Plasmodium falciparum DNA repair dynamics reveal unique roles for TLS polymerases and PfRad51 in genome diversification

Akshay Vishwanatha[1,#], Xu Zhang[1,3,#], Yi Jing Liu[1], Annie Leung[1], Mikayla Herring[1], Amanda Chan[1], Susanah Calhoun[1], Kirk Deitsch[2], Laura Kirkman[1,*]

Affiliations:

[1] Department of Medicine, Weill Cornell Medical College, New York, New York, USA
[2] Department of Microbiology and Immunology, Weill Cornell Medical College, New York, New York, USA
[3] Currently at Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[#]These authors contributed equally

* To whom correspondence should be addressed., lak9015@med.cornell.edu

Abstract

The haploid malaria parasite, *Plasmodium falciparum,* evolved a unique cohort of DNA repair pathways enabling the parasite to survive in a vertebrate host red blood cell and the mosquito vector. *P. falciparum* chromosomes are partitioned into a highly conserved core genome and remarkably diverse, largely subtelomeric regions that contain genes encoding important parasite virulence factors. The molecular mechanisms that maintain this chromosomal structure have not been identified. Here, we describe specific DNA repair pathways that differentiate between hypervariable subtelomeric and conserved core regions of the genome. Based on our previous work, we hypothesized that there are potentially important interactions between translesion (TLS) and homologous recombination (HR) pathways for the diversification of multicopy gene families in *P. falciparum.* Thus, we created knockout parasite lines of the DNA repair enzymes: *PfRad51* and the TLS polymerases *PfPolζ* and *PfRev1*. We identified that irradiation hypersensitivity varied across the cell cycle for TLSΔ parasites and was uniform across the erythrocytic cycle for *PfRad51*Δ parasites, highlighting the variable roles of these pathways. However, important interactions between these pathways were found when we studied directed double strand break (DSB) repair, which revealed a difference in the DNA damage response according to chromosomal location. *PfRad51* was essential for HR-mediated repair in the core genome. In contrast, we identified a Rad51 independent homology-directed repair in all three of our knockout lines when a DSB was made in the subtelomeric region of the chromosome. We propose that this differential DNA damage response maintains the distinction in diversification across the chromosome.

Introduction

Malaria inflicts a tremendous health and economic burden, with an estimated 263 million malaria cases in 2023 (an increase of 11 million compared to 2022) and a death toll of approximately 0.5 million globally.[1] The emergence of resistance to artemisinin and partner drugs used in combination therapy to treat malaria poses a significant risk to global efforts to control the disease [1]. Malaria is caused by infection with the unicellular obligate eukaryotic parasite of the genus *Plasmodium*. *Plasmodium falciparum* is the most virulent human infectious species. Unique features of the parasite's genome have been associated with virulence and its propensity to develop drug resistance. *P. falciparum* has a distinctly skewed genome with 80 % AT content, notable repetitive low complexity regions, and a propensity to accumulate small indels during mitotic division[2]. The parasite has a haploid genome for most of its lifecycle, including all stages in the human host, with a brief diploid stage in the mosquito vector. *P. falciparum* has 14 chromosomes displaying a high degree of genome plasticity, mostly confined to the subtelomeric regions that contain multicopy gene families that encode important virulence factors and variant surface antigens. These subtelomeric regions can extend to ~100 kb in length from the chromosome ends and include telomere-associated repeat elements (TAREs) and members of several hypervariable gene families such as *var*, *rifins* (repetitive interspersed family), *stevor* (subtelomeric variant open reading frame), and *Pfmc-2TM* (*P. falciparum* Maurer's cleft - 2 transmembrane domain protein)[3,4]. Each family contains multiple genes that display a high degree of sequence variation within the genome of a single parasite and between different geographical isolates[4-6]. We propose that in malaria parasites, DNA repair mechanisms play an important role in genetic diversification, particularly the diversification of the antigenic variant, erythrocytic membrane protein 1 (PfEMP1), encoded by the multicopy gene family *var,* which is present as clusters in the subtelomeric and several internal chromosomal regions (Supplementary Fig 1)[3]. Previous reports highlighted the high variability of these regions compared to the core genome[4] and possible mechanisms such as non-

allelic homologous recombination (NAHR), alternate end joining (Alt-EJ), gene conversion, and telomere healing (TH) contribute to this diversity during the asexual erythrocytic stages[7-10]. Therefore, understanding these different DNA repair pathways in the parasite will provide insight into the complex process of genome diversification, including this critical gene family of virulence factors.

P. falciparum possesses a complex yet unique genome maintenance system crucial for survival within the host. DNA repair mechanisms are required to overcome DNA damage induced by genotoxic stress, external damaging agents like mutagens, ionizing radiation, and drugs, including various antimalarials[11-14]. As the parasite proceeds through the intraerythrocytic development cycle (IDC), it undergoes schizogony, where one parasite replicates asynchronously to form 30-50 daughter cells. This development requires multiple rounds of DNA replication; thus, the parasite undergoes significant replication stress. Damaged bases or mistakes during DNA replication result in replication fork stalling. Hence, DNA damage must be repaired for replication to proceed. In addition to various repair mechanisms to remove damaged or mismatched bases, cells have developed specific processes that tolerate mistakes and allow replication to proceed. These processes are called replicative DNA damage bypass pathways or translesion synthesis (TLS)[15]. TLS is a highly conserved process mediated by specialized polymerases that replicate opposite and past damaged bases with lower fidelity, called TLS polymerases. The lower fidelity of these polymerases carries an increased risk of mutagenesis, which can be detrimental but also serve as a source of genomic diversification[15,16]. We recently reported the potential importance of TLS components in diversifying multigene families through ectopic recombination[17]. Rodent malaria parasites also harbor diverse multicopy gene families; however, there is a high degree of identity when comparing sequences between different isolates[18]. This high degree of sequence identity contrasts with much higher diversity between *P. falciparum* isolates. Of note, rodent parasites also lack the machinery for TLS[17]. In the human malaria parasite, components of the

TLS complex identified in silico include the scaffold protein subunit *PfRev1* and the catalytic subunit *PfPolζ*/*Rev3*, with two potential HORMA domain containing Rev 7 orthologs (Fig 1A).

Another lethal type of DNA damage is double-strand breaks (DSBs), which are repaired mainly by the non-homologous end joining (NHEJ) and the homologous recombination (HR) repair pathways. *P. falciparum* and all closely related blood-borne apicomplexan parasites lack the components of canonical NHEJ, making HR the major pathway for repairing DSBs[17,19]. Since Plasmodium is haploid throughout its IDC and lacks homologous pairs of chromosomes, it is surprising that it depends on HR for DSB repair. A relatively inefficient alt-EJ pathway was identified that might constitute a method of repairing DSBs in the absence of homologous templates[20,21]. In addition, we and others have reported how antigen diversification occurs through mitotic recombination mediated by HR in *P. falciparum* [7-10]. A defining feature of HR is the utilization of homologous DNA sequences as a repair template where the broken DNA molecule finds and invades the homologous DNA molecule. The Rad51 recombinase, assembled on the ssDNA at the site of the break as an oligomeric nucleoprotein filament, performs this function[22]. Therefore, in *P. falciparum*, HR is the main pathway by which DSBs are repaired, followed by alt-EJ in the absence of homologous templates.

Based on data suggesting the importance of the TLS polymerases and the HR pathway in the maintenance and diversification of the parasite genome, we set out to determine the general function of these proteins as well as their specific role in genome diversification.

Results

Creation of homologous recombination and translesion synthesis knockout mutants

To study the function of *PfRad51* (PF3D7_1107400) and TLS polymerases *PfRev1* (PF3D7_0910500) and *PfPolζ* (PF3D7_1037000) in *P. falciparum*, we first tested if these genes were essential for parasite viability. Rad51 is especially interesting because it is essential for survival in vertebrates where null mutation of Rad51 leads to embryonic lethality and death of mammalian cell lines[23], whereas Rad51 is not essential in other systems such as budding yeast[24], *T. brucei* [25], fission yeast *S. pombe*[26] and *S. japonicus* [27]. First, we generated a functional knockout of *PfRad51* using CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein-9 nuclease). As shown in Fig 1B (top cartoon), a repair block containing the selection marker hDHFR, encoding resistance to the drug WR99210 (WR) with flanking homologous regions upstream and downstream of the selected PAM (protospacer adjacent motif) site in the target gene, *PfRad51*, constituted the donor plasmid[28]. The donor plasmid also contained the single guide targeting the sequence coding for the ATPase domain of the protein (TTATTTGGTGAATTTCGTAC|AGG). A similar strategy was used to generate knockouts of the TLS polymerases *PfRev1* (ATATTAATAGTTCAAAACGG|AGG, BRCT domain) and *PfPolζ* (TTATCTTTTTAATGAAAGTG|AGG, upstream of DNA-directed DNA polB domain) using CRISPR/Cas9 as outlined in Fig 1B. Rev1 is not essential for the survival of higher eukaryotes, while Polζ is essential during development in mammalian systems[22,29,30] whereas both genes are non-essential in yeast[31]. The generated donor plasmid (Supplementary Table 1) along with a Cas9 plasmid (pUF1-Cas9)[28] were co-transfected into a wildtype (WT) 3D7 background to generate the three knockout mutants - *PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ. The successful transfections and comparable growth rates of the mutant lines to WT 3D7 line indicate that in *P. falciparum,* these three genes are not essential

for survival (Supplementary Fig 2). The knockouts were confirmed by Illumina whole genome sequencing, as shown in Fig 1C. The loss of the part of the gene replaced by the selection cassette is represented by the sudden loss of coverage towards the N-terminal region. Our attempts to generate double knockouts of *PfRad51* in each of the TLS polymerase knockout backgrounds did not yield viable parasites, indicating the essentiality of TLS polymerases in the absence of HR or that *PfRad51*, *PfRev1*, and *PfPolζ* function in parallel repair pathways.

P. falciparum **TLS polymerases are essential for the survival of irradiated ring-stage parasites**

The IDC of *P. falciparum* begins with the invasion of a red blood cell (RBC) by a single merozoite containing one copy of each chromosome. This merozoite develops into the 'ring' form from 0 h to 24 h post-invasion. At this stage of the IDC, the parasite contains a single copy of the genome. DNA replication at later stages of the cycle results in multiple copies inside an intact nucleus, potentially providing a template for the HR pathway to repair DSBs (Fig 2A)[32-35]. Therefore, in ring-stage parasites, given the absence of NHEJ in Plasmodium, there is greater reliance on alternate DSB repair pathways given a lack of template for repair required by HR. Different pathways are likely responsible for DNA repair in the ring and late IDC stages. Thus, sensitivity to DNA damage would also vary depending on the parasite's position within the replicative cycle[3,35]. Therefore, the knockouts of *PfRad51* and the TLS components described in the previous section were tested for their sensitivity to DNA damage using X-rays as a source of ionizing radiation (Fig 2B). Ionizing radiation such as X-rays is known to induce multiple kinds of DNA damage including base damage, single-strand breaks (SSB), double-strand breaks (DSB) that require different repair pathways such as nucleotide excision repair (NER), base excision repair (BER), mismatch repair (MMR) and DSB repair pathways for the exposed cell to recover [36].

*PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ, as well as WT 3D7 parasites were exposed to 100 Gy irradiation at both the ring- and late-stage of the replicative cycle. Cultures were then monitored for up to three weeks to determine how the loss of different DNA repair proteins affected the parasites' ability to repair DNA damage and reestablish growth. For both ring- and late-stages, WT 3D7 parasites recovered from an initial growth inhibition within 9 days of irradiation exposure (Fig 2C, green line), thereby establishing the time required for recovery when all repair pathways are active. Given that malaria parasites are thought to depend almost exclusively on HR for repairing DSBs, and since Rad51 is required for HR in model organisms, we anticipated a near-complete elimination of recovery in the *PfRad51*Δ line. We also predicted that the loss of Rad51 might cause a more pronounced effect in late-stage parasites, which contain multiple copies of the genome, and thus, DSB repair would be further dominated by HR. Remarkably, both ring- and late-stage parasites were able to recover from irradiation in the absence of Rad51 (15 days for ring-stage parasites vs. 19 days for late stages), and the loss of *PfRad51* did not appear to be associated with a substantial difference in survival between these two stages (Fig 2C, red line). These data indicate that despite the presumed dependence on HR for DSB repair in malaria parasites, sufficient alternative repair activity exists to enable reproducible recovery from extensive DNA damage without Rad51. Our previous work indicated that at this irradiation dose, virtually all parasites sustain damage and those parasites that survive display longer telomeres, a hallmark of recovery from DNA damage[37].

We then examined the effect of the loss of TLS polymerases on parasite recovery from the same irradiation dose. We observed that both *PfRev1* and *PfPolζ* are essential for the survival of ring-stage parasites when irradiated (Fig 2C, blue and orange lines) as the knockouts fail to recover after more than 20 days. The essentiality of these two TLS polymerases, even when the canonical HR pathway is intact, suggests they function in a Rad51-independent repair pathway required for DNA repair in

haploid, ring-stage parasites. Interestingly, the loss of TLS polymerases had little effect on the parasites' ability to recover from irradiation in later stages, when PfRad51-mediated HR presumably dominates (Fig 2C, blue and orange lines). These observations demonstrate distinct DNA damage sensitivity for our different knockout parasite lines.

PfRad51 and PfTLS polymerases differentially impact SNP and indel accumulation

To further validate the canonical functions of TLS polymerases and Rad51, we examined the genome sequence of parasites that had recovered from irradiation (Fig 3A). Briefly, we irradiated WT 3D7, *PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ late-stage parasites with 100 grays of X-rays three times, each time letting the parasites recover. Clonal lines were established from the recovered parasites, and genomic DNA was isolated, purified, and Illumina sequenced (NovaSeq 6000, Kapa HiFi library kit); the sequencing data were processed to identify single nucleotide polymorphisms (SNPs), insertions and deletions (Indels) using Freebayes [38] and SnpEff [39] as outlined in Fig 4A.

To determine the effect of the loss of TLS polymerases and Rad51 on the accumulation of mutations in *P. falciparum*, we analyzed the complete genome sequences of irradiated parasites from multiple clones from each knockout line and compared them to WT 3D7 parasites. As expected, the PfTLS knockout lines display a hypomutable phenotype (Fig 3D and E, 4A), with the WT 3D7 more prone to accumulating SNPs post irradiation (Fig 3B). The distribution of repaired damage was the same across the chromosome for all lines tested, with most SNPs and indels occurring in the core genome when looking at the percent of mutations (Fig 4B, E). The PfTLS knockout mutants accumulated considerably lower non-synonymous SNPs, indicating that any damage in coding regions was not repaired and, thus likely lethal (Fig 4C). In the case of indels, which are thought to originate from DSBs[40], disruption of the PfTLS pathway had little effect (Fig 4D). Although there are reports of the involvement of the TLS pathway in DSB repair in other systems[41-43], *PfRev1*Δ and *PfPolζ*Δ show

identical phenotypes where the total number of indels is unchanged compared to WT 3D7 (Fig 3D and E, 4D). Interestingly, both *PfRev1*Δ and *PfPolζ* Δ display a trend toward decreased accumulation of non-synonymous indels, which reaches statistical significance in the case of *PfPolζ*Δ (Fig. 4F), suggesting a greater role for TLS polymerases in repairing DNA damage within coding regions. Coding regions, specifically transcribed regions of the genome, can be more vulnerable to DNA damage due to the topological strain from transcription and a more open chromatin structure, perhaps leading to increased accumulation of DSBs. In *P. falciparum*, coding regions also have a pronounced increase in GC content compared to intergenic regions, which could affect DNA repair. Polζ can play a role in repairing DSBs through a post-replicative DNA recombination pathway observed in other model systems[44]. In addition, Polζ is enlisted as a DNA polymerase during the recombination process, and the action of Polζ leads to mutations in yeast [41]. Therefore, it is possible that in *P. falciparum*, the TLS polymerases might also have important cross functionality responding to the repair of excised damaged bases as well as DSB repair from stalled replication forks or direct DNA damage.

The loss of *PfRad51* did not affect the frequency and number of SNPs across the whole genome (Fig 3B, Fig 4A). We did, however, observe a small but statistically significant increased incidence of non-synonymous SNPs (Fig 4C, Supplementary Fig 3A), suggesting that in the absence of Rad51, repair in the coding regions is error-prone. In the case of indels, both the total number of indels and non-synonymous indels were drastically reduced (Fig 4D and F). This result indicates that Rad51 plays an important role in the repair of damage across coding and noncoding regions with different outcomes (an increase in SNP accumulations and a decrease in indels), supporting that different compensatory pathways respond to different types of damage in the absence of Rad51. This skew towards the accumulation of mutations in parasite coding regions was also seen in other work where DNA replication or repair proteins were modified or deleted[45,46].

***P. falciparum* exhibits differential DNA repair in the core genome and subtelomeric regions**

The Plasmodium core genome and the subtelomeric regions are considerably different in terms of sequence diversity. While the core genome is largely conserved when comparing different isolates, the subtelomeric regions are highly diverse, suggesting a greater potential for recombination within these regions[4]. Previous work analyzing parasites grown under standard culture conditions found that most small indels occurred in noncoding repetitive regions of the genome, and a relatively small percentage were found in the subtelomeric regions of the chromosome. In contrast, SNPs did not differ in distribution across the parasite chromosomes[45].

We were therefore interested to see if *PfRad51* or the TLS polymerases played any role in this diversification process. We took the SNPs and indels data we obtained from the irradiated parasites and grouped them into those located in the core genome and those in the subtelomeric *var* clusters (Fig 4B and E). We observed that the three knockouts displayed a slight reduction (statistically significant in *PfRad51*Δ) in the number of SNPs in the core genome (Fig 4B). Surprisingly, the subtelomeric regions displayed an opposite phenotype where the SNP accumulation was slightly higher than in the wild type (Fig 4B). In contrast, the loss of *PfRad51* resulted in reduced indels in the core genome (Fig 4E), while the TLS knockouts did not show any significant changes. However, in the subtelomeric regions, the loss of *PfRad51* led to a significant increase in indels (Fig 4E). Our *PfRad51*Δ data showed the opposite of what we saw in 3D7 WT and published data. Sequence analysis from parasite clone trees from various *P. falciparum* isolates found that indels were concentrated in the noncoding repetitive regions of the core genome[2]. Our data implicate *PfRad51* in the repair of damage in repetitive regions in the core genome only. We could not confidently determine if the same scenario was prevalent in the internal *var* clusters due

to the low incidence of SNPs and indels. Of note, the indels analyzed in this experiment are < 5 bases long, and though they could be included in larger recombination events, we instead saw a decrease in large recombinations as detailed below.

***PfRad51* loss abolishes radiation-induced translocations in *P. falciparum*, while TLS polymerases play a minor role in recombination events**

We performed a more detailed analysis of recombination events within the subtelomeric regions and the effects of the loss of HR or TLS pathways. To this end, we analyzed genomic DNA isolated from the three times irradiated clones and performed nanopore sequencing as outlined in Fig 5A. Nanopore sequencing provides read lengths that can extend over entire recombination events, including those within clusters or tandem arrays of multicopy gene families, which was not possible with Illumina sequencing. We expected most translocations to occur at the *var* clusters since they are more prone to recombination[8]. Our analysis revealed that the WT 3D7 exhibits some translocations after irradiation (~3-4 in a clone), with most translocations happening within the sub-telomeric regions (Fig 5B, Supplementary Fig 4), as expected[10]. We observed no verified translocations in a clone of irradiated *PfRad51*Δ (Fig 5B), consistent with the anticipated complete loss of recombination in the absence of HR. However, this *PfRad51*Δ clone displayed a large deletion in the AT rich region of the EMP1-trafficking protein gene (Pf3D7_1002100, Supplementary Fig 5). This deletion in the core genome indicates that alternative Rad51-independent HR repair pathways are present in the parasite, leading to large deletions but not translocations. The loss of TLS polymerases showed a similar phenotype to the WT 3D7, with 3-4 translocations per clone. Our results also show an increased propensity of the subtelomeric *var* clusters to undergo DNA breakage and recombination as most of the recombination events in all the lines happened between two ends of chromosomes, thus facilitating the generation of new *var* gene combinations as previously described [3,7-9,47].

***PfRad51* is essential for DSB repair in *P. falciparum*'s core genome, while TLS polymerases are dispensable**

Based on previous reports and our results showing differences in DNA repair between core and subtelomeric regions (Fig 4F), we tested repair dynamics at the core genome and subtelomeric regions[4]. Our results above were generated from X-ray irradiated parasites where damage was global. Therefore, to study the effects of the loss of *PfRad51* and PfTLS polymerases on DSB repair dynamics at specific loci in the genome, we designed a CRISPR/Cas9-based DSB repair experiment. We targeted the gene PF3D7_1005500, which codes for *PfUpf1*, a component of the nonsense-mediated decay pathway in *P. falciparum* that is non-essential[48]. In addition, we determined that the knockout of *PfUpf1* does not affect parasite growth or any of the DNA repair mechanisms in *P. falciparum* [49]. We added a variable in the distance from our directed cut to the homology blocks provided for repair by creating two plasmids containing repair templates for the *PfUpf1* gene, one with homology blocks that are ~425 bp apart from each other in terms of their location in the genomic DNA, which we called "p*Upf1*Δ" (Fig 6A) and the other with homology blocks that are ~3 kb apart from each other, called "p*Upf1*Δ-distant" (Fig 6C). The donor plasmids also supplied the single guide targeting the *PfUpf1* gene, as depicted in Figs 6A and 6C. Each plasmid was transfected into WT 3D7 and the knockout cell lines - *PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ, with a Cas9 expressing plasmid. Transfections were selected for 4 days with the required drugs for the Cas9 and donor plasmids. Each transfection was repeated 3 times.

In WT parasites and both TLS knockout lines, we observed successful and efficient DSB repair by HR (Fig 6B). Specifically, in the WT, *PfRev1*Δ, and *PfPolζ*Δ strains, we recovered parasites with the predicted double cross-over event, resulting in the replacement of the sequence targeted by both the p*Upf1*Δ and p*Upf1*Δ-distant constructs (Fig 6B, 6D). This was expected as we are targeting a dispensable housekeeping gene, and as *PfRad51* is intact in all these lines, HR

pathway should be functional. However, in the absence of *PfRad51*, the Cas9 induced double-strand break (DSB) proved lethal, with both donor plasmids, the p*Upf1*Δ and p*Upf1*Δ-distant, unable to serve as homology for repair in the *PfRad51*Δ cell line. Thus, *PfRad51* is essential for the repair of DSBs in the core genome, and TLS polymerases have a limited or no role in standard HR in *P. falciparum*.

A *PfRad51*-independent homologous recombination pathway operates in *P. falciparum* subtelomeric regions

To study repair dynamics within the subtelomeric regions, we generated two repair plasmid constructs targeting the *var2csa* locus near the end of chromosome 12: *var2csa*Δ and *var2csa*Δ-distant, similar to the donor plasmids made to study DSB repair in the core genome and shown in Fig 7A and C. We chose *var2csa* as it is more conserved between *P. falciparum* isolates and distinct enough from other *var* genes to reduce the likelihood of ectopic recombination events with other *var* genes[50,51]. The *var2csa*Δ and *var2csa*Δ-distant constructs contained repair templates ~1.5 kb and ~2.5 kb apart in the genome, respectively. In addition, the *var2csa*Δ-distant plasmid had no selection marker between the two homology blocks as opposed to the *var2csa*Δ plasmid (Fig 7, A and C). Similar to the core genome experiments, each plasmid was transfected into the WT 3D7 and the knockout cell lines - *PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ, with a Cas9 expressing plasmid. We selected for the repair plasmid for 4 days after transfection. Again, repair plasmids also supplied the single guide targeting the *var2csa* gene.

In subtelomeric regions, DSBs can be repaired either by *PfRad51* mediated HR or by telomere healing (TH) [8,10,37], a process in which the fragment between the site of the break and the chromosome end is lost, and a telomere is added de novo to the DSB site. Parasites can survive TH as the subtelomeric regions are not essential[8,10,37]. In the WT background, the *var2csa*Δ construct resulted in efficient HR in which the two homology blocks were utilized to repair the DSB (Fig 7B). In contrast, the loss of *PfRad51* abolished HR and forced the parasites to repair the DSB through TH (Fig 7B). As

observed in the core genome, the two TLS polymerase knockouts phenocopied the WT, indicating that the *PfRad51-mediated* HR is functional at all chromosome locations in these knockouts (Fig 7B).

Surprisingly, in the WT background, when the *var2csa*Δ-distant construct was provided as the template for repair, we observed that the repair machinery failed to recognize the homology blocks, resulting in a large (~35kb) resection and TH (Fig 7D). This result indicates that despite a functional HR pathway within the subtelomeric regions, TH outcompetes HR when there is a significant distance between homology blocks. In contrast, TH is not an option in the core genome, and therefore, HR proceeded as expected even when the distance between homology blocks was increased (Fig 6D). More surprisingly, we observed homology-mediated repair of a subtelomeric DSB in the *PfRad51*Δ line despite the loss of the HR pathway. This was unexpected and also observed in the two TLS knockouts (Fig 7D) indicating that when the distance between the two homology blocks was greater than 2 kb, in the absence of *PfRad51* or the two TLS polymerases, we saw the activation of an alternate Rad51-independent pathway that can perform homology-directed repair specifically within the subtelomeric region. In the absence of Rad51 or the TLS polymerases DSB repair was skewed towards a different result than what was seen in the WT 3D7. When we examined the individual reads that spanned the DSB site in all three knockouts, we observed a recombination mechanism that identified the first homology block but initially failed to recognize the second homology block and instead replicated the entire plasmid backbone multiple times before using the second homology block to complete recombination (Supplementary Fig 6). This resulted in the incorporation of multiple copies of the plasmid at the site of repair in all three of the knockout lines when our DSB targeted a subtelomeric location. The same result was seen in multiple repeated transfection experiments (3 transfections each). Previously, others have described similar integration events in non-CRISPR mediated genetic modifications in Plasmodium. It is distinct, however, that with the same guide and donor plasmid, we

noted very different repair events in our WT versus knockout parasite lines. Notably, we did not observe this type of repair within the core genome.

Discussion

We set out to identify mechanisms of malaria genome diversification with a particular interest in how DNA repair mechanisms would differ at different chromosome locations in *P. falciparum* and are now just beginning to understand how the parasite segregates its genome into a well conserved core genome and highly divergent subtelomeric regions[4]. We previously proposed that interactions between the TLS pathway and HR in malaria were critical to diversifying the parasite genome, and now, in this current work, we have pursued the mechanism behind these interactions and parasite genome diversification[17].

var gene diversity has been described by our lab and others, [9,47,52-54] and recent work is starting to address the possible mechanisms that underlie *var* gene diversification [7,8,17]. Analysis of recombination events showed that, in general, *var* gene recombination maintained open reading frames and preserved the domain architecture, resulting in functional chimeric *var* genes. This is consistent with gene conversion and NAHR as the primary mechanisms to generate novel *var* gene sequences during mitotic replication[7]. This idea is further supported by work describing recombination identified in parasites subjected to irradiation and targeted DSBs via CRISPR/Cas9[8,10]. Both modalities of DNA damage yielded subtelomeric gene recombination consistent with a cascade of ectopic HR events leading to the formation of chimeric *var* genes[8,10]. To date, these ectopic HR events have been primarily observed in the subtelomeric *var* clusters where the repair choice between TH and HR drives diversification[10]. We have yet to understand alternative diversification events that occur amongst the five *var* gene clusters located internally on chromosomes 4, 7, 8, and 12. HR is historically considered a non-diversifying DNA repair

pathway that maintains the overall gene structure, and we propose that this very characteristic is adapted by the parasite to recombine and produce chimeric *var* genes by using the redundancy in *var* gene sequences.

One of the most striking observations is that *PfRad51*, *PfRev1*, and *PfPolζ* are not essential for the survival of parasites and that there was no growth defect under standard culture conditions, even given the replication stress the parasite is under during erythrocytic development as it replicates from one ring to 30-50 infectious merozoites (Supplementary Fig 2). This contrasts with the essentiality of Rad51 in vertebrates[23] but aligns more closely with observations in lower eukaryotes such as yeast [24,26]. The stage-specific differential sensitivity to irradiation between the ring and late stages is primarily seen in the TLS knockout parasite lines and reveals stage-specific DNA repair mechanisms in *P. falciparum* similar to dominance of different DNA repair pathways at different stages of the cell cycle in higher eukaryotes[55] (Fig 2A, 0-24 hours post-invasion - hpi, and Fig 2C). Particularly noteworthy is the essential role of TLS polymerases only in ring-stage parasites' survival following irradiation (Fig 2C). Ring stage parasites are haploid and are not undergoing DNA replication and, therefore can be considered analogous to the G1 phase in higher eukaryotes such as fission yeast[56]. In the late stages (Fig 2A, ~24-36 hpi), which could be compared to the S-phase of higher eukaryotes [56], HR appears to dominate repair due to the availability of a second copy of the DNA, and in this setting, the TLS polymerases are dispensable. The differences in stage-specific requirements of TLS polymerases highlight the complex and dynamic nature of DNA repair in the parasite.

Since *P. falciparum* lacks the error-prone NHEJ pathway, we predicted that the subsequent loss of *PfRad51* would result in a large majority of DNA breaks being unrepaired and thus lethal. The slightly longer recovery for later stage *PfRad51*Δ parasites was not as pronounced as expected

and suggests that PfRad51 may have a non-HR related role in the ring stage parasite response to DNA damage that is worth exploring in future work. Analysis of whole genome sequencing (WGS) data or our irradiated knockout parasite did not highlight the dominance of an alternative DSB repair pathway to step in for HR. We found one example of a large deletion between repetitive regions consistent with a Rad51 independent HR termed Synthesis Dependent Strand Annealing (SDSA) which is characterized by deletions within a repetitive region and mediated by Rad52. However, to date, no Rad52 ortholog has been identified encoded in the parasite genome. Nor did we see an increase in insertions or deletions consistent with MMEJ, which has been described in *P. falciparum* but appears to be limited even in the setting of no C-NHEJ or HR (*PfRad51*Δ).

Irradiation results in large-scale genome damage that can lead to the accumulation of mutations in the form of SNPs, indels, and translocations. The decrease in SNPs in the TLSΔ lines was expected and is consistent with TLS polymerases' central role in the DNA damage response and knockouts leading to a hypomutable state. For Plasmodium, the increase in SNPs in our data suggests a more error prone alternative pathway is available to the parasite in the absence of *PfRad51*. An increase in SNPs in a Rad51Δ has been noted in other organisms such as Leishmania[57]. Interestingly, in engineered parasite lines missing the editing function of Polδ (created in both *P. falciparum* and the rodent malaria *P. berghei),* elevated mutation rates were seen and were increased specifically in coding regions[45,46]. In addition, the mutation rates reported were significantly higher in the *P. berghei* Polδ mutant. All rodent parasites lack the TLS pathway, which may have a role in the differential mutation rate between Plasmodium species[17]. When we exposed parasites to DNA damage in our system, the TLS Δ parasites could not survive and were, in contrast, hypomutable.

Genome analysis of parasites grown in standard culture conditions and then subcloned identified similar mutation rates between *P. falciparum* isolates of different geographical origins and an even distribution of SNPs across coding and noncoding regions[2,7,9]. This work also noted a predominance of new indels in noncoding AT-rich repetitive sequences in the parasite core genome[2]. The authors proposed that the sequence of exons and noncoding regions were associated with distinct forms of indel mutations. Our data show that the loss of *PfRad51* resulted in a significant reduction in the number of indels in the core genome, and thus, Rad51 is likely to play a role in indel formation. Interestingly, we did find that the subtelomeric regions had an increase in the number of indels in *PfRad51*Δ compared to the WT (Fig 4E), the opposite of what is seen in WT parasites[2]. This difference suggests that different DNA repair proteins may respond to damage at different chromosomal locations, perhaps due to epigenetic or other factors that might maintain a demarcation between core vs subtelomeric regions of the parasite chromosomes.

Our results show that *PfRad51* is important for structural variations (Fig 5B). We proposed that increased ectopic recombination occurring specifically in the subtelomeric regions plays an important role in this diversification [8,10]. This is also supported by others who observed that while SNPs were scattered across the genome, SVs were focused in and around *var* genes[7]. Hamilton et al. also found that while the SNP rate did not differ between parasite isolates, the frequency of *var* gene recombination through NAHR did, with more recent clinical isolates demonstrating a higher rate of chimeric *var* gene formation[2]. Potential influences, such as *var* gene transcription, remain to be studied. We utilized CRISPR/Cas9 directed cuts at a specific locus in the core genome and one of the *var* clusters in the subtelomeric region to determine the repair dynamics and how the loss of *PfRad51* and the two TLS polymerases affected DSB repair dynamics (Fig 6 and 7). Our results show a clear difference in the choice of DNA repair mechanisms depending on whether the

break occurs within the core genome or a subtelomeric region. In subtelomeric regions, parasites can either repair a DSB via HR or stabilize the chromosome end by TH, consistent with our results. In addition, the free fragment formed after the DSB within a subtelomeric region has been shown to be highly recombinogenic and lead to ectopic recombinations readily [58]. This would also explain why the multicopy gene families are clustered in the subtelomeric regions in apicomplexans, thus enabling a higher degree of spontaneous diversification.

Further, our experiments with the distant construct, which we expected to function as a 'non-perfect' repair template, showed a clear differential DNA repair between core vs subtelomeric region (Fig 6D and 7D). While *PfRad51* is essential for repairing DSBs in the core genome (Fig 6B), it is not essential in subtelomeric regions (Fig 7B) as the parasites can choose TH over HR. However, in the subtelomeric region when the distant repair template was provided, the choice of repair was HR, and to our surprise, this HR was *PfRad51* independent (Fig 7D). This difference between core and subtelomeric regions might contribute to the highly diverse nature of the subtelomeric regions. The question that remains unanswered is, how do parasites perform HR, specifically in the subtelomeric regions in the absence of Rad51. Further exploration into proteins that play a role in this process has the potential to identify novel mechanisms that enable human malaria parasites to diversify the *var* repertoire and thereby evade the human immune response more efficiently by switching to novel, chimeric *var* genes.

The *P. falciparum* genome is unique in many aspects and can serve as a model for the study of DNA repair under high oxidative and high replicative stress in a haploid organism that lacks NHEJ. The evolution and regulation of DNA repair pathways in the malaria parasite can serve as guides not just to antigenic variation and drug resistance in this parasite of global importance but also to DNA repair mechanisms in other eukaryotes.

Materials and Methods

Culture and genetic manipulation of parasites

Parasites were cultured in a standard culture system at 5% hematocrit in media containing RPMI 1640 (Corning Life Sciences, Tewksbury, MA, USA), 0.5% Albumax II (Invitrogen, Carlsbad, CA, USA), 0.25% sodium bicarbonate, and 0.1 mg/ml gentamicin in an atmosphere of 5% oxygen, 5% carbon dioxide, and 90% nitrogen. Culture media was supplemented with the required selections (WR, Blasticidin (BSD), or DSM1) as per the genotype of the cell lines. Clonal parasite lines were obtained by limiting dilution [59]. Parasites were transfected by electroporation[60] using derivatives of the plasmids pL6_eGFP and pUF1_Cas9 for CRISPR/Cas9-based genome editing as described by Ghorbal and colleagues[28]. Homology blocks for genome modification were amplified from parasite genomic DNA by PCR, single guide sequences were synthesized as oligos, annealed, and both were inserted into pL6 by infusion cloning (Clontech, Takara Bio USA, Mountain View, CA, USA). The plasmids generated are listed in Supplementary Table 1. Single guide sequences and sequences used as homology blocks for CRISPR modifications are listed in Supplementary Tables 2 and 3. Growth of the parasites was monitored using flow cytometry as described before[61,62]. The parasite cell lines used in this study are listed in Supplementary Table 4.

Genomic DNA extraction

Infected red blood cells from a 10-ml culture (~4-5% late stages) were centrifuged at 4,000 rpm. The supernatant was discarded, and the pellet was resuspended in 1 ml of phosphate-buffered saline and 15 µl 10% saponin. The parasite pellets were spun at 12,000 rpm, 3 mins, and washed twice in 1 ml of PBS. The pellet was then taken up in 400 µl Tris-sodium chloride-EDTA buffer, along with 80 µl of 10% SDS and 40 µl of 6 M NaClO4. This suspension was placed on a rocker at room temperature overnight. The following morning, an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) was added and gently vortexed (for Nanopore sequencing samples vortexing was skipped to maintain the length of the DNA and gently mixed by inverting the tube 5-10 times) and centrifuged at 16,000g for 5 min. The final aqueous phase was ethanol precipitated and resuspended in 50 µl of sterile distilled H2O. The final DNA concentration was determined by absorbance at 260 nm using a nanodrop machine.

Sequencing of parasite genomic DNA

Parasite genomic DNA was extracted using phenol-chloroform extraction, and sequencing libraries were prepared from 1 µg of unsheared gDNA (as measured using the Qubit dsDNA BR assay [Life Technologies, Carlsbad, CA, USA]) using the Oxford Nanopore Technologies ligation sequencing kit version 1D LSK 108 (Oxford, UK). This approach minimizes the potential shearing of DNA template fragments and entirely avoids PCR artifacts by ligation of a manufacturer-supplied adapter nucleoprotein complex to native gDNA to facilitate the loading of long library molecules into the protein nanopores. Nanopore sequencing was conducted with an Oxford

Nanopore Technologies GridION X5 instrument using MIN-106/R9.4.1 flowcells. All samples were sequenced using a standard 48 hr run controlled by Oxford Nanopore Technologies MinKnow software, and raw data were basecalled with Albacore version 2.2.4. FASTQ files were extracted from basecalled FAST5 files with Poretools [63].

Illumina whole genome sequencing was performed using NovaSeq 6000 system to get about 5 million reads per sample. Sample library was prepared using Kapa Hyper library prep kit (Roche) and run on an S2 flow cell with 2x100 cycles to obtain paired-end reads in the fastq format. All the sequencing data can be accessed at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1244413

Irradiation of parasites

Irradiation of the parasites was performed using RS 2000 Biological Research X-ray Irradiator (Rad Source Technologies). Briefly, parasite cultures were exposed to the required amount of irradiation utilizing the level 3 shelf, which supplied ~2.1 Gy/min. The culture was maintained by media change and by replenishing blood until the parasites that survived the irradiation came up. Once up, the parasites were irradiated again for the second time and repeated for a third time. After the third round of irradiation, the gDNA from the surviving parasites were isolated as mentioned above using the PCIA extraction, purified using Zymoresearch genomic DNA clean and concentrator kit (Cat.no - D4010), and Illumina sequenced as above for quantifying SNPs, Indels or used for determining *var* gene copy number variations using qPCR.

Quantification of SNPs, Indels, and translocations

To quantify SNPs and Indels from the Illumina sequencing datasets, reads were trimmed using Trimmomatic [64] to remove adapters, read quality was checked using FastQC[65], and the genome was indexed and aligned using BWA-MEM[66]. Duplicates were marked using sambamba [67] and the bam file was sorted and indexed using samtools[68]. Variants were called using freebayes[38] and the vcftools package was used to filter out variants having quality (QUAL) scores less than 30 and DP scores less than 10. To determine the translocations from nanopore reads, bam files from the raw fastq files were generated using minimap2[69], sorted and indexed using samtools[68], and variants called using sniffles[70]. SNPs, Indels, and translocations unique to the irradiated cell lines were identified using the vcf-isec command of vcftools [71] and were plotted into circos plots using the R package circlize[72]. Translocations identified by sniffles were visually confirmed using IGV[73] before circos plotting. All the codes for the analysis performed can be accessed at GitHub here: https://github.com/agnibhat/Characterization-of-Plasmodium-falciparum-DNA-repair-dynamics.

Figures

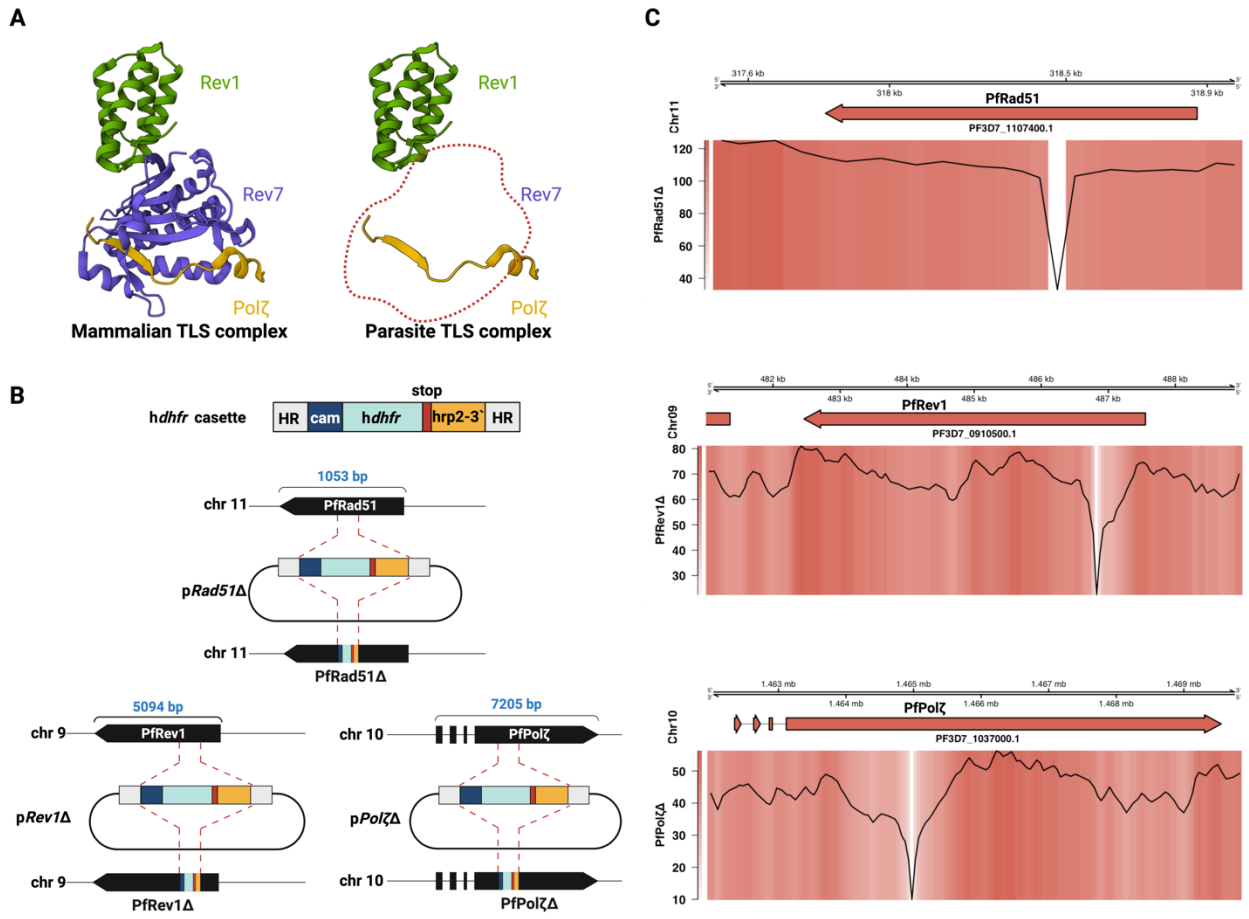

Figure 1: Construction and confirmation of knockouts of *PfRad51*, *PfRev1*, and *PfPolζ*. (A) Comparison between mammalian and parasite TLS complex (PDB:3VU7). The red dotted line indicates that the Rev7 subunit has not been identified to date in parasites. (B) Overview of the molecular approach used for creating the knockouts of *PfRad51*, *PfRev1,* and *PfPolζ*. The top cartoon shows the hDHFR (Dihydrofolate Reductase) cassette used for integration and selection. (C) Coverage plots of Nanopore sequencing results of the *PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ. The top panel denotes the genome region, chromosome, and gene. The bottom panel has both heatmap and lineplot to show the integration site where the hDHFR cassette has been integrated, which can be seen as a drastic drop in the coverage (Large dip in the lineplot coinciding with the white region in the coverage plot).

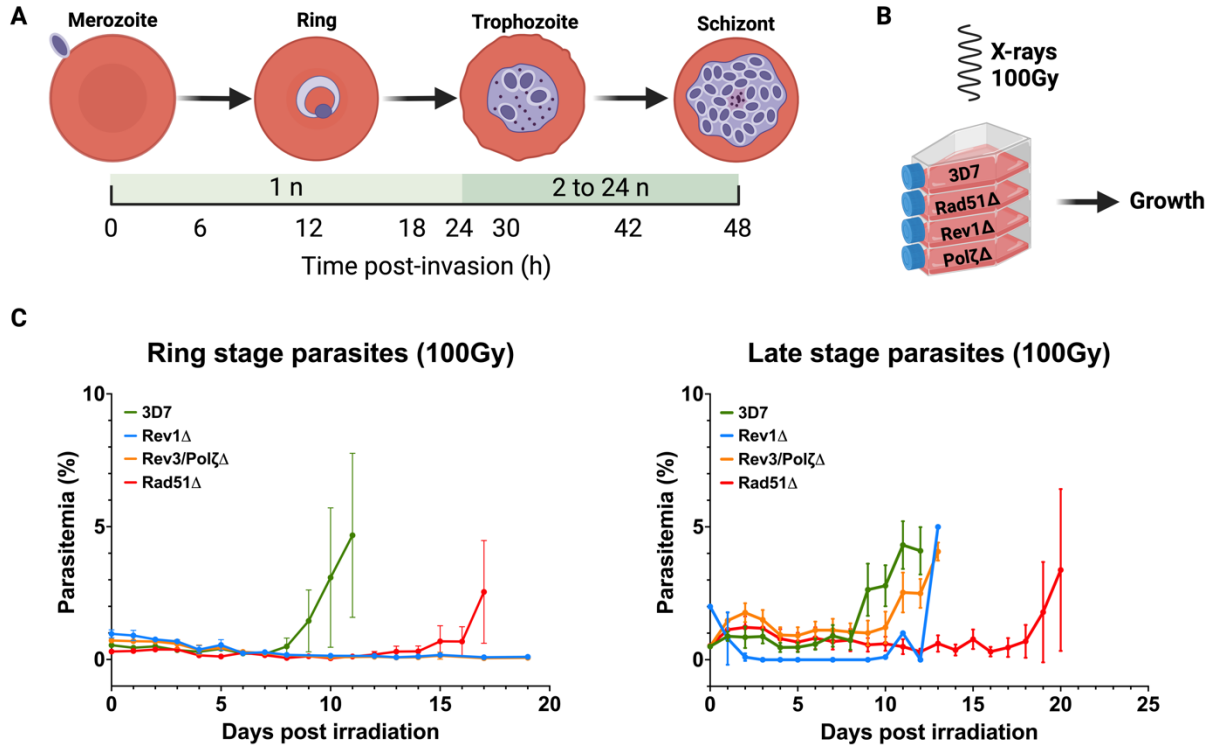

Figure 2: Growth analysis of knockout cell lines. (A) The top panel represents the main stages of the intraerythrocytic developmental cycle (IDC) of *P. falciparum*. The bottom panel shown the change in DNA content over the different stages of the IDC tracked as hours post-invasion (hpi). (B) Overview of the irradiation exposure of the 4 different cell lines - WT, *PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ. (C) Growth curve of 100 Gy irradiated ring stage (0-24 hpi) and late stage (24-48 hpi) parasites.

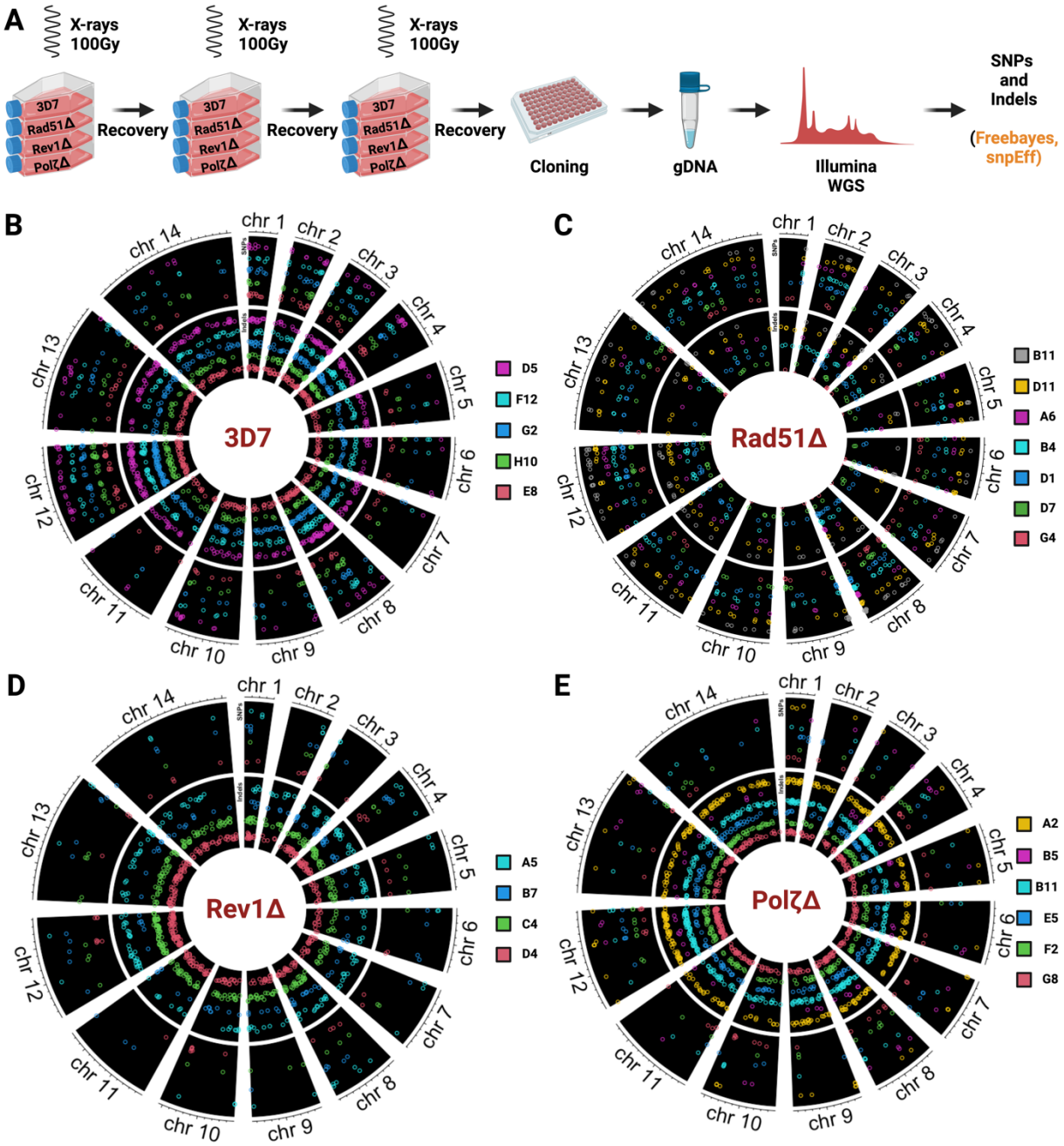

Figure 3: Irradiation induced SNPs and Indels. (A) Overview of irradiation experiment to analyze SNPs and Indels using Illumina whole genome sequencing. (B-E) Circos plots of SNPs (outer arcs) and Indels (inner arcs) distributed throughout the genome of *P. falciparum* cell lines – (B) WT-3D7, (C) *PfRad51*Δ, (D) *PfRev1*Δ, and (E) *PfPolζ*Δ after 3x irradiation with 100 Gy. The different colors represent the clones, and the arcs represent the chromosomes.

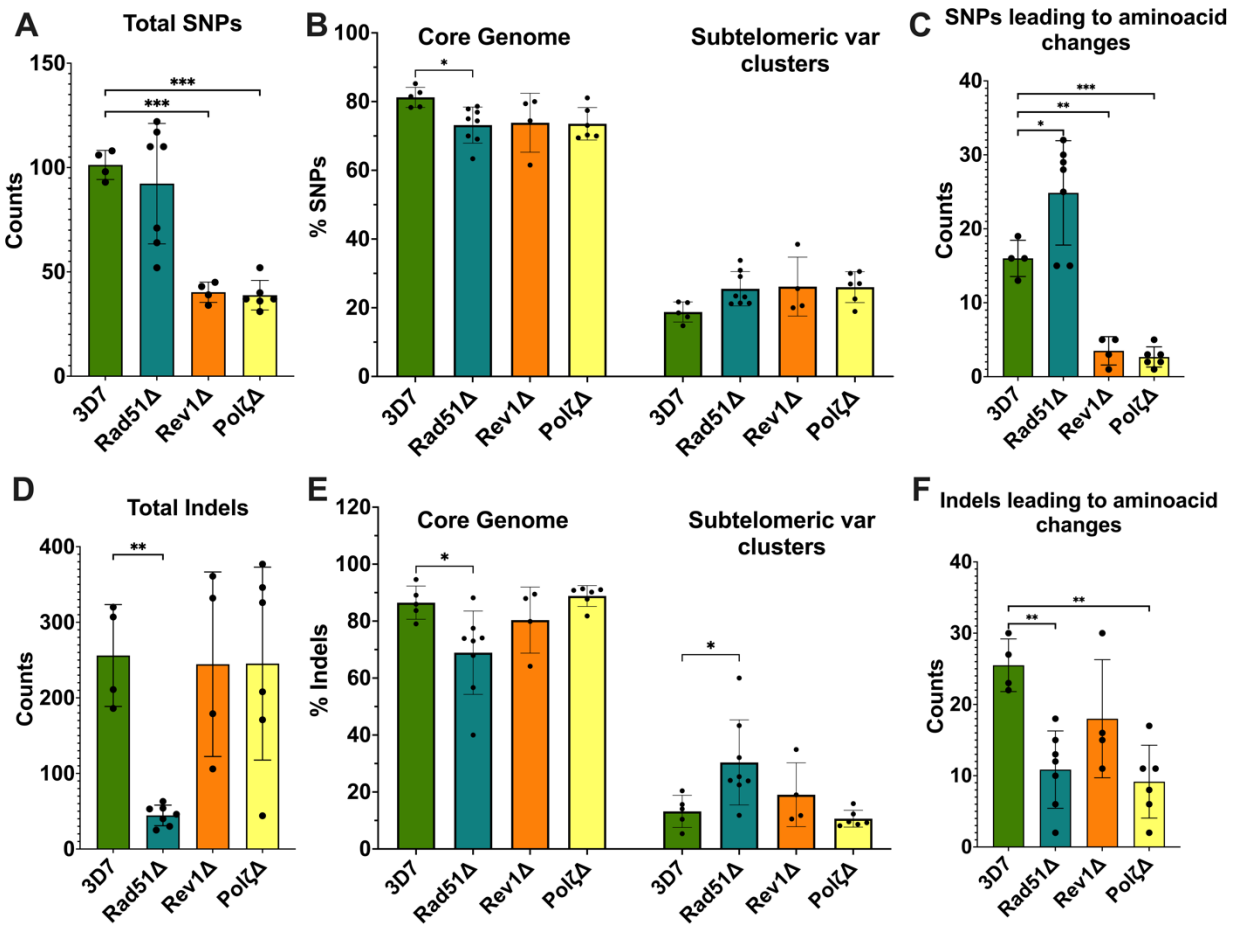

Figure 4: Analysis of SNPs and Indels determined from Figure 3: Quantification of total SNPs (A) and Non-synonymous SNPs (C); Quantification of total Indels (D) and Non-synonymous Indels (F) from Figure 3. Statistical significance was determined using ordinary one-way ANOVA followed by Šídák's multiple comparisons test *(p=0.05-0.01), **(p=0.01-0.001),***(p=0.001-0.0001). Distribution of SNPs (B) and Indels (E) across the core genome and the *var* clusters. Statistical significance was determined using Two-way ANOVA followed by Dunnett's multiple comparison test *(p=0.05-0.01), **(p=0.01-0.001), ***(p=0.001-0.0001).

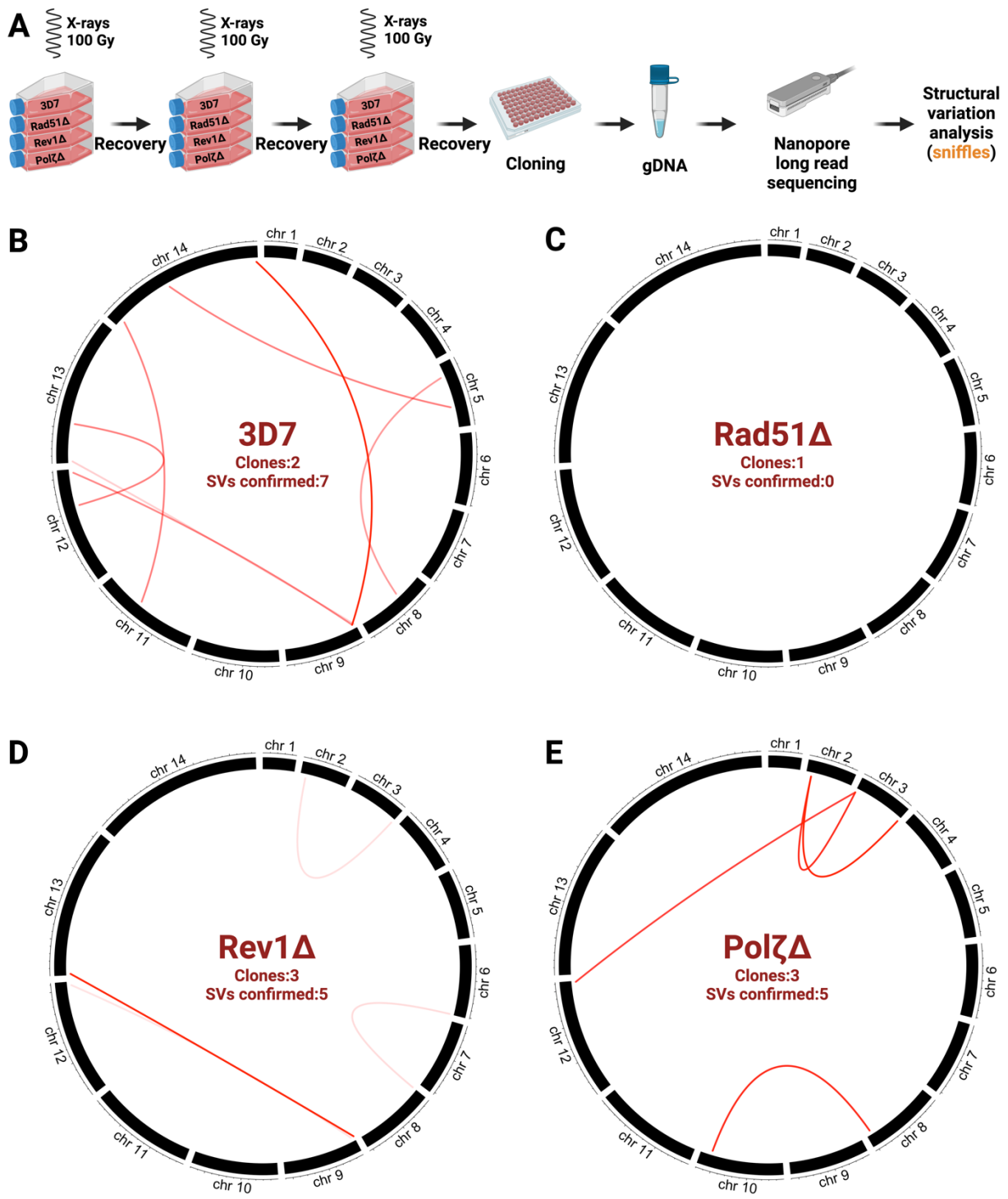

Figure 5: Irradiation induced translocations/recombination. (A) Overview of the experimental setup for the determination of recombination events. The translocations were confirmed by visual confirmation before plotting as described in methods. (B-E) Circos plots of translocations/recombinations distributed throughout the genome of *P. falciparum* cell lines – (B) WT-3D7, (C) *PfRad51*Δ, (D) *PfRev1*Δ, and (E) *PfPolζ*Δ after 3x irradiation with 100 Gy. Each line represents a recombination event between two genomic loci identified by the breakpoints. The

visual IGV screenshots of some translocations in the WT-3D7 cell line have been described in supplementary Fig 3.

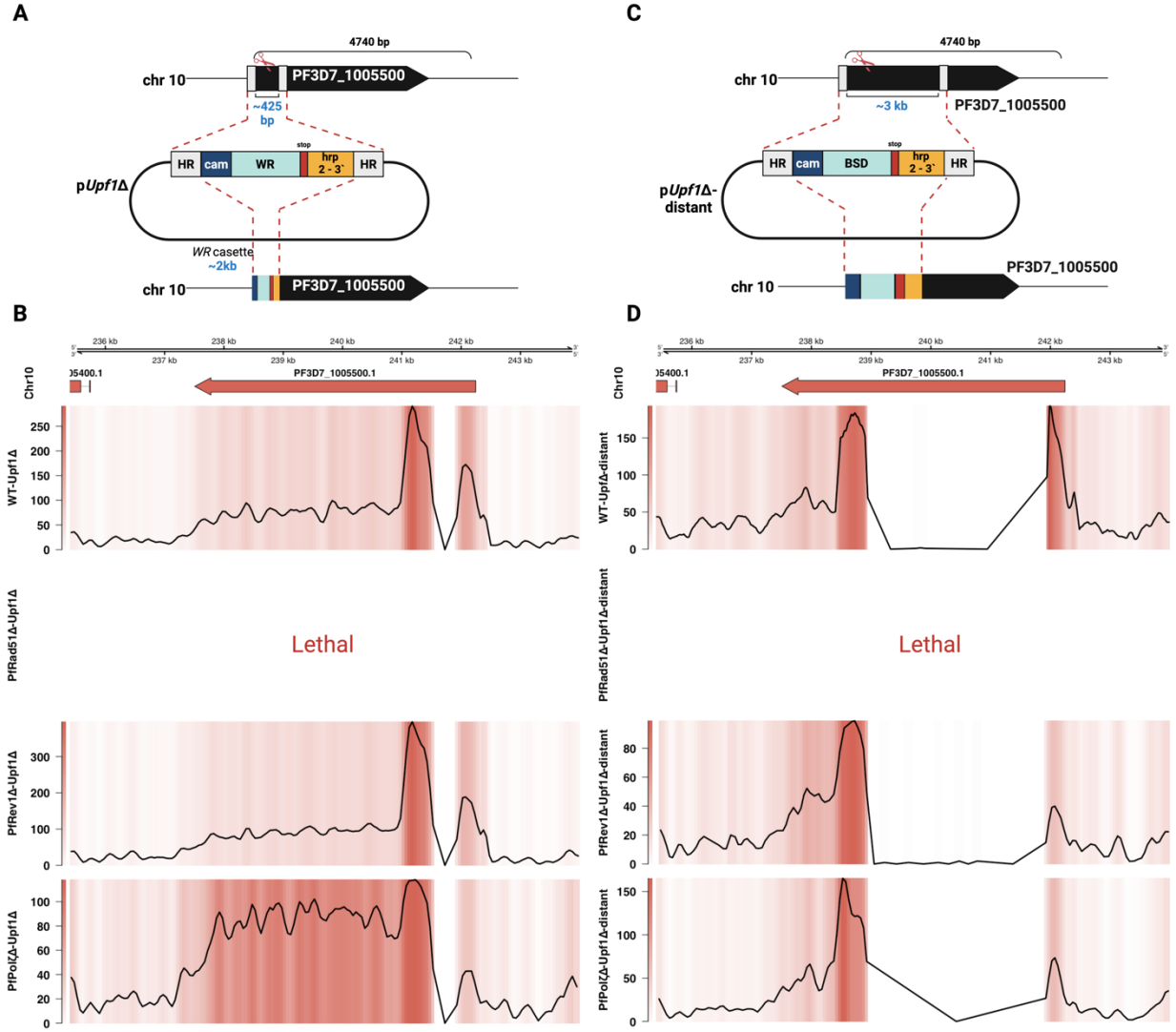

Figure 6: CRISPR/Cas9 mediated cuts to study homologous recombination in core genome of *P. falciparum*. (A and C) Overview of construction of *PfUpf1*Δ and *PfUpf1*Δ-distant plasmid constructs. (B and D) Representative coverage plots of Nanopore/Illumina sequencing results of *PfUpf1*Δ and *PfUpf1*Δ-distant in WT-3D7, *PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ. The top panel denotes the genome region, chromosome and the gene. The bottom panel has both heatmap and lineplot to show the integration site where the hDHFR/Blasticidin resistance cassette has been integrated which can be seen as a drastic drop in the coverage (Large dip in the lineplot coinciding with the white region in the coverage plot).

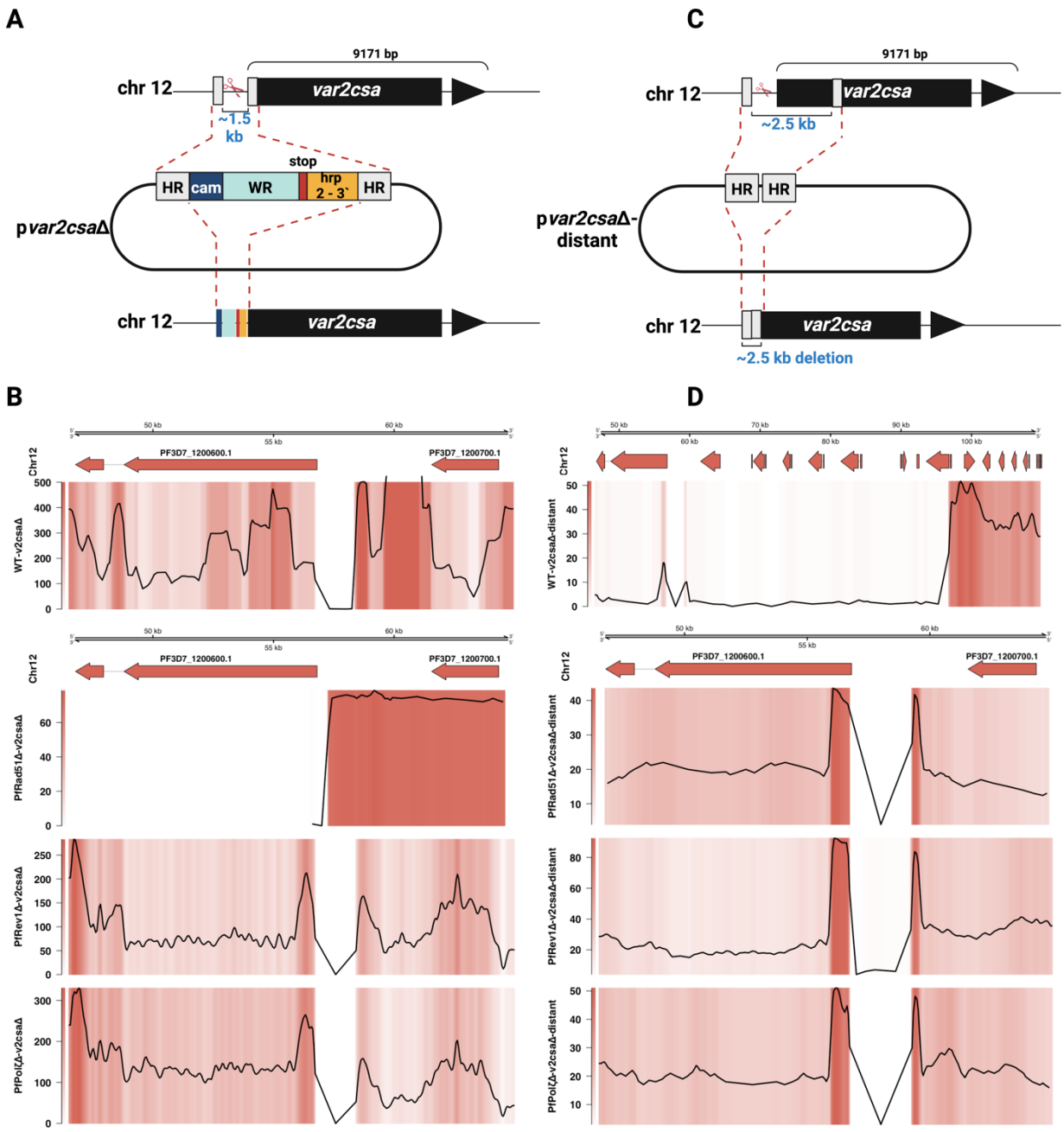

Figure 7: CRISPR/Cas9 mediated cuts to study homologous recombination in subtelomeric regions of *P. falciparum*. (A and C) Overview of construction of *Pfvar2csa*Δ and *Pfvar2csa*Δ-distant plasmid constructs. (B and D) Representative coverage plots of Nanopore/Illumina sequencing results of *Pfvar2csa*Δ and *Pfvar2csa*Δ-distant in WT-3D7, *PfRad51*Δ, *PfRev1*Δ, and *PfPolζ*Δ. Note that var2csa is abbreviated as v2csa in the figures. The top panel denotes the genome region, chromosome and gene. The bottom panel has both heatmap and lineplot to show the integration site where the hDHFR cassette has been integrated or the ~2.5kb DNA piece (as shown in C) has been lost due to HR, which can be seen as a drastic drop in the coverage (Large dip in the lineplot coinciding with the white region in the coverage plot). Please note that in D, the

WT- *Pfvar2csa*Δ-distant sequencing data shows a larger genomic region compared to the other samples to efficiently cover the large resection of ~35kb leading to TH.

References

1. World malaria report 2023.
2. Hamilton, W.L., Claessens, A., Otto, T.D., Kekre, M., Fairhurst, R.M., Rayner, J.C., and Kwiatkowski, D. (2017). Extreme mutation bias and high AT content in Plasmodium falciparum. Nucleic Acids Res *45*, 1889-1901. 10.1093/nar/gkw1259.
3. Kirkman, L.A., and Deitsch, K.W. (2014). Recombination and Diversification of the Variant Antigen Encoding Genes in the Malaria Parasite Plasmodium falciparum. Microbiol Spectr *2*. 10.1128/microbiolspec.MDNA3-0022-2014.
4. Otto, T.D., Böhme, U., Sanders, M., Reid, A., Bruske, E.I., Duffy, C.W., Bull, P.C., Pearson, R.D., Abdi, A., Dimonte, S., et al. (2018). Long read assemblies of geographically dispersed Plasmodium falciparum isolates reveal highly structured subtelomeres. Wellcome Open Res *3*, 52. 10.12688/wellcomeopenres.14571.1.
5. Chen, D.S., Barry, A.E., Leliwa-Sytek, A., Smith, T.-A., Peterson, I., Brown, S.M., Migot-Nabias, F., Deloron, P., Kortok, M.M., Marsh, K., et al. (2011). A Molecular Epidemiological Study of var Gene Diversity to Characterize the Reservoir of Plasmodium falciparum in Humans in Africa. PLOS ONE *6*, e16629. 10.1371/journal.pone.0016629.
6. Barry, A.E., Leliwa-Sytek, A., Tavul, L., Imrie, H., Migot-Nabias, F., Brown, S.M., McVean, G.A.V., and Day, K.P. (2007). Population Genomics of the Immune Evasion (var) Genes of Plasmodium falciparum. PLOS Pathogens *3*, e34. 10.1371/journal.ppat.0030034.
7. Claessens, A., Hamilton, W.L., Kekre, M., Otto, T.D., Faizullabhoy, A., Rayner, J.C., and Kwiatkowski, D. (2014). Generation of Antigenic Diversity in Plasmodium falciparum by Structured Rearrangement of Var Genes During Mitosis. PLOS Genetics *10*. 10.1371/journal.pgen.1004812.
8. Zhang, X., Alexander, N., Leonardi, I., Mason, C., Kirkman, L.A., and Deitsch, K.W. (2019). Rapid antigen diversification through mitotic recombination in the human malaria parasite Plasmodium falciparum. PLoS Biol *17*. 10.1371/journal.pbio.3000271.
9. Bopp, S.E.R., Manary, M.J., Bright, A.T., Johnston, G.L., Dharia, N.V., Luna, F.L., McCormack, S., Plouffe, D., McNamara, C.W., Walker, J.R., et al. (2013). Mitotic evolution of Plasmodium falciparum shows a stable core genome but recombination in antigen families. PLOS Genetics *9*, e1003293. 10.1371/journal.pgen.1003293.
10. Calhoun, S.F., Reed, J., Alexander, N., Mason, C.E., Deitsch, K.W., and Kirkman, L.A. (2017). Chromosome End Repair and Genome Stability in Plasmodium falciparum. mBio *8*. 10.1128/mBio.00547-17.
11. Gopalakrishnan, A.M., and Kumar, N. (2015). Antimalarial action of artesunate involves DNA damage mediated by reactive oxygen species. Antimicrobial Agents and Chemotherapy *59*, 317-325. 10.1128/AAC.03663-14.
12. Xiong, A., Prakash, P., Gao, X., Chew, M., Tay, I.J.J., Woodrow, C.J., Engelward, B.P., Han, J., and Preiser, P.R. (2020). K13-Mediated Reduced Susceptibility to Artemisinin in Plasmodium falciparum Is Overlaid on a Trait of Enhanced DNA Damage Repair. Cell Rep *32*, 107996. 10.1016/j.celrep.2020.107996.
13. Gupta, D.K., Patra, A.T., Zhu, L., Gupta, A.P., and Bozdech, Z. (2016). DNA damage regulation and its role in drug-related phenotypes in the malaria parasites. Scientific Reports *6*, 23603. 10.1038/srep23603.

14. Sheriff, O., Yaw, A., Lai, S.K., Loo, H.L., Sze, S.K., and Preiser, P.R. (2021). Plasmodium falciparum replication factor C subunit 1 is involved in genotoxic stress response. Cellular Microbiology *23*, e13277. 10.1111/cmi.13277.
15. Sale, J.E. (2013). Translesion DNA synthesis and mutagenesis in eukaryotes. Cold Spring Harbor Perspectives in Biology *5*, a012708. 10.1101/cshperspect.a012708.
16. Chatterjee, N., and Walker, G.C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. Environmental and Molecular Mutagenesis *58*, 235-263. 10.1002/em.22087.
17. Siao, M.C., Borner, J., Perkins, S.L., Deitsch, K.W., and Kirkman, L.A. (2020). Evolution of Host Specificity by Malaria Parasites through Altered Mechanisms Controlling Genome Maintenance. mBio *11*. 10.1128/mBio.03272-19.
18. Otto, T.D., Bohme, U., Jackson, A.P., Hunt, M., Franke-Fayard, B., Hoeijmakers, W.A., Religa, A.A., Robertson, L., Sanders, M., Ogun, S.A., et al. (2014). A comprehensive evaluation of rodent malaria parasite genomes and gene expression. BMC Biol *12*, 86. 10.1186/s12915-014-0086-0.
19. Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., et al. (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. Nature *419*, 498-511. 10.1038/nature01097.
20. Kirkman, L.A., Lawrence, E.A., and Deitsch, K.W. (2014). Malaria parasites utilize both homologous recombination and alternative end joining pathways to maintain genome integrity. Nucleic Acids Research *42*, 370-379. 10.1093/nar/gkt881.
21. Singer, M., Marshall, J., Heiss, K., Mair, G.R., Grimm, D., Mueller, A.K., and Frischknecht, F. (2015). Zinc finger nuclease-based double-strand breaks attenuate malaria parasites and reveal rare microhomology-mediated end joining. Genome Biol *16*, 249. 10.1186/s13059-015-0811-1.
22. Bonilla, B., Hengel, S.R., Grundy, M.K., and Bernstein, K.A. (2020). RAD51 gene family structure and function. Annual Review of Genetics *54*, 25-46. 10.1146/annurev-genet-021920-092410.
23. Wassing, I.E., and Esashi, F. (2021). RAD51: Beyond the break. Seminars in Cell & Developmental Biology *113*, 38-46. 10.1016/j.semcdb.2020.08.010.
24. Symington, L.S. (2002). Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. Microbiol Mol Biol Rev *66*, 630-670, table of contents. 10.1128/MMBR.66.4.630-670.2002.
25. McCulloch, R., and Barry, J.D. (1999). A role for RAD51 and homologous recombination in Trypanosoma brucei antigenic variation. Genes & Development *13*, 2875-2888.
26. Muris, D.F., Vreeken, K., Carr, A.M., Broughton, B.C., Lehmann, A.R., Lohman, P.H., and Pastink, A. (1993). Cloning the RAD51 homologue of Schizosaccharomyces pombe. Nucleic Acids Research *21*, 4586-4591. 10.1093/nar/21.19.4586.
27. Furuya, K., and Niki, H. (2010). The DNA damage checkpoint regulates a transition between yeast and hyphal growth in Schizosaccharomyces japonicus. Molecular and Cellular Biology *30*, 2909-2917. 10.1128/MCB.00049-10.
28. Ghorbal, M., Gorman, M., Macpherson, C.R., Martins, R.M., Scherf, A., and Lopez-Rubio, J.-J. (2014). Genome editing in the human malaria parasite Plasmodium falciparum using the CRISPR-Cas9 system. Nature Biotechnology *32*, 819-821. 10.1038/nbt.2925.
29. Jansen, J.G., Tsaalbi-Shtylik, A., Langerak, P., Calléja, F., Meijers, C.M., Jacobs, H., and de Wind, N. (2005). The BRCT domain of mammalian Rev1 is involved in regulating DNA translesion synthesis. Nucleic Acids Research *33*, 356-365. 10.1093/nar/gki189.

30. Simpson, L.J., and Sale, J.E. (2003). Rev1 is essential for DNA damage tolerance and non-templated immunoglobulin gene mutation in a vertebrate cell line. The EMBO Journal *22*, 1654-1664. 10.1093/emboj/cdg161.
31. Deshpande, G.P., Hayles, J., Hoe, K.-L., Kim, D.-U., Park, H.-O., and Hartsuiker, E. (2009). Screening a genome wide S. pombe deletion library identifies novel genes and pathways involved in the DNA damage response. DNA Repair *8*, 672-679. 10.1016/j.dnarep.2009.01.016.
32. Gerald, N., Mahajan, B., and Kumar, S. (2011). Mitosis in the Human Malaria Parasite Plasmodium falciparum. Eukaryotic Cell *10*, 474-482. 10.1128/EC.00314-10.
33. Liffner, B., and Absalon, S. (2021). Expansion Microscopy Reveals Plasmodium falciparum Blood-Stage Parasites Undergo Anaphase with A Chromatin Bridge in the Absence of Mini-Chromosome Maintenance Complex Binding Protein. Microorganisms *9*, 2306. 10.3390/microorganisms9112306.
34. Guttery, D.S., Zeeshan, M., Holder, A.A., and Tewari, R. (2024). The molecular mechanisms driving Plasmodium cell division. Biochemical Society Transactions *52*, 593. 10.1042/BST20230403.
35. Lee, A.H., Symington, L.S., and Fidock, D.A. (2014). DNA Repair Mechanisms and Their Biological Roles in the Malaria Parasite Plasmodium falciparum. Microbiol. Mol. Biol. Rev. *78*, 469-486. 10.1128/MMBR.00059-13.
36. Nakano, T., Akamatsu, K., Tsuda, M., Tujimoto, A., Hirayama, R., Hiromoto, T., Tamada, T., Ide, H., and Shikazono, N. (2022). Formation of clustered DNA damage in vivo upon irradiation with ionizing radiation: Visualization and analysis with atomic force microscopy. Proceedings of the National Academy of Sciences of the United States of America *119*, e2119132119. 10.1073/pnas.2119132119.
37. Reed, J., Kirkman, L.A., Kafsack, B.F., Mason, C.E., and Deitsch, K.W. (2021). Telomere length dynamics in response to DNA damage in malaria parasites. iScience *24*, 102082. 10.1016/j.isci.2021.102082.
38. Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv.
39. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly *6*, 80-92. 10.4161/fly.19695.
40. Cisneros-Aguirre, M., Ping, X., and Stark, J.M. (2022). To indel or not to indel: Factors influencing mutagenesis during chromosomal break end joining. DNA Repair *118*, 103380. 10.1016/j.dnarep.2022.103380.
41. Hirano, Y., and Sugimoto, K. (2006). ATR homolog Mec1 controls association of DNA polymerase zeta-Rev1 complex with regions near a double-strand break. Current Biology *16*, 586-590. 10.1016/j.cub.2006.01.063.
42. Covo, S., de Villartay, J.-P., Jeggo, P.A., and Livneh, Z. (2009). Translesion DNA synthesis-assisted non-homologous end-joining of complex double-strand breaks prevents loss of DNA sequences in mammalian cells. Nucleic Acids Research *37*, 6737-6745. 10.1093/nar/gkp703.
43. Sakofsky, C.J., Ayyar, S., Deem, A.K., Chung, W.H., Ira, G., and Malkova, A. (2015). Translesion Polymerases Drive Microhomology-Mediated Break-Induced Replication

Leading to Complex Chromosomal Rearrangements. Mol Cell *60*, 860-872. 10.1016/j.molcel.2015.10.041.
44. Gan, G.N., Wittschieben, J.P., Wittschieben, B.Ø., and Wood, R.D. (2008). DNA polymerase zeta (pol zeta) in higher eukaryotes. Cell Research *18*, 174-183. 10.1038/cr.2007.117.
45. Kumpornsin, K., Kochakarn, T., Yeo, T., Okombo, J., Luth, M.R., Hoshizaki, J., Rawat, M., Pearson, R.D., Schindler, K.A., Mok, S., et al. (2023). Generation of a mutator parasite to drive resistome discovery in Plasmodium falciparum. Nat Commun *14*, 3059. 10.1038/s41467-023-38774-1.
46. Honma, H., Niikura, M., Kobayashi, F., Horii, T., Mita, T., Endo, H., and Hirai, M. (2016). Mutation tendency of mutator Plasmodium berghei with proofreading-deficient DNA polymerase delta. Sci Rep *6*, 36971. 10.1038/srep36971.
47. Frank, M., Kirkman, L., Costantini, D., Sanyal, S., Lavazec, C., Templeton, T.J., and Deitsch, K.W. (2008). Frequent recombination events generate diversity within the multi-copy variant antigen gene families of Plasmodium falciparum. International Journal for Parasitology *38*, 1099-1109. 10.1016/j.ijpara.2008.01.010.
48. McHugh, E., Bulloch, M.S., Batinovic, S., Patrick, C.J., Sarna, D.K., and Ralph, S.A. (2023). Nonsense-mediated decay machinery in Plasmodium falciparum is inefficient and non-essential. mSphere *8*, e00233-00223. 10.1128/msphere.00233-23.
49. B, B., C, C., SP, M., DA, S., JC, K., MJ, F., and JH, A. (05/07/2009). piggyBac is an effective tool for functional analysis of the Plasmodium falciparum genome - PubMed. BMC microbiology *9*. 10.1186/1471-2180-9-83.
50. Dimonte, S., Bruske, E.I., Enderes, C., Otto, T.D., Turner, L., Kremsner, P., and Frank, M. (2020). Identification of a conserved var gene in different Plasmodium falciparum strains. Malaria Journal *19*, 194. 10.1186/s12936-020-03257-x.
51. Benavente, E.D., Oresegun, D.R., de Sessions, P.F., Walker, E.M., Roper, C., Dombrowski, J.G., de Souza, R.M., Marinho, C.R.F., Sutherland, C.J., Hibberd, M.L., et al. (2018). Global genetic diversity of var2csa in Plasmodium falciparum with implications for malaria in pregnancy and vaccine development. Scientific Reports *8*, 15429. 10.1038/s41598-018-33767-3.
52. Deitsch, K.W., del Pinal, A., and Wellems, T.E. (1999). Intra-cluster recombination and var transcription switches in the antigenic variation of Plasmodium falciparum. Molecular and Biochemical Parasitology *101*, 107-116. 10.1016/s0166-6851(99)00062-6.
53. Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature *407*, 1018-1022. 10.1038/35039531.
54. Duffy, M.F., Byrne, T.J., Carret, C., Ivens, A., and Brown, G.V. (2009). Ectopic recombination of a malaria var gene during mitosis associated with an altered var switch rate. J Mol Biol *389*, 453-469. 10.1016/j.jmb.2009.04.032.
55. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). DNA Repair. In Molecular Biology of the Cell. 4th edition, (Garland Science).
56. Voß, Y., Klaus, S., Guizetti, J., and Ganter, M. (2023). Plasmodium schizogony, a chronology of the parasite's cell cycle in the blood stage. PLOS Pathogens *19*, e1011157. 10.1371/journal.ppat.1011157.

57. Damasceno, J.D., Reis-Cunha, J., Crouch, K., Beraldi, D., Lapsley, C., Tosi, L.R.O., Bartholomeu, D., and McCulloch, R. (2020). Conditional knockout of RAD51-related genes in Leishmania major reveals a critical role for homologous recombination during genome replication. PLOS Genetics *16*, e1008828. 10.1371/journal.pgen.1008828.
58. Zhang, X., Deitsch, K.W., and Kirkman, L.A. (2021). The contribution of extrachromosomal DNA to genome plasticity in malaria parasites. Mol Microbiol *115*, 503-507. 10.1111/mmi.14632.
59. Kirkman, L.A., Su, X.-z., and Wellems, T.E. (1996). Plasmodium falciparum:Isolation of Large Numbers of Parasite Clones from Infected Blood Samples. Experimental Parasitology *83*, 147-149. 10.1006/expr.1996.0058.
60. Deitsch, K.W., Driskill, C.L., and Wellems, T.E. (2001). Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. Nucleic Acids Research *29*, 850-853.
61. Grimberg, B.T. (2011 Feb 4). Methodology and Application of Flow Cytometry for Investigation of Human Malaria Parasites. Journal of immunological methods *367*. 10.1016/j.jim.2011.01.015.
62. Zhang, X., Florini, F., Visone, J.E., Lionardi, I., Gross, M.R., Patel, V., and Deitsch, K.W. (2022). A coordinated transcriptional switching network mediates antigenic variation of human malaria parasites. eLife *11*, e83840. 10.7554/eLife.83840.
63. Loman, N.J., and Quinlan, A.R. (2014). Poretools: a toolkit for analyzing nanopore sequence data. Bioinformatics *30*, 3399-3401. 10.1093/bioinformatics/btu555.
64. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114. 10.1093/bioinformatics/btu170.
65. Andrew, S. (2010). FastQC A Quality Control tool for High Throughput Sequence Data.
66. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv.
67. Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032-2034. 10.1093/bioinformatics/btv098.
68. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008. 10.1093/gigascience/giab008.
69. Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics *34*, 3094-3100. 10.1093/bioinformatics/bty191.
70. Smolka, M., Paulin, L.F., Grochowski, C.M., Horner, D.W., Mahmoud, M., Behera, S., Kalef-Ezra, E., Gandhi, M., Hong, K., Pehlivan, D., et al. (2024). Detection of mosaic and population-level structural variants with Sniffles2. Nature Biotechnology *42*, 1571-1580. 10.1038/s41587-023-02024-y.
71. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156-2158. 10.1093/bioinformatics/btr330.
72. Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize Implements and enhances circular visualization in R. Bioinformatics *30*, 2811-2812. 10.1093/bioinformatics/btu393.

73. Robinson, J.T., Thorvaldsdottir, H., Turner, D., and Mesirov, J.P. (2023). igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). Bioinformatics *39*, btac830. 10.1093/bioinformatics/btac830.