Dual Roles for DNA Polymerase Theta in Alternative End-Joining Repair of Double-Strand Breaks in Drosophila

Sze Ham Chan1, Amy Marie Yu1, Mitch McVey1,2*

1 Department of Biology, Tufts University, Medford, Massachusetts, United States of America, 2 Program in Genetics, Tufts Sackler School of Graduate Biomedical Sciences, Boston, Massachusetts, United States of America

Abstract

DNA double-strand breaks are repaired by multiple mechanisms that are roughly grouped into the categories of homology-directed repair and non-homologous end joining. End-joining repair can be further classified as either classical non-homologous end joining, which requires DNA ligase 4, or “alternative” end joining, which does not. Alternative end joining has been associated with genomic deletions and translocations, but its molecular mechanisms are largely uncharacterized. Here, we report that Drosophila melanogaster DNA polymerase theta (pol theta), encoded by the mus308 gene and previously implicated in DNA interstrand crosslink repair, plays a crucial role in DNA ligase 4-independent alternative end joining. In the absence of pol theta, end joining is impaired and residual repair often creates large deletions flanking the break site. Analysis of break repair junctions from flies with mus308 separation-of-function alleles suggests that pol theta promotes the use of long microhomologies during alternative end joining and increases the likelihood of complex insertion events. Our results establish pol theta as a key protein in alternative end joining in Drosophila and suggest a potential mechanistic link between alternative end joining and interstrand crosslink repair.

Citation: Chan SH, Yu AM, McVey M (2010) Dual Roles for DNA Polymerase Theta in Alternative End-Joining Repair of Double-Strand Breaks in Drosophila. PLoS Genet 6(7): e1001005. doi:10.1371/journal.pgen.1001005

Editor: R. Scott Hawley, Stowers Institute for Medical Research, United States of America

Received September 15, 2009; Accepted May 27, 2010; Published July 1, 2010

Copyright: © 2010 Chan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by a grant from the NSF (MCB-0643253) and by a New Scholar in Aging award to MM from the Ellison Medical Foundation (www.ellisonfoundation.org). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: mitch.mcvey@tufts.edu

† These authors contributed equally to this work.

Introduction

DNA double-strand breaks (DSBs) and interstrand crosslinks pose serious threats to cell survival and genome stability. Because these lesions compromise both strands of the double helix, they impede DNA replication and transcription and therefore must be removed in a timely and coordinated manner. Interstrand crosslink repair has been shown to involve a DSB intermediate in some cases (reviewed in [1]). Therefore, there may be substantial mechanistic overlap in the processes used during repair of these two lesions.

Error-free repair of DSBs can be accomplished through homologous recombination (HR) with an undamaged homologous template (reviewed in [2]). However, in contexts where suitable templates for HR do not exist, error-prone repair mechanisms are also used. For example, non-homologous end joining (NHEJ) frequently creates small insertions and deletions during DSB repair, particularly in cases where the broken ends cannot be readily ligated (reviewed in [3]). Analogously, the use of translesion DNA polymerases during interstrand crosslink repair can result in mutations, due to the reduced fidelity of these polymerases [4,5].

Accumulating evidence suggests that NHEJ is composed of at least two genetically distinct mechanisms. Classical NHEJ (C-NHEJ) involves the sequential recruitment of two highly conserved core complexes (reviewed in [6]). First, the Ku70/80 heterodimer recognizes and binds to DNA ends in a sequence-independent manner, thereby protecting them from degradation. In many eukaryotes, Ku70/80 also recruits DNA-PKcs, forming a synaptic complex that can recruit additional processing enzymes such as the Artemis nuclease and the X family DNA polymerases mu and lambda. These proteins expand the spectrum of broken ends that can be rejoined. The second core complex, composed of DNA ligase 4, XRCC4, and XLF/Cernunnos, catalyzes ligation of the processed ends. Depending on the substrate, C-NHEJ can result in perfect repair of broken DNA, or it can result in small deletions of 1–10 nucleotides and/or insertions of 1–3 nucleotides [7]. Although C-NHEJ can repair blunt-ended substrates, a subset of C-NHEJ products appear to involve annealing at 1–4 nucleotide microhomologous sequences on either side of the break.

Alternative end-joining (alt-EJ) is defined as end-joining repair that is observed in cells or organisms lacking one or more C-NHEJ components (reviewed in [8]). Alt-EJ in yeast is associated with deletions larger than those typically created by C-NHEJ, together with an increased tendency to repair by annealing at microhomologous sequences. Ku and ligase 4-independent end joining observed in mammalian cells also displays an increased tendency towards use of short microhomologies compared to C-NHEJ [9,10]. Therefore, alt-EJ is sometimes called microhomology-mediated end joining (MMEJ) [11]. However, the relationship between MMEJ and alt-EJ is still unclear, and alt-EJ may comprise one or more C-NHEJ-independent repair mechanisms [8].

The importance of alt-EJ repair is highlighted by multiple studies that suggest it may promote chromosome instability and carcinogenesis. Alt-EJ produces chromosome translocations in...
Author Summary

DNA double-strand breaks, in which both strands of the DNA double helix are cut, must be recognized and accurately repaired in order to promote cell survival and prevent the accumulation of mutations. However, error-prone repair occasionally occurs, even when accurate repair is possible. We have investigated the genetic requirements of an error-prone break-repair mechanism called alternative end joining. We have previously shown that alternative end joining is frequently used in the fruit fly, Drosophila melanogaster. Here, we demonstrate that a fruit fly protein named DNA polymerase theta is a key player in this inaccurate repair mechanism. Genetic analysis suggests that polymerase theta may be important for two processes associated with alternative end joining: (1) annealing at short, complementary DNA sequences, and (2) DNA synthesis that creates small insertions at break-repair sites. In the absence of polymerase theta, a backup repair mechanism that frequently results in large chromosome deletions is revealed. Because DNA polymerase theta is highly expressed in many types of human cancers, our findings lay the groundwork for further investigations into how polymerase theta is involved in repair processes that may promote the development of cancer.

mouse embryonic stem cells lacking Ku70 [12] and the use of alt-EJ during V(D)J recombination in C-NHEJ-deficient murine lymphocytes causes complex chromosome translocations and progenitor B cell lymphomas [13]. Furthermore, alt-EJ has been implicated in various translocations associated with chronic myeloid leukemia and human bladder cancer [14,15]. Importantly, alt-EJ also operates during V(D)J rejoining in C-NHEJ-proficient B lymphocytes [16], suggesting that its role in DSB repair is not limited to situations where C-NHEJ is defective. However, alt-EJ is frequently masked by more dominant repair processes that are essential for vertebrate development, making it difficult to study. Therefore, its molecular mechanisms and the proteins involved remain largely unknown.

Several lines of evidence demonstrate that Drosophila is an excellent model system in which to study alt-EJ in a metazoan. The Drosophila genome lacks several mammalian C-NHEJ components, including DNA-PKcs, and Artemis. This may predispose flies towards non-C-NHEJ repair. Consistent with this, we have previously shown that a DSB caused by excision of a P element transposon in flies is readily repaired by a DNA ligase 4-independent end-joining process [17]. Interestingly, although Drosophila orthologs for the Pol X family DNA polymerases mu and lambda have not been identified [18], we and others have found evidence for polymerase activity in Drosophila end-joining repair [17,19,20]. Specifically, end joining in flies is often associated with the insertion of nucleotides at repair junctions, frequently involving imperfect repeats of 5–8 nucleotides. Full or partial templates for the insertions, occasionally possessing mismatches, can often be identified in adjacent sequences, suggesting the action of an error-prone polymerase. Similar templated nucleotides (T-nucleotides) have previously been identified at translocation breakpoints in human lymphomas [21–23]. Therefore, T-nucleotides could represent a signature of alt-EJ and may be informative regarding its molecular mechanisms.

Additional insight into alt-EJ is provided by recent reports suggesting a mechanistic link between alt-EJ and interstrand crosslink repair. For example, a study of two Chinese hamster ovary cell lines sensitive to the crosslinking agent mitomycin C found that they were also deficient in alt-EJ [24]. Furthermore, certain interstrand crosslink-sensitive cell lines from Fanconi Anemia patients are also impaired in DNA-PKcs-independent rejoining of linearized plasmids [25]. Based on these reports, we hypothesized that additional mechanistic insight into both interstrand crosslink repair and alt-EJ could be gained by searching for mutants defective in both processes. To test this, we have screened Drosophila mutants that are sensitive to DNA crosslinking agents for additional defects in alt-EJ repair. In this work, we describe our studies with one such mutant, mus308.

The mus308 (mutagen sensitive 308) mutant was originally identified by its extreme sensitivity to interstrand crosslinking agents but normal resistance to alkylating agents [26]. Subsequently, mus308 was found to code for DNA polymerase theta, which is most similar to a family DNA polymerases such as Escherichia coli Pol I [27]. Orthologs of polymerase theta (hereafter referred to as pol θ) are found in many metazoans, including Caenorhabditis elegans, Arabidopsis thaliana, and Homo sapiens, but not in unicellular eukaryotes, including the yeasts [28–30]. Several lines of evidence suggest that pol θ may play an important role in maintaining genome stability. Similar to flies, C. elegans with mutations in POLQ-1 are defective in repair of interstrand crosslinks [28]. Mice lacking pol θ (chaos I mutants) have a high frequency of spontaneous and mitomycin C-induced micronuclei in erythrocytes, consistent with genomic instability [31]. In addition, vertebrate pol θ orthologs have been implicated in a wide range of repair processes, including base excision repair, bypass of abasic sites, and somatic hypermutation of immunoglobulin genes [32–36]. Finally, upregulation of pol θ is observed in a variety of human tumors and is associated with a poor clinical outcome, suggesting that its overexpression may contribute to cancer progression [37].

Pol θ is unusual in possessing an N-terminal helicase-like domain and a C-terminal polymerase domain. Although pol θ purified from human cell lines and Drosophila has error-prone polymerase activity and single-stranded DNA-dependent ATPase activity, helicase activity has not been demonstrated in vitro [30,38,39]. Therefore, it remains unclear exactly how the structure of pol θ relates to its multiple functions in DNA repair in different organisms.

We report here that in addition to its role in DNA interstrand crosslink repair, Drosophila pol θ is involved in end-joining repair of DSBs. This alt-EJ mechanism operates independently of both Rad51-mediated HR and ligase 4-dependent C-NHEJ. Genetic analysis using separation-of-function alleles provides support for distinct roles of both the N- and C-terminal domains of pol θ in alt-EJ. Collectively, our data support a model in which helicase and polymerase activities of Drosophila pol θ cooperate to generate single-stranded microhomologous sequences that are utilized during end alignment in alt-EJ.

Results

Drosophila pol θ is important for both interstrand crosslink repair and end-joining repair of DNA double-strand breaks

Drosophila mus308 mutants were initially identified based on their sensitivity to low doses of chemicals that induce DNA interstrand crosslinks [26]. To confirm this phenotype, we assembled a collection of previously identified mus308 mutant alleles [40,41] and measured the ability of hemizygous mutant larvae to survive exposure to the crosslinking agent methylene blue (nitrogen mustard). Of the mutants that we tested, four were unable to survive exposure to 0.005% methylene blue: D2, D5, 2003, and 3294 (data not shown), consistent with their inability to repair interstrand crosslinks.
To determine the molecular lesions responsible for mechlorethamine sensitivity, we sequenced the entire mus308 coding region of flies hemizygous for each mutant allele. Pol θ possesses both a conserved N-terminal helicase-like domain and a C-terminal pol I-like polymerase domain (Figure 1, Figure S1) [30]. Three of the four alleles contain unique sequence changes that are predicted to affect pol θ primary structure (Figure 1, Figure S2, and Figure S3). The 2003 allele is a nonsense mutation upstream of the polymerase domain, while the D2 allele is a missense mutation that alters a highly conserved proline in the conserved N-terminus. The 3294 allele changes an invariant glycine in the helicase domain to serine. Interestingly, this residue is conserved in the related mus301 helicase, but not in other DNA helicases (data not shown). No mutations were found in the coding sequence of the D2 allele. Because homozygous D2 flies have undetectable levels of pol θ protein [38], the D2 mutation may affect a regulatory region of mus308.

One explanation for the extreme sensitivity of mus308 mutants to mechlorethamine could be a defect in the repair of certain types of DSB intermediates that are created during crosslink repair. To test this, we exposed flies hemizygous for each mutant allele to increasing doses of ionizing radiation (IR). Although IR creates many different types of lesions, unreplicated DSBs are thought to be the main cause of cell death following irradiation. All four mus308 mutants survived IR exposures as high as 4000 rads (Figure 2A), although bristle and wing defects characteristic of apoptotic cell death were frequently observed at high doses. Drosophila lig4 mutants, which are completely defective in C-NHEJ, also survive IR doses in excess of 4000 rads [17]. However, spn-A mutants, which lack the Rad51 recombinase required for strand invasion during homologous recombination initiation [42], are highly sensitive to IR [17]. Thus, in Drosophila, HR is the dominant mechanism used to repair IR-induced DSBs.

To test whether pol θ acts to repair IR damage in the absence of HR, we created mus308 spn-A double mutants and exposed them to doses of 125–1000 rads. Strikingly, doses as low as 125 rads resulted in almost complete killing of mus308 spn-A mutants (Figure 2B). In contrast, lig4 spn-A double mutants are only slightly more sensitive than spn-A single mutants to IR [17]. Thus, in the absence of HR, pol θ participates in a process crucial for repair of damage caused by ionizing radiation.

Because interstrand crosslink repair and alternative end joining have been shown to have partially overlapping genetic requirements in mammals [24,25], we hypothesized that the extreme sensitivity of mus308 spn-A mutants to IR might relate to a role of pol θ in an alternative end-joining mechanism. To explore this hypothesis, we tested each mus308 mutant allele using a site-specific double-strand break repair assay that can distinguish between synthesis-dependent strand annealing (SDSA, a specific type of HR) and end joining (EJ) (Figure 3A) [43]. We have previously shown that the majority of end joining observed in this assay occurs independently of DNA ligase 4, and is therefore a form of alt-EJ [17]. In this system, excision of a P element (P(Pα)) located on the X chromosome is catalyzed by P transposase, resulting in a 14-kilobase gap relative to an undamaged sister chromatid. The DNA ends remaining after excision each have 17-nucleotide non-complementary 3′ single-stranded overhangs [44].

These ends are highly recombinogenic and repair by SDSA is initially favored. However, because repair synthesis in this system is not highly processive, most repair products that are recovered from wild-type flies result from incomplete repair synthesis from one or both sides of the break, followed by end joining of the nascent DNA (SDSA+EJ events) [45]. To quantitate the percentage of repair events that derive from each mechanism, repair events are recovered from male pre-meiotic germline cells by mating individual males to females homozygous for the P(Pα) element. Each of the resulting female progeny represents a single repair event that can be classified by eye color. Red-eyed females inherit a repair event involving homology-dependent synthesis that generated complementary single-stranded regions that subsequently anneal (repair by SDSA). Yellow-eyed females inherit a chromosome that was repaired by EJ or SDSA+EJ mechanisms (these repair events are hereafter referred to as (SDSA)+EJ; for further details, see Materials and Methods).

Overall, the results from the P(Pα) assay indicated that mus308 mutants are defective in end-joining repair of DSBs. We observed no decrease in the percentage of red-eyed progeny recovered from mus308 mutant males (Figure 3B), suggesting that SDSA repair is not impaired when pol θ is missing or defective. In contrast, all four mus308 mutant alleles resulted in a significantly decreased percentage of yellow-eyed progeny relative to wild type (p<0.001, Kruskal-Wallis test). Because yellow-eyed progeny can only result from a repair mechanism involving end joining, these data suggest that pol θ is involved in an end-joining process.

The percentage of genomic deletions during repair of site-specific DSBs was determined by analyzing the frequency of genomic deletions during repair of site-specific DSBs

**Figure 1. Schematic of the mus308 gene.** Exons are represented by boxes. The helicase-like domain (light gray shading) and polymerase domain (dark gray shading) are shown. Locations of the D5, 3294, and 2003 point mutations are indicated with arrows (numbers correspond to amino acid positions in the protein). Not shown is the D2 mutation, which results in severely reduced levels of pol θ.

doi:10.1371/journal.pgen.1001005.g001
To determine whether mutation of mus308 also results in repair-associated deletions, we took advantage of the fact that deletions can be easily scored in the P element excision assay. Because P[w+] is inserted in the essential scalloped (sd) gene, repair events that delete into sd exons cause a scalloped-wing phenotype when recovered in heterozygous females and lethality in hemizygous males [48,49]. We observed a substantial increase in the percentage of deletion-associated repair events isolated from mus308 mutants relative to wild type (Figure 3D). Overall, the total percentage of end-joining repair events involving deletions recovered from mus308 mutants was elevated from 3- to 26-fold over wild type, depending on the mus308 allele tested.

Previously, we observed a similar deletion-prone phenotype in flies lacking the DmBltm protein, which is involved in repair of DNA Pol θ Promotes Alt-EJ

Figure 2. Drosophila pol θ is required for repair of IR–induced damage in the absence of Rad51.

(A) mus308 mutants are not sensitive to ionizing radiation. Female flies heterozygous for the indicated mus308 alleles were mated to Df(3R)Sz29/TM3 males and third instar larval progeny were irradiated with increasing doses of IR. The percent survival of mus308 hemizygous mutants was calculated relative to an unirradiated control. Each dose was repeated twice; the average of the two experiments is shown. (B) spn-A mus308 double mutants are extremely sensitive to ionizing radiation. Heterozygous spn-A057 mus308D5 females were mated to heterozygous spn-A093 mus308D5 males and third instar larval progeny were irradiated with increasing doses of IR. The percent survival of spn-A mus308 compound heterozygotes was calculated relative to an unirradiated control. Each dose was repeated at least twice. Error bars represent standard deviations.

doi:10.1371/journal.pgen.1001005.g002
Figure 3. Pol θ is involved in end-joining repair. (A) A P element excision assay monitors DSB repair outcomes. The P(wt) transposon is inserted in a 5.2 kb intron of the essential scalloped (sd) gene. A copia retrotransposon flanked by long terminal repeats (LTR) is inserted into an exon of the white gene, causing reduced expression of white. Excision of P(wt) results in a 14-kilobase gap relative to an intact sister chromatid. Following excision, repair events from the male pre-meiotic germline are recovered in female progeny and the method of repair is inferred from their eye color.

SDSA = synthesis-dependent strand annealing, SDSA + EJ = SDSA + end joining, EJ = end joining. Accompanying deletion into sd exons (EJ with deletion) results in a scalloped-wing phenotype and/or male lethality. (B) Pol θ mutants are defective specifically in end-joining repair of DSBs. Each bar represents the mean percentage of repair events recovered from independent males possessing the indicated mus308 allele in trans to DNA Pol θ Promotes Alt-EJ

PLoS Genetics | www.plosgenetics.org 5 July 2010 | Volume 6 | Issue 7 | e1001005
DSBs by SDSA [48]. Because our data did not support a role for pol θ in homologous recombination, we expected the deletion-prone phenotype of mus308 mutants to persist even in SDSA-deficient flies. To confirm this, we assayed repair following P[w +] excision in mus308 mutants lacking the Rad51 protein, which renders them unable to carry out HR repair [32,45].

As expected, PCR analysis of repair products showed that SDSA was abolished in both spn-A and spn-A mus308 mutants (data not shown). Approximately 17% of P[w +] chromosomes recovered from spn-A mutant males showed evidence for end joining at the 17-nucleotide overhangs that are created by P transposase (Figure 4A and Table 1); the other 83% of P[w +] chromosomes recovered were presumably uncut. We observed a 30–50% decrease in end-joining repair products in spn-A mus308 double mutants compared to spn-A mutants (p<0.001, Kruskal-Wallis test), confirming a unique role for pol θ in end joining when HR is absent. Importantly, mutation of mus308 still caused an increased percentage of deletions in the absence of Rad51 (Figure 4B). From these data, we conclude that the deletions formed during break repair in mus308 mutants are not the result of aborted SDSA. Rather, they are a consequence of a deletion-prone repair mechanism that operates in the absence of both SDSA and pol θ-dependent end joining.

During the course of these experiments, we made a number of observations suggesting that Rad51 and pol θ act in parallel and distinct DSB repair mechanisms. First, we recovered fewer spn-A and spn-A mus308 double mutants (data not shown) compared to single mutants. In contrast, we observed a further decrease in the percentage of end-joining repair products recovered from mus308 mutants (p<0.001, Kruskal-Wallis test), consistent with results obtained using mus308 single mutants. In contrast, we observed a further decrease in the percentage of end joining repair products recovered from mus308 double mutants relative to mus308 mutants, from 3.0% to 1.3% (P<0.01, Kruskal-Wallis test). Previously, we have shown that end-joining repair of DSBs induced by P[w +] excision is unaffected in mus308 mutants [17]. Therefore, the removal of pol θ-mediated end joining reveals a previously hidden role for DNA ligase 4 in the repair of DSBs created by P transposase. Strikingly, although only 50% of end-joining products isolated from mus308 mutants involved large, male-lethal deletions, 100% of end-joining products recovered from mus308 double mutant males were associated with large deletions (Figure 4D).

From these results, we conclude that at least three distinct mechanisms for end-joining repair exist in Drosophila. One, which corresponds to C-NHEJ, requires DNA ligase 4 and other canonical NHEJ proteins, including XRCC4, Ku70, and Ku80 [46,47,50]. Another mechanism, which is at least partially independent of DNA ligase 4, is defined by a requirement for pol θ and corresponds to alt-EJ. Interestingly, alt-EJ appears to be used more frequently than C-NHEJ, at least for the repair of P element-induced breaks. In the absence of these two repair processes, a Rad51-independent backup mechanism characterized by extensive genomic deletions operates at low efficiency.

Pol θ has two distinct functions in alternative end joining

Alt-EJ repair in Drosophila is frequently associated with annealing at microhomologous sequences of more than four nucleotides and with long DNA insertions at repair junctions [8]. To determine whether pol θ-dependent end joining involves either of these types of repair, we sequenced repair junctions obtained from spn-A and spn-A mus308 double mutants following P[w +] excision. Because we sequenced only one junction per male germline, each junction analyzed represents an independent repair event. Five distinct junction types were identified. Three of these types are characteristic of junctions arising from C-NHEJ in mammalian systems [7]: junctions involving small, 1–3 base pair insertions, junctions involving annealing at 1–3 nucleotide microhomologies, and junctions for which no microhomologies can be identified (apparent blunt end junctions). The other two types of junctions, characteristic of alt-EJ [8], involve annealing at 5–10 nucleotide microhomologous sequences or insertions of more than three base pairs.

Approximately 38% of junctions from spn-A mus308 mutants showed structures considered typical of C-NHEJ repair, while 29% involved annealing at 5–10 nucleotide microhomologies and 13% had insertions of greater than three base pairs (Figure 6A and Table 1). Potential templates for the larger insertions could almost always be identified in flanking sequences. These insertions may be analogous to T-nucleotides that have been observed at translocation breakpoint junctions isolated from certain human cancers [21–23].
When we sequenced repair junctions from spn-A mus308 mutants, we observed two distinct patterns, depending on the mus308 alleles used. For both the D2/2003 and D5/2003 allele combinations, the percentage of junctions involving annealing at long microhomologies was significantly decreased ($P < 0.01$, Fisher’s exact test; Figure 6A, Table 2, and Table 3). Only 12% of D2/2003 junctions possessed an insertion greater than three base pairs, compared to 44% of junctions recovered from males with the D5 and 2003 alleles. In addition, most insertions isolated from D5/2003 males were highly complex and had multiple copies of imperfect repeats of T-nucleotides. Similar results were obtained with the D5/3294 allele combination (data not shown).

Figure 4. Pol θ-dependent end joining acts independently of HR and C-NHEJ. (A) Decreased end joining in mus308 mutants does not depend on homologous recombination. Each bar represents the mean percentage of end-joining events recovered from independent males of each genotype. Number of independent males assayed: spn-A = 63; spn-A057mus308D2/spn-A057mus3082003 = 216; spn-A057mus308D2/spn-A057mus3083294 = 33; spn-A057mus308D5/spn-A057mus3082003 = 80; spn-A057mus308D5/spn-A057mus3083294 = 128. Error bars represent SEM. ** indicates $P < 0.01$, *** indicates $P < 0.001$ (Kruskal-Wallis non-parametric ANOVA). (B) Genomic deletions during DSB repair in mus308 mutants do not depend on initiation of HR. Each bar represents the percentage of yellow-eyed female progeny with scalloped wings (representing large flanking deletions) recovered from males of indicated genotype. (C) Mutation of mus308 in a lig4169a background further reduces end-joining repair relative to mus308 single mutants. Each bar represents the mean percentage of repair events recovered from wild type, mus3082003/Df(3R)kar3l, or lig4169a; mus3082003/Df(3R)kar3l males. Number of independent males assayed: wild type = 101, mus308 = 182, lig4 mus308 = 57. Error bars represent SEM. ** indicates $P < 0.01$, *** indicates $P < 0.001$ (Kruskal-Wallis non-parametric ANOVA). (D) Genomic deletions always occur during DSB repair in lig4 mus308 mutants. All yellow-eyed repair events recovered from lig4169a; mus3082003/Df(3R)kar3l males had large deletions that resulted in male lethality and/or female progeny with scalloped wings. Number of independent repair events assayed: wild type = 101, mus308 = 167, lig4 mus308 = 22.

doi:10.1371/journal.pgen.1001005.g004
In summary, both mus308 mutant combinations significantly abrogated annealing at long microhomologies during alt-EJ repair. However, we observed a distinct difference in repair junctions recovered from males harboring the D2 allele, which greatly reduces overall pol θ protein levels [38], compared to flies with the D5 allele, which alters a conserved residue near the helicase-like domain. These results suggest that pol θ has two distinct functions in alt-EJ: one that promotes the annealing of long microhomologous sequences during end alignment, and another that is responsible for complex T-nucleotide insertions. Flies with the D2 allele are impaired in their ability to carry out both the annealing and insertion functions, whereas flies possessing the D5 separation-of-function allele cannot perform the microhomology annealing function but can still produce complex insertions.

Table 1. P(w*) repair junctions recovered from spn-A093-057 mutants.

| Type of repair event | Sequence 5' of break* | Microhomology/inserted sequence | Sequence 3' of break* | Number of isolates |
|----------------------|------------------------|---------------------------------|-----------------------|-------------------|
| Original sequence    | accccagacCATGATGAAATAACATA | -                               | TATGTATTTCATCATGacccagac | -                 |
| Long microhomology   | accccagac (CATgATGA)\textsuperscript{b} | cccagac                       | 14                    |
|                      | accccagac (CATGA)        | cccagac                        | 4                     |
|                      | none\textsuperscript{d}  | (TGACCCAGAC)                    | 2                     |
| Short microhomology  | accccagacCATG (ATG)      | TTATTCATCATGacccagac           | 1                     |
|                      | accccagacCA (TGA)        | cccagac                        | 1                     |
|                      | accccagacCATGATGAAATAAAAC (CAT) | Gaccccagac                  | 3                     |
|                      | accccagacCATGATGAAATAACA (AT) | GTATAATTCATCATGacccagac       | 14                    |
|                      | accccagacCATGATGAAATAAACAC (TA) | TGTATTTTCATCATGacccagac       | 5                     |
|                      | accccagacCATGATGAAAA (TA) | TGTATTTTCATCATGacccagac       | 1                     |
|                      | accccagacCATGATGAA (AT)  | GTATTTTCATCATGacccagac        | 1                     |
|                      | none\textsuperscript{d}  | (T)                            | 1                     |
| Blunt join           | accccagacCATGATGAATAATA | -                              | TTATTCATCATGacccagac  | 1                 |
| Insertion\textsuperscript{e} | accccagacCATGATGAAATAACATA | A                              | CATCATGacccagac       | 1                 |
|                      | accccagacCATGATGAAATAACATA | G                              | CATCATGacccagac       | 1                 |
|                      | accccagacCATGATGAAATAACATA | AC                             | ATGTATTTTCATCATGacccagac | 3               |
|                      | accccagacCATGATGAAATAACATA | GT                            | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | AC                            | ATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCA AG            | ATGacccagac                   | 1                     |
|                      | accccagacCATGATGAAATAACATA | TTC                           | ATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | GTTA                          | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TTTA                          | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TTATCA                        | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TTAACATAAC                    | ATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TTAATTTTATTA                  | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | GAAATATAAC                    | ATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | GTATACATAAC                   | ATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TAAATAATAATTTA                | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TAATAATAATAATA                | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TCAAATAATAATAAT              | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TACAAATAATAATA                | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TAAATAATAATAATTA             | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TCAAATAATAATAATATTA          | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TACAAATAATAATAATTTA          | TATGTATTTTCATCATGacccagac | 1               |
|                      | accccagacCATGATGAAATAACATA | TACAAATAATAATAATTAATTA       | TATGTATTTTCATCATGacccagac | 1               |

*a* Uppercase letters represent the 17-nucleotide 3' single-stranded tails that remain following transposase action, lowercase letters correspond to the 8 base pair target sequence duplicated upon P element insertion.

*b* Microhomologies (in parentheses) are sequences that could have been derived from either side of the break site. Long microhomologies were five or more nucleotides, while short microhomologies were three or fewer nucleotides.

*c* Represents an 8-nucleotide imperfect microhomology.

*d* Indicates a deletion that extends past the 8 base pair target sequence.

*e* Insertions were identified as any sequence not present at the original break site. Templated insertions and corresponding potential templates in flanking sequences are underlined.

doi:10.1371/journal.pgen.1001005.t001
and males with the phenotypic defects. Experiment. Error bars represent SEM. * indicates p represents three independent experiments with at least 300 progeny per P
percentage of viable compound heterozygous progeny was determined. The crosses were repeated in the presence of spn-A093 or
spn-A mus308 (A) spn-A mus308 mutants are sub-viable. Females and males with the P(w+) transposon and heterozygous for the spn-A093 or
spn-A mus308 alleles were interbred and the percentage of viable compound heterozygous progeny was determined. The crosses were repeated in the presence of P transposase. Each bar represents one experiment. Error bars represent SEM. * indicates p < 0.05, ** indicates p < 0.01 (unpaired T test). (B) Increased male sterility in spn-A mus308 or
mus308 (A) spn-A mus308 (B) Increased male sterility in spn-A mus308 or
spn-A mus308 mutants. Because microhomologies of greater than four base pairs are not present near the I-
endonuclease and the previously characterized [Iw]7 reporter construct [50] to create site-specific DSBs in wild-type flies and flies lacking either DNA ligase 4 or pol θ. I-Sol produces a DSB with 4-nucleotide complementary overhangs that can be directly ligated through a C-NHEJ mechanism [50,51]. Accurate repair regenerates the original I-Sol recognition sequence, which can then be cut again, while inaccurate end-joining repair abolishes further cutting. We utilized an hsp70 or ubiquitin-driven I-Sol construct integrated on chromosome 2 to drive high levels of I-Sol expression [50,52]. Nearly 100% of repair events that we recovered involved gene conversion (HR repair from the homologous chromosome) or inaccurate end-joining (data not shown). In the [Iw]7 system, both gene conversion events and large deletions that remove the white marker gene are phenotypically indistinguishable. PCR analysis confirmed that many repair events recovered from mus308 mutants involved large deletions (>700 base pairs, data not shown). Our subsequent analysis focused on the characterization of repair events involving smaller deletions.

Twenty-three percent of I-Sol repair junctions isolated from wild-type flies possessed insertions of more than 3 base pairs (Figure 6B). This percentage was significantly increased to 46% in lig4 mutants (P < 0.01, Fisher’s exact test), consistent with increased use of alt-EJ in the absence of C-NHEJ. If pol θ plays a general role in insertional mutagenesis during alt-EJ repair, one would predict that the frequency and length of insertions following I-Sol cutting should decrease in mus308 mutants. Indeed, the percentage of large insertions decreased to 9% in mus308 mutants (P = 0.03, Fisher’s exact test). Wild-type flies had an average insertion length of 7.6 base pairs, compared to 4.2 base pairs for mus308 mutants. Strikingly, no mus308 insertion was longer than twelve base pairs, while insertions of more than twenty base pairs occurred in both wild type and lig4 mutants. Because microhomologies of greater than four base pairs are not present near the I-Sol cut site in this construct, repair involving annealing at long microhomologies was not observed.

Surprisingly, the total percentage of repair junctions with short, 1–3 base pair insertions was not decreased in lig4 mutants relative to wild type (17% vs. 13%, respectively). Furthermore, the percentage of junctions involving annealing at 1–3 nucleotide microhomologies was also similar between the two genotypes (25% for lig4 mutants vs. 34% for wild type). These two types of junctions have historically been associated with ligase 4-dependent C-NHEJ repair. Our results suggest that this may not be the case. Indeed, a fine-level sequence analysis of I-Sol repair junctions that we have recently conducted suggests that alt-EJ may produce C-NHEJ-like junctions in certain sequence contexts [53]. Nevertheless, our data obtained using two independent site-specific DSB repair assays strongly suggest that C-NHEJ and alt-EJ represent at least partially independent mechanisms for the repair of DSBs and that pol θ plays an important role in the generation of T-nucleotide insertions during alt-EJ repair of both P element and I-Sol-induced breaks.

Figure 5. Flies lacking Rad51 and Pol θ have synergistic phenotypic defects. (A) spn-A mus308 mutants are sub-viable. Females and males with the P(w+) transposon and heterozygous for the spn-A093 or
spn-A mus308 alleles were interbred and the percentage of viable compound heterozygous progeny was determined. The crosses were repeated in the presence of P transposase. Each bar represents three independent experiments with at least 300 progeny per experiment. Error bars represent SEM. * indicates p < 0.05, ** indicates p < 0.01 (unpaired T test). (B) Increased male sterility in spn-A mus308 double mutant males. Each bar represents the percentage of spn-A or
spn-A mus308 males that produced no progeny. Number of independent males assayed: spn-A = 57; spn-A093mus308/spn-A093mus308 = 305; spn-A093mus308/spn-A093mus308 = 50; spn-A093mus308/spn-A093mus308 = 102; spn-A093mus308/spn-A093mus308 = 100. (C) Mutation of both spn-A and mus308 results in external morphological defects such as abdominal cuticle mispatterning (note the severe disruption of normal segmental banding patterns). Pictured are representative wild type and spn-A093mus308/spn-A093mus308 males in which P element excision is occurring. doi:10.1371/journal.pgen.1001005.g005

C-NHEJ. To test whether the results obtained with I-elements can be generalized to other types of breaks, we used the I-Sol endonuclease and the previously characterized [Iw]7 reporter construct [50] to create site-specific DSBs in wild-type flies and flies lacking either DNA ligase 4 or pol θ. I-Sol produces a DSB with 4-nucleotide complementary overhangs that can be directly ligated through a C-NHEJ mechanism [50,51]. Accurate repair regenerates the original I-Sol recognition sequence, which can then be cut again, while inaccurate end-joining repair abolishes further cutting. We utilized an hsp70 or ubiquitin-driven I-Sol construct integrated on chromosome 2 to drive high levels of I-Sol expression [50,52]. Nearly 100% of repair events that we recovered involved gene conversion (HR repair from the homologous chromosome) or inaccurate end-joining (data not shown). In the [Iw]7 system, both gene conversion events and large deletions that remove the white marker gene are phenotypically indistinguishable. PCR analysis confirmed that many repair events recovered from mus308 mutants involved large deletions (>700 base pairs, data not shown). Our subsequent analysis focused on the characterization of repair events involving smaller deletions.

Twenty-three percent of I-Sol repair junctions isolated from wild-type flies possessed insertions of more than 3 base pairs (Figure 6B). This percentage was significantly increased to 46% in lig4 mutants (P < 0.01, Fisher’s exact test), consistent with increased use of alt-EJ in the absence of C-NHEJ. If pol θ plays a general role in insertional mutagenesis during alt-EJ repair, one would predict that the frequency and length of insertions following I-Sol cutting should decrease in mus308 mutants. Indeed, the percentage of large insertions decreased to 9% in mus308 mutants (P = 0.03, Fisher’s exact test). Wild-type flies had an average insertion length of 7.6 base pairs, compared to 4.2 base pairs for mus308 mutants. Strikingly, no mus308 insertion was longer than twelve base pairs, while insertions of more than twenty base pairs occurred in both wild type and lig4 mutants. Because microhomologies of greater than four base pairs are not present near the I-Sol cut site in this construct, repair involving annealing at long microhomologies was not observed.

Surprisingly, the total percentage of repair junctions with short, 1–3 base pair insertions was not decreased in lig4 mutants relative to wild type (17% vs. 13%, respectively). Furthermore, the percentage of junctions involving annealing at 1–3 nucleotide microhomologies was also similar between the two genotypes (25% for lig4 mutants vs. 34% for wild type). These two types of junctions have historically been associated with ligase 4-dependent C-NHEJ repair. Our results suggest that this may not be the case. Indeed, a fine-level sequence analysis of I-Sol repair junctions that we have recently conducted suggests that alt-EJ may produce C-NHEJ-like junctions in certain sequence contexts [53]. Nevertheless, our data obtained using two independent site-specific DSB repair assays strongly suggest that C-NHEJ and alt-EJ represent at least partially independent mechanisms for the repair of DSBs and that pol θ plays an important role in the generation of T-nucleotide insertions during alt-EJ repair of both P element and I-Sol-induced breaks.
both spn-A mus308 mutants (P<0.01, Fisher’s exact test). (B) Reduced frequency of ≥4 bp insertions during inaccurate end-joining repair of a chromosomal I-SceI DSB in mus308 mutants. Complementary ended DSBs were induced by expression of I-SceI endonuclease in the pre-meiotic germline of wild type (n = 70), lig4Δ Δ (n = 83), and mus308ΔΔDf(3R) kar-Sz29 (n = 57) males, and independent inaccurate end-joining repair junctions were sequenced. White, blunt joins; light gray, 1–4 bp microhomologies; dark gray, 1–3 bp insertions; black, ≥4 bp insertions. No ≥4 bp microhomologies are available near the DSB in this system. * indicates P = 0.03, ** indicates P<0.01, relative to wild type (Fisher’s exact test).

doi:10.1371/journal.pgen.1001005.g006

Discussion

Several studies have identified proteins important for end-joining repair of DSBs in the absence of C-NHEJ factors in yeasts [11,54,55]. More recently, the Mre11 protein has been identified as an important alt-EJ component in vertebrate systems [10,56–59]. However, much remains uncertain about the genetics or mechanisms of alt-EJ in metazoans. Our previous work demonstrated that end-joining repair of P element-induced breaks can occur independently of DNA ligase 4, suggesting the presence of a highly active alternative end-joining mechanism in Drosophila [17]. We have now identified the mus308 gene, encoding DNA polymerase theta, as a critical component of alt-EJ. Pol θ-dependent alt-EJ operates in parallel to C-NHEJ to promote repair of both P element and I-SceI-induced breaks. Because we observed similar alt-EJ defects with four different mus308 mutant alleles, several of which were studied in trans to at least two independently-derived deficiencies, we consider it highly unlikely that the phenotypes we observed are due to second-site mutations or other differences in genetic background.

Importantly, pol θ does not appear to participate in homology-directed repair. HR repair of DSBs following P(au) excision is thought to proceed largely through synthesis-dependent strand annealing (SSA) [60]. We observed that SDSA frequencies in the P(au) assay were similar in wild type and mus308 mutants. Although we did not formally test for a role of pol θ in single-strand annealing (SSA), a non-conservative HR pathway that involves annealing at direct repeats larger than 25 nucleotides, it seems unlikely that the repair observed in mus308 mutants arose through an SSA mechanism.

Three findings suggest that mus308-dependent alt-EJ is an important repair option for both cell and organism survival in flies, particularly in the absence of homologous recombination. First, spn-A mus308 double mutants are sub-viable and have severe defects in adult abdominal cuticle closure, consistent with a high level of apoptosis in rapidly proliferating histoblasts during pupariation. Second, spn-A mus308 double mutant males undergoing P(au) excision have up to 30-fold increased sterility relative to spn-A mutants. Third, mus308 mutant males undergoing I-SceI cutting show premature sterility and produce few progeny. The few germinal repair events that are recovered from each male are frequently clonal, suggesting extensive germinal apoptosis (A. Yu, unpublished data).

Evidence for two functions of pol θ in alternative end joining

Pol θ orthologs characterized from a variety of metazoans possess both helicase-like and DNA polymerase domains...
be used during the annealing stage of alt-EJ when suitable insertions may be ‘synthesis footprints’ that are formed during the percentage of repair junctions with templated insertions. These protein A and rewind single-stranded DNA bubbles [61].

HARP protein, which is able to displace stably bound replication forks, is also implicated in the unwinding or strand displacement function. Flies with the D5 and 3294 mutations (located in or near the conserved helicase domain) produce repair products with complex T-nucleotide insertions but not products involving annealing at long microhomologies. The D5 and 3294 alleles may therefore encode proteins that retain polymerase activity but lack unwinding activity, resulting in an inability to expose internal microhomologous sequences. Because the microhomologies used in repair following P element excision are often located in the 17-nucleotide 3’ single-stranded tails, pol θ may also be important for the unwinding of secondary structures that form in single-stranded DNA. Alternatively, the DNA-dependent ATPase activity demonstrated by pol θ might represent an annealing function of the protein that is required during alt-EJ. Such an annealing activity was recently described for the human HARP protein, which is able to displace stably bound replication protein A and rewind single-stranded DNA bubbles [61].

One notable aspect of alt-EJ in Drosophila is the large percentage of repair junctions with templated insertions. These insertions may be ‘synthesis footprints’ that are formed during the cell’s attempt to create microhomologous sequences that can be used during the annealing stage of alt-EJ when suitable endogenous microhomologies are not present or are not long enough to allow for stable end alignment. Indeed, analysis of the insertions from I-SceI repair junctions suggests a model involving local unwinding of double-stranded DNA and iterative synthesis of 3–8 nucleotide runs [53]. The P[w+] repair junctions isolated from mus308D5/mus3082003 mutants are consistent with an important (but not exclusive) role for the polymerase domain of pol θ in the synthesis of T-nucleotides.

We speculate that pol θ may be involved in both DNA unwinding and repair synthesis during alt-EJ (Figure 7). Linking these two activities in one protein would provide a convenient mechanism for creating longer microhomologies that could increase the thermodynamic stability of aligned ends prior to the action of a DNA ligase. Studies based on the crystal structure of a dual function NHEJ polymerase-ligase protein found in Mycobacterium tuberculosis suggest that a synaptic function for an NHEJ polymerase is plausible [62]. Because ligase 4 is not involved in alt-EJ in Drosophila, another ligase must be involved in the ligation step. Studies from mammalian systems have identified DNA ligase 5 as a likely candidate [63,64].

Potential roles of pol θ in DNA interstrand crosslink repair

Pol θ was originally identified in Drosophila based on the inability of mus308 mutants to survive exposure to chemicals that induce DNA interstrand crosslinks. A crucial question posed by our findings is whether pol θ performs a common function during the repair of both DSBs and interstrand crosslinks. The C. elegans
Table 3. P[w\textsuperscript{+}] sequences recovered from spn-A\textsuperscript{093/057} mus308\textsuperscript{02/003} mutants.

| Type of repair event | Sequence to left of break | Microhomology/inserted sequence | Sequence to right of break | Number of isolates |
|----------------------|---------------------------|---------------------------------|---------------------------|-------------------|
| Original sequence    | acccagacCATGATGAAATAACATA  | -                               | TATGTTATTTTCATCATGacccagac | -                 |
| Long microhomology\textsuperscript{a} | acccagac                         | (CATgATGA)                        | cccagac                     | 1                 |
| Short microhomology  | acccagacCATGATGAAATAACAC      | (AT)                            | GTATTTTCATCATGacccagac      | 7                 |
|                      | acccagacCATGATGAAATAACACTA    | (TA)                            | TGATTTTCATCATGacccagac      | 3                 |
|                      | acccagacCATGATGAAATAACATCA    | (AT)                            | GTATTTTCATCATGacccagac      | 1                 |
|                      | acccagacCATGATGAAATAACATCA    | (T)                             | TATTTTCATCATGacccagac       | 1                 |
| Blunt join           | acccagacCATGATGAAATAACATA     | -                               | TATGTTATTTTCATCATGacccagac  | 1                 |
|                      | acccagacCATGATGAAATAACATA     | -                               | TATGTTATTTTCATCATGacccagac  | 1                 |
| Insertion\textsuperscript{b} | acccagacCATGATGAAATAACATTC    | T                               | TATGTTATTTTCATCATGacccagac  | 1                 |
|                      | acccagacCATGATGAAATAACATA     | CA                              | ATGTTATTTTCATCATGacccagac   | 1                 |
|                      | acccagacCATGATGAAATAACATA     | AC                              | ATGTTATTTTCATCATGacccagac   | 1                 |
|                      | acccagacCATGATGAAATAACATA     | TA                              | TATGTTATTTTCATCATGacccagac  | 1                 |
|                      | acccagacCATGATGAAATAACATA     | TG                              | TATGTTATTTTCATCATGacccagac  | 1                 |
|                      | acccagacCATGATGAAATAACATA     | TG                              | TATGTTATTTTCATCATGacccagac  | 2                 |
|                      | acccagacCATGATGAAATAACATA     | ACAATA                          | ATGTTATTTTCATCATGacccagac   | 1                 |
|                      | acccagacCATGATGAAATAACATA     | TATACCG                         | ATGTTATTTTCATCATGacccagac   | 1                 |
|                      | acccagacCATGATGAAATAACATA     | TGTTTAAAC                      | ATGTTATTTTCATCATGacccagac   | 1                 |

\textsuperscript{a} Microhomologies (in parentheses) are sequences that could have been derived from either side of the break site.

\textsuperscript{b} Insertions were identified as any sequence not present at the original break site. Templated insertions and corresponding potential templates in flanking sequences are underlined.

doi:10.1371/journal.pgen.1001005.t003

pol \theta ortholog, POLQ-1, is also required for resistance to interstrand crosslinks and acts in a pathway that is distinct from HR but depends on CeBRCA1 [28]. In S. cerevisiae, several repair mechanisms are utilized during interstrand crosslink repair, including nucleotide excision repair (NER), HR, and translesion synthesis [65,66]. Given our results and the findings from C. elegans, it seems unlikely that the role of pol \theta in interstrand crosslink repair involves a function in HR.

In human cells, exposure to agents that induce interstrand crosslinks causes a shift in repair mechanisms that leads to increased use of non-conservative pathways associated with complex insertions and deletions [67]. Furthermore, interstrand crosslinks can cause frequent recombination between direct repeats [68,69], suggesting that single-strand annealing may provide a viable mechanism for interstrand crosslink repair. The single-strand annealing model of interstrand crosslink repair posits that NER-independent recognition and processing of the crosslinked DNA is followed by generation of single-stranded regions flanking the crosslink and annealing at repeated sequences. Because alt-EJ frequently proceeds through annealing at short direct repeats, it is tempting to speculate that the role of pol \theta in interstrand crosslink repair might be to expose and/or promote the annealing of microhomologous single-stranded regions that flank the crosslinked DNA. Consistent with this model, the initial incision step made after recognition of the interstrand crosslink remains normal in mus308 mutants [26]. Alternatively, pol \theta might utilize its polymerase activity and nearby flanking sequences as a template to synthesize short stretches of DNA that could be used to span a single-stranded gap opposite of a partially excised crosslink.

Such a model has been proposed to explain the formation of microindels in human cancers [70]. We are currently testing these two models using helicase- and polymerase-specific mus308 mutant alleles.

Pol \theta and alternative end joining: promoting genome (in)stability

Although it seems counterintuitive, alt-EJ likely functions in some situations to promote genome stability. As evidence of this, we found that DSB repair following P element excision in mus308 mutant flies frequently results in genomic deletions of multiple kilobases. A similar deletion-prone phenotype was previously observed in mus309 mutants, which lack the Drosophila BLM ortholog [43,71]. Epistasis analysis demonstrated that the mus309 deletion phenotype depends on Rad51, implying that DmBlm acts after strand invasion during HR and that the deletions observed in mus309 mutants are likely a result of failed SDSA [48]. In contrast, the deletions observed in mus308 mutants do not depend on Rad51, demonstrating that the function of pol \theta in DSB repair is independent of HR. The deletion phenotype is exacerbated in double mutants, suggesting that C-NHEJ and alt-EJ represent two parallel mechanisms to prevent deletions. In the absence of these two end-joining options, resection at the broken ends may continue unchecked, resulting in extensive genomic deletions that are generated by an unknown Rad51-independent repair mechanism. Therefore, both C-NHEJ and alt-EJ function to prevent overprocessing of broken DNA ends and extreme degradation of the genome. Microhomology-mediated end
joining, which shares many features with alt-EJ, has been proposed to perform a similar function in urothelial cells [72]. Nonetheless, alternative end-joining repair can also be genome destabilizing, as demonstrated by an increasing number of reports linking it to cancer. We have shown that complex insertions observed in alternative end-joining products are more frequent in flies possessing pol

These microhomologies (MH) are used to align the broken ends to provide a template for pol θ polymerase activity. The unwinding activity could also serve to make the polymerase more processive. (C) Processing of the ends and ligation results in repair accompanied by a deletion. (D) In cases where the ends do not remain stably aligned, annealing at other microhomologies closer to the break site could result in the insertion of T-nucleotides. Multiple rounds of unwinding, synthesis, and alignment could result in the complex insertions that are often observed in alt-EJ in flies.

doi:10.1371/journal.pgen.1001005.g007

Conclusions
The results described here establish that Drosophila pol θ plays two distinct roles in an alternative end-joining mechanism operating in parallel to canonical DNA ligase 4-mediated C-NHEJ. This novel finding lays the groundwork for future studies focusing on the specific roles of the pol θ helicase-like and polymerase domains in alt-EJ and DNA interstrand crosslink repair. Whether pol θ plays a similar role in alt-EJ in other organisms, including mammals, remains to be determined. Regardless, these studies reveal an unexpected role for DNA polymerase θ that is required for genomic integrity in Drosophila and possibly other metazoans.

Materials and Methods
Drosophila stocks and mus308 alleles
All flies were maintained on standard cornmeal-based agar food and reared at 25 °C. The mus308 D2 and D5 stocks were obtained from the Bloomington Stock Center and the 2003 and 3294 stocks were from the Zuker collection [76]. To identify mutations in these stocks, genomic DNA was isolated from flies harboring each allele in trans to Df(3R)Exel6166 and PCR was performed with primers specific to overlapping regions of the entire coding sequence. PCR products were sequenced and the sequence was compared to the Drosophila reference sequence release 5.10. Sequence changes unique to each allele were verified by sequencing in both orientations. The lig4<sup>com</sup> [17], spn-A<sup>4932</sup> and spn-A<sup>357</sup> [42] stocks harbor null alleles of DNA ligase 4 and Rad5l, respectively.
Mutagen sensitivity studies

For mechlorethamine sensitivity assays, balanced, heterozygous parents were crossed to Df(3R)Exel166 and allowed to lay eggs in vials containing 10mL of food for three days, after which they were moved to new vials for two additional days. The first vials were treated with 250μL of 0.005% mechlorethamine dissolved in ddH2O, while the second vials were treated only with ddH2O. Survival was calculated as the number of homozygous mutant adults divided by the total number of adults that eclosed within 10 days of treatment. Ratios were normalized to untreated controls for each set of vials (five to eight sets of vials were counted for each experiment).

For ionizing radiation sensitivity assays, heterozygous parents laid eggs on grape-juice agar plates for 12 hr. Embryos developed at 25°C until larvae reached third-instar stage, at which point they were irradiated in a Gammarator 1000 irradiator at a dose rate of 800 rads/min and larvae were transferred to food-containing bottles. Relative survival rates were calculated as above.

P[w\textsuperscript{+}] assay

Repair of DNA double-strand breaks was monitored after excision of the P[w\textsuperscript{+}] transposon as described previously [43,77]. P[w\textsuperscript{+}] was excised in males using a second chromosome transposase source (CyO, H[w\textsuperscript{+}A2–3]) and individual repair events were recovered in female progeny over an intact copy of P[w\textsuperscript{+}]. Females with two copies of P[w\textsuperscript{+}] have apricot eyes [78]. Preny with red eyes possess a repair event involving HR with annealing of the copia LTRs. A fraction of apricot-eyed females also possess HR repair events, but these cannot be distinguished from chromosomes in which no excision event occurred (using the CyO, H[w\textsuperscript{+}A2–3] transposase source, ~80% of apricot-eyed female progeny inherit a non-excised P[w\textsuperscript{+}] element). Yellow-eyed females harbor a repair event in which repair is completed by end joining.

For each genotype, at least 50 individual male crosses were scored for eye color of female progeny lacking transposase. The percentage of progeny from each repair class was calculated on a per vial basis, and the vial representing a separate experiment. Statistical comparisons were done using a Kruskal-Wallis non-parametric ANOVA followed by Dunn’s multiple comparisons test using InStat3 (GraphPad).

Analysis of HR synthesis tract lengths, genomic DNA was purified and PCR reactions were performed as in [43], using primer pairs with the internal primer located 250, 2420, and 4674 base pairs from the cut site at the 5’ end of P[w\textsuperscript{+}].

For deletion analysis, the percentage of females with scalloped eyes possessing transposase was calculated relative to all yellow-eyed females counted. The percentage of male lethal and small (0.1–3.6 kb) deletions was calculated based on a subset of yellow-eyed females (one from each original male parent) that were individually crossed to males bearing the FM7w balancer. Vials for which no white-eyed male progeny were recovered were scored as male lethal. Some of the male lethal events also caused a scalloped-wing phenotype in heterozygous females. For those that did not, testing to ensure that the male lethality was due to deletion of scalloped coding sequence was performed by recovering the repaired chromosomes in trans to the hypomorphic sd\textsuperscript{d} mutation [79] and scoring for a scalloped-wing phenotype. Repair events which could be recovered in males were subjected to PCR analysis, using primers internal to P[w\textsuperscript{+}] [43], to detect small deletions into one or both introns of sd.

I-SceI break repair assay

Repair of I-SceI mediated DNA double strand breaks was studied in the context of the chromosomally integrated [Iw]7 construct [52], which contains a single target site for the I-SceI endonuclease. DSBs were induced in the male pre-meiotic germline by crossing females harboring [Iw]7 to males expressing the I-SceI endonuclease from a second chromosomal location under the control of either the hsp70 promoter (70[I-SceI]1A) [52] or the ubiquitin promoter (UIE[I-SceI]2R) [50]. Independent inaccurate end-joining repair events from the male pre-meiotic germline were recovered in male progeny and DNA was isolated for analysis [80]. PCR was performed using primers PES5’ (GATAGCCGACAATTACGAAGGTC) and jn3’b (GGACATTGACGCTATCGGACAGT) to amplify a 1.3 kb fragment of the [Iw]7 construct including the I-SceI target site. Products were gel purified (GenScript) and sequencing of PCR products was performed using the PES5’ primer. Sequences were aligned using ClustalW or by manual inspection against sequence obtained from an uncut [Iw]7 construct. Statistical comparisons were done using Excel and SPSS.

Supporting Information

Figure S1 Polymerase theta orthologues from various metazoans. The conserved helicase-like (blue oval) and polymerase (pink box) domains are indicated. All of the orthologs have additional conserved regions in the N and C-termini (white boxes), separated by a variable-length linker region.

Figure S2 Sequence changes in different mus308 mutant alleles, compared to the Drosophila reference genome sequence. Allele-specific changes are highlighted in yellow. +1 corresponds to the ‘A’ in the start codon of mus308.

Figure S3 Alignment of conserved N-terminus of mus308 orthologs. Dmel, Drosophila melanogaster; Agam, Anopheles gambiae; Mmus, Mus musculus; Hsap, Homo sapiens; Drer, Danio rerio; Atha, Arabidopsis thaliana; Cele, Caenorhabditis elegans. Conserved amino acids are indicated below each alignment. The red arrow corresponds to the G621S substitution in the 3294 allele, the blue arrow corresponds to the P781L substitution in the D5 allele.

Acknowledgments

We are grateful to members of the McVey, Mirkin, and Freudenreich labs for discussions that informed this work and to Jeff Sekelsky and Graham Walker for helpful comments. We would also like to thank the three anonymous reviewers whose suggestions greatly improved the paper. We thank the Bloomington stock center and Zuker labs for mus308 mutant alleles, Bill Engels for the UIE-I-SceI stock, and Yikang Rong for the hsp70-I-SceI and [Iw]7 stocks.

Author Contributions

Conceived and designed the experiments: SHC AMY MM. Performed the experiments: SHC AMY MM. Contributed reagents/materials/analysis tools: SHC MM. Wrote the paper: AMY MM.
3. Pardo B, González-González B, Aguilera A (2009) DNA repair in mammalian cells: DNA double-strand break repair: how to fix a broken relationship. Cell Mol Life Sci 66: 1039–1056.

4. Nohmi T, Masumura K (2005) Molecular nature of intrachromosomal deletions and base substitution-induced by environmental mutagens. Environ Mol Mutagen 45: 150–161.

5. Raschle M, Knipper P, Enois M, Angelov T, Sun J, et al. (2008) Mechanism of replication-coupled DNA interstrand crosslink repair. Cell 134: 969–980.

6. Wetters F, Chen JJ (2008) The endless tale of non-homologous end-joining. Cell 138: 114–121.

7. Lieber MR, La H, Gu J, Schwartz K (2008) Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligation of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune response. Cell 135: 125–133.

8. McVey M, Lee SE (2008) MMEJ repair of double-strand breaks (director's cut), deleted sequences and alternative ends, Trends Genet 24: 529–538.

9. Simsek D, Jain M (2010) Alternative end-joining is suppressed by the canonical NHEJ component Xrc5/ligase IV during chromosomal translocation formation. Nat Struct Mol Biol 17: 410–416.

10. Rass E, Grabarz A, Plot J, Bertrand P, et al. (2009) Role of Mre11 in mammalian chromosomal non-homologous end joining in mammalian cells. Nat Struct Mol Biol 16: 191–204.

11. Ma JL, Kim EM, Haber JE, Lee SE (2003) Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. Mol Cell Biol 23: 8012–8028.

12. Weinstock DM, Brunet E, Jain M (2007) Formation of NHEJ-derived reciprocal chromosomal translocations does not require Ku70. Nat Cell Biol 9: 976–981.

13. Zhu C, Mills KD, Ferguson DO, Lee C, Manis J, et al. (2002) Unrepaired DNA breaks in Drosophila melanogaster is largely DNA ligase IV independent. Proc Natl Acad Sci USA 99: 811–816.

14. Bender J, Diggel CP, Harnden P, Knowles MA, Kiltie AE (2004) DNA double-strand break repair in human bladder cancer is error prone and involves microhomology-associated end-joining. Nucleic Acids Res 32: 5249–5259.

15. Mattaruzzo E, Guerini V, Rambaldi A, Campiotti L, Venco A, et al. (2008) Microhomologies and interseed repeat elements at genomic breakpoints in chronic myeloid leukemia. Genes Chromosomes Cancer 47: 625–632.

16. Corneo B, Wendland RL, Deriano L, Cui X, Klein IA, et al. (2007) Rag mutations reveal robust alternative end joining. Nature 449: 463–466.

17. McVey M, Radar D, Sekelsky JJ (2004) End-joining repair of double-strand breaks in Drosophila melanogaster is largely DNA ligase IV independent. Genetics 168: 2067–2076.

18. Sekelsky JJ, Brodsky MH, Burcis KC (2000) DNA repair in Drosophila: insights from the Drosophila genome sequence. J Cell Biol 150: F1–36.

19. Taka-Ishikawa E, Yoshihara M, Hotta Y (1992) Extra sequences found at P element excision sites in Drosophila melanogaster. Mol Gen Genet 232: 17–23.

20. Ducay J, Breglino JC, de La Roche Saint-andre C (2000) Gamma-irradiation stimulates homology-directed DNA double-strand break repair in Drosophila embryonic, Mutat Res 469: 69–80.

21. Welsel N, Le T, Marculescu R, Mutterbaier G, Chott A, et al. (2001) Templated nucleotide addition and immunoglobulin JH-gene utilization in ()14;14) junctions: implications for the mechanism of translocation and the origin of human B-cell lymphomas. Cancer Res 61: 1029–1036.

22. Jager U, Beckor S, Le T, Mutterbaier G, Bolz I, et al. (2000) Follicular lymphomas’ BCL-2/IgH junctions contain templated nucleotide insertions: novel insights into the mechanism of ()14;14) translocation. Blood 95: 3250–3259.

23. Denny CT, Holtin GF, Magraith IT, Kirkir IR (1985) Burkitt lymphoma cell line carrying a variant translocation creates new DNA at the breakpoint and violates the hierarchy of immunoglobulin gene rearrangement. Mol Cell Biol 5: 329–330.

24. Verkade NS, Esveld-van Langen RE, van Heemst D, Bruggeworth HT, Hoijmakers JH, et al. (2002) Different types of VDJ recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. Eur J Immunol 32: 701–709.

25. Landberg R, Mavakot M, Campbell C (2001) Deficient DNA end joining activity exists in extracts from fanci anemia fibroblasts. J Biol Chem 276: 9435–9449.

26. Boyd JB, Sakaguchi K, Harris PV (1999) mu306 mutants of Drosophila exhibit hypomorphic DNA sensitivity to DNA linking agents and are defective in a deoxyribo nuclease. Genetics 128: 813–819.

27. Harris PV, Mazina OM, Leonhardt EA, Case RB, Boyd JB, et al. (1996) Molecular cloning of Drosophila mu306, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes. MCB 16: 5704–5711.

28. Mazina DM, Plass P, Boulton SJ, Casasanta G, Marini F (2000) Carboxytail-elegant PolQ-1 and HEL-308 function in two distinct DNA interstrand crosslink repair pathways. DNA Repair (Amst) 7: 941–950.

29. Inagaki S, Suzuki T, Ohno MA, Urawa H, Horiuchi T, et al. (2006) Arabidopsis TEL1/CHC1, with helicase and DNA polymerase domains, is required for regulated cell division and differentiation in meristems. Plant Cell 18: 879–892.

30. Seki M, Marini F, Wood RD (2003) POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells. Nuclear Acids Res 31: 6117–6126.
60. Nassif N, Penney J, Pal S, Engels WR, Gloor GB (1994) Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Mol Cell Biol 14: 1613–1625.

61. Yusufzai T, Kadonaga JT (2008) HARP is an ATP-driven annealing helicase. Science 322: 748–750.

62. Brissett NC, Pitcher RS, Juarez R, Picher AJ, Green AJ, et al. (2007) Structure of a NHEJ polymerase-mediated DNA synaptic complex. Science 318: 456–459.

63. Aulebert M, Salles B, Caliou P (2004) Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. J Biol Chem 279: 35117–35126.

64. Wang H, Rosidi B, Perrault R, Wang M, Zhang L, et al. (2005) DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. Cancer Res 65: 4020–4030.

65. Grossmann KF, Ward AM, Markovic ME, Folias AE, Moses RE (2001) S. cerevisiae has three pathways for DNA interstrand crosslink repair. Mutat Res 487: 80–90.

66. Sarkar S, Davies AA, Ulrich HD, McHugh PJ (2006) DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase eta. EMBO J 25: 1285–1294.

67. Jonnalagadda VS, Matsuguchi T, Engelward BP (2005) Interstrand crosslink-induced homologous recombination carries an increased risk of deletions and insertions. DNA Repair (Amst) 4: 594–605.

68. Liu Y, Nairn RS, Vasquez KM (2008) Processing of triplex-directed psoralen DNA interstrand crosslinks by recombination mechanisms. Nucleic Acids Res 36: 4680–4688.

69. Zheng H, Wang X, Legerken RJ, Glazer PM, Li L (2006) Repair of DNA interstrand cross-links: interactions between homology-dependent and homology-independent pathways. DNA Repair (Amst) 5: 566–574.

70. Scarinse WA, Li K, Gu D, Gonzalez KD, Chen Z, et al. (2006) Somatic microindels in human cancer: the insertions are highly error-prone and derive from nearby but not adjacent sense and antisense templates. Hum Mol Genet 17: 2910–2918.

71. Kunano K, Johnson-Schlitz DM, Engels WR (2001) Sterility of Drosophila with mutations in the Bloom syndrome gene—complementation by Ku70. Science 291: 2600–2602.

72. Windhofer F, Krause S, Hader C, Schalz WA, Florl AR (2008) Distinctive differences in DNA double-strand break repair between normal urothelial and urothelial carcinoma cells. Mutat Res 658: 56–63.

73. Marculescu R, Vanura K, Montpellier B, Roulland S, Le T, et al. (2006) Recombination, chromosomal translocations and lymphoid neoplasia: Targeting mistakes and repair failures. DNA Repair (Amst) 5: 1246–1258.

74. Okamura K, Chuang WJ, Ruby JG, Guo H, Barrett DP, et al. (2008) The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature 453: 803–806.

75. Czech B, Malone CD, Zhou R, Stark A, Schlinghebyde C, et al. (2008) An endogenous small interfering RNA pathway in Drosophila. Nature 453: 798–802.

76. Koundakjian EJ, Cowan DM, Hardy RW, Becker AH (2004) The Zuker Collection: A Resource for the Analysis of Autosomal Gene Function in Drosophila melanogaster. Genetics 167: 203–206.

77. McVey M, Andersen SL, Broze Y, Sekelsky J (2007) Multiple functions of Drosophila BLM helicase in maintenance of genome stability. Genetics 176: 1979–1992.

78. Kurkdjian EJ, Keene J, Borkowski JK, Green JD (1994) P-element-mediated in vivo deletion analysis of white-apricot: deletions between direct repeats are strongly favored. Genetics 136: 1001–1011.

79. Zider A, Paumard-Rigal S, Frouin I, Silber J (1998) The vestigial gene of Drosophila melanogaster is involved in the formation of the peripheral nervous system: genetic interactions with the su(sce) gene. J Neurogenet 12: 87–99.

80. Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillips RW, et al. (1993) Type I repressors of P element mobility. Genetics 135: 81–95.