*, *PRKDC*, *BRCC3*, *BABAM2*, *ABRAXAS1*, *UIMC1* abrogated ecDNA production, as measured by either qPCR or ddPCR assay (Figure 2A-2C).

To finally validate the screen outcome, we performed the rescue experiment by re-introducing either wild-type gene of *LIG4*, *BRCC3*, *BABAM2*, *ABRAXAS1*, and *UIMC1* via a plasmid vector into corresponding mutant cells (Figure 2D, 2E, S2C, and S2D). For *LIG4*, we also generated catalytically-dead *LIG4* transgene for the rescue experiment to examine its ligase function in ecDNA formation (Figure S2D). Meanwhile, since *XRCC4* stabilizes *LIG4* by directly interacting with its two BRCT domains^{39,40}, we additionally included a mutant version of *LIG4* by removing the BRCT domains—termed as *LIG4-ΔBRCT*—for the rescue experiment (Figure S2D). As expected, re-introducing wild-type *LIG4*, *BRCC3*, *BABAM2*, *ABRAXAS1*, and *UIMC1* restored ecDNA production in the mutant cells, as evidenced by eGFP expression from the reporter (Figure 2D, 2E, S2C, and

S2D). However, either mutating LIG4 ligase activity or suppressing its interaction with XRCC4 abolished the rescue outcome (Figure 2E and S2D). These findings indicate a pivotal role of LIG4 and BRCA1-A complex in circularizing the reporter.

On the other end, we sought to validate the MRN complex as a suppressor of ecDNA production via two independent approaches. First, we generated two *RAD50* mutant cell lines using two distinct sgRNAs (Figure S4A). Upon mutating *RAD50*, the ecDNA-producing cells significantly increased from 21.2% to 37.4% (P -value = 0.030, sgRNA-1) and 36% (P -value = 0.046, sgRNA-2, Figure S4A and S4B). Next, we further validated the antagonizing function of the MRN complex by using a small molecule drug, Mirin, which serves as an MRN inhibitor to block the end-resection function of MRE11^{41,42}. Applying this inhibitor significantly increased the eGFP-positive population by 2.7-fold (P -value = 0.004, Figure S4C). Altogether, our results robustly validated the screen findings and support a model that suppressing end-resection by BRCA1-A complex promotes LIG4-mediated ecDNA biogenesis.

BRCA1-A and LIG4 drive ecDNA formation in cancer cell lines

All data described above were collected based on the CRISPR-C reporter, which relies on CRISPR/Cas9 to generate a DNA segment for ecDNA formation. Although unlikely, it is possible that LIG4 or BRCA1-A complex is essential for DNA fragmentation by Cas9, thus leading to no ecDNA production upon its mutation. To exclude this possibility, we adapted the original CRISPR-C system to directly introduce linear DNA molecules that mimic the fragments generated by CRISPR/Cas9 cleavage (Figure 3A). Under this design, the sequences of linear DNA can be easily modified. Moreover, another advantage is that these linear molecules can be introduced into any cells without a pre-integration event, allowing us to monitor ecDNA production in different cell types. As

such, we named our design as “Versatile biosensor”. As a proof-of-principal, we initially designed the biosensor with just a promoter and eGFP (Figure S5A). Transfecting these DNA molecules into HEK293T cells gave eGFP positive cells, suggesting the formation of ecDNA (Figure S5B).

To validate that eGFP expression from our versatile biosensor reflects ecDNA production, and is not due to the introduced eGFP sequences integrating into the host cell genome with a local promoter driving eGFP expression, we designed two other control constructs (Figure S5B). One is just the eGFP sequence without a promoter (Figure S5B). The other is a circle harboring the designed biosensor (Figure S5B), thus preventing the joining of the eGFP and promoter. Introducing either construct into cultured cells could not give eGFP expression (Figure S5B), suggesting eGFP signals from our biosensor reflect circle formation. Meanwhile, given that cells can take up multiple biosensor molecules, their concatenation can form linear DNA and bring a promoter upstream of eGFP. To test this possibility, we treated DNA from the eGFP-positive cells with exonuclease, which removes linear DNA but leaves ecDNA intact. By using a pair of primers that spans the junction site of eGFP and promoter (Figure S5A), we observed that the PCR amplicon signals were preserved upon exonuclease treatment (Figure S5C), indicating that our biosensor mainly forms ecDNA.

Since eGFP expression reflects the circularization process from our versatile biosensor, we sought to test the function of LIG4 and BRCA1-A complex for ecDNA biogenesis in different cancer lines when Cas9-mediated DNA fragmentation is bypassed. Given that these cancer lines have different efficiency on uptake our engineered biosensor, we introduced a DsRed cassette that is directly driven by its own promoter into the middle of the biosensor (Figure 3A). This allows us to measure the percentage of cells harboring biosensors upon delivery. With this design, we next tested

the function of LIG4 and BRCA1-A complex in three cancer cell lines: cervical cancer HeLa cells, non-small cell lung cancer PC9 cells, and colon cancer HCT116 cells. In all cell lines, mutating *LIG4* or the BRCA1-A complex key component *UIMC1* led to the abolishment of ecDNA production (Figure 3B, 3C, S2E, and S2F), suggesting a generic function of LIG4 and BRCA1-A complex in mediating ecDNA biogenesis across different cell types.

Lastly, while our findings suggest LIG4 and BRCA1-A complex as the key factors driving ecDNA formation, these findings were based on the reporter system that produced DNA fragments with fixed DNA sequences at the two ends. Are LIG4 and BRCA1-A still essential if the DNA sequences at the two ends are altered? To address this question, we harnessed the versatility of our ecDNA biosensor design by including 6 random nucleotides at the two ends (Figure S5D). In all cells tested (HCT116, PC9, and HeLa cells), loss the function of BRCA1-A by mutating *UIMC1* or mutating *LIG4* led to abrogation of ecDNA formation from the biosensor with random end sequences (Figure S5E). Altogether, our findings suggest that when cells experience DNA fragmentation, LIG4 and BRCA1-A complex drives ecDNA biogenesis regardless of the nucleotide composition at the two ends.

LIG4 mediates natural ecDNA formation in vivo

Given that LIG4 is conserved in *Drosophila*, we sought to test the function of LIG4 in driving natural ecDNA formation in vivo. During *Drosophila* egg development, two chorion genomic clusters undergo massive amplification in the oocyte-surrounding follicle cells to immensely produce chorion protein for egg encapsulation⁴³. Previous studies showed that the uneven amplification of chorion clusters often leads to DNA breaks with variable length (Figure 4A)⁴⁴. We previously sequenced all circular DNA in

Drosophila ovaries to examine retrotransposon-derived ecDNA²¹. Mining our dataset, we found sequencing reads with head-to-tail junctions that support ecDNA generated from the two chorion clusters (Figure S6A). To investigate whether LIG4 is required for the formation of chorion ecDNA, we sequenced ovarian ecDNA from *LIG4* homozygotes (*Lig4*^{-/-}) and heterozygotes (*Lig4*^{+/-}). While the *LIG4* heterozygotes still produced ecDNA from both chorion regions, loss of LIG4 resulted in a nearly 100% abolishment of ecDNA production (Figure 4B and S6B). To further validate our findings, we designed a pair of primers that can only yield PCR products when ecDNA is generated (Figure 4A and 4C). Similarly, while ecDNA with variable length can be detected from *Lig4*^{+/-} ovaries, *Lig4*^{-/-} flies lost chorion-ecDNA production (Figure 4C).

While our findings suggest LIG4 directly drives chorion ecDNA production in vivo, we are cognizant of the possibility that LIG4 is essential for chorion region amplification, thus its depletion would abolish ecDNA biogenesis. To test this possibility, we sequenced the ovarian genomic DNA from both *Lig4*^{-/-} and *Lig4*^{+/-} flies. Regardless of the genotype, the two chorion clusters amplified to the same magnitude (Figure 4B and S6B), suggesting that LIG4 is dispensable for the genomic amplification of these two regions. The amplification process of these two regions can incorporate a large amount of EdU into the replicated DNA and be visualized by EdU staining. Consistently, we observed comparable EdU foci in the follicle cells from both *Lig4*^{-/-} and *Lig4*^{+/-} flies (Figure 4D), further indicating that mutating *LIG4* does not alter chorion cluster amplification but does abolish ecDNA production. In summary, our data suggest that LIG4 plays a conserved function in driving natural ecDNA production during *Drosophila* egg development. Given that *Drosophila* does not possess BRCA1-A complex, we were unable to test its function under this oogenesis setting.

Glioblastoma ecDNA harbor end-end junctions with a LIG4 signature

ecDNA are frequently produced in cancer cells and have been extensively shown to contribute to tumor heterogeneity and cancer cell adaptation³⁻¹⁰. Since 0-3 bp microhomology is the strong signature of LIG4-mediated ligation process^{45,46}, this can be used to test whether LIG4 likely mediates ecDNA formation in patient tumors by examining junction site sequences. Previous research assessed the possible contribution of distinct DNA repair pathways in mediating ecDNA formation by using datasets from The Cancer Genome Atlas, which often lack paired germline genome from the same patients to serve as controls⁸. Given glioblastoma shows the highest frequency for ecDNA production⁸, we sequenced tumor DNA from 36 patients with 100X genome coverage to reconstruct ecDNA and to identify the junction sites (Figure 5A). Meanwhile, we sequenced blood DNA from the same patients to build the personalized germline genome (Figure 5A), which serves as the reference for ecDNA and mutation detection. Among these patients, 21 (58%) harbor cancer-related ecDNA within their tumors (Figure 5B). Similar as previously reported⁸, EGFR is the gene that most frequently amplified via ecDNA formation: EGFR ecDNA are abundantly produced in the tumors of 14 patients (39%, Figure 5B). From all ecDNA, we in total detected 190 junction sites for microhomology analysis (Figure 5C-5E). Analyzing the sequences of junction sites revealed a strong LIG4 signature: 43 sites without homology sequences from the two ends (supporting direct ligation events); 51 sites with 1 bp homology; 41 sites with 2 bp homology; and 26 sites with 3 bp homology (Figure 5D and 5E). In total, 85% of the junction sites are highly likely to be catalyzed by LIG4 (Figure 5E). Meanwhile, we also detected 15 sites with 4 bp homology; 9 sites with 5 bp; and 5 sites with homology sequences ≥ 6 bp (Figure 5E). These 15% events were potentially driven by either LIG4 or Pol θ , as both can mediate the repair process that leaves a footprint with > 3 bp homology

^{47,48}. In summary, our data support LIG4 to be the major driver for ecDNA formation within glioblastoma tumor cells.

BRCA1-A complex protects the DNA ends from resection

We next aimed to dissect the function of BRCA1-A complex in ecDNA formation. The aforementioned data demonstrated the function of BRCA1-A complex in circularizing DNA less than 10 kb. Given that the ecDNA driving cancer cell evolution often reach megabases, we next examined the function of BRCA1-A complex in large circle formation. Previous research established the CRISPR-C approach to generate megabase-sized ecDNA species¹¹. We employed similar approach to generate a 1.8 mega-base segment in PC9 cells harboring *EGFR* (Figure 6A), an oncogene frequently forming massive amount of ecDNA in cancer patients, such as glioblastoma^{8,49}. Wild-type PC9 cells generated 1.8 mega-bases *EGFR* ecDNA (Figure 6B). Treating these cells with Osimertinib, the 3rd-generation EGFR inhibitor that has been widely used in clinics, led to accumulation of a massive amount of *EGFR* ecDNA (Figure 6C and 6D), suggesting that these circles confer drug resistance. However, in *UIMC1* mutant cells with BRCA1-A function disrupted, the amount of ecDNA produced significantly dropped to 15% of the wild-type level (Figure 6B; $P < 0.0001$, measured by ddPCR). Correspondingly, loss of BRCA1-A complex led to complete cell-death upon Osimertinib exposure (Figure 6C).

Does BRCA1-A complex protect the two ends of DNA segments from resection, thus promoting ecDNA formation? To seek an answer, we sought to examine the sequences of the end-end junction site of ecDNA to search the “footprint” of the resection events—nucleotide deletion. For this purpose, we amplified the junction region and employed high-throughput sequencing to survey the junction sequences of *EGFR* ecDNA generated in PC9 cells from precisely defined DNA segment. In wild-type cells, direct

ligation of the two free ends is the dominant form: taking 25.2% of total events (Figure 6E). Meanwhile, we also observed ecDNA with deletions at variable lengths (Figure 6E), suggesting that end-resection is an active process even in wild-type cells.

Given the collapse of ecDNA biogenesis upon loss of BRCA1-A complex (Figure 6B), examining the end-end junction sites from trace amount of ecDNA produced under this condition is technically challenging. By using a 5-times more *UIMC1* mutant cells to perform the experiment, we managed to obtain enough ecDNA for sequencing-library preparation. As expected, the ecDNA species from this condition dropped to only 49% of wild-type level, indicating severe deficiency in ecDNA formation. From these ecDNA, the percentage of having direct ligation dropped from 25.2% in wild-type to 3.7% (Figure 6E and S6C). Meanwhile, we also observed dramatic increase of ecDNA with large deletions at the junction site upon loss of BRCA1-A complex (Figure 6E and S6C). The junction sites with > 50 bp deletions increased from 27.4% in wild-type to 42% upon loss of RAP80 by mutating *UIMC1* (Figure S6C). For the sites with > 100 bp deletions, the frequency increased from 20.6% to 42% (Figure S6C). This suggests the occurrence of excessive resection without BRCA1-A complex. In summary, our data suggest without the protection of BRCA1-A complex, DNA segments are more prone for resection at the ends and followed by degradation, as reflected by the abolishment of ecDNA production.

BRCA1-A and LIG4 drive ecDNA-mediated cancer cell adaptation

The CRISPR-C approach relies on CRISPR/Cas9 to generate pre-defined DNA fragments for ecDNA biogenesis. Without CRISPR/Cas9, is BRCA1-A complex still required for natural ecDNA biogenesis in cancer cells and enabling drug resistance? To answer this question, we tested the function of BRCA1-A complex in two systems: (1) methotrexate (MTX) induced spontaneous *DHFR* ecDNA production in HeLa cells; and (2) EGFR

inhibitor induced natural *EGFR* ecDNA biogenesis in PC9 cells for drug resistance (Figure 7A). Pioneering research in 1970s established the system that upon methotrexate treatment, HeLa cells rely on ecDNA—at that time termed “double minutes”—harboring the *DHFR* gene to evolve resistance for this chemotherapy drug⁵⁰⁻⁵³. We adapted this system and tested the function of BRCA1-A complex during this natural ecDNA-mediated evolution process (Figure 7B-7D). Wild-type cells evolved resistance to methotrexate by producing *DHFR* ecDNA, which enable the expression of *DHFR* increased from 53 fpkm in parental cells to 1748 fpkm in methotrexate resistant cells (Figure 7B, 7C, S7A, and S7B). Mutating *UIMC1* to disrupt BRCA1-A function led to the abolishment of resistance development (Figure 7D).

To examine whether BRCA1-A complex is broadly required for natural ecDNA production, we established PC9 lung cancer cells as a new system to study spontaneous *EGFR* ecDNA formation. After treating PC9 cells with EGFR inhibitor Osimertinib, we observed the production of natural *EGFR* ecDNA (Figure 7E), which drives a 12.9-fold increase of the expression of *EGFR* (33 fpkm in parental cells *vs* 424 fpkm in Osimertinib resistant cells, Figure 7F, S7C and S7D). Similar to HeLa cells, while wild-type cells evolved resistance to Osimertinib, mutating *UIMC1* resulted in the abrogation of adaptation (Figure 7G). Notably, we used similar approach to examine the function of *LIG4* and found that mutating it also led to the abolishment of ecDNA production and the incapability of cancer cells to evolve drug resistance in both HeLa-methotrexate and PC9-Osimertinib models (Figure S7E). In summary, our data provide strong evidence to suggest the essentiality of BRCA1-A complex and *LIG4* for the process of ecDNA-mediated adaptation that allows cancer cells to acquire drug resistance.

DISCUSSION

DNA repair factors mediate ecDNA biogenesis

The biogenesis of ecDNA is a crucial biological process that occurs across various species. Previous studies have explored the involvement of specific DNA repair pathways in this process, with suggestive data leading to models proposing distinct roles for different repair mechanisms^{3,4,25,31,54}. In this study, we performed a genome-wide CRISPR screening and identified LIG4 and BRCA1-A complexes play key roles in mediating ecDNA formation. While LIG4 complex serves as core components from NHEJ DNA repair, BRCA1-A acts at the upstream of DNA break signaling and is linked to determine the Homologous Recombination (HR) pathway choice³⁴⁻³⁷. Thus, our findings indicate that ecDNA biogenesis is not driven by a single DNA repair pathway. Instead, it is orchestrated by selective factors from various steps of the DNA repair process. Notably, in an ATM-deficient background, BRCA1-A and LIG4 have been reported to work synergistically to repair DNA breaks⁵⁵. Our findings highlight that their cooperative relationship is also pivotal underlying the mechanism of ecDNA formation.

Upon DNA fragmentation, the upstream BRCA1-A complex protects the DNA ends from resection, likely funneling the un-resected ends towards downstream ligation to form a circle. Meanwhile, from the screen we uncovered MRN complex, which mediates end-resection, as a suppressor of ecDNA formation. Our findings thus highlight upon DNA fragmentation, the end-processing step is a highly dynamic process, which may ultimately determine the fate of ecDNA formation. At downstream, a few NHEJ core factors—LIG4, XRCC4, and DNA-PKcs—mediate ecDNA formation with LIG4 as the key enzyme to catalyze the process. However, the other NHEJ factors were not identified as essential for ecDNA biogenesis. Ku80 and Ku70 are essential for the viability of the human cells^{56,57}, confounding the efforts to discern them as ecDNA biogenesis factors. Additional NHEJ factors (ARTEMIS, PAXX, POLL, and POLM) are required when the

ends possess non-complementary overhangs⁵⁸, and this could explain why they are dispensable for ecDNA biogenesis when blunt ends are generated during DNA fragmentation.

Targeting ecDNA formation for cancer therapy

Given that ecDNA is frequently generated in cancer cells to drive tumor evolution and adaptation^{3,7,8,11,16-18}, targeting ecDNA biogenesis presents a promising strategy for novel cancer therapies. Here, we report that both the BRCA1-A and LIG4 complexes play critical roles in ecDNA formation and contribute to drug resistance, highlighting new avenues for therapeutic intervention.

We showed that the LIG4 complex directly catalyzes ecDNA formation and thus represents ideal target. Humans appear to be able to tolerate the loss of LIG4 at a high level⁵⁹. Mutations in *LIG4* leads to LIG4 syndrome, a disease from which patients show developmental microcephaly and growth retardation but only have manageable immune deficiency in adult life⁵⁹. Additionally, our findings show that LIG4 is generally not essential for cell viability but becomes critical for cancer cells undergoing drug treatment, suggesting minimal and controllable toxicity from LIG4-targeted therapies. In fact, with the potential usage to sensitize cancer cells for radiotherapy, efforts were spent in the past to develop LIG4 drugs⁶⁰⁻⁶³, but failed to reach a fruition. Uncovering its essential function in driving ecDNA biogenesis and cancer cell adaptation, our work calls for innovative approaches to develop LIG4 inhibitors for cancer therapy.

In parallel, the BRCA1-A complex, a key player in safeguarding ecDNA formation and mediating drug resistance, represents another therapeutic target. Notably, BRCA1-A shares certain components with the BRISC complex³³, a deubiquitination complex involved in interferon responses, spindle assembly during mitosis, and hematopoiesis⁶⁴.

⁶⁶. Therefore, targeting proteins that are exclusive to the BRCA1-A complex, such as ABRAXAS and RAP80, may help avoid broader side effects. Encouragingly, loss of either ABRAXAS or RAP80 in mice does not lead to developmental defects^{67,68}, and our data indicate that these proteins are not normally required for cell viability but are essential when cancer cells adapt to drug treatment. Since these proteins lack enzymatic activities, developing compounds that block their interactions with partner proteins or induce their degradation is a promising path forward. In summary, by elucidating the essential roles of BRCA1-A and LIG4 complexes in ecDNA formation and drug resistance, our work underscores the potential for innovative, selective inhibitors of these complexes as cancer therapies.

Limitations of the study

In this study, we identified the critical roles of BRCA1-A and LIG4 complexes in ecDNA formation and drug resistance. Although we provide compelling in vivo data for LIG4 in the context of ecDNA formation during *Drosophila* oogenesis, we are limited by the absence of murine models that would enable testing the function of these two complexes in vivo in a cancer setting. Additionally, alternative explanations exist for our data showing that LIG4 mutations led to HeLa cell death under methotrexate treatment. Methotrexate is a genotoxin capable of inducing chromothripsis^{16,17}. Beyond its role in promoting DHFR ecDNA formation, LIG4 may also be required to repair genomic damage to support cancer cell survival under methotrexate stress. Lastly, while our findings advance understanding of ecDNA biogenesis in the context of genome fragmentation, other DNA repair factors and pathways can potentially contribute to ecDNA formation during development, disease, or in cases where the factors we

identified are dysfunctional. Future studies are necessary to explore these additional players in ecDNA formation and their roles across various biological contexts.

REFERENCES

1. Brown, D.D., and Dawid, I.B. (1968). Specific gene amplification in oocytes. Oocyte nuclei contain extrachromosomal replicas of the genes for ribosomal RNA. *Science* 160, 272-280. 10.1126/science.160.3825.272.
2. Hotta, Y., and Bassel, A. (1965). Molecular Size and Circularity of DNA in Cells of Mammals and Higher Plants. *Proceedings of the National Academy of Sciences of the United States of America* 53, 356-362. 10.1073/pnas.53.2.356.
3. Wu, S., Bafna, V., Chang, H.Y., and Mischel, P.S. (2022). Extrachromosomal DNA: An Emerging Hallmark in Human Cancer. *Annu Rev Pathol* 17, 367-386. 10.1146/annurev-pathmechdis-051821-114223.
4. Noer, J.B., Horsdal, O.K., Xiang, X., Luo, Y., and Regenbreg, B. (2022). Extrachromosomal circular DNA in cancer: history, current knowledge, and methods. *Trends in genetics : TIG*. 10.1016/j.tig.2022.02.007.
5. Zhu, Y., Gujar, A.D., Wong, C.H., Tjong, H., Ngan, C.Y., Gong, L., Chen, Y.A., Kim, H., Liu, J., Li, M., et al. (2021). Oncogenic extrachromosomal DNA functions as mobile enhancers to globally amplify chromosomal transcription. *Cancer Cell* 39, 694-707 e697. 10.1016/j.ccell.2021.03.006.
6. Hung, K.L., Yost, K.E., Xie, L., Shi, Q., Helmsauer, K., Luebeck, J., Schopflin, R., Lange, J.T., Chamorro Gonzalez, R., Weiser, N.E., et al. (2021). ecDNA hubs drive cooperative intermolecular oncogene expression. *Nature*. 10.1038/s41586-021-04116-8.
7. Koche, R.P., Rodriguez-Fos, E., Helmsauer, K., Burkert, M., MacArthur, I.C., Maag, J., Chamorro, R., Munoz-Perez, N., Puiggros, M., Dorado Garcia, H., et al. (2020). Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. *Nature genetics* 52, 29-34. 10.1038/s41588-019-0547-z.
8. Kim, H., Nguyen, N.P., Turner, K., Wu, S., Gujar, A.D., Luebeck, J., Liu, J., Deshpande, V., Rajkumar, U., Namburi, S., et al. (2020). Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. *Nature genetics*. 10.1038/s41588-020-0678-2.
9. Wu, S., Turner, K.M., Nguyen, N., Raviram, R., Erb, M., Santini, J., Luebeck, J., Rajkumar, U., Diao, Y., Li, B., et al. (2019). Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* 575, 699-703. 10.1038/s41586-019-1763-5.
10. Morton, A.R., Dogan-Artun, N., Faber, Z.J., MacLeod, G., Bartels, C.F., Piazza, M.S., Allan, K.C., Mack, S.C., Wang, X., Gimple, R.C., et al. (2019). Functional Enhancers Shape Extrachromosomal Oncogene Amplifications. *Cell* 179, 1330-1341 e1313. 10.1016/j.cell.2019.10.039.
11. Lange, J.T., Rose, J.C., Chen, C.Y., Pichugin, Y., Xie, L., Tang, J., Hung, K.L., Yost, K.E., Shi, Q., Erb, M.L., et al. (2022). The evolutionary dynamics of extrachromosomal DNA in human cancers. *Nature genetics* 54, 1527-1533. 10.1038/s41588-022-01177-x.

12. Durkin, K., Coppieters, W., Drogemuller, C., Ahariz, N., Cambisano, N., Druet, T., Fasquelle, C., Haile, A., Horin, P., Huang, L., et al. (2012). Serial translocation by means of circular intermediates underlies colour sidedness in cattle. *Nature* 482, 81-84. 10.1038/nature10757.
13. Libuda, D.E., and Winston, F. (2006). Amplification of histone genes by circular chromosome formation in *Saccharomyces cerevisiae*. *Nature* 443, 1003-1007. 10.1038/nature05205.
14. Kalt, M.R., and Gall, J.G. (1974). Observations on early germ cell development and premeiotic ribosomal DNA amplification in *Xenopus laevis*. *The Journal of cell biology* 62, 460-472. 10.1083/jcb.62.2.460.
15. Ilic, M., Zaalberg, I.C., Raaijmakers, J.A., and Medema, R.H. (2022). Life of double minutes: generation, maintenance, and elimination. *Chromosoma* 131, 107-125. 10.1007/s00412-022-00773-4.
16. 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. 10.1038/s41586-020-03064-z.
17. Rosswog, C., Bartenhagen, C., Welte, A., Kahlert, Y., Hemstedt, N., Lorenz, W., Cartolano, M., Ackermann, S., Perner, S., Vogel, W., et al. (2021). Chromothripsis followed by circular recombination drives oncogene amplification in human cancer. *Nature genetics* 53, 1673-1685. 10.1038/s41588-021-00951-7.
18. Turner, K.M., Deshpande, V., Beyter, D., Koga, T., Ruser, 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. 10.1038/nature21356.
19. Bailey, C., Pich, O., Thol, K., Watkins, T.B.K., Luebeck, J., Rowan, A., Stavrou, G., Weiser, N.E., Damercharla, B., Bentham, R., et al. (2024). Origins and impact of extrachromosomal DNA. *Nature* 635, 193-200. 10.1038/s41586-024-08107-3.
20. Nathanson, D.A., Gini, B., Mottahedeh, J., Visnyei, K., Koga, T., Gomez, G., Eskin, A., Hwang, K., Wang, J., Masui, K., et al. (2014). Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science* 343, 72-76. 10.1126/science.1241328.
21. Yang, F., Su, W., Chung, O.W., Tracy, L., Wang, L., Ramsden, D.A., and Zhang, Z.Z.Z. (2023). Retrotransposons hijack alt-EJ for DNA replication and eccDNA biogenesis. *Nature*. 10.1038/s41586-023-06327-7.
22. Wells, J.N., and Feschotte, C. (2020). A Field Guide to Eukaryotic Transposable Elements. *Annual review of genetics*. 10.1146/annurev-genet-040620-022145.
23. Kazazian, H.H., Jr., and Moran, J.V. (2017). Mobile DNA in Health and Disease. *The New England journal of medicine* 377, 361-370. 10.1056/NEJMra1510092.
24. Burns, K.H. (2017). Transposable elements in cancer. *Nature reviews. Cancer* 17, 415-424. 10.1038/nrc.2017.35.
25. Moller, H.D., Lin, L., Xiang, X., Petersen, T.S., Huang, J., Yang, L., Kjeldsen, E., Jensen, U.B., Zhang, X., Liu, X., et al. (2018). CRISPR-C: circularization of genes and chromosome by CRISPR in human cells. *Nucleic acids research* 46, e131. 10.1093/nar/gky767.
26. Goncalves, E., Thomas, M., Behan, F.M., Picco, G., Pacini, C., Allen, F., Vinceti, A., Sharma, M., Jackson, D.A., Price, S., et al. (2021). Minimal genome-wide human CRISPR-Cas9 library. *Genome biology* 22, 40. 10.1186/s13059-021-02268-4.

27. Dharanipragada, P., Zhang, X., Liu, S., Lomeli, S.H., Hong, A., Wang, Y., Yang, Z., Lo, K.Z., Vega-Crespo, A., Ribas, A., et al. (2023). Blocking Genomic Instability Prevents Acquired Resistance to MAPK Inhibitor Therapy in Melanoma. *Cancer Discov* 13, 880-909. 10.1158/2159-8290.CD-22-0787.
28. Rose, J.C., Belk, J.A., Wong, I.T., Luebeck, J., Horn, H.T., Daniel, B., Jones, M.G., Yost, K.E., Hung, K.L., Kolahi, K.S., et al. (2024). Disparate pathways for extrachromosomal DNA biogenesis and genomic DNA repair. *Cancer Discov.* 10.1158/2159-8290.CD-23-1117.
29. 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. 10.1146/annurev.biochem.052308.093131.
30. Scully, R., Panday, A., Elango, R., and Willis, N.A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature reviews. Molecular cell biology* 20, 698-714. 10.1038/s41580-019-0152-0.
31. 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 research* 49, 11787-11799. 10.1093/nar/gkab984.
32. 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*. 10.1038/s41586-021-04009-w.
33. Rabl, J., Bunker, R.D., Schenk, A.D., Cavadini, S., Gill, M.E., Abdulrahman, W., Andres-Pons, A., Luijsterburg, M.S., Ibrahim, A.F.M., Branigan, E., et al. (2019). Structural Basis of BRCC36 Function in DNA Repair and Immune Regulation. *Molecular cell* 75, 483-497 e489. 10.1016/j.molcel.2019.06.002.
34. Hu, Y., Scully, R., Sobhian, B., Xie, A., Shestakova, E., and Livingston, D.M. (2011). RAP80-directed tuning of BRCA1 homologous recombination function at ionizing radiation-induced nuclear foci. *Genes & development* 25, 685-700. 10.1101/gad.2011011.
35. Vohhodina, J., Toomire, K.J., Petit, S.A., Micevic, G., Kumari, G., Botchkarev, V.V., Jr., Li, Z., Livingston, D.M., and Hu, Y. (2020). RAP80 and BRCA1 PARsylation protect chromosome integrity by preventing retention of BRCA1-B/C complexes in DNA repair foci. *Proceedings of the National Academy of Sciences of the United States of America* 117, 2084-2091. 10.1073/pnas.1908003117.
36. Sobhian, B., Shao, G., Lilli, D.R., Culhane, A.C., Moreau, L.A., Xia, B., Livingston, D.M., and Greenberg, R.A. (2007). RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* 316, 1198-1202. 10.1126/science.1139516.
37. Her, J., Soo Lee, N., Kim, Y., and Kim, H. (2016). Factors forming the BRCA1-A complex orchestrate BRCA1 recruitment to the sites of DNA damage. *Acta Biochim Biophys Sin (Shanghai)* 48, 658-664. 10.1093/abbs/gmw047.
38. Blackford, A.N., and Jackson, S.P. (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Molecular cell* 66, 801-817. 10.1016/j.molcel.2017.05.015.
39. Wu, P.Y., Frit, P., Meesala, S., Dauvillier, S., Modesti, M., Andres, S.N., Huang, Y., Sekiguchi, J., Calsou, P., Salles, B., and Junop, M.S. (2009). Structural and functional interaction between the human DNA repair proteins DNA ligase IV and XRCC4. *Molecular and cellular biology* 29, 3163-3172. 10.1128/MCB.01895-08.

40. Sibanda, B.L., Critchlow, S.E., Begun, J., Pei, X.Y., Jackson, S.P., Blundell, T.L., and Pellegrini, L. (2001). Crystal structure of an Xrcc4-DNA ligase IV complex. *Nat Struct Biol* 8, 1015-1019. 10.1038/nsb725.
41. Dupre, A., Boyer-Chatenet, L., Sattler, R.M., Modi, A.P., Lee, J.H., Nicolette, M.L., Kopelovich, L., Jasin, M., Baer, R., Paull, T.T., and Gautier, J. (2008). A forward chemical genetic screen reveals an inhibitor of the Mre11-Rad50-Nbs1 complex. *Nat Chem Biol* 4, 119-125. 10.1038/nchembio.63.
42. Garner, K.M., Pletnev, A.A., and Eastman, A. (2009). Corrected structure of mirin, a small-molecule inhibitor of the Mre11-Rad50-Nbs1 complex. *Nat Chem Biol* 5, 129-130; author reply 130. 10.1038/nchembio0309-129.
43. Spradling, A.C., and Mahowald, A.P. (1980). Amplification of genes for chorion proteins during oogenesis in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* 77, 1096-1100. 10.1073/pnas.77.2.1096.
44. Yarosh, W., and Spradling, A.C. (2014). Incomplete replication generates somatic DNA alterations within polytene salivary gland cells. *Genes & development* 28, 1840-1855. 10.1101/gad.245811.114.
45. Wyatt, D.W., Feng, W., Conlin, M.P., Yousefzadeh, M.J., Roberts, S.A., Mieczkowski, P., Wood, R.D., Gupta, G.P., and Ramsden, D.A. (2016). Essential Roles for Polymerase theta-Mediated End Joining in the Repair of Chromosome Breaks. *Molecular cell* 63, 662-673. 10.1016/j.molcel.2016.06.020.
46. Pannunzio, N.R., Li, S., Watanabe, G., and Lieber, M.R. (2014). Non-homologous end joining often uses microhomology: implications for alternative end joining. *DNA Repair (Amst)* 17, 74-80. 10.1016/j.dnarep.2014.02.006.
47. Ramsden, D.A., Carvajal-Garcia, J., and Gupta, G.P. (2022). Mechanism, cellular functions and cancer roles of polymerase-theta-mediated DNA end joining. *Nature reviews. Molecular cell biology* 23, 125-140. 10.1038/s41580-021-00405-2.
48. Hussmann, J.A., Ling, J., Ravisankar, P., Yan, J., Cirincione, A., Xu, A., Simpson, D., Yang, D., Bothmer, A., Cotta-Ramusino, C., et al. (2021). Mapping the genetic landscape of DNA double-strand break repair. *Cell* 184, 5653-5669 e5625. 10.1016/j.cell.2021.10.002.
49. Bigner, S.H., Humphrey, P.A., Wong, A.J., Vogelstein, B., Mark, J., Friedman, H.S., and Bigner, D.D. (1990). Characterization of the epidermal growth factor receptor in human glioma cell lines and xenografts. *Cancer research* 50, 8017-8022.
50. Snapka, R.M., and Varshavsky, A. (1983). Loss of unstably amplified dihydrofolate reductase genes from mouse cells is greatly accelerated by hydroxyurea. *Proceedings of the National Academy of Sciences of the United States of America* 80, 7533-7537. 10.1073/pnas.80.24.7533.
51. Brown, P.C., Beverley, S.M., and Schimke, R.T. (1981). Relationship of amplified dihydrofolate reductase genes to double minute chromosomes in unstably resistant mouse fibroblast cell lines. *Molecular and cellular biology* 1, 1077-1083. 10.1128/mcb.1.12.1077-1083.1981.
52. Kaufman, R.J., Brown, P.C., and Schimke, R.T. (1979). Amplified dihydrofolate reductase genes in unstably methotrexate-resistant cells are associated with double minute chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* 76, 5669-5673. 10.1073/pnas.76.11.5669.

53. Balaban-Malenbaum, G., and Gilbert, F. (1977). Double minute chromosomes and the homogeneously staining regions in chromosomes of a human neuroblastoma cell line. *Science* 198, 739-741. 10.1126/science.71759.
54. Dutta, A., Eckelmann, B., Adhikari, S., Ahmed, K.M., Sengupta, S., Pandey, A., Hegde, P.M., Tsai, M.S., Tainer, J.A., Weinfeld, M., et al. (2017). Microhomology-mediated end joining is activated in irradiated human cells due to phosphorylation-dependent formation of the XRCC1 repair complex. *Nucleic acids research* 45, 2585-2599. 10.1093/nar/gkw1262.
55. Balmus, G., Pilger, D., Coates, J., Demir, M., Sczaniecka-Clift, M., Barros, A.C., Woods, M., Fu, B., Yang, F., Chen, E., et al. (2019). ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks. *Nature communications* 10, 87. 10.1038/s41467-018-07729-2.
56. Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and Sabatini, D.M. (2015). Identification and characterization of essential genes in the human genome. *Science* 350, 1096-1101. 10.1126/science.aac7041.
57. Blomen, V.A., Majek, P., Jae, L.T., Bigenzahn, J.W., Nieuwenhuis, J., Staring, J., Sacco, R., van Diemen, F.R., Olk, N., Stukalov, A., et al. (2015). Gene essentiality and synthetic lethality in haploid human cells. *Science* 350, 1092-1096. 10.1126/science.aac7557.
58. Zhao, B., Rothenberg, E., Ramsden, D.A., and Lieber, M.R. (2020). The molecular basis and disease relevance of non-homologous DNA end joining. *Nature reviews. Molecular cell biology* 21, 765-781. 10.1038/s41580-020-00297-8.
59. Altmann, T., and Gennery, A.R. (2016). DNA ligase IV syndrome; a review. *Orphanet J Rare Dis* 11, 137. 10.1186/s13023-016-0520-1.
60. Menchon, G., Bombarde, O., Trivedi, M., Negrel, A., Inard, C., Giudetti, B., Baltas, M., Milon, A., Modesti, M., Czaplicki, G., and Calsou, P. (2016). Structure-Based Virtual Ligand Screening on the XRCC4/DNA Ligase IV Interface. *Sci Rep* 6, 22878. 10.1038/srep22878.
61. Tomkinson, A.E., Howes, T.R., and Wiest, N.E. (2013). DNA ligases as therapeutic targets. *Transl Cancer Res* 2.
62. Srivastava, M., Nambiar, M., Sharma, S., Karki, S.S., Goldsmith, G., Hegde, M., Kumar, S., Pandey, M., Singh, R.K., Ray, P., et al. (2012). An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell* 151, 1474-1487. 10.1016/j.cell.2012.11.054.
63. Zhong, S., Chen, X., Zhu, X., Dziegielewska, B., Bachman, K.E., Ellenberger, T., Ballin, J.D., Wilson, G.M., Tomkinson, A.E., and MacKerell, A.D., Jr. (2008). Identification and validation of human DNA ligase inhibitors using computer-aided drug design. *J Med Chem* 51, 4553-4562. 10.1021/jm8001668.
64. Donaghy, R., Han, X., Rozenova, K., Lv, K., Jiang, Q., Doepner, M., Greenberg, R.A., and Tong, W. (2019). The BRISC deubiquitinating enzyme complex limits hematopoietic stem cell expansion by regulating JAK2 K63-ubiquitination. *Blood* 133, 1560-1571. 10.1182/blood-2018-10-877563.
65. Zheng, H., Gupta, V., Patterson-Fortin, J., Bhattacharya, S., Katlinski, K., Wu, J., Varghese, B., Carbone, C.J., Aressy, B., Fuchs, S.Y., and Greenberg, R.A. (2013). A BRISC-SHMT complex deubiquitinates IFNAR1 and regulates interferon responses. *Cell reports* 5, 180-193. 10.1016/j.celrep.2013.08.025.
66. Rabl, J. (2020). BRCA1-A and BRISC: Multifunctional Molecular Machines for Ubiquitin Signaling. *Biomolecules* 10. 10.3390/biom10111503.
67. Wu, J., Liu, C., Chen, J., and Yu, X. (2012). RAP80 protein is important for genomic stability and is required for stabilizing BRCA1-A complex at DNA

- damage sites in vivo. *The Journal of biological chemistry* 287, 22919-22926.
10.1074/jbc.M112.351007.
68. Castillo, A., Paul, A., Sun, B., Huang, T.H., Wang, Y., Yazinski, S.A., Tyler, J., Li, L., You, M.J., Zou, L., et al. (2014). The BRCA1-interacting protein Abraxas is required for genomic stability and tumor suppression. *Cell reports* 8, 807-817.
10.1016/j.celrep.2014.06.050.

RESOURCE AVAILABILITY:

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Zhao Zhang (z.z@duke.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The sequencing data were deposited to the National Center for Biotechnology Information (NCBI) with accession number PRJNA1068592.

ACKNOWLEDGMENTS

We thank Roberto Zoncu and Claire Goul for sharing scripts on CRISPR network analysis. We thank Charlie Gersbach and his lab for assisting with ddPCR. We thank David MacAlpine for suggestions on detecting chorion ecDNA. We thank members from ZZ lab for constructive suggestions during the project development and for reading the manuscript. This work was supported by the grants to Z.Z. from the Pew Biomedical

Scholars Program, the Sontag Distinguished Scientist Award program, and the National Institutes of Health (R01 GM141018 and R01 GM152423), and to K.C.W. from the National Institutes of Health (R01CA263593), to Christian Cerda-Smith from the National Institutes of Health (F30CA247323), to Lee Zou from National Cancer Institute (R01 CA263394), and to D.A.R from the National Cancer Institute (R01 CA097096), and to M.K. from The Cooperative Trials Group for Neuro-Oncology.

AUTHOR CONTRIBUTIONS

Z.Z., O.C., S.Y., and F.Y. conceived the project. L.Z. and D.A.R. provided critical comments on CRISPR screen data interpretation. All authors designed the experiments. M.K. sequenced glioblastoma DNA. F.Y. performed ecDNA versatile biosensor-related experiments. L.W. performed computational analysis. S.Y. performed cancer cell-related experiments. C.C.S., H.M.H., and K.C.W. provided guidance on CRISPR screen and performed related data analysis. O.C. performed all the rest of experiments. Z.Z. and O.C. wrote the manuscript. All authors read and approved the manuscript.

DECLARATION OF INTERESTS:

Z.Z., O.C., F.Y., and S.Y. are co-inventors on two US provisional patent applications filed by Duke University related to this work. Z.Z. is the co-founder of Muye Therapeutics.

Figure 1

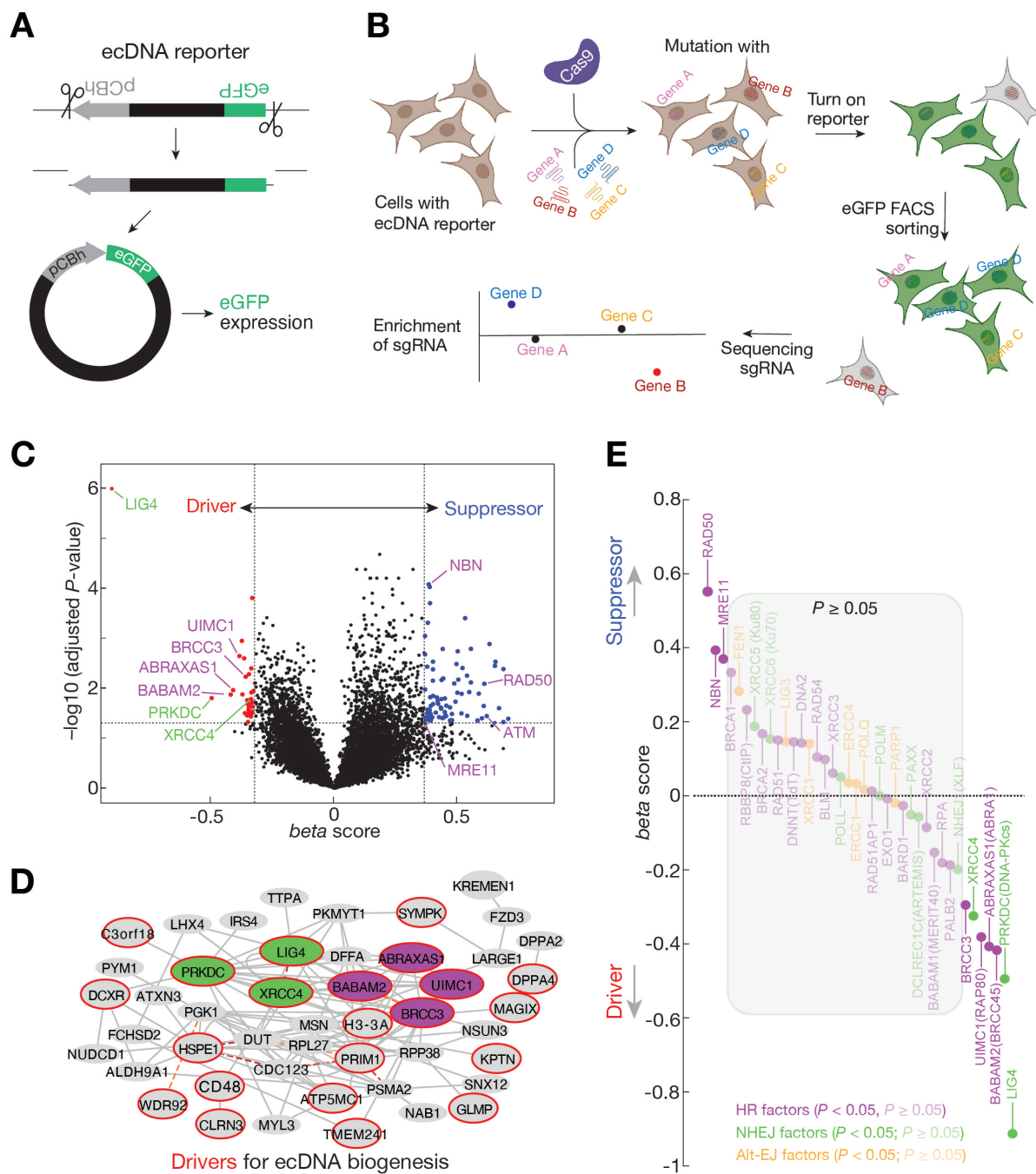


Figure 1. CRISPR screening to identify factors regulating ecDNA biogenesis

(A) Schematic design of the reporter used for monitoring ecDNA production.

(B) Schematic design of the genome-wide CRISPR screens, which were performed in 3 biological replicates.

(C) Volcano plot to show the regulators of ecDNA biogenesis. The factors further investigated in this study are labelled.

(D) Interactome analysis of the factors that drive ecDNA biogenesis. LIG4 and BRCA1-A complexes are highlighted by green and magenta respectively. Factors identified as drivers in panel C are circled with red. Factors that show co-essentiality are connected with dashed red lines.

(E) Snake plot to show the potential function of factors from distinct DNA break repair pathways in ecDNA biogenesis.

See also Figure S1.

Figure 2

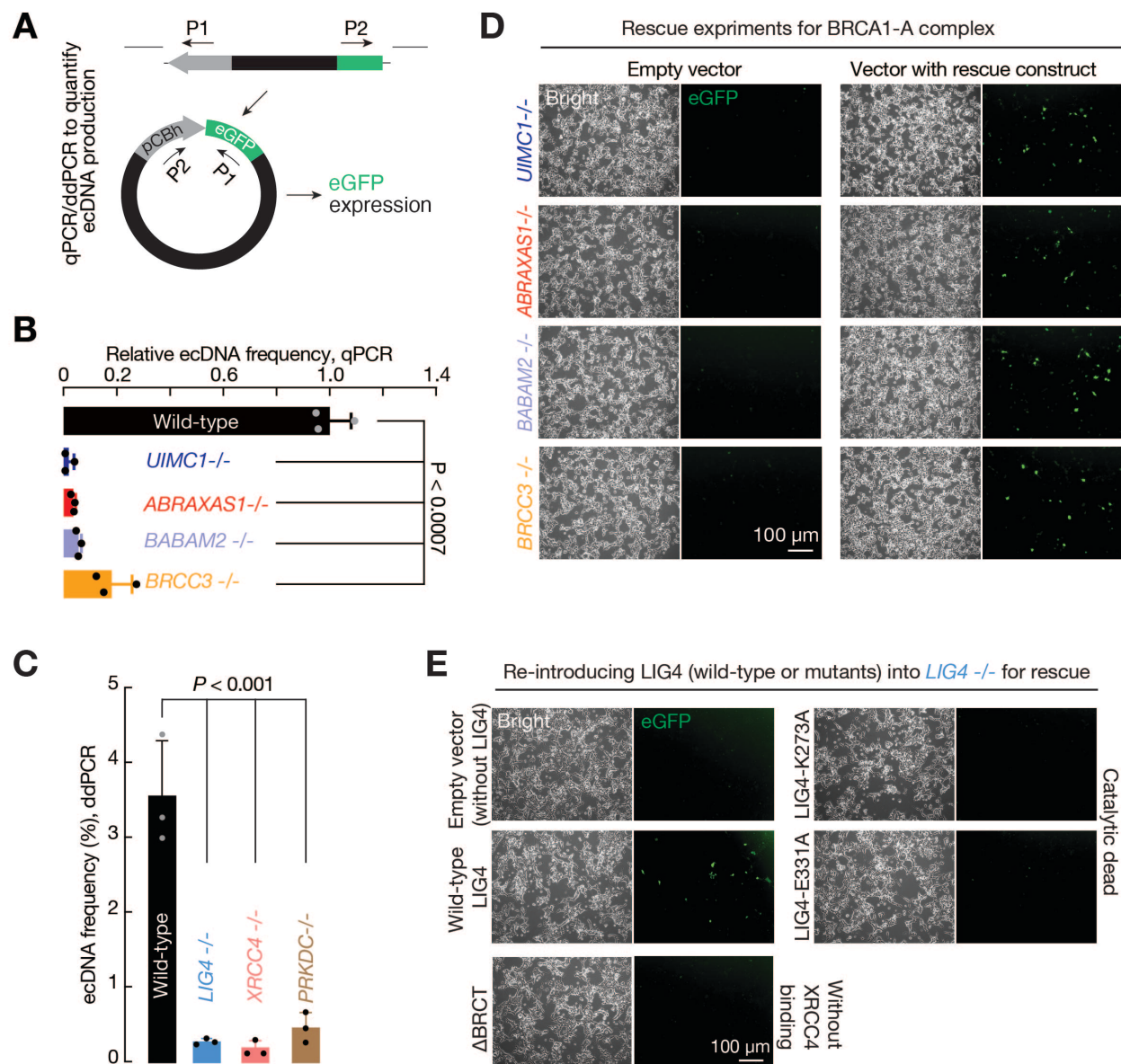


Figure 2. BRCA1-A and LIG4 complexes drives ecDNA biogenesis from the reporter

(A) Schematic design of the qPCR and Droplet digital PCR (ddPCR) assays to quantify ecDNA production from the pre-integrated reporter.

(B) TaqMan qPCR to quantify ecDNA production when BRCA1-A function is disrupted.

(C) ddPCR to quantify ecDNA production when LIG4 complex is mutated. For both panel B and C, the bars report mean \pm standard deviation from three biological replicates ($n=3$). P -values were calculated with a two-tailed, two-sample unequal variance t test. For each gene, the mutant cell line generated by sgRNA-1 was picked for qPCR.

(D) Re-introducing wild-type genes into corresponding BRCA1-A mutants rescues ecDNA production from the pre-integrated reporter, as evidenced by the restoration of eGFP expression.

(E) Re-introducing wild-type *LIG4* rescues ecDNA production from the pre-integrated reporter. Mutant versions of *LIG4*—either catalytically dead (E331A or K273A) or with the XRCC4-binding BRCT domains deleted—fail to rescue. Each rescue construct achieves a similar level of protein expression (Figure S2D). See also Figure S2A-2D, S3, and S4.

Figure 3

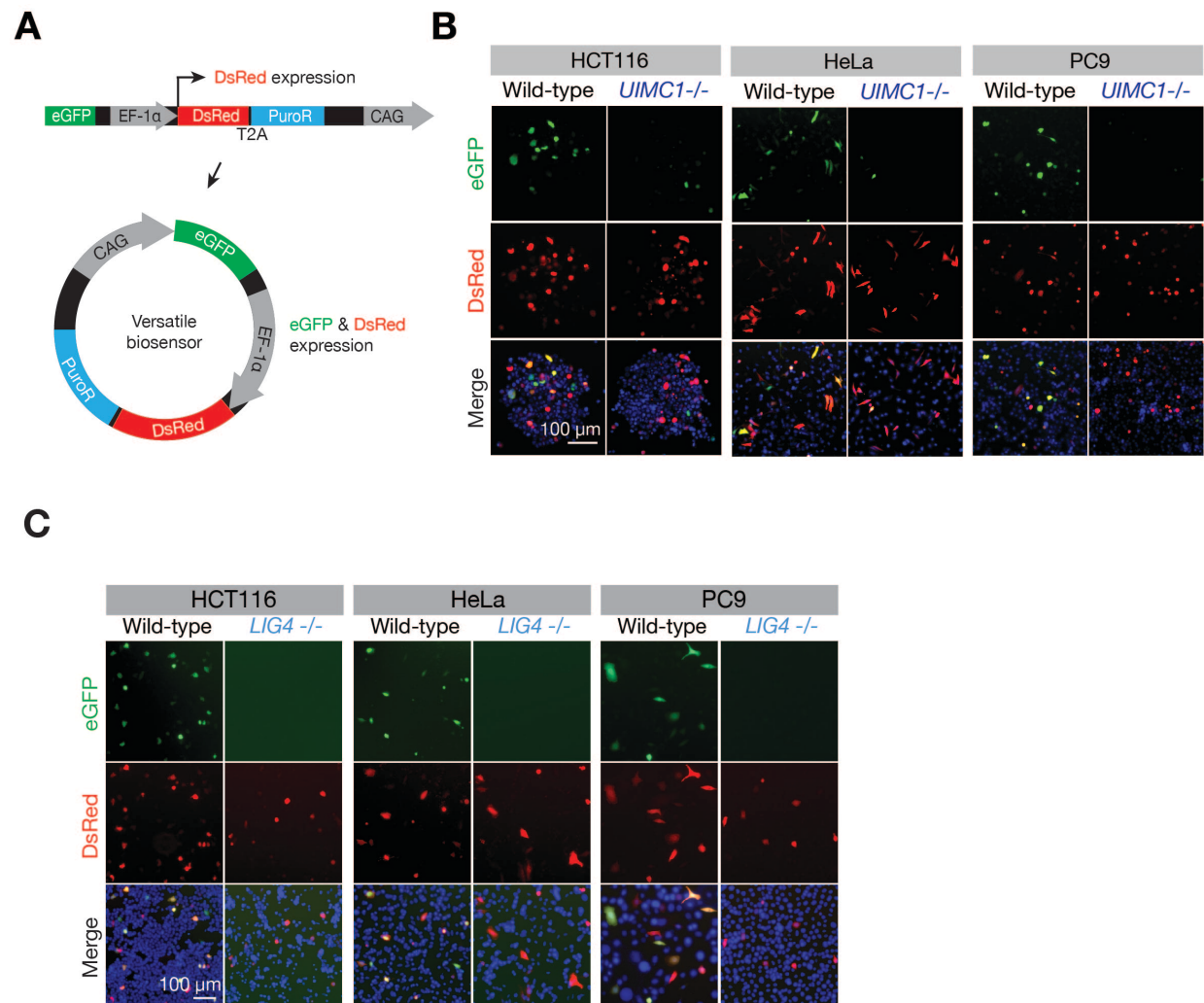


Figure 3. BRCA1-A and LIG4 complexes drives ecDNA biogenesis in cancer cells

(A) Schematic design of the versatile biosensor. Circularization brings the CAG promoter upstream of the eGFP sequences to initiate eGFP expression. EF-1 α promoter drives the expression of DsRed and the puromycin resistance gene, enabling selection of the cells harboring the biosensors.

(B) BRCA1-A complex drives ecDNA production in cancer cells: HCT116, HeLa, and PC9.

(C) LIG4 drives ecDNA production in cancer cells: HCT116, HeLa, and PC9.

See also Figure S2E, S2F, and S5.

Figure 4

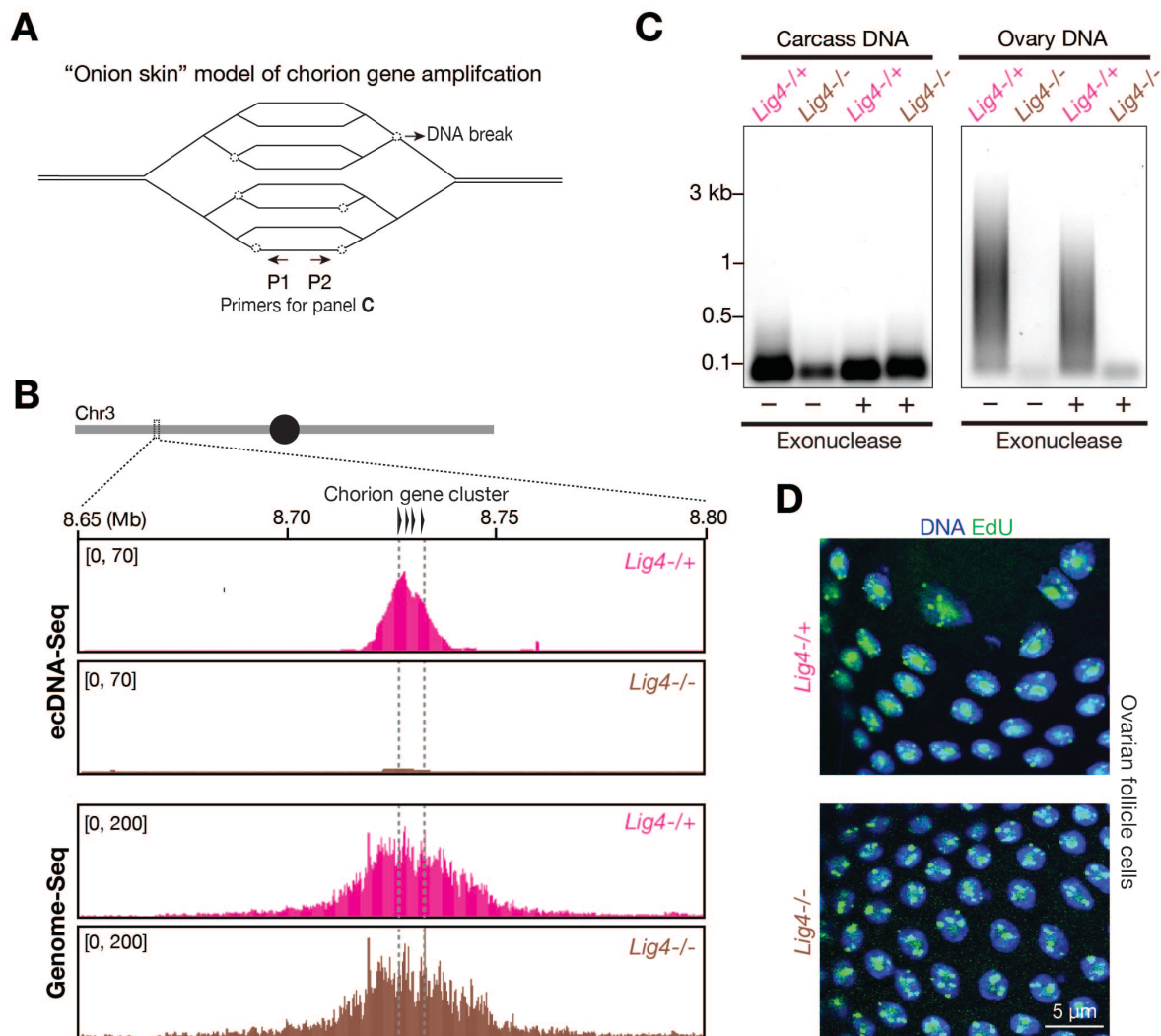


Figure 4. LIG4 drives natural ecDNA biogenesis in vivo

(A) Cartoon depicting the "Onion skin" model of chorion gene amplification occurring in *Drosophila* ovarian follicle cells. Dashed circles stand for DNA breaks.

(B) ecDNA-Seq and Genome-Seq to measure the amplification of the chorion locus on the 3rd chromosome and ecDNA production from this region.

(C) PCR-based assay to measure the production of chorion ecDNA from *Drosophila* ovaries. Fly bodies without ovaries were used as the carcass for DNA extraction, serving as a negative control.

(D) EdU incorporating into the replicated DNA in *Drosophila* ovarian follicle cells. *Lig4* mutation has no impact on DNA amplification, as indicated by EdU immunostaining.

See also Figure S6A and S6B.

Figure 5. LIG4 likely mediates ecDNA formation in glioblastoma

(A) Cartoon depicting the experiment design to sequence ecDNA from tumor tissues from 36 glioblastoma patients.

(B) Copy number of cancer-related ecDNA being detected in glioblastoma from each patient. Color indicates the oncogene detected on the circle. Bottom black box serves as the color key, and the total number of patients harboring the labelled oncogene on ecDNA was shown in the parenthesis. For example, 14 patients having EGFR on their ecDNA.

(C) Copy number and breakpoints of the EGFR amplicon from patient #26, which serves as a representative example. The brown arc labels the junction site for circularization. The purple arc illustrates EGFR vIII mutation occurs in tumor tissues, but not in blood DNA. The left red arc labels a germline deletion, which can be detected in blood DNA. The right red arc indicates a 905 bp deletion occurs in tumor tissues.

(D) Reconstructed ecDNA from the EGFR amplicon showing in panel C. The junction sequences from the two ends share 1-bp homology.

(E) The bar graph to show the percentage of junction sites harboring distinct length of homologies.

Figure 6

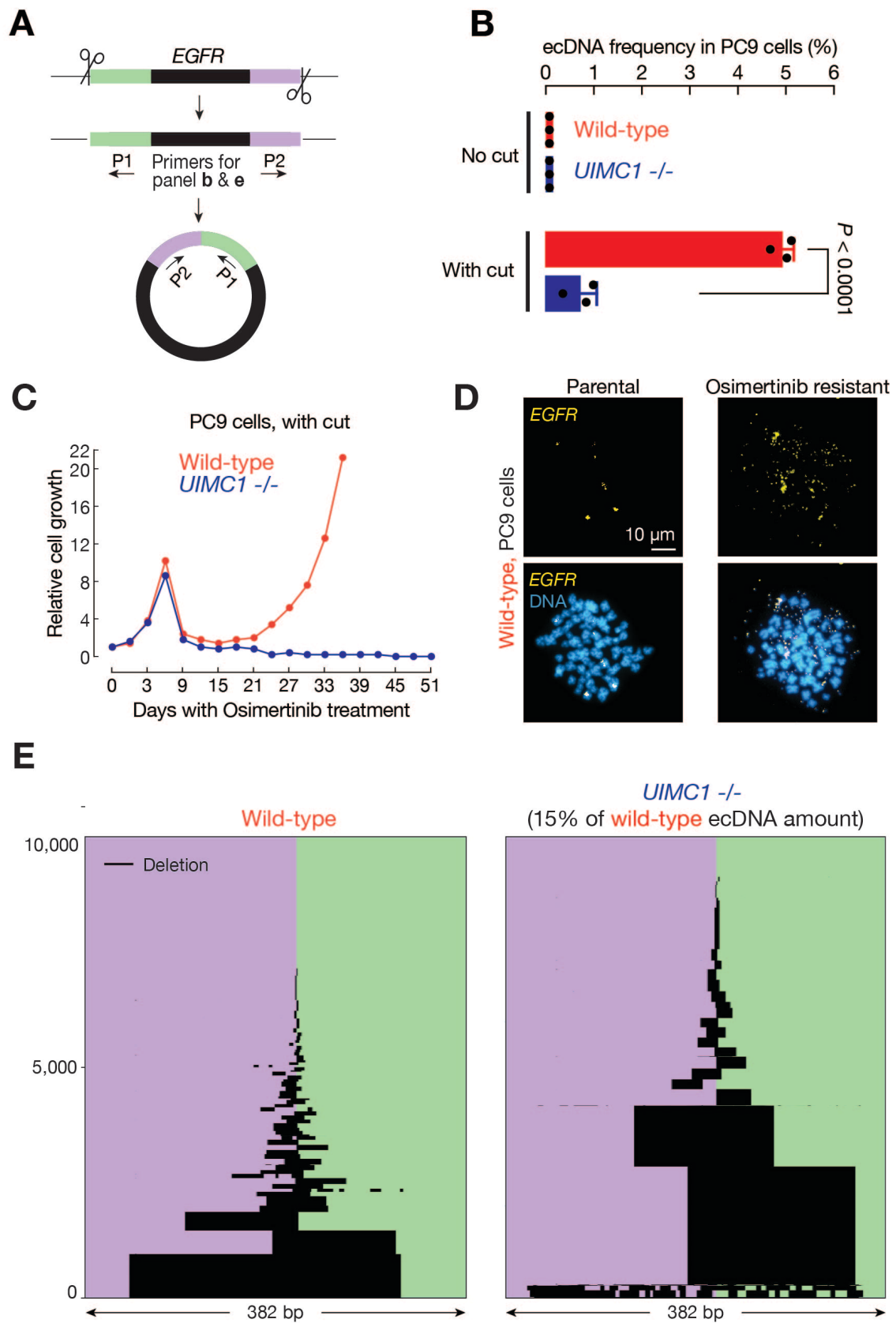


Figure 6. BRCA1-A complex prevents end-resection to drive ecDNA formation

- (A) Schematic design of the CRISPR-C approach to generate mega-bases *EGFR* ecDNA in PC9 lung cancer cells.
- (B) ddPCR to quantify ecDNA production from the *EGFR* region in PC9 cells. The bars report mean \pm standard deviation from three biological replicates ($n = 3$). P -values were calculated with a two-tailed, two-sample unequal variance t test.
- (C) Cell growth curve upon Osimertinib treatment. CRISPR-C approach was applied to generate *EGFR* ecDNA in PC9 cells before Osimertinib exposure.
- (D) DNA-FISH to measure *EGFR* ecDNA in parental and Osimertinib-resistant cells.
- (E) High-throughput sequencing to survey the end-resection events at the end-end junction site. Ten thousands reads were randomly selected for figure display. Black lines reflect the deletion regions from each read.
- See also Figure S6C.

Figure 7

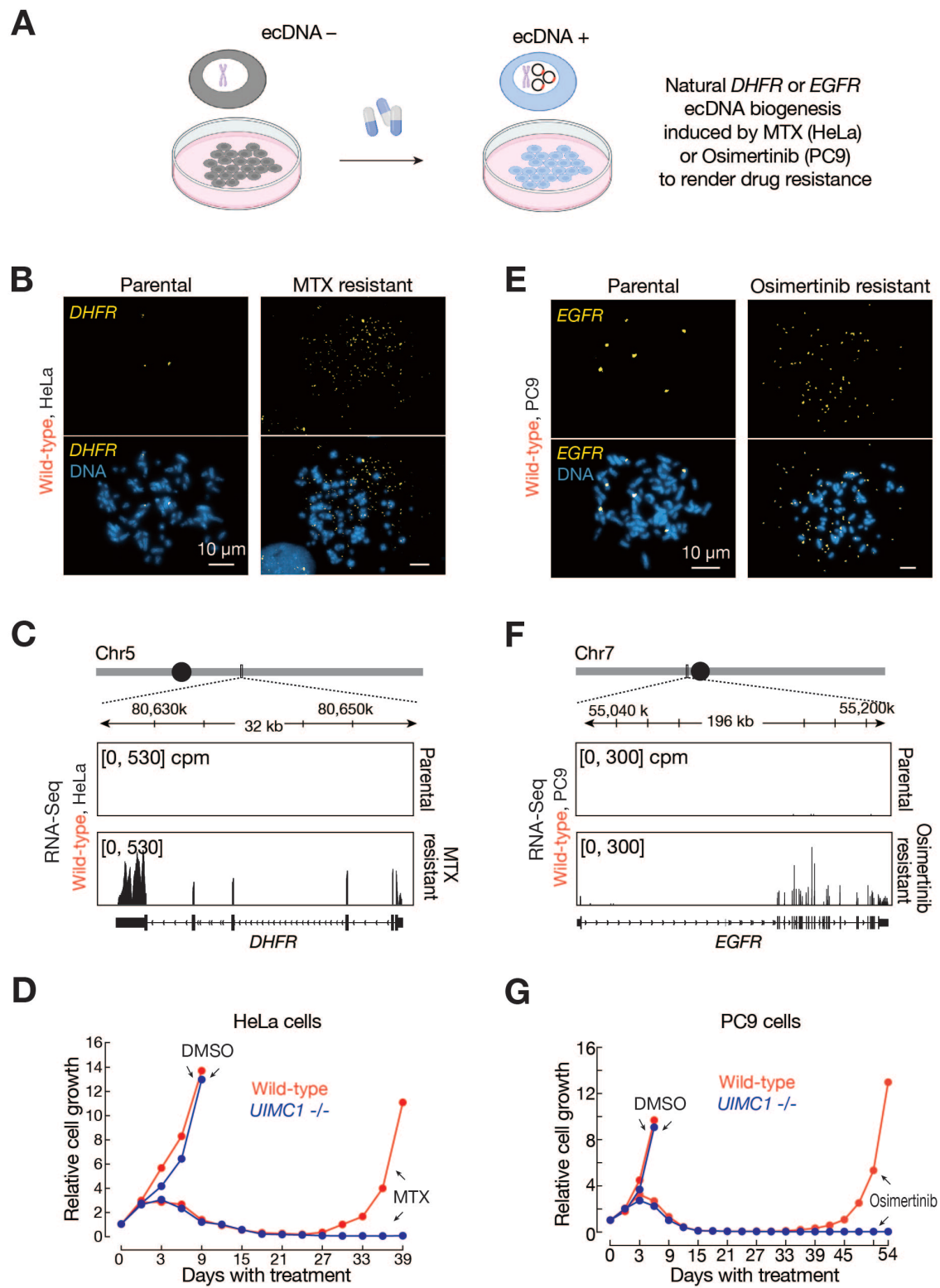


Figure 7. BRCA1-A complex drives ecDNA-mediated cancer cell evolution to acquire drug resistance

(A) Schematic design of using cancer treatment drugs to induce natural ecDNA production in cell culture. Methotrexate =MTX.

(B) DNA-FISH to measure *DHFR* ecDNA in HeLa cells. Upon drug treatment, wild-type cells formed and accumulated ecDNA to acquire drug resistance.

(C) RNA-Seq to show the expression of *DHFR* in parental and MTX resistant cells from the biological replicate #1. (replicate #2 and #3 were shown as Figure S7A and S7B). cpm = counts per million.

(D) Cell growth curve upon MTX treatment. Mutating *UIMC1* abolishes HeLa cells evolving to MTX resistance.

(E) DNA-FISH to measure *EGFR* ecDNA in PC9 cells. Upon Osimertinib treatment, wild-type cells formed and accumulated *EGFR* ecDNA to acquire drug resistance.

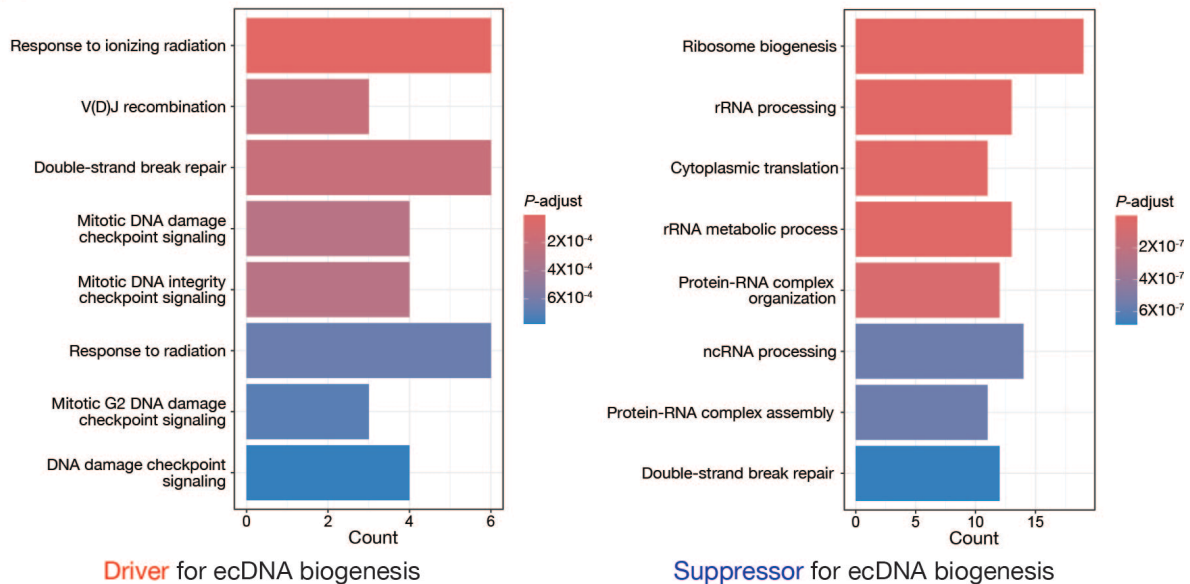
(F) RNA-Seq to show the expression of *EGFR* in parental and Osimertinib resistant cells from the biological replicate #1. (replicate #2 was shown as Figure S7C and S7D). cpm = counts per million.

(G) Cell growth curve upon Osimertinib treatment. Without *UIMC1*, PC9 cells failed to evolution resistance to the third generation EGFR inhibitor. Note, without CRISPR cut, it took one extra month for acquiring resistance, comparing with Figure 6C.

See also Figure S7.

Figure S1

A



B

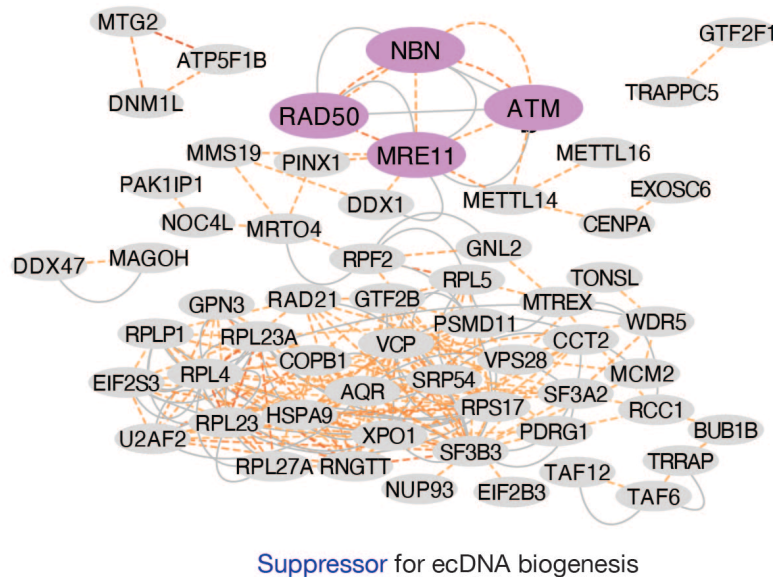


Figure S1. Bioinformatic analysis of CRISPR screen hits, related to Figure 1

(A) GO analysis to characterize the pathways that potentially regulate ecDNA biogenesis. (B) Interactome analysis of the factors that suppress ecDNA biogenesis. MRN complex components and ATM are highlighted by magenta. Factors that show co-essentiality are connected with dashed orange lines.

Figure S2

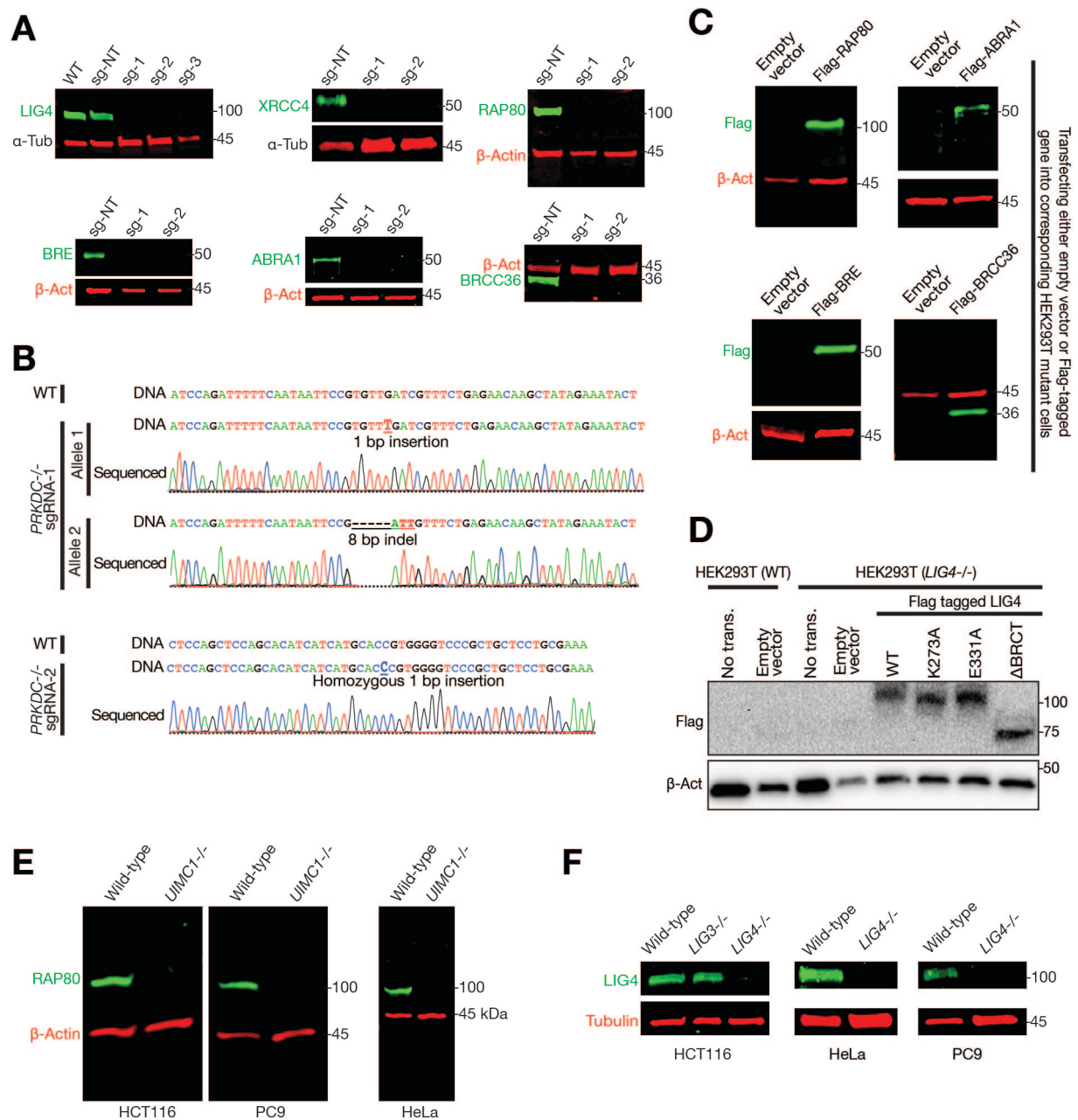


Figure S2. Validate mutation and rescue expression of genes from LIG4 and BRCA1-A complexes, related to Figure 2 and 3

- (A) Western blotting to validate the mutation of *LIG4*, *XRCC4*, *UIMC1* (encodes RAP80 protein), *BABAM2* (encodes BRE protein), *ABRAXAS1* (encode ABRA1 protein), and *BRCC3* (encode BRCC36 protein) in HEK293T cells.
- (B) Sanger sequencing to validate *PRKDC* mutation in the HEK293T cells. The mutant allele generated by sgRNA-1 is trans-heterozygotes. The mutant allele generated by sgRNA-2 is a homozygote allele. The sequencing traces were manually divided to represent deletions.
- (C) Western blotting to examine the expression of BRCA1-A complex rescue constructs. Empty vector = transfecting the vector that does not harbor the rescue gene.
- (D) Western blotting to examine the expression of LIG4 rescue constructs. No trans. = no plasmid transfection. Empty vector = transfecting the vector that does not harbor LIG4 sequences.
- (E) Western blotting to validate the mutation of *UIMC1* in different cancer cell lines.
- (F) Western blotting to validate the mutation of *LIG4* in different cancer cell lines.

Figure S3

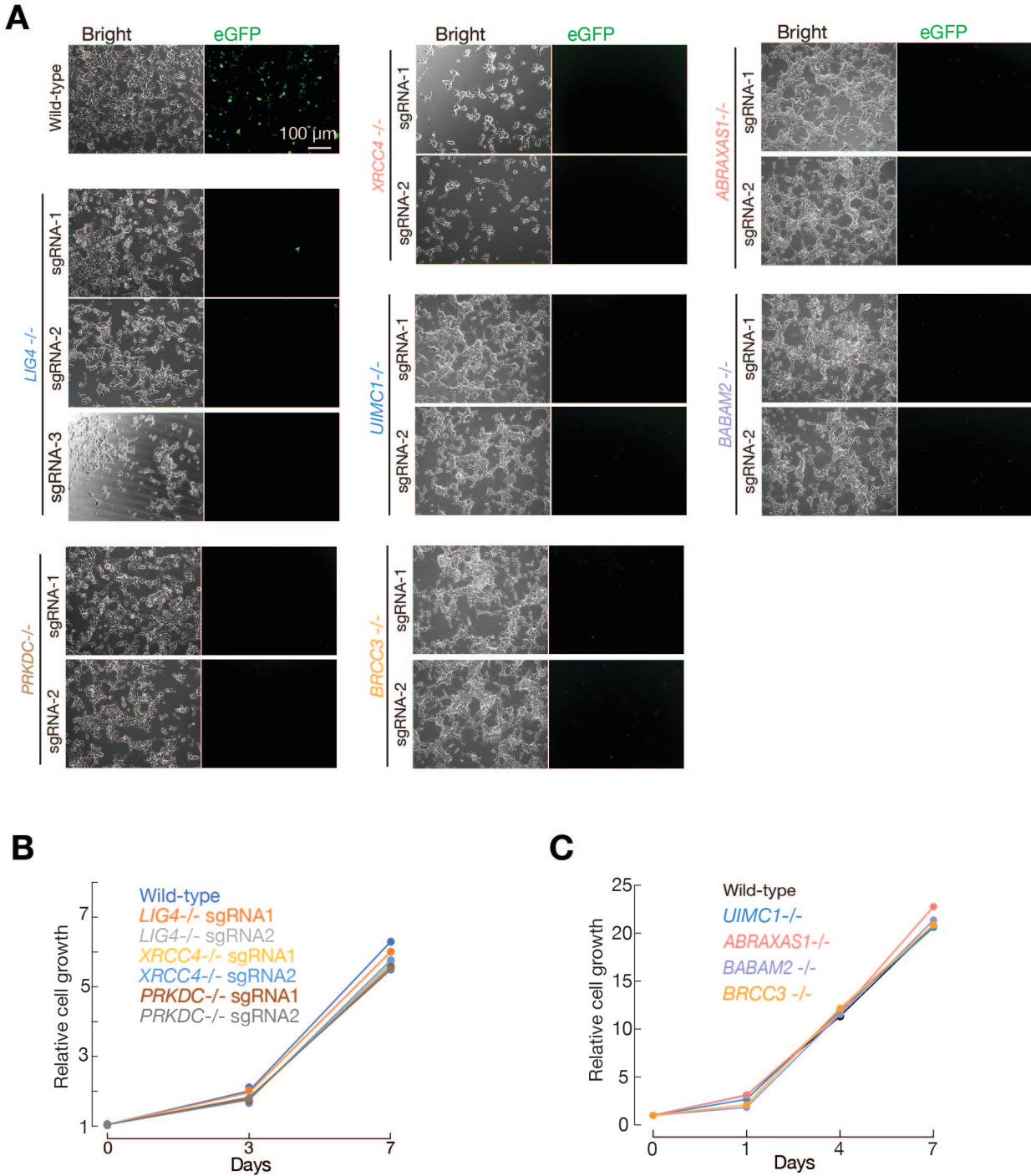


Figure S3. Mutating genes in LIG4 and BRCA1-A complexes abolishes ecDNA production from the pre-integrated reporter, related to Figure 2

(A) Individually mutating *LIG4*, *XRCC4*, *PRKDC*, *UIMC1*, *ABRAXAS1*, *BABAM2*, and *BRCC3* abolishes ecDNA production, as indicated by the loss of eGFP expression from the pre-integrated reporter.

(B) Cell growth curve upon mutating *LIG4*, *XRCC4*, and *PRKDC* individually.

(C) Cell growth curve upon mutating *UIMC1*, *ABRAXAS1*, *BABAM2*, and *BRCC3* individually. Mutants generated by sgRNA-1 for each gene were used for data display.

Figure S4

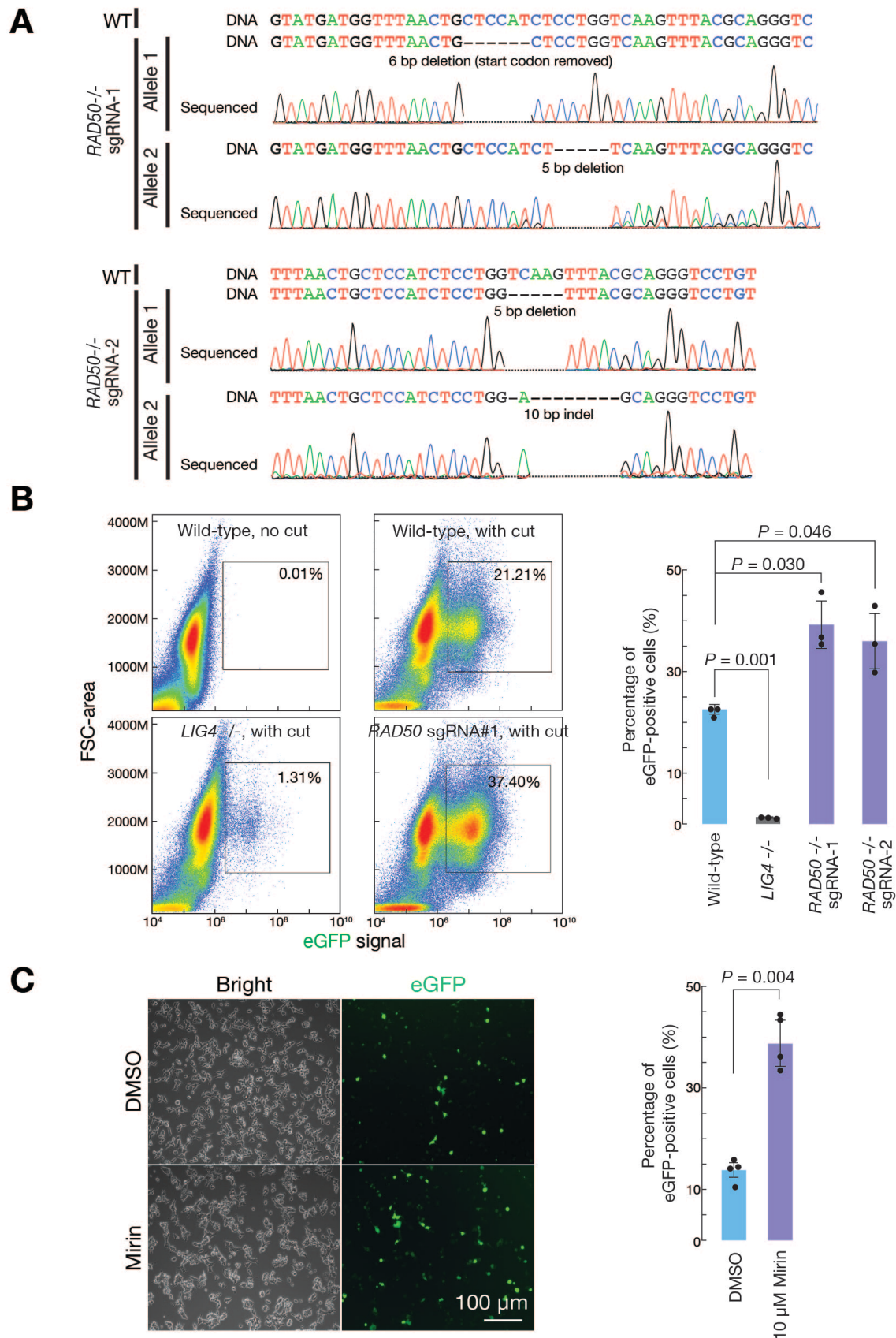


Figure S4. Suppressing the function of MRN complex enhances ecDNA production, related to Figure 2

- (A) Sanger sequencing to validate *RAD50* mutation of the HEK293T cells. The sequencing traces were manually divided to represent deletions.
- (B) FACS to quantify ecDNA positive cells. The bar graph shows the quantification of three biological replicates. The bars report mean \pm standard deviation from three biological replicates ($n=3$). *P*-values were calculated with a two-tailed, two-sample unequal variance *t* test.
- (C) Blocking the MRN complex function by using Mirin enhances ecDNA production. The bar graph shows the quantification of three biological replicates. The bars report mean \pm standard deviation from three biological replicates ($n=3$). *P*-values were calculated with a two-tailed, two-sample unequal variance *t* test.

Figure S5

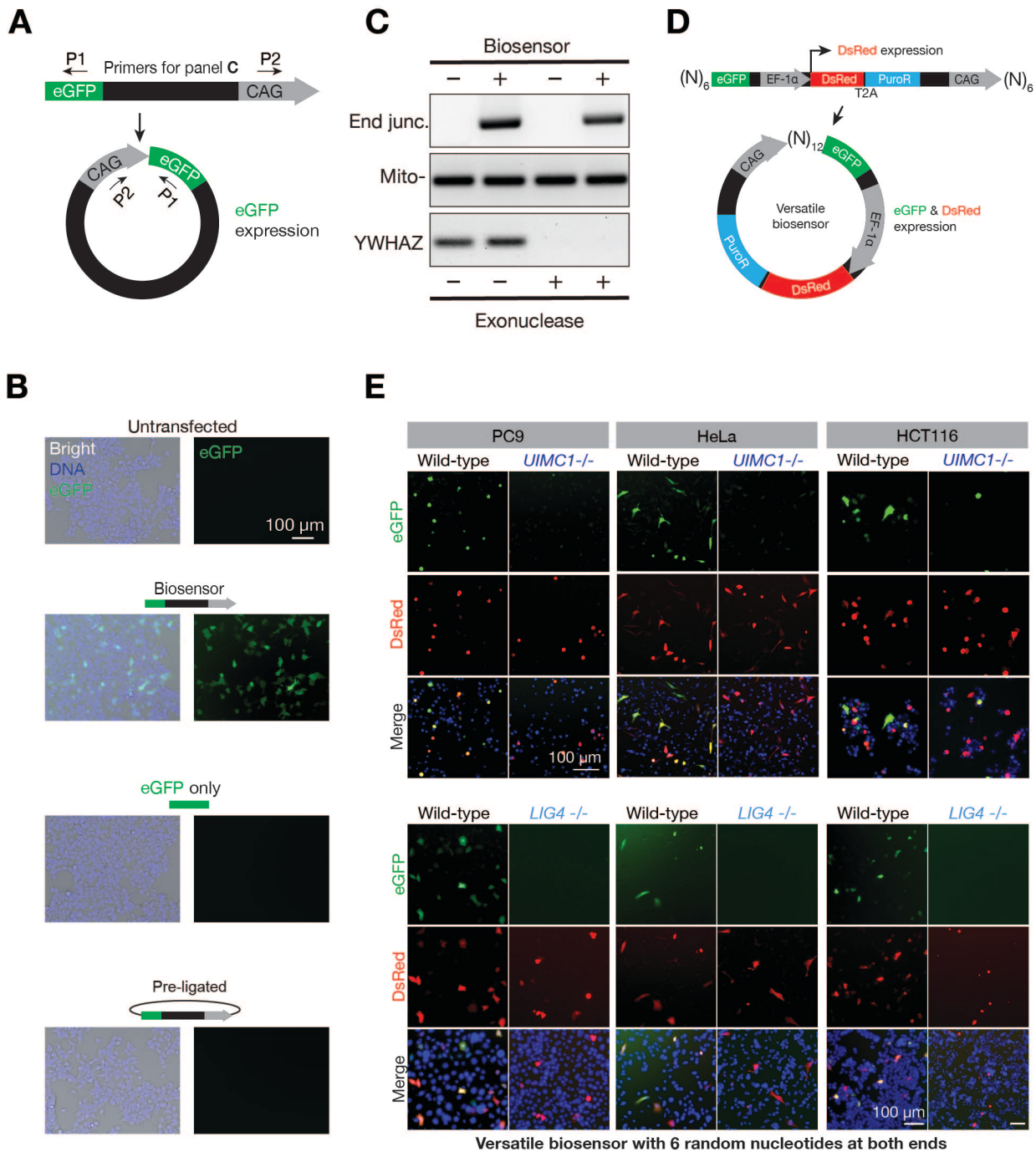


Figure S5. Versatile and robust biosensors to monitor ecDNA biogenesis, related to Figure 3

- (A) Schematic design of the biosensor. Circularization brings the CAG promoter upstream of the eGFP sequences to initiate eGFP expression.
- (B) Testing the biosensor in HEK293T cells. Introducing the eGFP sequence only or pre-ligating the biosensor into a circle prevents eGFP expression, suggesting the eGFP expression from the biosensor relies on the circularization process, not random integration of the biosensor into the genome.
- (C) PCR assay to measure the production of ecDNA from the biosensor. Exonuclease treatment removes linear DNA, as evidenced by the signals from the YWHAZ gene.
- (D) Schematic design of the biosensor flanked by six random nucleotides.
- (E) Mutating *UIMC1* or *LIG4* abolishes ecDNA formation from the biosensor flanked by six random nucleotides.

Figure S6

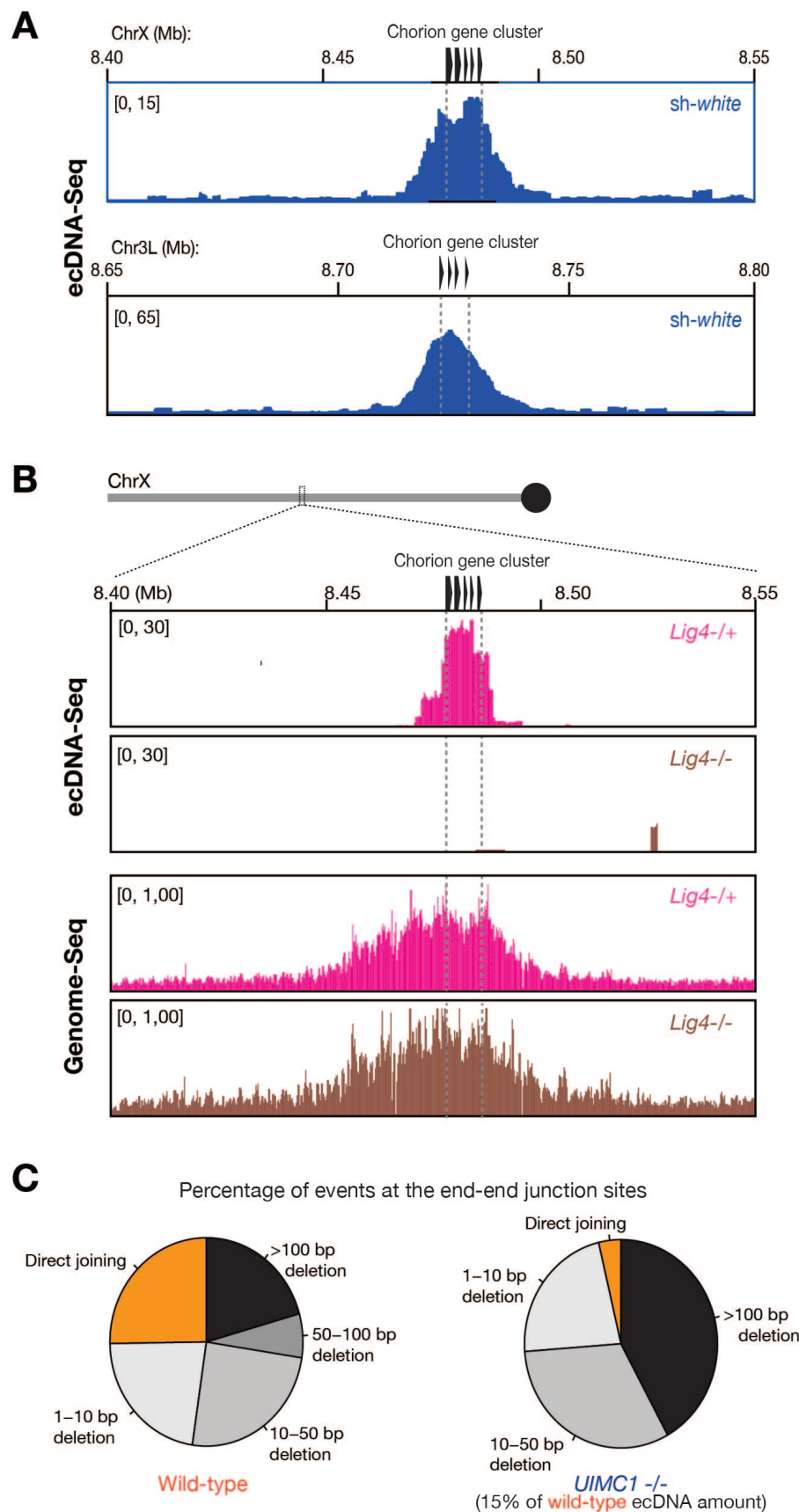


Figure S6. LIG4 drives natural ecDNA production during *Drosophila* oogenesis, related to Figure 4 and 6

- (A) Mining the published ecDNA-Seq data from the *Drosophila* ovary shows ecDNA production from the two chorion loci.
- (B) ecDNA-Seq and Genome-Seq to measure the amplification of the chorion locus on the X chromosome and ecDNA production from this region.
- (C) Pie chart to show the percentage of sequencing reads harboring each event category (Related to Figure 6).

Figure S7

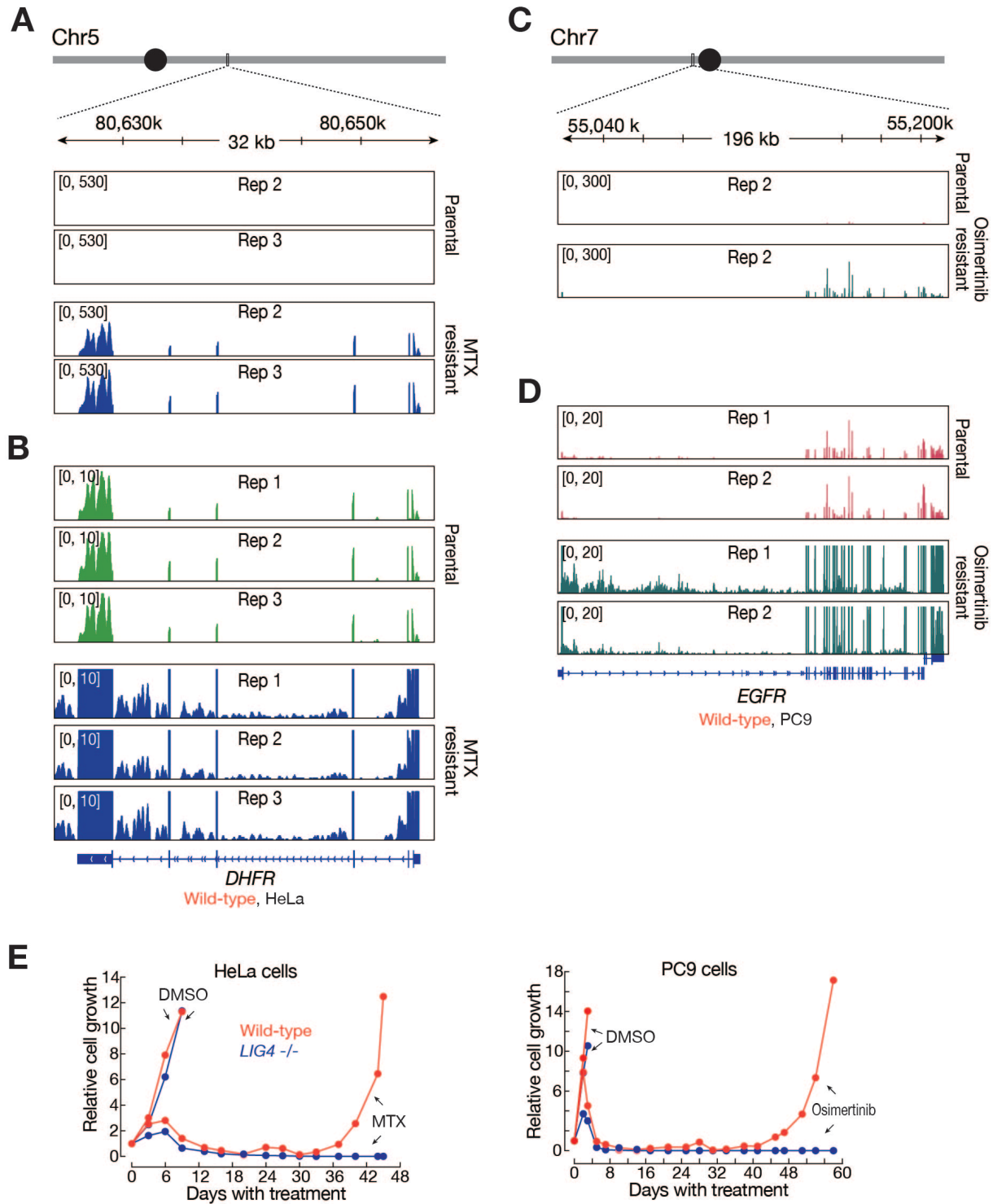


Figure S7. LIG4 drives ecDNA-mediated cancer cell evolution to acquire drug resistance, related to Figure 7

- (A) Genome browser to show the expression of *DHFR* in parental and methotrexate (MTX) resistant cells from the biological replicate #2 and #3 (replicate #1 was shown as Figure 7C).
- (B) Genome browser to show the expression of *DHFR* in parental and methotrexate (MTX) resistant cells from three biological replicates. The y-axis was adjusted to display signals from parental cells.
- (C) Genome browser to show the expression of *EGFR* in parental and Osimertinib resistant cells from the biological replicate #2 (replicate #1 was shown as Figure 7F).
- (D) Genome browser to show the expression of *EGFR* in parental and Osimertinib resistant cells from two biological replicates. The y-axis was adjusted to display signals from parental cells. In all plots, the unit of y-axis is counts per million (cpm).
- (E) Cell growth curve upon Methotrexate or Osimertinib treatment. Without *LIG4*, HeLa cells failed to adapt to Methotrexate and PC9 cells failed to evolution resistance to the third generation EGFR inhibitor.