Efficient Repair of DNA Breaks in Drosophila: Evidence for Single-Strand Annealing and Competition With Other Repair Pathways

Christine R. Preston, William Engels and Carlos Flores[1]

Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin 53706

Manuscript received January 24, 2002
Accepted for publication March 8, 2002

ABSTRACT

We show evidence that DNA double-strand breaks induced in the Drosophila germ line can be repaired very efficiently by the single-strand annealing (SSA) mechanism. A double-strand break was made between two copies of a 1290-bp direct repeat by mobilizing a *P* transposon. In >80% of the progeny that acquired this chromosome, repair resulted in loss of the *P* element and loss of one copy of the repeat, as observed in SSA. The frequency of this repair was much greater than seen for gene conversion using an allelic template, which is only ~7%. A similar structure, but with a smaller duplication of only 158 bp, also yielded SSA-like repair events, but at a reduced frequency, and gave rise to some products by repair pathways other than SSA. The 1290-bp repeats carried two sequence polymorphisms that were examined in the products. The allele nearest to a nick in the putative heteroduplex intermediate was lost most often. This bias is predicted by the SSA model, although other models could account for it. We conclude that SSA is the preferred repair pathway in Drosophila for DNA breaks between sequence repeats, and it competes with gene conversion by the synthesis-dependent strand annealing (SDSA) pathway.

A clearer understanding of how eukaryotic cells repair DNA double-strand breaks (DSBs) should aid cancer research and lead to improved gene replacement techniques (Jasin 2000). Most cells have multiple pathways capable of repairing DSBs (Lankenau and Gloor 1998; Lambert *et al.* 1999; Paques and Haber 1999). One set of pathways, referred to as nonhomologous end joining (NHEJ), leads to the direct rejoining of the broken ends. In another group of pathways, repair of the broken ends is guided by an intact molecule of identical or similar DNA sequence. These processes, known collectively as "homologous recombination," include gene conversion with crossing over, conversion without crossing over, single-strand annealing (SSA), and break-induced replication.

Although DNA breaks are eventually resealed by all of these pathways, some restore the original sequence and chromosome structure much more faithfully than others. For example, repair by NHEJ usually results in small deletions or insertions at the break site. When a less faithful pathway is used, genome instability and cancer predisposition can result (Difilippantonio *et al.* 2000; Gao *et al.* 2000; Saintigny *et al.* 2001b; Tutt *et al.* 2001). Therefore, an important question is, What governs the "choice" of repair pathways?

Evidence of competition between pathways has been seen (Fishman-Lobell *et al.* 1992; Ivanov *et al.* 1996; Lambert and Lopez 2000; Fiorenza *et al.* 2001; Saintigny *et al.* 2001a). Which pathway predominates may be influenced by (1) the cell (*e.g.*, the cell type, p53 expression status, phase of cell cycle, and growth rate; Fishman-Lobell *et al.* 1992; Hagmann *et al.* 1996, 1998; Moore and Haber 1996; Willers *et al.* 2000); (2) template availability (*e.g.*, whether a homologous sequence exists on the sister chromatid, the homolog, or elsewhere; Moore and Haber 1996); and (3) the nature of the broken ends (*e.g.*, whether the ends are flush *vs.* with complementary overhangs *vs.* jagged; Pastwa *et al.* 2001; Smith *et al.* 2001; whether a protein is covalently bound, if a terminal "hairpin" is formed, and whether the break is flanked by sequence repeats).

Breaks induced by mobilization of *P*-element transposons in the developing germ line of Drosophila are most often repaired by gene conversion without crossing over (Engels *et al.* 1990; Gloor *et al.* 1991). Data suggest that synthesis-dependent strand annealing (SDSA) is the predominant mechanism (Nassif *et al.* 1994; Flores 2001). In SDSA, a broken DNA end invades a homologous template and primes DNA synthesis, producing long, 3′ single-strand extensions. Since the template is homologous, the sequence of the 3′ extension is complementary to the other broken end. Therefore, after synthesis, strands from both ends of the break may anneal together. Subsequent sequence trimming and/or gap filling, followed by ligation, complete the repair of the break.

Genetic and biochemical evidence demonstrates that in many systems, shortly after DSB formation, the bro-

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AF449103 and AF449104.

[1]*Corresponding author:* Lab of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706. E-mail: ccflores@facstaff.wisc.edu

ken ends are processed by exonuclease(s) to produce 3′ single-stranded tails (Maryon and Carroll 1989; Sugawara and Haber 1992). The 3′ single-strand regions facilitate the search for a homologous template and are necessary for invasion into homologous duplex DNA to prime synthesis. However, a unique opportunity exists when a direct sequence repeat is present on both sides of the DNA break. In this case, the single-stranded tails formed by exonuclease action can simply anneal with each other, leading to break repair with loss of one repeat. This is the essence of the SSA mechanism (Lin *et al.* 1984).

Studies of SSA have been performed in yeast, Xenopus oocytes, and mammalian cells (Lin *et al.* 1984; Maryon and Carroll 1991; Fishman-Lobell *et al.* 1992). Many of these studies used engineered substrates in which the repeats were separated by a central region containing a site for a break-inducing enzyme. The effects of varying the length of the repeats and the length of the central spacer region have been addressed. In yeast, repeats as small as 29 bp can elicit SSA, although the rate is very low, while ~415 bp are required for maximal efficiency (Sugawara *et al.* 2000). SSA can occur even if the repeats are separated from each other by several kilobases. When a central spacer region is present additional gene products are required to trim the nonhomologous ends (*e.g.*, Rad1p, Rad10p, Msh2p, Msh3p, and Srs2p; Fishman-Lobell and Haber 1992; Sugawara *et al.* 1997; Colaiacovo *et al.* 1999).

SSA requires a subset of the genes that are essential for gene conversion. Whereas *RAD52* is required for both mechanisms, *RAD51* and *RAD54* are necessary only for conversion, not SSA (Ivanov *et al.* 1996). This is probably because gene conversion involves a wide search for homology and strand invasion, which are not required with SSA. Extensive DNA synthesis is another feature that may distinguish gene conversion from SSA.

In this report we show that when a direct sequence repeat flanks the DSB site there is a very high rate of *P*-element loss and reversion to the unduplicated structure (see bottom of Figure 1). This repair outcome and the remarkable efficiency suggest that rather than gene conversion or NHEJ, an alternative mechanism of repair is at work. The very high frequency is dependent on the length of the repeat. Since SSA has been shown to have a similar dependence on repeat length and can be equally proficient in other organisms, it most likely accounts for the extreme rate of *P*-element loss.

MATERIALS AND METHODS

Drosophila crosses: Flies were raised on standard cornmeal-molasses-agar media at 25° to produce the test males and at 21° or 25° to obtain their progeny. Genetic symbols not described here can be found at the Drosophila database (FlyBase 1999).

Screens for *P*-element loss: DSBs occurred in the germ line of males of the genotype

TABLE 1

Primers used

Sequence	Name
CGTCCGCACACAACCTTTCC	2223
TCGCTGTCTCACTCAGACTC	2231
AATCACAGCTCCCTTGGTGC	D1-50C
CTCCTCCTCCTGGCACTTGT	D2a-50C
TCATTTTGCTGCAAGTCTCG	G1-50C
ACGTCGTTGGTTCTCGTTGC	D2-50C
ATTTCGATTTCGGCGATACG	G2-50C
GCCATCCTCCAAATTGGTTG	G0-50C
GCAAACGCAAACTCAATCGC	D0-50C
TGCAGCAACAACAACAACCTG	G2b-cf
CGATGCACAGCTATCGTCCTC	D2b-cf
GAGACGGGCGGTAATAGCAAA	G2a-cf
CCAGACCACATCCACCGATAC	D2a-cf
CAACGAAGTTTCCTAAGCCACA	G2c-cf
CAATGTGCCAAGTACCCTGGA	G1a-cf
ATGAGGAAATGCGCTCCATCT	D1a-cf
CCAGGGTACTTGGCACATTGA	D1b-cf
CAACAACAACCTGTTCCACATT	G2Asp
CAACAACAACCTGTTCCACAAG	G2Csp
GAAGTGGCCGCATAGGAGGAA	D2Asp
GAAGTGGCCGCATAGGAGGCT	D2Csp

$$\frac{w}{Y};\ \frac{[dup]P\{w^+\}[dup]\ bw}{al\ b\ cn\ [deletion]};\ \frac{Dr\ \Delta 2\text{-}3}{+},$$

where [*dup*] indicates a sequence of 8, 158, or 1290 bp directly duplicated on either side of the $P\{w^+\}$ element at 50C, and [*deletion*] refers to a deficiency of at least 100 kb immediately to the right of the $P\{w^+\}$ insertion point. This deletion, designated Df(2R)50C-101, is described in Preston *et al.* (1996). These males were crossed individually to females of the genotype *w/w*; *SM1*, al^2 *Cy* cn^2 sp^2*/Pm*. Crosses where the homolog carried an intact 50C template were similar, but lacked the deletion. Sons of phenotype Cy, al^+, Dr^+ were scored for eye color. Those with white eyes were considered candidates for *P* loss, and a sample of these were crossed to *CyO, Roi cn bw/cn bw* females to confirm the presence of the *bw* allele, which indicates a noncrossover second chromosome. The same progeny were used to establish homozygous stocks of noncrossover, *P*-loss chromosomes for further analysis.

Molecular analysis of repair products: PCR tests of the homozygous lines were used to determine loss of the $P\{w^+\}$ element and/or loss of one copy of the duplication. The first amplification used primers that flank the duplicated region (primers G0 and D1a-cf; see Table 1 and Figure 2), to test if the duplication had collapsed into a single unit of the repeat. Amplification of a wild-type-sized fragment with primers G0 and D1a-cf indicated loss of one copy of the duplicated segment. Loss of the $P\{w^+\}$ element was confirmed by amplification of a 145-bp product with primers D0 and G0, located on either side of the *P* element (Table 1 and Figure 2).

A sample of w^+ progeny was analyzed by PCR to test whether a *P* element remained in the original site at 50C. The presence of the two junctions was tested using the *P*-element/chromosome primer combinations 2231/D0 and 2223/G0 (Table 1).

Sample bias correction: From experiment B, 413 out of 1019 total progeny were not w^-; *i.e.*, they had some degree of eye color, from pale yellow to deep red (similar to wild-

type eyes), or had speckled eye color. A sample of 79 of these from three eye color categories were analyzed by PCR. Out of 261 flies that had eye color indistinguishable from the parental line, 53 were tested. From 148 that had eye color distinctly darker or paler than the parental line, 22 were tested, and all four flies that had speckled eye color were tested. If PCR showed that both of the original *P*-element-chromosome junctions were intact, the chromosome was classified as unchanged. If one *P*-element-chromosome junction was intact but not the other junction, the event was classified as faulty gene conversion. If neither of the original *P*-element-chromosome junctions could be amplified, and if primers D0 and G0 amplified a normal-sized fragment, the event was scored as precise reversion (*i.e.*, SSA). Nine of the sampled progeny with parental eye color appeared to have undergone SSA at the break site. Only one of the sampled flies from the darker/paler category and none of the speckled-eyed flies appeared to have undergone SSA. To estimate the fraction of all the progeny that had colored eyes yet had undergone SSA it was necessary to correct for our sampling bias. The corrected frequency used was $(\Sigma C_i p_i)/(\text{total scored})$, where C_i is the number of flies with the ith eye color (unchanged, paler/darker, speckled) and p_i is the proportion of SSA among the sample tested with that eye color. Thus, for experiment B, our estimate is

$$\frac{261(9/53) + 148(1/22) + 4(0/4)}{1019} = 0.05.$$

Therefore we estimate that 5% of all the flies scored in experiment B had undergone SSA at 50C yet had colored eyes.

DNA sequencing: The sequence of the 1290-bp repeats was determined after amplifying each repeat by PCR: primer G2 to *P* element (GenBank accession no. AF449103) and *P* element to D1 (GenBank accession no. AF449104). Purified fragments were sequenced using BigDye fluorescent dideoxy-terminator reactions according to the manufacturer's protocol (ABI; Applied Biosystems, Foster City, CA).

RESULTS

Origin and structure of duplications: The duplications used in this study are by-products of an earlier experiment addressing the nature of *P*-element-induced crossing over (Preston and Engels 1996; Preston *et al.* 1996). The initial Drosophila line carried *P*{CaSpeR}*Cp1*^{50C}, a *P*-element marked with the eye-color gene, *white* (w^+), transposed into an intron of the *Cp1* gene at cytological position 50C on the second chromosome. From this line, *P*-element-induced crossover events were selected, (as in Figure 1C). Approximately one-third of the recombinant chromosomes contained duplications of the sequence adjacent to the *P* element. Another one-third had adjacent deletions, such as Df(2R)50C-101 used in this study. The remaining recombinants had neither gained nor lost sequences. The majority of recombinants retained a *P* element and had at least one of the original transposon/chromosome junctions intact. Such structures are consistent with the hybrid element insertion (HEI) model (Figure 1, A–C), in which recombinant chromosomes are produced by an aberrant transposition reaction rather than homologous recombination (Gray *et al.* 1996; Preston *et al.* 1996).

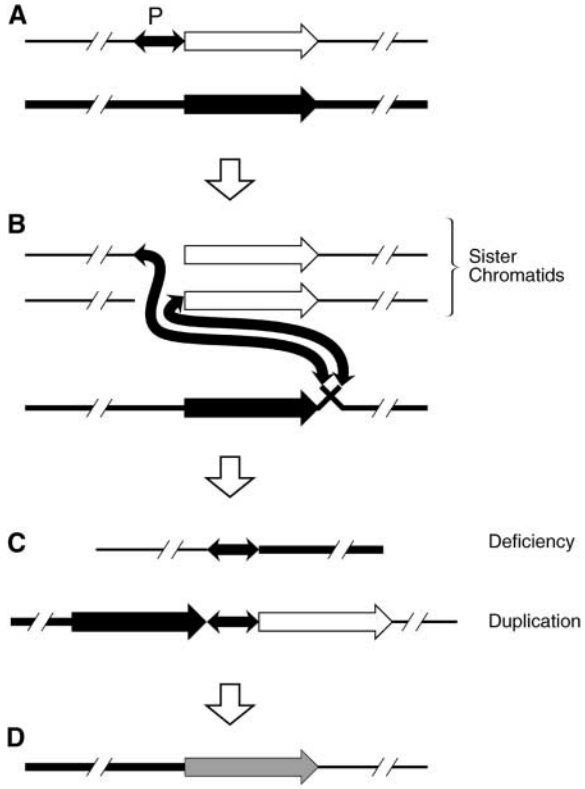

Figure 1.—Proposed mechanism for the formation of duplications (Preston *et al.* 1996) and collapse by SSA. (A) Parental chromosomes. Two homologous chromosomes (distinguished by thin and thick lines) are shown in a male heterozygous for a *P* element. Wide arrows indicate the sequence that becomes duplicated, and double arrows represent *P* elements throughout Figures 1–4. (B) HEI intermediate. A hybrid *P* element forms between the two sister chromatids and inserts into the homolog at a point to the right of the original insertion site. (C) Recombinant products. The resulting reciprocal products are recovered as recombinant for outside markers. (D) SSA product. Finally, SSA can result in reversion through collapse of the duplication (shaded arrow).

One isolate, Dp(2R)P{w^+}50C-125 (hereafter designated Dp1290), had a 1290-bp duplication and retained the *P* element between the sequence repeats. This type of structure was quite common among the products, but Dp1290 was chosen for this study because its duplication was an intermediate size. We determined the sequence of the duplicated region and found two sites where the sequence of one repeat differed from the other. The location of these polymorphisms confirmed that the copy to the right derived from the *P*-bearing chromosome (designated C) and the copy to the left derived from its homolog (designated A) as depicted in Figure 2.

Another chromosome, Dp(2R)P{w^+}50C-22 (hereafter designated Dp158), was produced in the same experiment and its structure is similar to Dp1290, except it

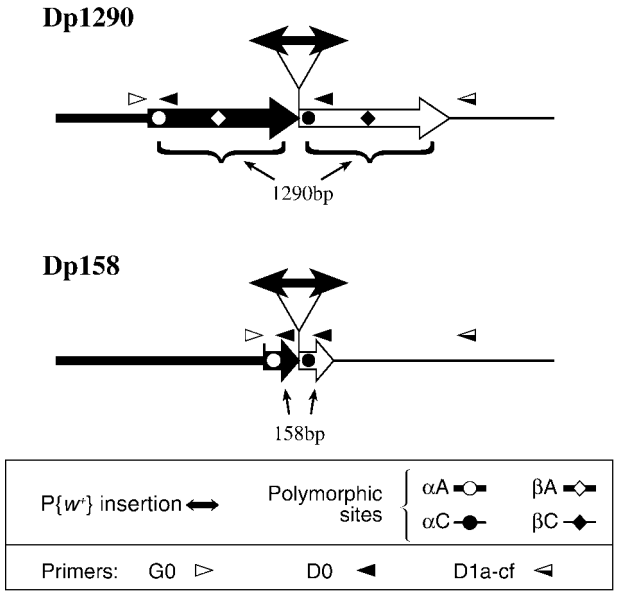

FIGURE 2.—Structure of duplications Dp1290 and Dp158: The P{w^+} "CaSpeR" element is the site of DSB formation. The duplicated sequence is represented by wide arrows. Solid arrows and thick lines depict sequence originating from the A chromosome, while open arrows and thin lines represent C chromosome sequence. The positions of polymorphic sites are symbolized by circles and diamonds within the duplicated region. The positions of primers (D0, G0, and D1a-cf) used to analyze the repair products are shown as arrowheads.

has a duplication of only 158 bp instead of 1290 bp (Figure 2).

Mobilization of the *P* element in Dp1290 leads to a remarkably high rate of element loss: Transposition of a *P* element does not necessarily lead to its loss at the original (donor) site, because it is often replaced by gene conversion (ENGELS *et al.* 1990; GLOOR and LANKENAU 1998; LANKENAU and GLOOR 1998). When a *P*{w^+} element is mobilized by a source of active transposase (abbreviated Δ2-3 in the following text), it undergoes transposition in somatic and germ cells by cutting itself out of the donor site and inserting into a new site. The DNA mechanics of *P* transposition are as follows:

1. At each end of the *P* element, transposase cleaves the two DNA strands with a 17-base stagger. Thus the chromosome is left with 5′ strands ending where the transposon began and 3′ strands extended by 17 bases of *P*-element sequence (BEALL and RIO 1997).
2. At the insertion site, an 8-bp staggered cut is made.
3. The 3′ (longer) strands of the *P* element are ligated to the overhanging strands at the new insertion site. The result is that the *P* element loses no sequence from its termini and a new 8-bp duplication of chromosomal sequence is created.

In the germ line, the broken ends that remain at the donor site are usually repaired by gene conversion using the recently replicated sister chromatid as a template. This results in the replacement of the *P* element at the original site, since it is copied back into the gap. Alternatively, but at a lower frequency, the allelic sequence residing at the same site on the homolog can be copied into the break. When the homolog does not carry a *P* element at this site, this leads to precise reversion with loss of the *P* element.

We hypothesized that *P*-element-induced breaks are not normally repaired via the SSA mechanism because the 8-bp flanking duplication is too short. However, a structure such as Dp1290 should be an appropriate substrate for SSA. Therefore, we combined Dp1290 with a transposase source and screened the progeny for loss of *white* expression, signifying loss of the *P* element. We found that 82% of the progeny had lost w^+ expression from a cross where Dp1290 was opposite a deletion on the homolog (experiment C in Table 2 and Figure 3). Furthermore, when a normal homolog was used, the frequency of w^+ loss rose to nearly 92% (experiment E in Table 2 and Figure 3.) To identify the structural changes underlying the loss of w^+, we then used PCR to analyze samples of 94 and 50 independent w^- progeny, respectively, from experiments C and E. In all 144 cases, the chromosome had been restored to the wild-type structure. That is, each chromosome had lost its *P* element and one of the two copies of the 1290-bp duplicated segment. Evidently Dp1290 undergoes SSA repair at an extremely high frequency in the presence of *P* transposase. The additional events seen in the presence of a normal homolog (92% *vs.* 82%) suggested that another repair pathway involving the homolog is also available.

For comparison, we invoked data from a similar cross using the progenitor of Dp1290 (PRESTON and ENGELS 1996) plus additional data (our unpublished results). This chromosome is identical to Dp1290 except that the duplication is only 8 bp instead of 1290 bp. As shown in Table 2 and Figure 3 (experiments A and D) the frequency of w^- progeny is much lower in these crosses. Moreover, PCR analysis showed that a large proportion of the w^- progeny had structures different from that expected from SSA repair. Combining the phenotypic and PCR data yields estimates of 1.3 and 7.2% precise loss, depending on whether the homolog carried a deletion at 50C. Therefore the larger duplication present in Dp1290 is necessary for the high frequency of SSA-like repair. As in the previous experiment, some homolog-dependent repair is also indicated.

***P* transposition in Dp158 leads to an intermediate rate of element loss:** Dp158 was also tested for its rate of *P* loss to determine whether the shorter duplication affects SSA (experiment B, Figure 3 and Table 2). It produced a high level of w^+ loss, but substantially less than Dp1290: only 59.5% *vs.* 82.4% (Table 2, B *vs.* C). Furthermore, molecular analysis of a sample of 57 of the w^- products revealed that only 89.5% of these resulted from precise reversion as opposed to 100% with

TABLE 2

Analysis of *P*-loss progeny

Expt.[a]	Duplication length (bp)	Template on homolog	Progeny scored	White eyes	Sampled for PCR[b]	Precise loss among white-eyed progeny	% precise loss[c]
A	8	No	830	131 (15.8%)	47	4 (8.5%)	1.3
B	158	No	1019	606 (59.5%)	57	51 (89.5%)	53.3
C	1290	No	1000	824 (82.4%)	94	94 (100.0%)	82.4
D[d]	8	Yes	2497	385 (15.4%)	186	87 (46.5%)	7.2
E	1290	Yes	1018	933 (91.7%)	50	50 (100.0%)	91.7

[a] Letters refer to the experiments shown in Figure 3.

[b] Only one white-eyed son per cross was used for the samples in column 6, to ensure that only independent events were analyzed.

[c] Estimates of precise loss are the products of percentages in columns 5 and 7. These values represent the estimated proportion of all flies scored that were both w^- and had reverted precisely to the original structure.

[d] Data in row D are from experiments reported previously (Preston and Engels 1996) plus additional data (our unpublished results).

the larger duplication (Table 2, B *vs.* C). Apparently SSA is only moderately efficient when presented with a duplication of 158 bp.

This assay underestimates the rate of w^+ loss at the original site: It is expected that some of the *P* excision events will coincide with reinsertion of the element elsewhere in the genome. Most new insertions will express w^+, thus masking the loss of $P\{w^+\}$ at 50C. This process will result in an underestimate of the SSA frequency. We can assess the magnitude of the underestimate by analyzing loss of $P\{w^+\}$ at the original site within a sample of the w^+ progeny. Out of a total of 413 w^+ flies from experiment B, a sample of 79 independent events was analyzed by PCR, testing for both of the chromosome/*P*-element junctions. After correcting for sampling bias (see materials and methods), we calculate that 12.4% of the w^+ progeny from this experiment had precisely lost $P\{w^+\}$ at the original site. Given that w^+ flies were 40.5% of the total progeny scored, ~5% of the total had undergone precise $P\{w^+\}$50C loss with a new $P\{w^+\}$ insertion elsewhere. Therefore the true rate of precise loss is 53.2% + 5% = 58.2%. This analysis confirmed that the purely phenotypic assay underestimates the rate of precise loss, but that the magnitude of the error is reasonably small.

Inclusion of sequence polymorphisms in the repaired products: The 1290-bp duplication includes two polymorphic sites, designated α and β. We used PCR and restriction site analysis to determine which form was present at each site in a sample of the repair products. Figure 4 shows the results for experiments C and E, *i.e.*, with and without a deletion opposite the 1290-bp duplication-bearing homolog. The results allow interpretations in terms of the SSA mechanism.

The SSA repair process includes an intermediate structure with a long heteroduplex region, as shown in Figure 4. Sequence polymorphisms within the duplication will lead to mismatches in this heteroduplex, which are often repaired by the cell's mismatch repair (MMR) system. Several studies (Holmes *et al.* 1990; Carroll *et al.* 1994; Deng and Nickoloff 1994; Lehman *et al.* 1994; Miller *et al.* 1997; Taghian *et al.* 1998) indicate that the MMR process has a strong bias such that the repaired product tends to reflect the sequence of whichever allele lies farthest from any nick. For experiment C, where we hypothesize that virtually all repair events occur by SSA, we expect a bias in favor of the A form at site α and the C form at site β. Furthermore, we expect this bias to be more pronounced in the case of site α, since it lies much closer to a duplication boundary and therefore closer to a nick in the proposed heteroduplex intermediate. The results (Figure 4, experiment C) are in good agreement with this expectation, considering that 92.6% (41.5 + 51.1%) of the α sites and 58.5% (51.1 + 7.4%) of the β sites have the favored sequence. Experiment E also showed a strong bias for the A form at site α (90%), whereas the two forms were nearly equal at site β. A similar bias was clear in experiment B, where only the α site was duplicated. The favored form (αA) occurred in 88.5% of the repair products (Figure 4).

DISCUSSION

SSA is efficient in Drosophila: We found that a sequence repeat of 1290 bp at the site of a *P*-induced DSB leads to an exceptionally high rate of *P*-element loss with reversion to the unduplicated structure. These events have the hallmarks of repair by SSA. It should be noted that these precise reversion products could have been produced by other mechanisms such as SDSA-type gene conversion. However, this would require a more complex series of steps and could not explain why a repeat would elevate the frequency so dramatically.

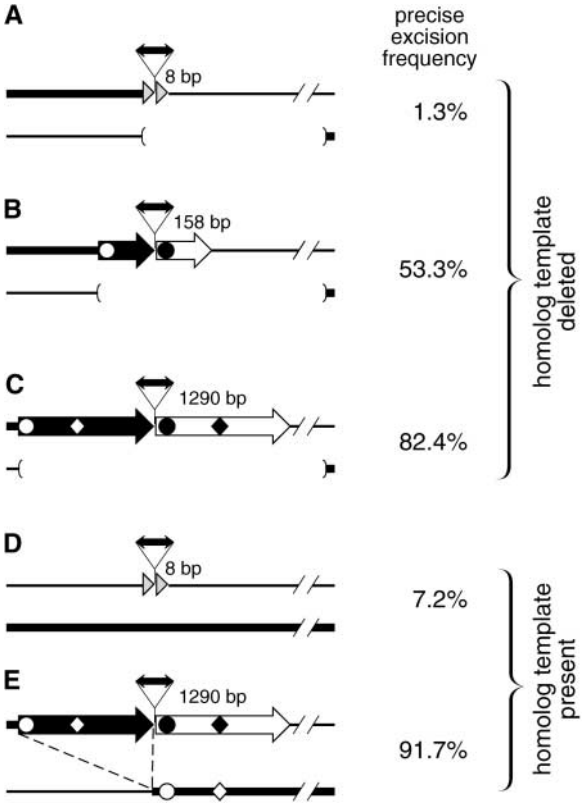

Figure 3.—Frequency of precise reversion depends on sequence flanking the break and the allelic site on the homolog. We measured the frequency of precise reversion for each of the five genotypes shown. Males with the indicated configuration at the 50C site plus a transposase source, Δ2-3(99B) (Robertson *et al.* 1988), were crossed individually to tester females. The symbols used are the same as in Figure 2. The detailed genotypes for experiment D are shown in Preston and Engels (1996; see Figure 2 of that article), and experiments A–C and E are analogous. Progeny receiving the upper homolog were then scored for eye pigmentation. A sample of independent white-eyed progeny was further tested by PCR, and the combined data were used to estimate the frequency of precise reversion, as in Table 2. The lower homolog in experiments A–C carries a long deletion extending rightward from the original P{*w*+} CaSpeR insertion point. The origin and mapping of this deletion is described in Preston *et al.* (1996; see deletion no. 101 in Figure 5A of that report).

Neither could it account for the homogeneity of products—all having lost one repeat.

Polymorphic sites were recovered in the repaired products as expected from a fully heteroduplex intermediate as shown in Figure 4, but we have not directly addressed whether the most common SSA intermediate includes the entire 1290 bp of both repeats. Some intermediates may have less heteroduplex due to either gap widening or incomplete exposure of one repeat by exonuclease, and this would affect the recovery of polymorphisms.

Evidence for SSA has been seen in *Escherichia coli*, *Saccharomyces cerevisiae*, Xenopus oocytes, mouse cells, and tobacco cells (de Groot *et al.* 1992; Dronkert *et al.* 2000; Sugawara *et al.* 2000; Tomso and Kreuzer 2000; Bibikova *et al.* 2001). One previous report suggested that SSA was efficient in Drosophila. A DSB induced in the germ line between two 3.5-kb repeats led to the loss of one copy in 85% of the progeny (Rong and Golic 2000).

Another apparently similar phenomenon has been investigated in Drosophila whereby a sequence repeat *within* a *P* element can collapse upon mobilization of the element (Kurkulos *et al.* 1994; Delattre *et al.* 1995). Though this process is also very efficient, it is distinct from SSA since it involves recopying of the element from the sister chromatid and undoubtedly occurs by SDSA with misalignment during the annealing step.

SSA efficiency depends on repeat length: On the basis of our analysis of three different-sized repeats, we conclude that an 8-bp repeat is not long enough to participate in SSA, and a 158-bp repeat is not sufficient for fully effective SSA. Our experiments cannot determine whether the 1290-bp repeat is above or below the minimum size required for maximal efficiency. One study in yeast determined that repeats of 415 bp were the minimum required for full SSA efficiency. Although SSA was detectable with repeats as small as 29 bp, it was 500-fold less efficient (Sugawara *et al.* 2000).

Interpathway competition in repair of DSBs: There is some evidence of competition between SSA and other repair pathways in our data. When the DSB was made at the simple *P*-element insertion (having an 8-bp flanking duplication), the majority of repair events that could be identified had retained or recopied some *P*-element sequence. Products that retain 17 bases or less from either side of *P* are most likely produced by NHEJ since transposase leaves behind 17 bases from one strand of the *P* element at each side of the donor site (Beall and Rio 1997). Products that included >17 bp of the *P* element from either side contained partial copies of the original element. These larger products were interpreted as incomplete or aborted gene conversion events, which used the sister chromatid as a template. These NHEJ and aborted gene conversion products together make up 14.5% (15.8 − 1.3% in Table 2) of the relevant progeny in experiment A. By comparison none of the products from Dp1290 (experiment C) appear to derive from NHEJ or aborted gene conversion. If there were no competition between SSA and NHEJ/aborted conversion, then we would expect ~13 of the 94 products analyzed to have imprecise loss. Instead, we found no examples of imprecise loss among the 94 flies sampled from experiment C. Thus SSA outcompetes NHEJ and aborted gene conversion.

We cannot say whether SSA competes with successful gene conversion using the sister chromatid since those events are undetectable and produce a product that can undergo further rounds of DSB repair. However, if conversion using the sister chromatid occurs at a significant rate from Dp1290 (experiment C), the ratio of

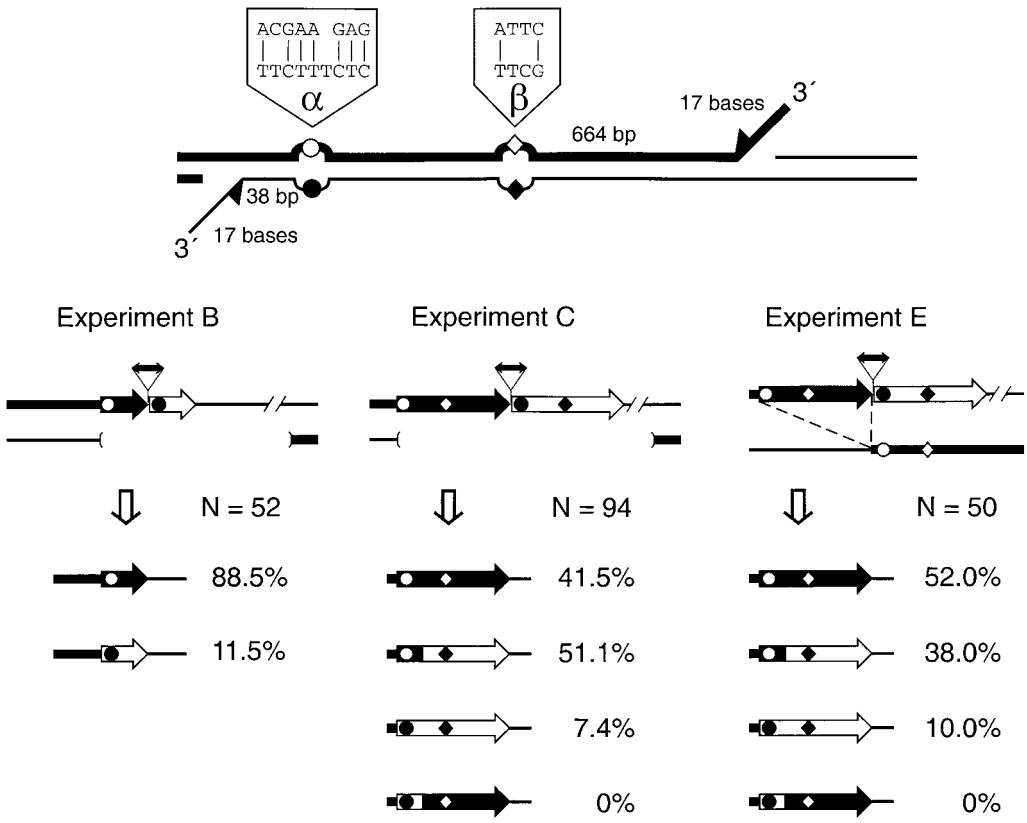

FIGURE 4.—Sequence polymorphisms in the repair intermediate and products. The nucleotide sequence of the polymorphic sites in the Dp1290 repeats is shown as paired heteroduplex in the theoretical intermediate (top). The positions of the polymorphisms are designated α and β. The symbols used are the same as in Figure 2. Polymorphisms were analyzed in the repair products (bottom) from three experiments: experiment B in which Dp158 was paired with a homolog bearing a deletion of the region, experiment C in which Dp1290 was paired with the same deletion-bearing homolog, and experiment E in which Dp1290 was paired with an undeleted homolog and therefore a repair template was present. All events analyzed were independent. Site α was assayed using the restriction enzyme *Ear*I, which cuts only the A form. For site β we employed a pair of specific PCR primers (Table 1), each of which amplified only one of the two forms.

P-element excisions to successful insertions, which is already surprisingly high, must be higher still. Competition between SSA and other repair mechanisms has also been observed in yeast and vertebrate cells (FISHMAN-LOBELL *et al.* 1992; HAGMANN *et al.* 1996; FIORENZA *et al.* 2001).

***P* elements appear to be inefficient at insertion:** One might assume that most *P*-element excisions would be accompanied by a successful insertion elsewhere. Surprisingly, this assay suggests that *P* elements are not particularly efficient in the insertion step of transposition. It has not been possible to measure the true excision rate because the *P* element is frequently copied back into the "donor" site using the sister chromatid as a template. Our observation that >80% of the progeny in experiment C lost w^+ expression suggests that Dp1290 allows SSA to compete more favorably with templated repair. Successful transposition events would spread the w^+ gene around the genome, lowering the recovery of w^- flies. In fact, if the *P* element excised from Dp1290 and resulted in SSA repair together with a new random w^+ *P* insertion in every germ cell precursor, we would expect to score only 50% w^- progeny. Therefore a large proportion of the *P*-excision events are not associated with an insertion. This conclusion would not be valid if a significant fraction of *P* insertions were phenotypically w^-; however, there is no evidence to suggest this is the case.

Is SSA prevalent in other organisms? Conflicting results have been obtained concerning whether SSA is the predominant mechanism of DSB repair at duplications in various other organisms. For example, in one assay in mammalian cells SSA was very efficient, accounting for ~70% of the products (DRONKERT *et al.* 2000), while in a very similar assay SSA constituted only 3% of the events (TAGHIAN and NICKOLOFF 1997). Numerous other reports have used DSB repair assays that differ to varying degrees. Many variables have been indicted as likely causes of the different efficiencies of SSA reported, including the length of repeats, the distance between repeats, integrated (chromosomal) *vs.* extrachromosomal substrates, location of the DSB within one repeat *vs.* between repeats, heterology within the repeats, length of terminal heterology, p53 expression status, and cell type. One particularly important factor may be the point within the cell cycle at which the DSB is repaired. Most spontaneous DSBs would likely arise during S-phase due to replication-fork collapse. The

timing of a DSB induced by endonuclease may be governed by accessibility of the cleavage site, which could be dictated by the cell type and the chromosomal location (or the fact of being extrachromosomal). Therefore, perhaps if a sister chromatid is present at the time of DSB formation, gene conversion predominates, whereas SSA is more frequent when there is no sister chromatid.

Is SSA biologically relevant? Most double-strand breaks are not repairable by SSA, which requires either a flanking duplication or two breaks within regions of sequence similarity (Haber and Leung 1996; Richardson and Jasin 2000). It is possible that this pathway exists only because the necessary machinery is present in the cell as components of other repair pathways. However, SSA may be a functional repair mechanism in certain circumstances.

Recombination substrates that enlist SSA have been valuable tools in the study of homologous recombination and DSB repair, partly because SSA utilizes some of the same steps as gene conversion. Since SSA requires a subset of the genes used in gene conversion, testing mutants for SSA proficiency can reveal in what part of the conversion pathway a particular gene functions. For example, we find that mutants for the Drosophila homolog of the Bloom syndrome gene, *mus309* (Kusano *et al.* 2001), are proficient for SSA but deficient for SDSA (our unpublished data). Our interpretation is that this gene is involved in a component of SDSA that is not shared by SSA, such as specialized DNA synthesis. SSA is also of interest to investigators trying to mutate genes using "ends-in" gene targeting constructs as performed in Drosophila (Rong and Golic 2000, 2001). Ends-in targeting leads to a duplication that may include one functional and one mutated copy of the gene. Such structures can be converted into "knock-outs" when a DSB between the repeats is repaired by SSA. Informed positioning of mutations in the targeting construct should maximize the recovery of mutants by this method.

One way SSA substrates can occur naturally is by the HEI process (Figure 1). Any DNA-based transposable element that is prone to form "hybrid" elements between sister chromatids is likely to yield structures similar to Dp1290 and Dp158 at reasonably high frequencies. As these studies show, these structures will readily degrade in the presence of transposase into what appears to be precise excisions of the element. This two-step process, *i.e.*, HEI followed by SSA (Figure 1, A–D), can explain some published observations. For example, Preston *et al.* (1996) recovered among their recombinant chromosomes 43 deletions, 35 duplications, and 45 of what appeared to be precise excisions. One interpretation is that there were fewer duplications than deletions because some of the HEI-induced duplications underwent SSA in a subsequent cell generation to yield what appeared to be a precise excision. This interpretation is strengthened by the observation of biased conversion of flanking polymorphisms, as seen here in Figure 4, within their precise excision category (see Figure 5A of Preston and Engels 1996).

Finally, there may be ample opportunity for SSA to repair spontaneous breaks, especially when two or more breaks arise simultaneously. Note that >40% of the human genome is repetitive sequence (Venter *et al.* 2001) and that best estimates suggest that mammalian cells suffer ~10 DSBs during each S-phase (Haber 1999). In fact, deletions between intrachromosomal repeats and recombination between interchromosomal repeats have already been found to underlie many genetic diseases and cancer lines (Onno *et al.* 1992; Burwinkel and Kilimann 1998; Deininger and Batzer 1999; Segal *et al.* 1999). Therefore it may be imperative that the cells of higher eukaryotes minimize their use of SSA. Testament to this constraint is the familial early onset breast and ovarian cancer syndrome caused by defects in the *BRCA2* gene. A crucial role of BRCA2 seems to be promoting gene conversion and/or preventing SSA (Tutt *et al.* 2001).

We thank Dena Johnson-Schlitz and Judith Ducau for critical reading of the manuscript. This work was supported by National Institutes of Health grant GM30948 to W.E. This is publication no. 3595 of the Laboratory of Genetics.

LITERATURE CITED

Beall, E. L., and D. C. Rio, 1997 Drosophila P-element transposase is a novel site-specific endonuclease. Genes Dev. **11:** 2137–2151.

Bibikova, M., D. Carroll, D. J. Segal, J. K. Trautman, J. Smith *et al.*, 2001 Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol. Cell. Biol. **21:** 289–297.

Burwinkel, B., and M. W. Kilimann, 1998 Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. J. Mol. Biol. **277:** 513–517.

Carroll, D., C. W. Lehman, S. Jeong-Yu, P. Dohrmann, R. J. Dawson *et al.*, 1994 Distribution of exchanges upon homologous recombination of exogenous DNA in *Xenopus laevis* oocytes. Genetics **138:** 445–457.

Colaiacovo, M. P., F. Paques and J. E. Haber, 1999 Removal of one nonhomologous DNA end during gene conversion by a *RAD1*- and *MSH2*-independent pathway. Genetics **151:** 1409–1423.

de Groot, M. J., R. Offringa, M. P. Does, P. J. Hooykaas and P. J. van den Elzen, 1992 Mechanisms of intermolecular homologous recombination in plants as studied with single- and double-stranded DNA molecules. Nucleic Acids Res. **20:** 2785–2794.

Deininger, P. L., and M. A. Batzer, 1999 Alu repeats and human disease. Mol. Genet. Metab. **67:** 183–193.

Delattre, M., D. Anxolabehere and D. Coen, 1995 Prevalence of localized rearrangements vs. transpositions among events induced by Drosophila P element transposase on a P transgene. Genetics **141:** 1407–1424.

Deng, W. P., and J. A. Nickoloff, 1994 Mismatch repair of heteroduplex DNA intermediates of extrachromosomal recombination in mammalian cells. Mol. Cell. Biol. **14:** 400–406.

Difilippantonio, M. J., J. Zhu, H. T. Chen, E. Meffre, M. C. Nussenzweig *et al.*, 2000 DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. Nature **404:** 510–514.

Dronkert, M. L., H. B. Beverloo, R. D. Johnson, J. H. Hoeijmakers, M. Jasin *et al.*, 2000 Mouse RAD54 affects DNA double-strand break repair and sister chromatid exchange. Mol. Cell. Biol. **20:** 3147–3156.

Engels, W. R., D. M. Johnson-Schlitz, W. B. Eggleston and J. Sved, 1990 High-frequency P element loss in Drosophila is homolog dependent. Cell **62:** 515–525.

Fiorenza, M. T., A. Bevilacqua, S. Bevilacqua and F. Mangia, 2001 Growing dictyate oocytes, but not early preimplantation embryos, of the mouse display high levels of DNA homologous recombination by single-strand annealing and lack DNA nonhomologous end joining. Dev. Biol. **233:** 214–224.

Fishman-Lobell, J., and J. E. Haber, 1992 Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene *RAD1*. Science **258:** 480–484.

Fishman-Lobell, J., N. Rudin and J. E. Haber, 1992 Two alternative pathways of double-strand break repair that are kinetically separable and independently modulated. Mol. Cell. Biol. **12:** 1292–1303.

Flores, C. C., 2001 Repair of DNA double-strand breaks and mismatches in Drosophila, pp. 173–206 in *DNA Damage and Repair: Vol. III. Advances From Phage to Humans*, edited by J. A. Nickoloff and M. F. Hoekstra. Humana Press, Totowa, NJ.

FlyBase, 1999 The FlyBase database of the Drosophila Genome Projects and community literature. The FlyBase Consortium—http://cbbridges.harvard.edu:7081/. Nucleic Acids Res. **27:** 85–88.

Gao, Y., D. O. Ferguson, W. Xie, J. P. Manis, J. Sekiguchi *et al.*, 2000 Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. Nature **404:** 897–900.

Gloor, G. B., and D. H. Lankenau, 1998 Gene conversion in mitotically dividing cells: a view from Drosophila. Trends Genet. **14:** 43–46.

Gloor, G. B., N. A. Nassif, D. M. Johnson-Schlitz, C. R. Preston and W. R. Engels, 1991 Targeted gene replacement in Drosophila via P element-induced gap repair. Science **253:** 1110–1117.

Gray, Y. H., M. M. Tanaka and J. A. Sved, 1996 P-element-induced recombination in *Drosophila melanogaster*: hybrid element insertion. Genetics **144:** 1601–1610.

Haber, J. E., 1999 DNA recombination: the replication connection. Trends Biochem. Sci. **24:** 271–275.

Haber, J. E., and W. Y. Leung, 1996 Lack of chromosome territoriality in yeast: promiscuous rejoining of broken chromosome ends. Proc. Natl. Acad. Sci. USA **93:** 13949–13954.

Hagmann, M., K. Adlkofer, P. Pfeiffer, R. Bruggmann, O. Georgiev *et al.*, 1996 Dramatic changes in the ratio of homologous recombination to nonhomologous DNA-end joining in oocytes and early embryos of *Xenopus laevis*. Biol. Chem. Hoppe-Seyler **377:** 239–250.

Hagmann, M., R. Bruggmann, L. Xue, O. Georgiev, W. Schaffner *et al.*, 1998 Homologous recombination and DNA-end joining reactions in zygotes and early embryos of zebrafish (*Danio rerio*) and *Drosophila melanogaster*. Biol. Chem. **379:** 673–681.

Holmes, J., Jr., S. Clark and P. Modrich, 1990 Strand-specific mismatch correction in nuclear extracts of human and *Drosophila melanogaster* cell lines. Proc. Natl. Acad. Sci. USA **87:** 5837–5841.

Ivanov, E. L., N. Sugawara, J. Fishman-Lobell and J. E. Haber, 1996 Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae*. Genetics **142:** 693–704.

Jasin, M., 2000 Chromosome breaks and genomic instability. Cancer Invest. **18:** 78–86.

Kurkulos, M., J. M. Weinberg, D. Roy and S. M. Mount, 1994 P element-mediated in vivo deletion analysis of white-apricot: deletions between direct repeats are strongly favored. Genetics **136:** 1001–1011.

Kusano, K., D. M. Johnson-Schlitz and W. R. Engels, 2001 Sterility of Drosophila with mutations in the Bloom syndrome gene—complementation by *Ku70*. Science **291:** 2600–2602.

Lambert, S., and B. S. Lopez, 2000 Characterization of mammalian RAD51 double strand break repair using non-lethal dominant-negative forms. EMBO J. **19:** 3090–3099.

Lambert, S., Y. Saintigny, F. Delacote, F. Amiot, B. Chaput *et al.*, 1999 Analysis of intrachromosomal homologous recombination in mammalian cell, using tandem repeat sequences. Mutat. Res. **433:** 159–168.

Lankenau, D. H., and G. B. Gloor, 1998 In vivo gap repair in Drosophila: a one-way street with many destinations. Bioessays **20:** 317–327.

Lehman, C. W., S. Jeong-Yu, J. K. Trautman and D. Carroll, 1994 Repair of heteroduplex DNA in *Xenopus laevis* oocytes. Genetics **138:** 459–470.

Lin, F. L., K. Sperle and N. Sternberg, 1984 Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. Mol. Cell. Biol. **4:** 1020–1034.

Maryon, E., and D. Carroll, 1989 Degradation of linear DNA by a strand-specific exonuclease activity in *Xenopus laevis* oocytes. Mol. Cell. Biol. **9:** 4862–4871.

Maryon, E., and D. Carroll, 1991 Involvement of single-stranded tails in homologous recombination of DNA injected into Xenopus laevis oocyte nuclei. Mol. Cell. Biol. **11:** 3268–3277.

Miller, E. M., H. L. Hough, J. W. Cho and J. A. Nickoloff, 1997 Mismatch repair by efficient nick-directed, and less efficient mismatch-specific, mechanisms in homologous recombination intermediates in Chinese hamster ovary cells. Genetics **147:** 743–753.

Moore, J. K., and J. E. Haber, 1996 Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **16:** 2164–2173.

Nassif, N., J. Penney, S. Pal, W. R. Engels and G. B. Gloor, 1994 Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Mol. Cell. Biol. **14:** 1613–1625.

Onno, M., T. Nakamura, J. Hillova and M. Hill, 1992 Rearrangement of the human tre oncogene by homologous recombination between Alu repeats of nucleotide sequences from two different chromosomes. Oncogene **7:** 2519–2523.

Paques, F., and J. E. Haber, 1999 Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. Microbiol. Mol. Biol. Rev. **63:** 349–404.

Pastwa, E., R. D. Neumann and T. A. Winters, 2001 In vitro repair of complex unligatable oxidatively induced DNA double-strand breaks by human cell extracts. Nucleic Acids Res. **29:** E78.

Preston, C. R., and W. R. Engels, 1996 P-element-induced male recombination and gene conversion in Drosophila. Genetics **144:** 1611–1622.

Preston, C. R., J. A. Sved and W. R. Engels, 1996 Flanking duplications and deletions associated with P-induced male recombination in Drosophila. Genetics **144:** 1623–1638.

Richardson, C., and M. Jasin, 2000 Frequent chromosomal translocations induced by DNA double-strand breaks. Nature **405:** 697–700.

Robertson, H. M., C. R. Preston, R. W. Phillis, D. M. Johnson-Schlitz, W. K. Benz *et al.*, 1988 A stable genomic source of P element transposase in *Drosophila melanogaster*. Genetics **118:** 461–470.

Rong, Y. S., and K. G. Golic, 2000 Gene targeting by homologous recombination in Drosophila. Science **288:** 2013–2018.

Rong, Y. S., and K. G. Golic, 2001 A targeted gene knockout in Drosophila. Genetics **157:** 1307–1312.

Saintigny, Y., F. Delacote, G. Vares, F. Petitot, S. Lambert *et al.*, 2001a Characterization of homologous recombination induced by replication inhibition in mammalian cells. EMBO J. **20:** 3861–3870.

Saintigny, Y., A. Dumay, S. Lambert and B. S. Lopez, 2001b A novel role for the Bcl-2 protein family: specific suppression of the RAD51 recombination pathway. EMBO J. **20:** 2596–2607.

Segal, Y., B. Peissel, A. Renieri, M. de Marchi, A. Ballabio *et al.*, 1999 LINE-1 elements at the sites of molecular rearrangements in Alport syndrome-diffuse leiomyomatosis. Am. J. Hum. Genet. **64:** 62–69.

Smith, J., C. Baldeyron, I. De Oliveira, M. Sala-Trepat and D. Papadopoulo, 2001 The influence of DNA double-strand break structure on end-joining in human cells. Nucleic Acids Res. **29:** 4783–4792.

Sugawara, N., and J. E. Haber, 1992 Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. Mol. Cell. Biol. **12:** 563–575.

Sugawara, N., F. Paques, M. Colaiacovo and J. E. Haber, 1997 Role of *Saccharomyces cerevisiae* Msh2 and Msh3 repair proteins in double-strand break-induced recombination. Proc. Natl. Acad. Sci. USA **94:** 9214–9219.

Sugawara, N., G. Ira and J. E. Haber, 2000 DNA length dependence of the single-strand annealing pathway and the role of *Saccharomyces cerevisiae RAD59* in double-strand break repair. Mol. Cell. Biol. **20:** 5300–5309.

Taghian, D. G., and J. A. Nickoloff, 1997 Chromosomal double-strand breaks induce gene conversion at high frequency in mammalian cells. Mol. Cell. Biol. **17:** 6386–6393.

Taghian, D. G., H. Hough and J. A. Nickoloff, 1998 Biased short tract repair of palindromic loop mismatches in mammalian cells. Genetics **148:** 1257–1268.

Tomso, D. J., and K. N. Kreuzer, 2000 Double-strand break repair in tandem repeats during bacteriophage T4 infection. Genetics **155:** 1493–1504.

Tutt, A., D. Bertwistle, J. Valentine, A. Gabriel, S. Swift *et al.*, 2001 Mutation in *Brca2* stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. EMBO J. **20:** 4704–4716.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural *et al.*, 2001 The sequence of the human genome. Science **291:** 1304–1351.

Willers, H., E. E. McCarthy, W. Alberti, J. Dahm-Daphi and S. N. Powell, 2000 Loss of wild-type p53 function is responsible for upregulated homologous recombination in immortal rodent fibroblasts. Int. J. Radiat. Biol. **76:** 1055–1062.

Communicating editor: M. Lichten