Mutagenic-end joining results in smaller deletions in heterochromatin relative to euchromatin

Jacob M Miller*[1], Sydney Prange*[1], Huanding Ji*[1], Alesandra R. Rau*[1], Nadejda Butova[1], Avery Lutter[1], Helen Chung[1], Chiara Merigliano[1], Chetan C. Rawal[1], Mitch McVey[2], Irene Chiolo[1]#.

*Equal contribution

#Corresponding author. chiolo@usc.edu

[1]Department of Molecular and Computational Biology, University of Southern California, CA, USA

[2]Department of Biology, Tufts University, MA, USA

ABSTRACT

Pericentromeric heterochromatin is highly enriched for repetitive sequences prone to aberrant recombination. Previous studies showed that homologous recombination (HR) repair is uniquely regulated in this domain to enable 'safe' repair while preventing aberrant recombination. In *Drosophila* cells, DNA double-strand breaks (DSBs) relocalize to the nuclear periphery through nuclear actin-driven directed motions before recruiting the strand invasion protein Rad51 and completing HR. End-joining (EJ) repair also occurs with high frequency in heterochromatin of fly tissues, but how different EJ pathways operate in heterochromatin remains uncharacterized. Here, we induce DSBs in single euchromatic and heterochromatic sites using the DR-*white* reporter and I-SceI expression in spermatogonia. We detect higher frequency of HR repair in heterochromatic insertions, relative to euchromatin. Sequencing of repair outcomes reveals the use of distinct EJ pathways across different euchromatic and heterochromatic sites. Interestingly, synthesis-dependent michrohomology-mediated end joining (SD-MMEJ) appears differentially regulated in the two domains, with a preferential use of motifs close to the cut site in heterochromatin relative to euchromatin, resulting in smaller deletions. Together, these studies establish a new approach to study repair outcomes in fly tissues, and support the conclusion that heterochromatin uses more HR and less mutagenic EJ repair relative to euchromatin.

INTRODUCTION

Pericentromeric heterochromatin (hereby 'heterochromatin') occupies about 30% of the fly genome and is largely composed of highly repeated DNA sequences. In *Drosophila* about half of these sequences are satellite repeats (mostly 5-bp sequences repeated in tandem for hundreds of kilobases to megabases), and the rest are transposons, other scrambled repeats, and about 120 genes. Given the abundance of repeated sequences, heterochromatin is prone to aberrant recombination during DSB repair.

DSBs can be repaired through different pathways, mostly: HR, single-strand annealing (SSA), non-homologous end joining (NHEJ), and alternative-EJ (alt-EJ). During HR, DSBs are resected to generate long stretches of single-stranded DNA (ssDNA) that invade homologous sequences on the sister chromatid or the homologous chromosome, which are used as templates for DNA synthesis and repair. HR is typically error-free in single copy sequences. In heterochromatin, however, the presence of thousands to millions of alternative donor sequences associated with different chromosomes can lead to strand invasion of non-allelic (ectopic) templates, resulting in chromosome rearrangements and genome instability.

Previous studies in *Drosophila* cells showed that HR repair of heterochromatic DSBs occurs through a unique pathway that relocalizes repair sites to the nuclear periphery. Resection starts inside the heterochromatin domain [1], but recruitment of the strand invasion protein Rad51 is temporarily halted and resumes only after relocalization [1, 2]. This movement occurs by directed motion along a striking network of nuclear actin filaments (F-actin), and is mediated by nuclear myosins associated with the repair sites [3]. F-actin-dependent relocalization of heterochromatic DSBs has also been detected in mouse cells [3, 4], revealing conserved pathways. This mechanism likely promotes 'safe' repair by moving the damage away from the bulk of ectopic repeated sequences inside the heterochromatin domain, before promoting strand invasion with the homologous templates on the sister chromatid or the homologous chromosome. In flies, homologous chromosome pairing occurs throughout interphase, providing alternative templates for repair. Accordingly, both the sister and the homolog can be used as repair templates for HR, although the sister chromatid is used at higher frequency [5]. Studies in fly and human cells suggest that HR is the preferred pathway used for heterochromatin repair when both HR and NHEJ are available, such as in the S and G2 phases of the cell cycle [1, 6].

In addition to HR, long-range resection can promote SSA when ssDNA unmasks identical sequences that can pair with each other, resulting in large deletions. Despite the abundance of repeated sequences in heterochromatin, this pathway seems rarely used in this domain [5]. Alternatively, short-range resection and small regions of homology (1-8bp) can be used in alt-EJ pathways: microhomology-mediated EJ (MMEJ), which results in small deletions, or pol-theta dependent synthesis-dependent MMEJ (SD-MMEJ), which results in deletions or indels (deletions plus insertions) [7-9]. Finally, canonical NHEJ (cNHEJ) simply rejoins the two ends of the DSB together, with small mutations, deletions, or indels around the cut site [10]. Given the different mutagenic outcomes of these repair pathways, it is possible to detect their activity based on the sequence of the repair product. In *Drosophila melanogaster*, alt-EJ is commonly used to repair DSBs even in the presence of fully functional cNHEJ [11-13], making this an ideal model system to study the impact of different chromatin environments on alt-EJ repair.

While HR appears to be the preferred mechanism for heterochromatin repair in cultured cells, a study in fly tissues revealed higher frequency of EJ rather than HR repair in heterochromatin. This is likely a consequence of the high number of G1 cells in these tissues, and might also result from the inability of the DR-*white* cassette to detect intersister HR resulting in perfect repair [5]. This study described high frequency of NHEJ-associated deletions in somatic cells, and did not investigate other EJ repair mechanisms occurring in euchromatin and heterochromatin. Another study systematically looked at the balance between MMEJ and cNHEJ in different epigenetic environments using a sequencing-based reporter cut with Cas9 inserted in a large number of genomic locations of human cells [14]. This study identified a higher ratio of MMEJ/cNHEJ usage in H3K9me2-enriched lamina-associated silenced regions, and in polycomb silenced regions (H3K27me3-enriched), relative to non-silenced regions [14]. This is the result of more efficient usage of cNHEJ in non-silenced regions, suggesting that silencing interferes with cNHEJ while enabling MMEJ [14]. It also did not examine DSB repair in pericentromeric heterochromatin. This nuclear domain is unique in that it is not only enriched for H3K9me3, but also mostly composed of repeated sequences, is phase separated, and is not typically associated with the nuclear periphery.

Importantly, DNA sequence composition affects the balance between repair pathways, such as by generating SD-MMEJ-promoting secondary structures in the DNA flanking the break site [15, 16]. A major advantage of using the same reporter inserted in different genomic locations to

study repair is that repair outcomes are not affected by differences in the local DNA sequence, thus enabling a more direct focus on the chromatin-associated effects.

Here, we use a DR-*white* reporter inserted in euchromatic and heterochromatic sites, and we directly investigate repair outcomes in fly spermatogonia and primary spermatocytes as a model for mitotically dividing cells of adult flies. Contrary to previous conclusions, we detect higher frequency of HR repair in all heterochromatic insertions relative to euchromatin. In addition, different alt-EJ repair mechanisms prevail at distinct heterochromatic locations, with SD-MMEJ typically resulting in smaller deletions in heterochromatin relative to euchromatin. This is linked to a higher usage of motifs closer to the cut site, suggesting that farther motifs are less accessible for secondary structure formation in heterochromatin. Together, this study establishes new tools to study heterochromatin repair in flies, reveals unique features of alt-EJ in heterochromatin, and suggests that alt-EJ is less mutagenic in this domain.

RESULTS

A new approach to induce single DSBs in euchromatin or heterochromatin of spermatogonia

We generated new fly lines to examine repair outcomes with the DR-*white* reporter. This reporter is characterized by two non-functional repeats of the *white* gene (*Sce.white* and *iwhite*), the first of which contains a restriction site for I-SceI (Fig. 1A) [17]. Exposure to the I-SceI endonuclease results in DSB induction in *Sce.white*. Intra-chromosomal or inter-sister chromatid HR repair using the *iwhite* sequence as a template restores the *wild-type white* gene. SSA repair results in the loss of the DsRed marker. Mutagenic EJ repair results in the loss of the I-SceI cut site and can be characterized through PCR amplification across the cut site, enzymatic digestion, and DNA sequencing [17].

We induced I-SceI expression using the *Bam* promoter (BamP898-CFI [18]), which is induced in spermatogonia and suppressed in spermatocytes. This results in I-SceI accumulation in spermatogonia and likely early spermatocytes [19] (Fig. 1B). This enables the induction of DSBs in the germ line of F1 males, and the quantification of repair outcomes in the progeny. Placing I-SceI under the *Bam* promoter is an important change from previous studies that expressed I-SceI broadly in the germ line or in somatic cells with a heat-shock or a ubiquitin promoter [5, 17, 20]. Given the specificity of the *Bam* promoter for spermatogonia, damage induction and repair occurs in a relatively homogeneous population of mitotically dividing cells, reducing

confounding effects from multiple cell types and stages of differentiation. Additionally, each individual of the progeny contains a unique repair product that can be scored for the acquisition of *white+* or *DsRed-* phenotypes and further analyzed by sequencing.

We integrated DR-*white* in one euchromatic location and four pericentromeric heterochromatic sites of chromosome 2 (Fig. 1C). The euchromatic site is located in transcriptionally silenced euchromatin ('gray' chromatin state [21]), and was previously studied with the DR-*white* cassette (Fig. 1D) [5]. The yellow marker of the MiMIC lines used to generate these insertions and the Ds-Red and *white+* marker of the DR-*white* are fully expressed in this location, indicating a transcriptionally active state for the reporter. The heterochromatic sites are heavily silenced based on H3K9me3 levels (Fig. 1E), and are classified as 'blue' chromatin state in previous studies [21], corresponding to silenced pericentromeric heterochromatin. Accordingly, the yellow marker of the MiMIC lines used to generate these insertions is nearly fully silenced resulting in a faint 'mosaic' phenotype, while the Ds-Red or the *white+* marker of the DR-*white* is completely silenced.

Given the silent state of the heterochromatic insertions, F1 males were crossed with females carrying a mutation in the heterochromatin silencing component HP1a (*su(var)205+/-*) [22] (Fig. 1F). HP1 heterozygosity in the F2 generation results in the partial reactivation of DsRed and *white* markers carried by the reporter (Fig. 1F), leading to a mosaic phenotype that can easily be scored in the F2 generation.

HR repair occurs more frequently in heterochromatin

First, we investigated HR repair of adult F1 males containing DR-*white* and *Bam*-ISceI by quantifying the fraction of individuals with red eyes (w+, DsRed+) in the F2 progeny (Fig. 1F).

This analysis revealed that 3.6% of the flies repair DSBs using HR when DR-*white* is placed in euchromatin, and this percentage remains remarkably consistent across replicates revealing very little experimental variability. This is a significantly lower frequency than observed in previous studies with I-SceI under a heat-shock promoter and DR-*white* in the same genomic location [5], consistent with *Bam* being active only in a subset of cells in the germline and thus inducing DSBs less frequently.

Next, we analyzed HR repair in the heterochromatic insertions of DR-*white* by scoring variegated red-eyed flies in the F2 progeny. All heterochromatic insertions used HR at significantly higher frequency relative to the euchromatic insertion, ranging from 5.5% to 9.2% of the total DRW-containing flies (Fig. 2). We conclude that HR repair is more frequently used in heterochromatin relative to euchromatin.

Alt-EJ repair pathways operate with different frequencies at distinct DR-white insertions

We then focused on characterizing EJ events by analyzing repair products in white-eyed flies. White-eyed F2 individuals can derive from SSA (w- DsRed-) or mutagenic EJ repair (w- DsRed+), resulting in loss of the I-SceI cut site (Fig. 1A). Alternatively, white-eyed flies can derive from germ line cells where the I-SceI site was not cut, from perfect NHEJ repair, or from HR with the sister chromatid, and each of these outcomes retain a wild-type I-SceI site. Mutagenic and non-mutagenic events were distinguished through PCR amplification across the cut site followed by digestion with I-SceI (Fig. 1A).

Consistent with previous studies [5, 17, 20], we observed that SSA is rarely used in euchromatin and heterochromatin (Suppl. Fig. 1A) and its frequency is not significantly different across DR-*white* insertions. Conversely, EJ repair occurs with high frequency across all DR-*white* insertions (Suppl. Fig. 1B). We also identified a large number of flies where the I-SceI site was still intact (nearly 50% of the white-eyed flies), consistent with either low cut efficiency or non-mutagenic repair. The frequency of mutagenic and non-mutagenic events is similar across euchromatic and heterochromatic DR-white insertions, with the exception of HC1 where non-mutagenic outcomes are more frequent, potentially reflecting lower cut efficiency at this site (Suppl. Fig. 1B).

We further characterized mutagenic EJ events by Sanger sequencing of the PCR product across the cut site of individual F2 flies and subsequent analysis of the repair junction. MMEJ typically results from the annealing of direct repeats positioned on each side of the cut sites after resection, and deletion of the intervening sequence [23, 24]. SD-MMEJ requires the presence of two 'repeat motifs', each containing a 'primer' (P) and a nearby 'microhomology' (MH) sequence [15]. Duplex unwinding or resection exposes the primer repeats that anneal to form loops or transient hairpins (Fig. 3A) [15]. Primer annealing can occur between two direct repeats ('loop-out' mechanism) or two inverted repeats ('snap-back' mechanism) [15]. The 3′ end of the DNA in these structures then uses ssDNA as a template for DNA synthesis and

repair. Unwinding of the secondary structure and re-annealing of the nascent DNA with a MH sequence on the other side of the break enables repair completion, typically resulting in deletions or indels (insertions with or without deletions).

Most repair junctions are characterized by deletions of fewer than 23 bp. The mean deletion size for all five insertions is between 9 and 16 bp (Supplementary Fig. 2), suggesting frequent repair by alt-EJ mechanisms.

To establish the frequency of different mutagenic EJ events, we used previously published programs in Python and R that identify repeat motifs and primer and microhomology repeats potentially used during SD-MMEJ [15]. We classified repair products based on the presence of deletions or indels, which revealed that indels are rare for all DR-*white* insertions (less than 10%) (Fig. 3B). This is consistent with previous studies that used DR-*white* as a repair reporter [5, 17]. We further classified repair events as deletions with apparent blunt joins (ABJs) and deletions with microhomology joins (MHJs) [8, 15]. This revealed that deletions with microhomologies are very common (~70%) across all DR-*white* insertions, consistent with frequent MMEJ or SD-MMEJ repair (Fig. 3B).

Next, we established whether each repair event was consistent with SD-MMEJ mechanisms (Fig. 3C). To be classified as SD-MMEJ-consistent, the repair junctions need to: i) contain a repeat motif of at least 4 bp; ii) contain a primer repeat of at least 1 bp and a microhomology repeat of at least 1 bp within the repeat motif; and iii) the break-distal end of the repeat motif has to occur within 30 bp of the DSB [15]. MHJ repair events that are SD-MMEJ-inconsistent were classified as MMEJ. ABJ repair events or indels that are SD-MMEJ-inconsistent and result in a 1-5 deletions, respectively, were classified as NHEJ [15].

This analysis suggests significant differences in the frequencies of usage of distinct mutagenic EJ pathways across the DR-*white* insertions (Fig. 3C). Specifically, HC3 is characterized by less SD-MMEJ-consistent and more MMEJ-consistent events (Fig. 3C). Within SD-MMEJ, loop-out mechanisms are significantly reduced in HC2 and HC4 relative to the euchromatic DR-*white* insertion (Fig. 3D). We conclude that there is variation in the frequency of use of distinct mutagenic EJ pathways across different DR-*white* insertions, without prevalence of a specific pathway in heterochromatin relative to euchromatin.

SD-MMEJ repair results in smaller deletions in heterochromatin

Interestingly, the analysis of repair deletion size revealed differences between euchromatic and heterochromatic DR-*white* insertions (Supplementary Fig. 2). HC3 and HC4 are characterized by a significantly lower deletion size relative to the euchromatic site. HC1 and HC2 also show a lower median size of deletions relative to the euchromatic site (Supplementary Fig. 2), although the difference in size distribution compared to the euchromatic site did not reach significance.

To understand these differences in greater detail, we first analyzed the position of the deletion boundaries associated with different repair mechanisms across all DR-*white* insertions. This analysis reveals a higher frequency of boundaries close to the cut site for all heterochromatic insertions relative to the euchromatic one, resulting from SD-MMEJ and/or MMEJ-consistent events (Fig. 4). Notably, for events close to the cut site and resulting in very small deletions (1-5bp) we cannot exclude that cNHEJ contributes to these products.

In agreement, a heat map plot displaying the frequency of deletion spreads associated with SD-MMEJ events highlights a prevalence of smaller deletions for all heterochromatic DR-*white* insertions relative to the euchromatic one (Fig. 5A). Conversely, the spread of deletions for SD-MMEJ-inconsistent repair events does not significantly change across different DR-*white* insertions, and mostly corresponds to MMEJ repair using the CCCT/CCCT direct repeat around the cut site (Fig. 5B). This indicates that, within SD-MMEJ-consistent events all heterochromatic DR-*white* insertions result in smaller deletions in heterochromatin relative to euchromatin, while deletions resulting from MMEJ-consistent events remain of the same size in the two domains.

Smaller deletions from SD-MMEJ repair could derive from the usage of repeat motifs positioned closer to the cut site in heterochromatin. We analyzed this directly by comparing the frequency of use of repair motifs associated with the different SD-MMEJ repair events (Fig. 6). This analysis shows that the smaller deletion sizes for SD-MMEJ events in heterochromatin is associated with the use of repair motifs closer to the cut site for all heterochromatic DR-white insertions relative to the euchromatic site (Fig. 6).

Together, we conclude that alt-EJ repair relies on microhomologies closer to the cut site in heterochromatin, resulting in SD-MMEJ repair with smaller deletions or higher frequency of MMEJ using the CCCT/CCCT direct repeat proximal to the cut site.

DISCUSSION

Previous studies identified a clear role of the DNA sequence in biasing alt-EJ repair, but the role of the chromatin state in influencing alt-EJ remains poorly explored. Recent studies showed that lamin-associated silenced chromatin and H3K27me3-enriched chromatin interfere with cNHEJ while enabling MMEJ. However, there is no previous study directly addressing how alt-EJ repair operates in pericentromeric heterochromatin, a large region of the genome mostly composed of repeated DNA sequences, subject to epigenetic silencing, characterized by phase separation, and mostly not associated with the nuclear periphery. There is also no clear understanding of how chromatin states affect SD-MMEJ repair. Here, we use DR-*white* repair reporter inserted in a previously characterized euchromatic locus and several new locations in pericentromeric heterochromatin to investigate DSB repair responses. We establish an experimental approach where damage occurs in a uniform population of pre-meiotic cells that are actively dividing. With this setup, DNA sequence and cell cycle phases remain the same for all the DR-*white* insertions analyzed, removing these factors as potential sources of variability in repair pathways. Our studies revealed a higher frequency of HR usage in heterochromatin relative to euchromatin. In addition, we identify a higher frequency of usage of microhomologies close to the cut site for alt-EJ repair in heterochromatin relative to euchromatin, resulting in smaller deletions. Together, these studies suggest that DSB repair is less mutagenic in heterochromatin than in euchromatin.

Detecting a higher frequency of HR repair in heterochromatin relative to euchromatin is unexpected, given that previous studies with DR-*white* detected a similar frequency of HR usage between euchromatin and heterochromatin. A possible reason for this discrepancy lies in the amount of silencing present in the DR-*white* cassette across different insertions. The DR-*white* reporter was exceptionally silenced in the heterochromatic locations used for the present study, such that HR products could only be revealed by crossing the F1 progeny with su(var)205+/- flies. This was not the case for heterochromatic DR-*white* insertions used in previous studies, consistent with a higher level of expression of the cassette.

Additionally, the induction of DSBs in pre-meiotic cells provides unique advantages to the study of HR repair. First, these cells are mitotically dividing, providing a high number of cells in S/G2 with active HR repair. In addition, having a uniform population of cells introduces less variability in the system. In agreement, we found more consistency in the frequency of HR outcomes

across different experiments relative to what was previously shown with a ubiquitin-driven I-SceI in the germline. This might further facilitate the identification of differences across distinct DR-*white* insertions.

What mechanisms promote HR repair in heterochromatin still needs to be defined. An interesting possibility is that high levels of HP1 promote resection in heterochromatin, similar to the role of transient HP1 deposition in euchromatin. Consistent with this hypothesis, live imaging studies in cultured cells showed that ionizing-radiation (IR)-induced repair foci of the ssDNA/RPA-associated proteins ATRIP and TopBP1 form faster and appear brighter in heterochromatin relative to euchromatin [1].

Our analysis of mutagenic-EJ repair revealed a high frequency of SD-MMEJ-consistent events, in agreement with previous studies with DR-*white* in the germline [17]. However, we detected a much broader set of SD-MMEJ and MMEJ-consistent repair junctions than in previous studies, suggesting that our experimental setup with *Bam*-dependent induction of I-SceI is particularly suitable for detecting micro homology-driven mutagenic EJ.

Interestingly, we detected significant variability in the type of repair pathway prevailing in each heterochromatic site, with MMEJ prevailing at HC3 and snap-back SD-MMEJ prevailing at HC2 and HC4. This suggests that factors other than the epigenetic composition, sequence composition, cell cycle phase, and association with the heterochromatin domain and its phase separated state might be important to balance SD-MMEJ and MMEJ mechanisms.

All the heterochromatic sites investigated here have been selected because of their heavily silenced state throughout development and differentiation. However, we cannot exclude that transient variations of the chromatin state occur in spermatogonia for some of these sites, skewing the balance between alt-EJ mechanisms during the assay. For example, HC1 and HC2 insertion sites span recently-identified lncRNAs that might be transiently expressed, and HC4 is proximal to a gene weakly expressed in testes. In this light, the completely silenced state of HC3 might be the reason why MMEJ is particularly efficient at this site, suggesting MMEJ as the main mutagenic-EJ pathway available in highly repeated and highly silenced heterochromatin.

Regardless of these differences, we also uncovered a general preference for the use of repeat motifs closer to the cut site in heterochromatin relative to euchromatin, resulting in smaller

deletions. The reasons for this distinctive behavior remains unknown. It is tempting to speculate that features linked to a more compact chromatin or the phase separated state of heterochromatin affect end-processing and/or repair protein accessibility to facilitate the use of microhomologies close to the cut site for SD-MMEJ or MMEJ repair.

A likely reason for the use of microhomologies closer to the cut site is less availability of long stretches of 'naked' ssDNA that can form secondary structures or engage in re-annealing events, e.g., as a result of less resection or more RPA association. We tend to favor the second hypothesis, given that heterochromatic DSBs also experience more efficient HR repair relative to euchromatin, which relies on resection (Fig. 7). However, it is also possible that short-range resection initiates at a slower rate in heterochromatin, perhaps because of the tighter chromatin compaction, favoring the use of homologies close to the cut site for repair by alt-EJ. Alternatively, RPA could bind ssDNA with higher efficiency in heterochromatin, or pol-theta could be less promptly available to remove it, interfering with the usage of more distant microhomologies for repair.

Together, our study highlights the preferential use of HR or alt-EJ pathways relying on microhomologies close to the cut site in heterochromatin, both of which reduce the mutagenic potential of repair mechanisms in heterochromatin. This could be advantageous, because about half of fly heterochromatin is composed of repeated DNA sequences where mutagenic events can quickly spread across the satellite affecting large genomic domains. In addition, given the short size of the repeated unit on most satellites, alt-EJ mechanisms using sequences close to the cut site could be particularly effective.

In sum, this study generated new experimental approaches to study repair responses to site-specific DSBs, and particularly alt-EJ repair, which are broadly applicable in the field. It also provides the first systematic characterization of alt-EJ pathways in euchromatin and heterochromatin, identifying unique responses in heterochromatin. Future studies are needed to establish which properties of the heterochromatin domain are responsible for these differences, and how disease states can affect these responses threatening genome stability.

METHODS

Drosophila stocks and maintenance

Drosophila were maintained on standard media at 25°C, which was prepared as in [25]. We used the Minos-mediated integration cassette (MiMIC) system [26] to integrate the DR-*white* reporter [5] (DR-white-RFP) into four pericentromeric heterochromatic and one euchromatic loci. Embryo injection and validation of the insertions were performed by Bestgene Inc. (Chino Hills, CA) and Genetivision Corporation (Houstin, TX). An overview of the MiMIC integration sites is in Suppl. Table 1. BamP898-CFI (*Bam*-I-SceI) was described in [18]. All insertions were verified by genotyping with locus-specific oligos. Su(var)205[1545] was described in [27].

Crosses and experimental setup for the DR-*white* repair assay

Homozygous females containing DR-*white* on chromosome 2 were crossed to males containing BamP898-CFI (*Bam*-I-SceI) on chromosome 3, for 3 days (F0). After enclosure, 0-1 day old males containing DR-*white* and *Bam*-I-SceI were individually crossed to Su(var)205[1545]/Cyo females (F1). The parents were removed after 3 days and the F2 progeny containing DR-*white* and Su(var)205[1545] was collected for visual scoring and sequence analysis. We obtained between 21 and 47 viable F1 crosses for each experiment, resulting in 200-926 DRW-containing F2 flies that were used for the different analyses.

DR-*white* data analysis

F1 male flies with phenotype w+DsRed+ were classified as HR, while flies with phenotype w-DsRed- were classified as SSA and confirmed by PCR as previously described [5, 17]. Notably, only about 40% of the flies scored as DsRed- was confirmed as SSA by PCR, the rest was identified as an EJ product after sequencing. This indicates that DsRed is not fully penetrant or easy to score in the F1 progeny with our experimental setup. w-DsRed+ flies were used for molecular analyses of EJ-products as follows. Individual F2 males were processed for DNA extraction, PCR amplification, and I-SceI digestion, as previously described [17]. PCR products were obtained using Q5 Polymerase (NEB) or Expand Long Template PCR system (Roche/Sigma) and the oligos: 5'GCGTCAATTCGGCACGAG (DR-white2.2) and 5' CTTAGAGCCAGATATGCGAGCAC (DR-white2.2a). Sanger sequencing of PCR products that lost the I-SceI cut site was performed by Laragen (Culver City, CA) using the oligo TGGCAACCATCGTTGTCTG 420 bp from the I-SceI recognition site, after an additional purification step.

Analysis of EJ events

Sequences were trimmed to common start and end positions, and mapped to the reference DR-white sequence using Clustal Multiple Sequence Alignment. This alignment revealed deleted and inserted sequences (Supplementary Fig. 3). Using a previously published Python script [16], deletion junctions were characterized as apparent blunt joins (ABJ) or microhomology junctions (MHJ), and further analyzed for SD-MMEJ consistency. SD-MMEJ consistency was defined as the presence of direct (loop-out) or inverted (snap-back) repeat motifs of 4 or more bp in the repair product. Insertion sequences (containing novel sequences absent in the original construct, with or without deletions) were analyzed using a previously described R-script to determine SD-MMEJ consistency and identify motifs, primers, and microhomologies [16]. For SD-MMEJ consistent sequences with ambiguous repeat motif usage, motifs and their respective deletions were assigned a weight assuming a uniform probability for each motif-deletion pair.

Quantification of H3K9me3 enrichments at different insertion sites

H3K9me3 ChIP-seq data in Kc cells was obtained from SRA accession SRX2559027 [28]. Two additional replicates were done in our lab. For all replicates and input, sequencing reads from illumina were trimmed of their adapters and low quality reads were removed using trim galore. Reads were aligned to the Dm6 reference genome using the bwa-mem alignment algorithm. Optical duplicates were removed using MarkDuplicates (Picard). Coverage within +/- 0.5 Mb of each insert site was determined using Rsamtools. H3K9me3 and input read counts within +/- 0.5 Mb of each insert were normalized for sequencing depth using RPM normalization.

Acknowledments

This research was supported by NIH R01GM117376 and NSF Career 1751197 to I.C, MCB1716039 to M.M. We are grateful to Tzumin Lee for sharing BamP898-CFI flies before publication, and Gary Karpen for sharing Su(var)205[1545] flies. We thank Cuiyi Zhang for helping with fly genotyping and Haseeb Khan for helping with PCRs of F1 flies. MiMIC stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) were used in this study.

Competing interests

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1: Experimental setup to induce single DSBs in the male germline. A) Schematic representation of the DR-*white* reporter, containing a mutated *white* gene with an I-SceI site inserted, the dsRed sequence, and a truncated *white* gene. HR repair of I-SceI induced breaks using the *iwhite* sequence results in *w+*, red-eyed (in euchromatin) or variegated-eyed flies (in heterochromatin). SSA repair results in the loss of the fluorescent marker dsRed in the eyes. Uncut cassettes, mutagenic repair, or perfect NHEJ results in *w-* flies that can be further characterized by PCR amplification and sequencing of the repair junction. B) Schematic representation of the cells affected by *Bam* expression in the testis. Under the control of the Bam promoter, I-SceI is expressed in induced pre-meiotic spermatogonia. Each sperm will contain a single repair event that can be tracked in the progeny. GSCs: germline stem cells. C) Schematic view of the genomic location of the DR-*white* insertions used in this study. D) Description of some of the most distinctive features of these sites. Colors indicate the chromatin state described by the 9-state model [21]. Blue: Pericentromeric heterochromatin. Grey: intergenic silenced regions. HC1 and HC2 insertion sites are inside of an annotated lncRNA (+), while EC, HC3, and HC4 are not (-). TE density and repeat density are qualitatively illustrated by the number of light and dark purple bars respectively. E) H3K9me3 levels at the insertion sites determined by ChIP-seq ([21] and our data). H3K9me3 enrichment was determined by the ratio of the amount of reads +/- 0.5 Mb from the insert site per million mapped reads between IP and input samples. Error bars: SD of 3 replicates. F) Scheme of the fly crosses done to analyze repair outcomes.

Figure 2: Percentage of HR products detected at different DR-*white* insertions. Frequency of HR events calculated as number of red-eyed F2 flies that still retain Ds-Red expression over the total number of flies (n). *$p<0.05$; **$p<0.005$; ***$p<0.001$ by t-test; n=1352 flies for EC; n=485 flies for HC1; n=1267 flies for HC2; n=679 flies for HC3; n=880 flies for HC4; Error bars: SD of two independent experiments.

Figure 3: Frequency of Alt-EJ repair pathways at different DR-white insertions A) Schematic representation of SD-MMEJ pathways (adapted from [16]), indicating the position of repeat motifs and the corresponding primers (P1, P2) and microhomologies (MH1, MH2). For the loop-out mechanism, the DNA is unwinded before loop formation. For the snap-back mechanism, resection occurs before hairpin formation. In both mechanisms, the annealing of the break-proximal primer (P2) to the break-distal primer (P1), initiates DNA synthesis (red arrows) that can generate insertions and the creation of new microhomology sequences (MH1). P1 and P2 are direct repeats for loop-out and inverted repeats for snap-back. Repair continues through secondary structure unwinding, annealing of the newly-synthesized microhomologies with MH2 sequences on the opposite side of the break, fill-in DNA synthesis, and ligation. When P2 and MH2 are not adjacent to the break site DNA flaps can form and get trimmed, resulting in long deletions. When P1 and MH1 are adjacent to each other, insertions do not occur. B) Frequency of indels, MHJ, and ABJ deletions in mutagenic-EJ repair across indicated DR-*white* insertions. Differences across DR-*white* insertions are not significant.. EC n = 86; HC1 n = 14; HC2 n = 280; HC3 n = 104; HC4 = 74. C) Frequency of SD-MMEJ-consistent, MMEJ-consistent (i.e., MHJ and SD-MMEJ-inconsistent), NHEJ (i.e, ABJ and SD-MMEJ-inconsistent, less than 5bp deletions or indels), or other repair outcomes, across different DR-*white* insertions. Other repair outcomes include point mutations as well as ABJ and SD-MMEJ-inconsistent deletions or indels of greater than 5bp. **$p<0.005$; ***$p<0.001$ by two-proportion z-test, EC n = 86; HC1 n = 14; HC2 n = 280; HC3 n = 104; HC4 = 74. D) Frequency of loop-out and snap-back-consistent SD-MMEJ outcomes (deletions in darker colors and indels in lighter colors), across different DR-*white* insertions, EC n = 86; HC1 n = 14; HC2 n = 280; HC3 n = 104; HC4 = 74. *$p<0.05$ by two-proportion z-test for the comparison between loop-out consistent deletion events of HC2 and HC4 vs HC1. EC n = 59, HC1 n = 9, HC2 n = 211, HC3 n = 52, HC4 n = 61. All other comparisons are not significant.

Figure 4: Distribution of deletion boundaries. Deletion boundaries associated with different Alt-EJ pathways for the DR-*white* insertions EU (A), HC1 (B), HC2 (C), HC3 (D), HC4 (E) are shown. Deletion boundaries were defined as the first base proximal and first base distal to the deleted bases, with ambiguous bases aligned to the left of the cut site. Diagonal lines highlight short deletions that could also result from cNHEJ. + indicates all deletion boundaries to the left or right of the sequence shown.

Figure 5: Spread of deletions for Mutagenic EJ repair. Heat map plots show the spread of deletions for SD-MMEJ-consistent (A) and SD-MMEJ inconsistent (B) repair events across different DR-*white* insertions. The frequency of deletion at each bp (color-coded vertical legend) was calculated as the number of events with that bp deleted divided by the total number of events. Red text highlights the TTAT/AATA overhangs produced by I-SceI cutting. Frequencies of deletions were compared between the heterochromatic inserts and euchromatic insert by two-proportion z-test at each bp. * $p < 0.05$; ** $p < 0.005$; *** $p < 0.001$. (A) EC: n = 58; HC1 n = 9; HC2 n = 198; HC3 n = 52; HC4 n = 58. (B) EC n = 19; HC1 n = 5; HC2 n = 56; HC3 n = 48; HC4 n = 7.

Figure 6: Position of repeat motifs for SD-MMEJ. Heat map plots show the frequency of usage of different repeat motifs 2 in SD-MMEJ repair resulting in deletions across different DR-*white* insertions. Color corresponds to frequency as indicated by the y axis. Arrows highlight motifs used more frequently (red arrows) or less frequently (black arrows) in heterochromatic DR-*white* insertions, relative to the euchromatic one. EC n = 58; HC1 n = 9; HC2 n = 198; HC3 n = 52; HC4 n = 58.

Figure 7: Model for increased frequency of HR and reduced deletion size for mutagenic repair in heterochromatin. In heterochromatin (A) RPA association promotes HR repair and prevents secondary structure accumulation. In euchromatin (B) increased availability for secondary structure formation allows larger deletions to occur.

Supplementary Figure Legends:

Supplementary Fig. 1: Mutagenic repair frequency. The frequency of A) SSA, or B) mutagenic-EJ and non-mutagenic repair/uncut, in white-eyed F2 flies is shown. HC4 showed statistically significant differences between mutagenic-EJ and non-mutagenic repair/uncut when compared with the euchromatic insert by a two proportion z-test. ***$p < 0.001$ All other differences between the euchromatic and each heterochromatic DR-*white* insertions were not significant. Mutagenic repair results in a PCR product across the cut site that cannot be digested by I-SceI. Non-mutagenic repair results in a PCR product across the cut site that can be digested by I-SceI.

Supplementary Fig. 2: Average size of deletions. Violin plots showing the distribution of deletion sizes across all DR-white insertions. The mean and median deletion size for each insert is indicated by a blue and red line respectively.

Supplementary Fig. 3: Common SD-MMEJ and MMEJ repair outcomes. The most frequent motifs used in all insertions and the corresponding deletions are shown. Red rectangle: Overhangs from I-SceI cut. Repair outcomes for alt-EJ repair are shown above their respective possible deletions and repeat motifs. Repeat motifs are shown in red and underlined. Ambiguous bases that could align to the left or right of the break are marked by brackets. The respective SD-MMEJ repair mechanism (snap-back or loop-out) is shown to the right of the deletion. Microhomology used for MMEJ repair is shown in red.

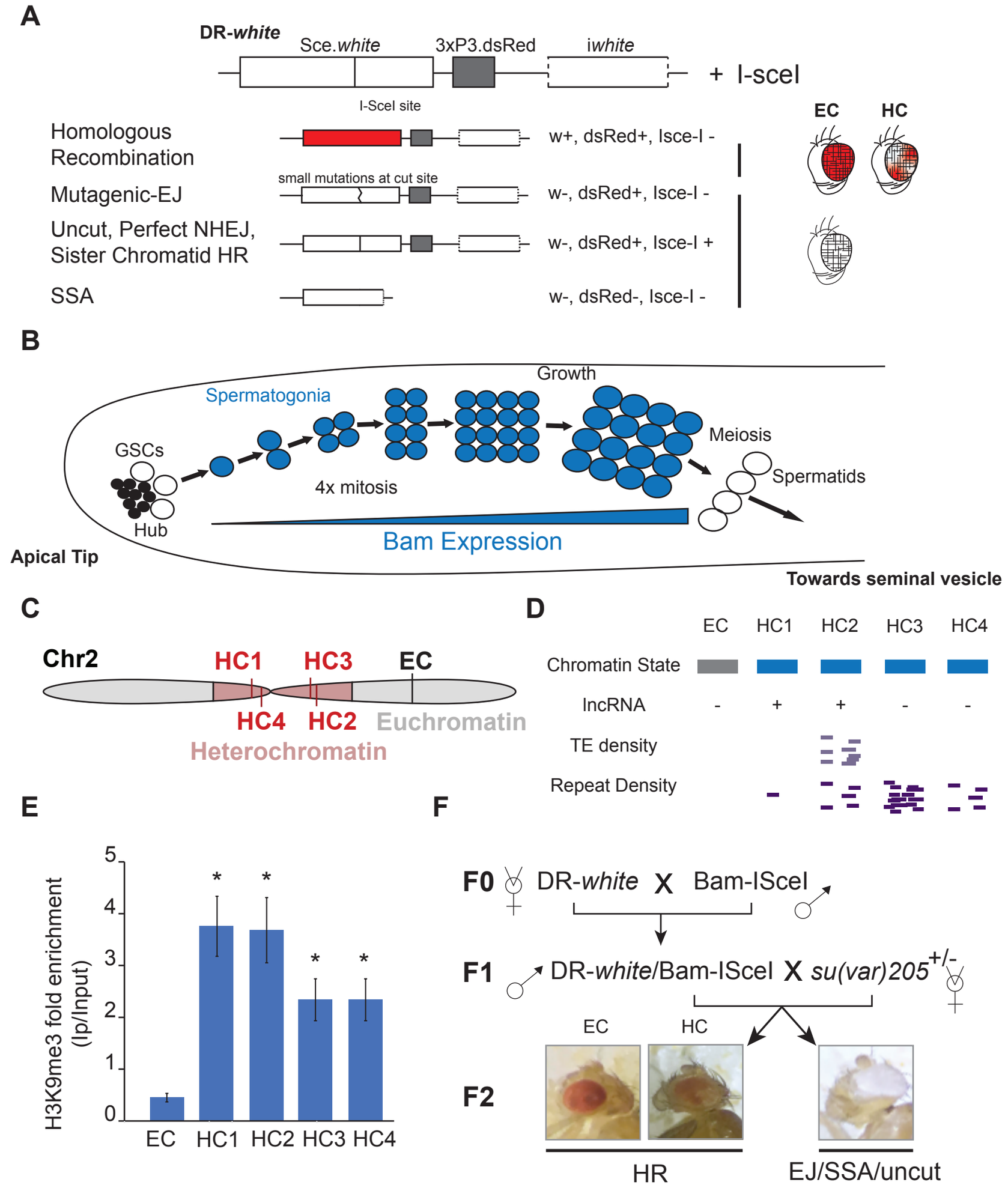
-

References

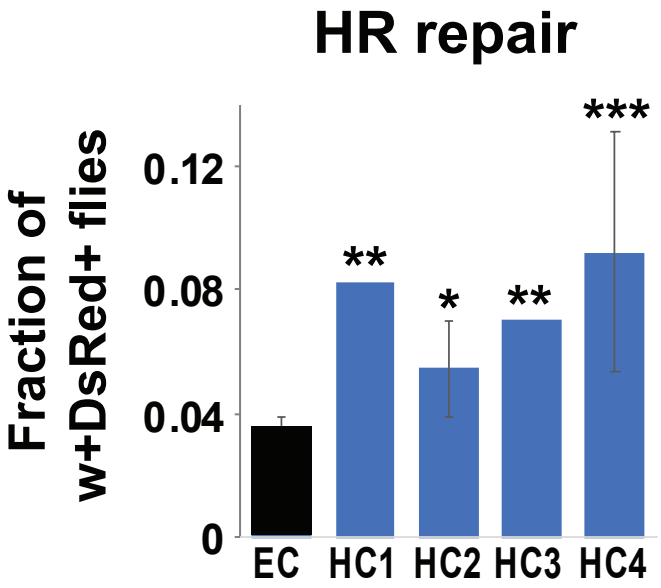
1. Chiolo, I., et al., Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. Cell, 2011. **144**(5): p. 732-44.
2. Ryu, T., et al., Heterochromatic breaks move to the nuclear periphery to continue recombinational repair. Nat Cell Biol, 2015. **17**(11): p. 1401-11.
3. Caridi, C.P., et al., Nuclear F-actin and myosins drive relocalization of heterochromatic breaks. Nature, 2018. **559**(7712): p. 54-60.
4. Tsouroula, K., et al., Temporal and Spatial Uncoupling of DNA Double Strand Break Repair Pathways within Mammalian Heterochromatin. Mol Cell, 2016. **63**(2): p. 293-305.
5. Janssen, A., et al., A single double-strand break system reveals repair dynamics and mechanisms in heterochromatin and euchromatin. Genes Dev, 2016. **30**(14): p. 1645-57.
6. Beucher, A., et al., ATM and Artemis promote homologous recombination of radiation-induced DNA double-strand breaks in G2. EMBO J, 2009. **28**(21): p. 3413-27.
7. Chan, S.H., A.M. Yu, and M. McVey, Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in Drosophila. PLoS Genet, 2010. **6**(7): p. e1001005.
8. Yu, A.M. and M. McVey, Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. Nucleic Acids Res, 2010. **38**(17): p. 5706-17.
9. Hanscom, T. and M. McVey, Regulation of Error-Prone DNA Double-Strand Break Repair and Its Impact on Genome Evolution. Cells, 2020. **9**(7).
10. Chang, H.H.Y., et al., Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.
11. McVey, M., D. Radut, and J.J. Sekelsky, End-joining repair of double-strand breaks in Drosophila melanogaster is largely DNA ligase IV independent. Genetics, 2004. **168**(4): p. 2067-76.
12. Preston, C.R., C.C. Flores, and W.R. Engels, Differential usage of alternative pathways of double-strand break repair in Drosophila. Genetics, 2006. **172**(2): p. 1055-68.
13. Wei, D.S. and Y.S. Rong, A genetic screen for DNA double-strand break repair mutations in Drosophila. Genetics, 2007. **177**(1): p. 63-77.
14. Schep, R., et al., Impact of chromatin context on Cas9-induced DNA double-strand break repair pathway balance. Molecular Cell, 2021. **81**(10): p. 2216-2230.e10.

15. Khodaverdian, V.Y., et al., Secondary structure forming sequences drive SD-MMEJ repair of DNA double-strand breaks. Nucleic Acids Res, 2017. **45**(22): p. 12848-12861.
16. Hanscom, T., et al., Characterization of sequence contexts that favor alternative end joining at Cas9-induced double-strand breaks. Nucleic Acids Res, 2022. **50**(13): p. 7465-7478.
17. Do, A.T., et al., Double-strand break repair assays determine pathway choice and structure of gene conversion events in Drosophila melanogaster. G3 (Bethesda), 2014. **4**(3): p. 425-32.
18. Chen, H.M., et al., An enhanced gene targeting toolkit for Drosophila: Golic+. Genetics, 2015. **199**(3): p. 683-94.
19. White-Cooper, H., Tissue, cell type and stage-specific ectopic gene expression and RNAi induction in the Drosophila testis. Spermatogenesis, 2012. **2**(1): p. 11-22.
20. Delabaere, L., et al., Aging impairs double-strand break repair by homologous recombination in Drosophila germ cells. Aging Cell, 2017. **16**(2): p. 320-328.
21. Kharchenko, P.V., et al., Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature, 2011. **471**(7339): p. 480-5.
22. Donaldson, K.M., A. Lui, and G.H. Karpen, Modifiers of terminal deficiency-associated position effect variegation in Drosophila. Genetics, 2002. **160**(3): p. 995-1009.
23. Decottignies, A., Alternative end-joining mechanisms: a historical perspective. Front Genet, 2013. **4**: p. 48.
24. Sfeir, A. and L.S. Symington, Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? Trends Biochem Sci, 2015. **40**(11): p. 701-714.
25. Ren, C., S.E. Finkel, and J. Tower, Conditional inhibition of autophagy genes in adult Drosophila impairs immunity without compromising longevity. Exp Gerontol, 2009. **44**(3): p. 228-35.
26. Venken, K.J., et al., MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. Nat Methods, 2011. **8**(9): p. 737-43.
27. Le, H.D., et al., A high proportion of genes involved in position effect variegation also affect chromosome inheritance. Chromosoma, 2004. **112**(6): p. 269-76.

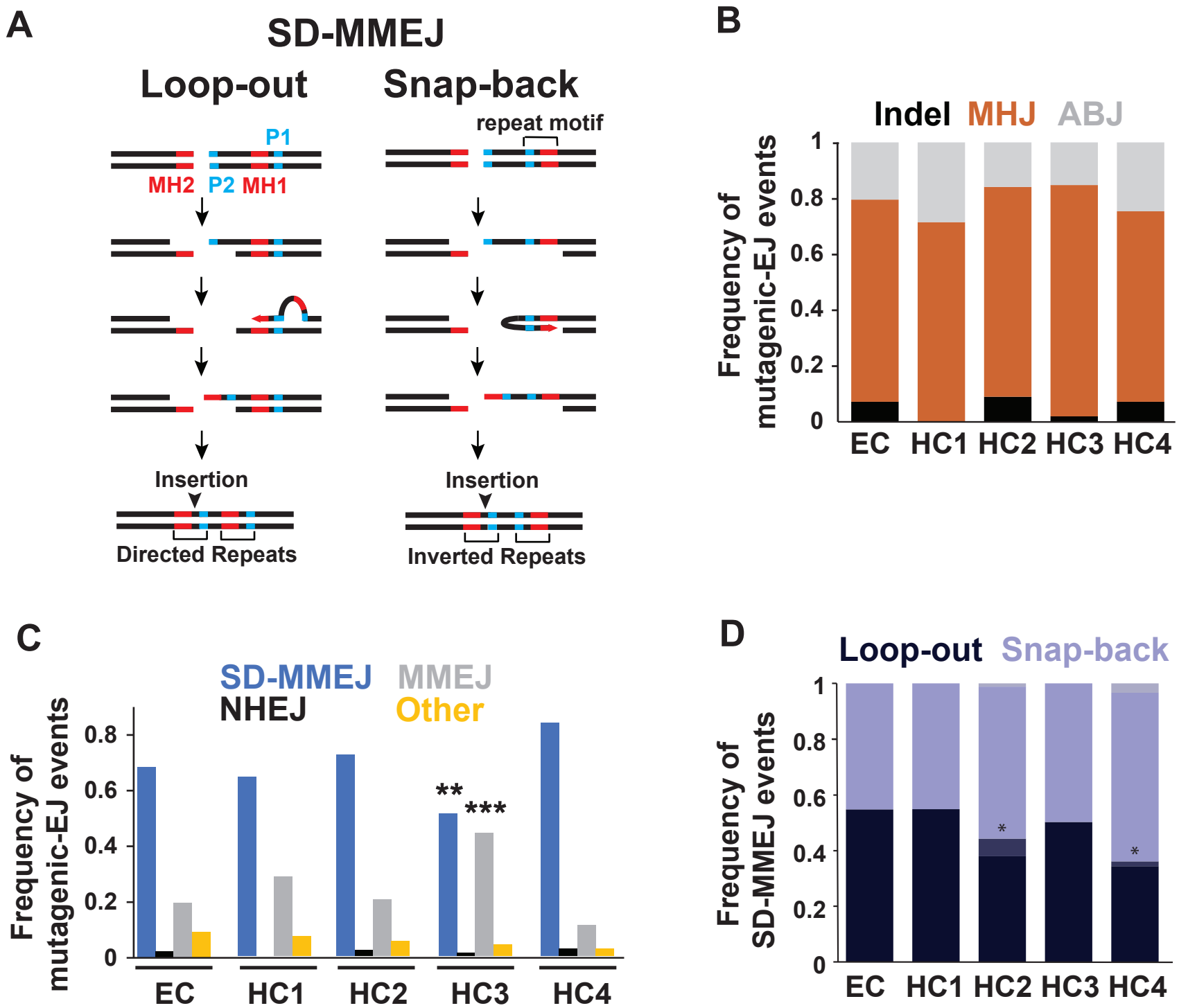
Miller et al., Figure 1



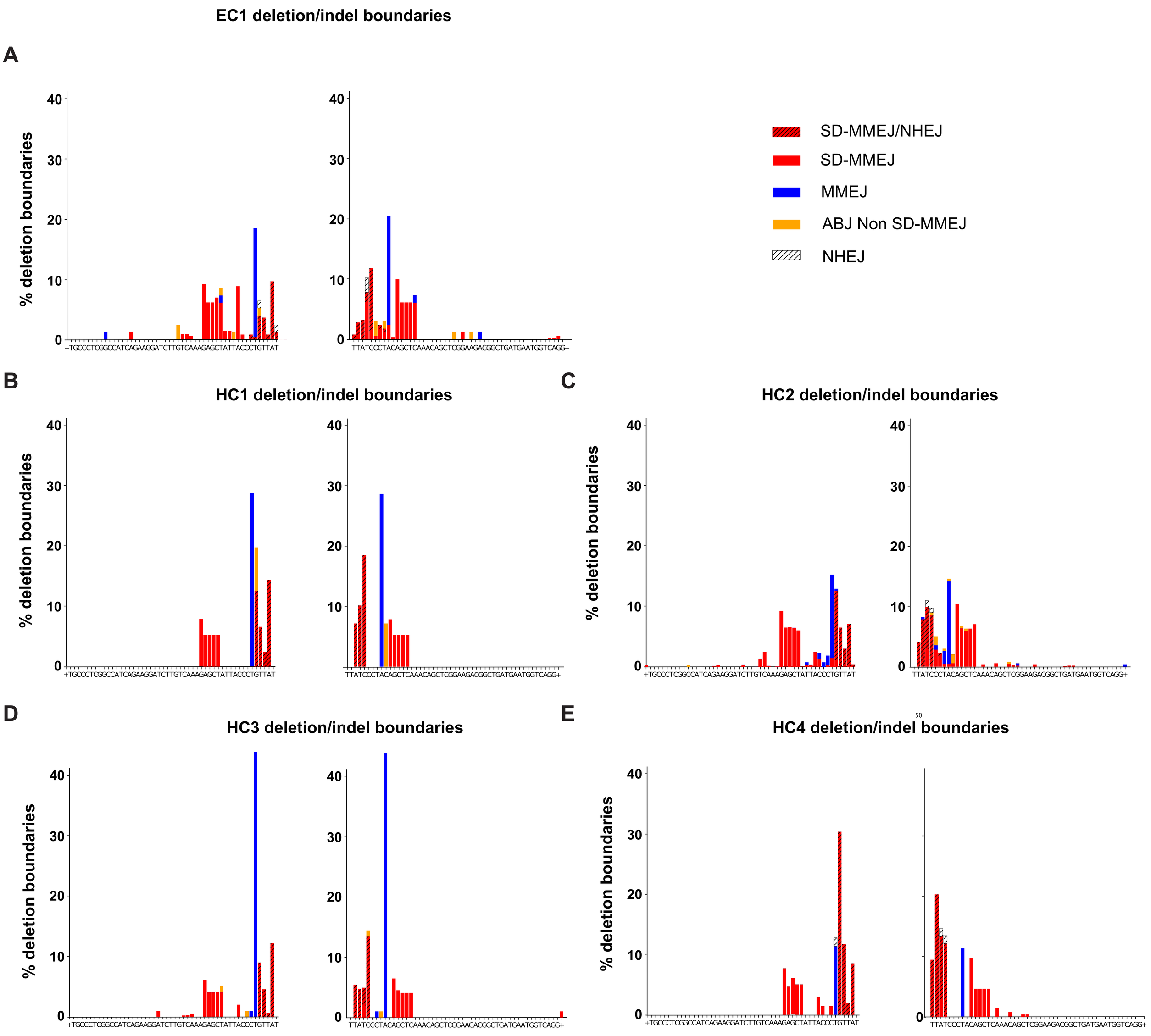
Miller et al., Figure 2



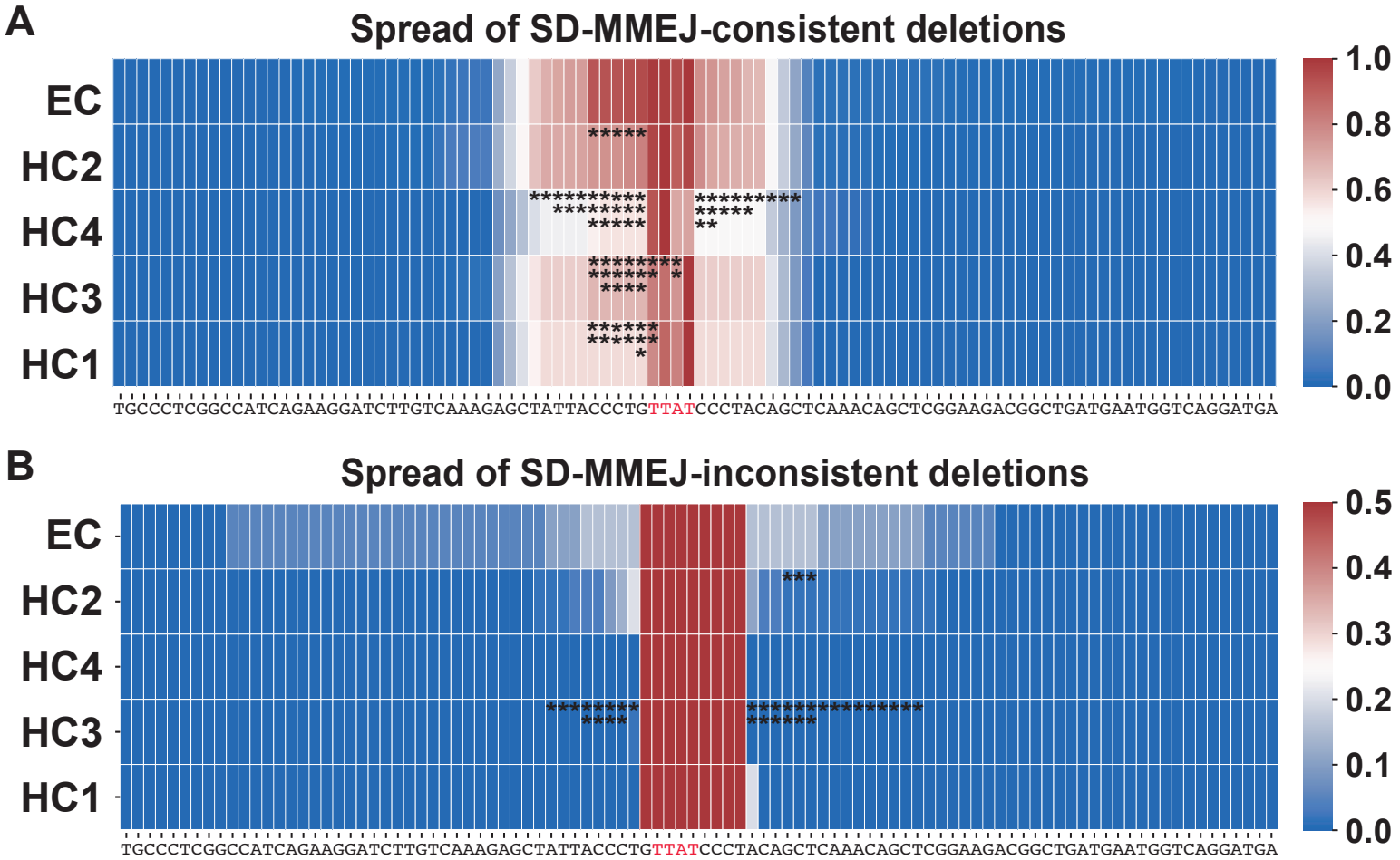
Miller et al., Figure 3



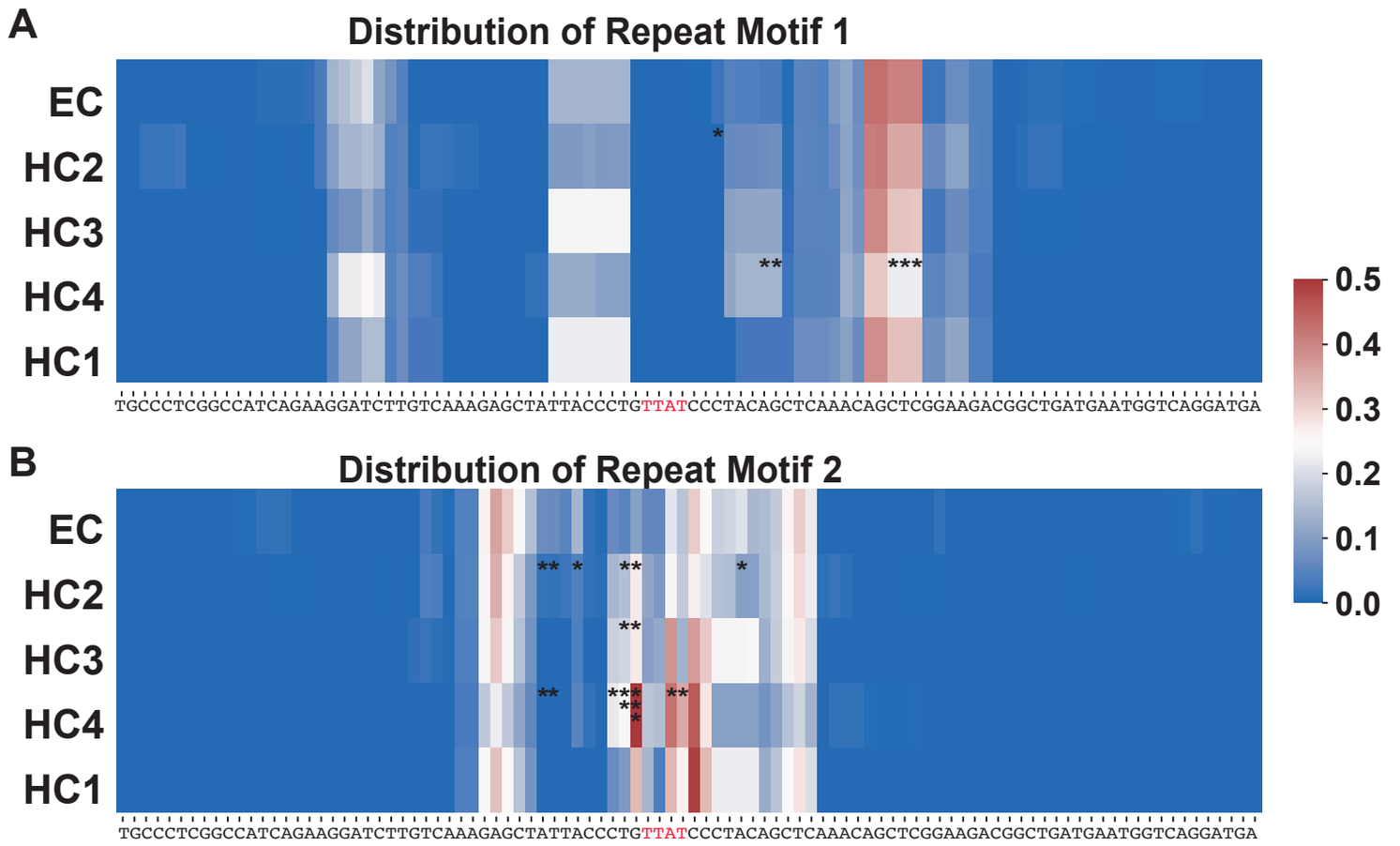
Miller et al., Figure 4



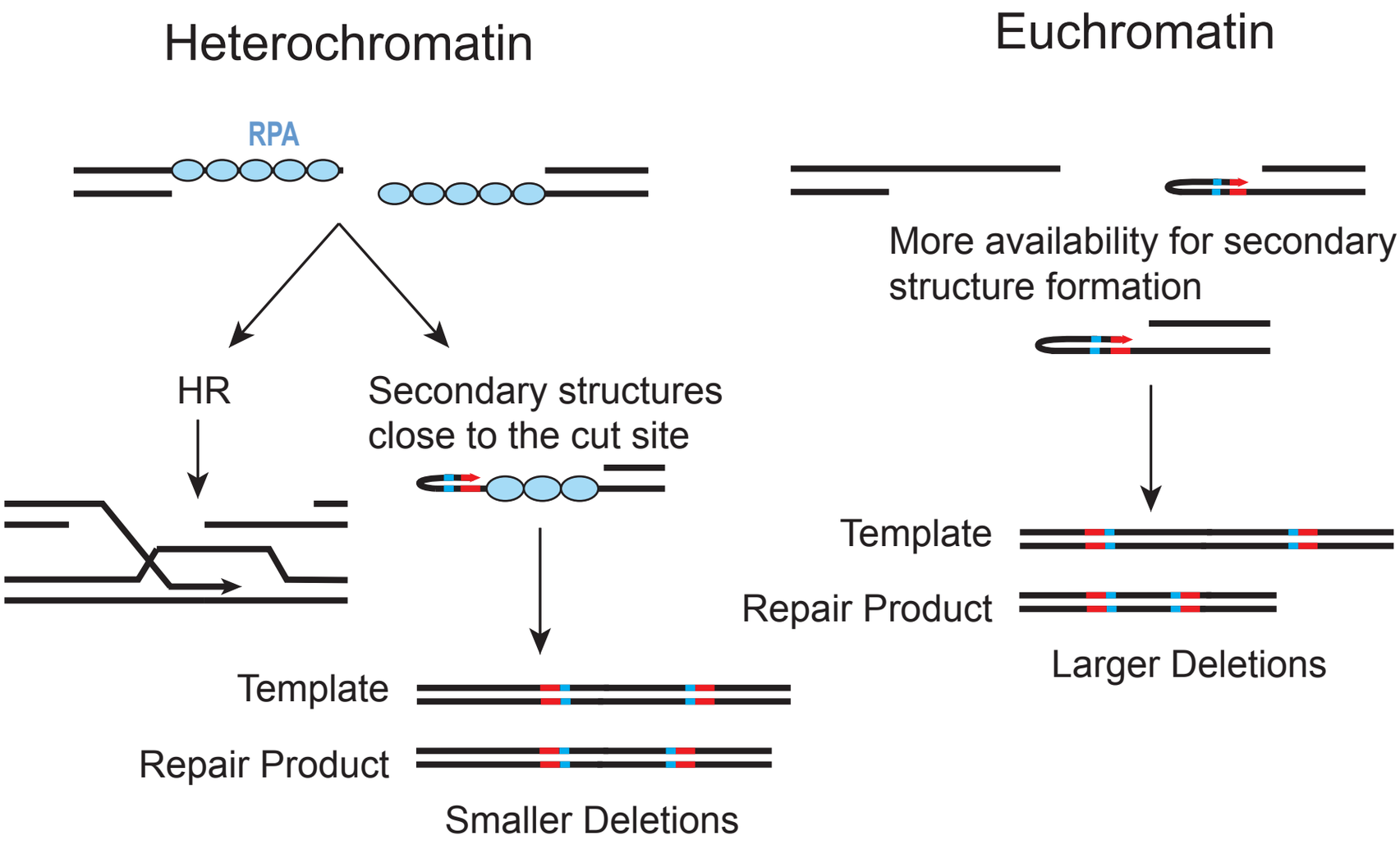
Miller et al., Figure 5



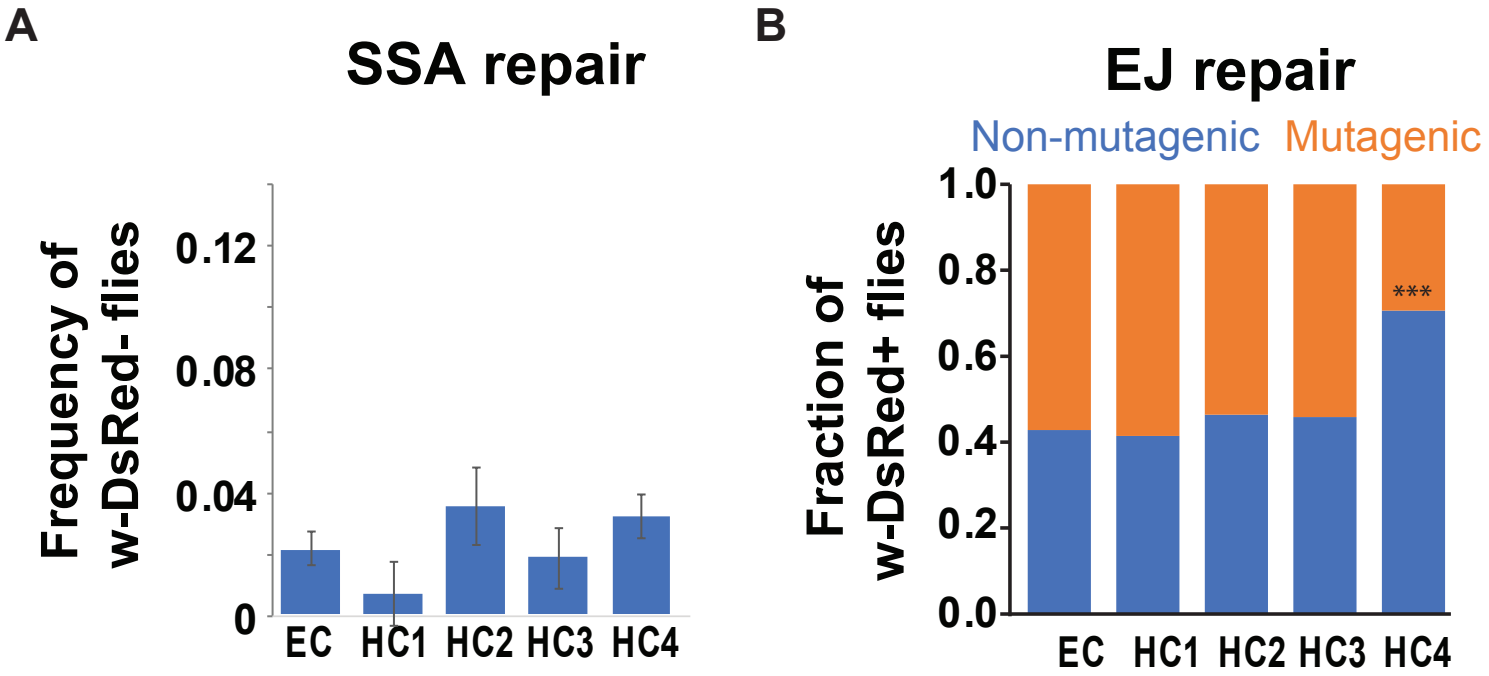
Miller et al., Figure 6



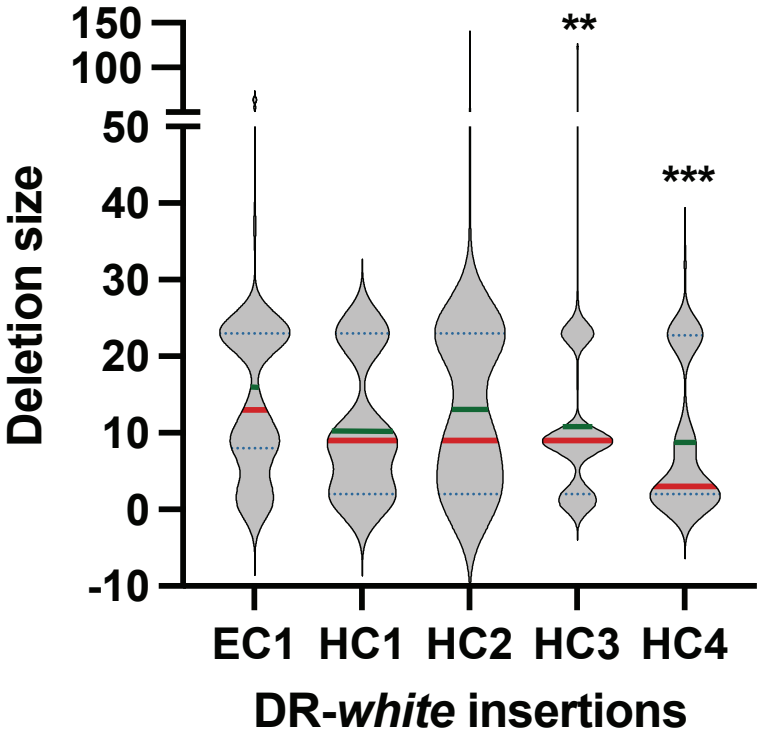
Miller et al., Figure 7



Miller et al., Supplementary Figure 1:



Miller et al., Supplementary Figure 2:



Miller et al., Supplementary Figure 3:

Template

ATCTTGTCAAAGAGCTATTACCCTGTTATCCCTACAGCTCAAACAGCTCGGAAGACGG

Recurrent SD-MMEJ Repair Outcomes and Repeat Motifs

Repair Outcome 1

ATCTTGTCAAAG[AGCT]----------CAAACAGCTCGGAAGACGG

Potential Deletion	Mechanism
ATCTTGTCAAAG----------[AGCT]CAAACAGCTCGGAAGACGG	loop-out
ATCTTGTCAAAG----------[AGCT]CAAACAGCTCGGAAGACGG	snap-back
ATCTTGTCAAAG[AG----------CT]CAAACAGCTCGGAAGACGG	snap-back
ATCTTGTCAAAG[AG----------CT]CAAACAGCTCGGAAGACGG	loop-out
ATCTTGTCAAAG[AGCT]----------CAAACAGCTCGGAAGACGG	loop-out

Repair Outcome 2

ATCTTGTCAAAGAGCTATTACCCTG[T]-ATCCCTACAGCTCAAACAGCTCGGAAGACGG

Potential Deletion	Mechanism
ATCTTGTCAAAGAGCTATTACCCTG[T]-ATCCCTACAGCTCAAACAGCTCGGAAGACGG	snap-back
ATCTTGTCAAAGAGCTATTACCCTG-[T]ATCCCTACAGCTCAAACAGCTCGGAAGACGG	snap-back

Repair Outcome 3

ATCTTGTCAAAGAGCTATTACCCTGTTA-CCCTACAGCTCAAACAGCTCGGAAGACGG

Potential Deletion	Mechanism
ATCTTGTCAAAGAGCTATTACCCTGTTA-CCCTACAGCTCAAACAGCTCGGAAGACGG	loop-out

Recurrent MMEJ Repair Outcomes and Microhomologies

Repair Outcome 4

ATCTTGTCAAAGAGCTATTA[CCCT]-------ACAGCTCAAACAGCTCGGAAGACGG

Miller et al, Supplementary Table 1

Annota	Chromatin region	Integration Site (Release 6 coordinates)	MiMIC line	Reference
EC	Euchromatin	51C1 (2R:14737578..14739242)	**MI00794**	Remade as in Jannsen et al., 2016
HC1	Heterochromatin	40F7 (2L:22538230..22548452)	**MI07270**	This paper
HC2	Heterochromatin	41C2 (2R:4,792,318..4,792,318)	**MI01294**	This paper
HC3	Heterochromatin	41A1 (2R:4,201,370..4,201,370)	**MI00234**	This paper
HC4	Heterochromatin	40F7 (2L:23,049,836..23,049,836)	**MI04108**	This paper