Distinct *P*-Element Excision Products in Somatic and Germline Cells of *Drosophila melanogaster*

Gregory B. Gloor, Jessica Moretti, Joanne Mouyal[1] and Katherine J. Keeler[2]

Department of Biochemistry, Faculty of Medicine and Dentistry, The University of Western Ontario, London, Ontario N6A 5C1, Canada

Manuscript received October 5, 1999
Accepted for publication April 24, 2000

ABSTRACT

The footprints remaining following somatic *P*-element excision from the Drosophila *white* locus were recovered and characterized. Two different types of footprints were observed. Over 75% of the footprints were short, composed of 4 or 7 nucleotides of the *P*-element inverted terminal repeat, and were similar to those found in a previously described plasmid excision assay. The remaining footprints were composed of 14–18 nucleotides of both inverted terminal repeats. These large footprints were indistinguishable from those recovered following germline *P*-element excision. Enhanced expression of the Drosophila homologue of the Ku70 protein did not affect the structure of the somatic footprints. Therefore, this protein is not a limiting factor for double-strand break repair by nonhomologous end-joining in Drosophila somatic cells.

P elements in *Drosophila melanogaster* are among the best-studied eukaryotic transposons. In part, this scrutiny is due to their utility; many genetic tools in Drosophila are based on *P* elements (Engels 1997). However, *P* elements are interesting in their own right. For example, *P* elements only recently invaded the genome of *D. melanogaster* (Clark and Kidwell 1997; Engels 1997). As a consequence, Drosophila strains collected prior to 1950 lack *P* elements but most strains collected after 1970 contain *P* elements. In addition, the P transposase protein appears to be unrelated to other transposases. This suggests an origin for *P* elements that is independent of many other transposable elements (Capy *et al.* 1996).

P elements are bounded by 31-bp inverted terminal repeats and make an 8-bp target site duplication upon their insertion (O'Hare and Rubin 1983). *P* elements are DNA-intermediate transposons that move via cut-and-paste transposition (Kaufman and Rio 1992). Their excision generates a double-strand break in the chromosome (Engels *et al.* 1990; Gloor *et al.* 1991; Kaufman and Rio 1992). The *P*-element transposase cuts the DNA to generate a 17-nucleotide (nt) 3′ single-stranded end (Beall and Rio 1997). Such a cut is apparently unique among transposons.

The protein complex that is required for *P*-element transposition probably includes the P transposase protein and host proteins (Kaufman *et al.* 1989; Beall and Rio 1997). However, P transposase appears to be the only protein required for integration (Beall and Rio 1998). One host protein that is implicated in *P*-element transposition is the Drosophila homologue of the Ku70 protein (Rio and Rubin 1988; Beall and Rio 1996). This protein is the product of the *mus309* locus and is also called (inverted repeat binding protein IRBP; Beall and Rio 1996). We refer to it as Ku70 and to the gene as *mus309*. The Ku70 protein is a common nuclear protein found in eukaryotes that is involved in DNA repair and recombination. The Ku70 protein forms a heterodimer with the Ku80 protein and binds very tightly to the ends of DNA. Binding of the Ku heterodimer to the ends of DNA is involved in signaling DNA damage and in telomere maintenance (Featherstone and Jackson 1999). The Drosophila homologue of the Ku70 protein was found originally because of its ability to bind tightly to the 17-nt single-strand ends formed following *P*-element excision (Rio and Rubin 1988).

The double-strand breaks made by *P*-element excision are repaired by the host cell to give a variety of excision products. They can be repaired either by homologous recombination or by nonhomologous end-joining. Homologous recombination primarily produces gene conversions when the double-strand break is made in premeiotic germline cells (Geyer *et al.* 1988; Engels *et al.* 1990; Gloor *et al.* 1991; Nassif *et al.* 1994). The most common gene conversion product in this instance is the replacement of the excised *P* element with another *P* element copied from the same site on the sister chromatid (Johnson-Schlitz and Engels 1993). This copying results in the exact regeneration of the *P* element at the excision site. Gene conversion can also result in precise loss of the *P* element if the allelic site on the

Corresponding author: Gregory B. Gloor, Department of Biochemistry, Faculty of Medicine and Dentistry, The University of Western Ontario, London, ON N6A 5C1, Canada.
E-mail: ggloor@julian.uwo.ca

[1] *Present address:* Department of Psychiatry, Faculty of Medicine and Dentistry, The University of Western Ontario, London, ON N6A 5C1, Canada.

[2] *Present address:* PE Biosystems, Foster City, CA 94404.

homologue that is copied lacks a *P* element (Engels *et al.* 1990; Nassif and Engels 1993).

Nonhomologous end-joining occurs by fusion of the broken DNA ends with little processing (Critchlow and Jackson 1998). Repair by this pathway usually leaves a few base pairs of the *P* element at the excision site. Thus, these types of repair products are classified as imprecise *P*-element excisions. Nonhomologous end-joining accounts for less than one-third of the double-strand break repair products following *P*-element excision in premeiotic germ cells (Johnson-Schlitz and Engels 1993). However, these types of repair products can be selected if they result in a phenotypic change (Staveley *et al.* 1995). When double-strand breaks are made by *P*-element excision in plasmids injected into developing embryos only nonhomologous end-joining products are recovered (O'Brochta *et al.* 1991).

Although nonhomologous end-joining products are recovered following *P*-element excision in both premeiotic germline cells and from plasmids injected into embryos, their structures are radically different. The nonhomologous end-joining products recovered from premeiotic germ cells typically retain 15–18 bp of both *P*-element inverted terminal repeats (Takasu-Ishikawa *et al.* 1992; Staveley *et al.* 1995). The repeats may be separated by a few base pairs of sequence derived from an unknown origin or by a partially copied *P* element. The sequences derived from the *P*-element ends are termed footprints and, by convention, an excision product that retains sequence from each terminal inverted repeat is said to have two distinct footprints. In contrast, the nonhomologous end-joining products recovered from the plasmid excision assay conducted in embryos typically retain only 4 or 7 bp of one *P*-element inverted repeat and lack any extraneous DNA (O'Brochta *et al.* 1991). By convention, these excision products are said to contain only one footprint.

P elements in nature are normally active only in the germline because the third intron of the P transposase mRNA is spliced out in that tissue, but not in the soma (Laski *et al.* 1986; Siebel and Rio 1990). Our goal was to study the excision of *P* elements in somatic tissues to see if they were repaired primarily by gene conversion (as in the developing germline) or by nonhomologous end-joining. Additionally, we were interested to characterize somatic imprecise excision products. To excise a *P* element inserted in exon 6 of the *white* locus we used the stable transposase source Δ2-3(99B), which expresses the P transposase protein in both somatic and germline tissues (Robertson *et al.* 1988). These excision products were cloned and sequenced to determine their structure. It should be noted that a co-injected Δ2-3 plasmid was the transposase source for the plasmid excision assay conducted in embryos. Thus some of these footprints were derived from somatic *P*-element excision and others were derived from germline excisions. The experiment reported here unambiguously characterizes the products of *P*-element excision in somatic cells. Our results show that somatic *P*-element excisions are repaired to produce both the short and the long type of footprint, with the short footprint predominating. Enhanced expression of the Ku70 protein in somatic tissues had no effect on the structure of the excision products.

MATERIALS AND METHODS

Genetic techniques: Flies were reared on standard cornmeal-sugar-agar medium at room temperature (20°–23°) unless noted. Mating schemes are described in the text. Genetic symbols not described here are in Lindsley and Zimm (1992).

Drosophila strains: The *y* w^{hd} *f* ; + ; + stock carries a single *P*-element insertion in exon 6 of the *white* locus (Engels *et al.* 1990). It is the parent stock for all the flies that carry the w^{hd} allele. The *w* ; *CyO* / *SP* ; *ry Sb* Δ2-3(99B) / *TM6* (*Ubx*) and *C(1)DX, y w f* ; *ry Sb* Δ2-3(99B) / *TM3* (*Ser*) stocks were used to introduce the transposase source Δ2-3(99B) (Robertson *et al.* 1988) into the crosses to mobilize the *P* element in the w^{hd} allele. The wild-type *mus309* gene was carried as a transgene in two stocks that were a generous gift of Dr. D. Rio (University of California, Berkeley): w^{118} ; *P*{w^{+mC} gIRBP}23.1 / *CyO* (Beall and Rio 1996) and w^{118} ; *P*{w^{+mC} *hsP70* IRBP}b-11. The constructs in these stocks are referred to in the text as *P*{w^{+} gKu70} and *P*{w^{+} H·Ku70}, respectively. Both of the *mus309* transgenes express the Ku70 protein and rescue the *mus309* DNA repair defect (Beall and Rio 1996; D. Rio, personal communication).

DNA preparation for PCR: To separate fly heads from their bodies about 50 flies were placed in a 1.5-ml centrifuge tube and frozen in dry ice for at least 1 hr. The frozen flies were then shaken vigorously and dumped onto a clean surface. Individual heads were collected for DNA preparation. DNA was isolated from single flies or single fly heads as described (Gloor *et al.* 1993), except that the grinding solution contained 0.1% Triton X-100 and the first incubation was at 50° for 20 min. Incubating at 50° instead of 37° in the presence of detergent gave more reproducible DNA preparations.

PCR: PCR amplification was done as described previously (Gloor *et al.* 1991; Nassif *et al.* 1994). The oligonucleotide primers used in these amplifications were: ef+, GGTTG TCGTACCTCTCATGG [*white* locus nt 1845–1864 using the standard *white* gene coordinates established by O'Hare *et al.* (1984)]; Hi−, ACAGCGAAAGAGCAACTACG (*white* locus nt 2446–2427); and wRR, CAACACAACTTATGCCGCGT (*white* locus nt 3149–3130; T. Dray and G. Gloor, unpublished data). The standard thermal profile for PCR amplifications with these primers was to denature at 95° for 1 min, to anneal at 55° for 1 min, and to extend at 72° for 1.5 min. Nonspecific amplification products were minimized in two ways: first, by adding the primers after the reaction mix was prewarmed to 85° (D'Aquila *et al.* 1991) and, second, by carrying out the first annealing segment of the PCR at 65° with the temperature of each subsequent annealing segment being reduced by 1° every cycle until a final annealing temperature of 55° was achieved (Don *et al.* 1991). Reactions were always amplified for a further 20 cycles after reaching the target annealing temperature.

DNA quantitation: Agarose gels were stained in ethidium bromide for 45 min, and a TIFF image was captured with an AlphaImager gel documentation system (Alpha Innotech, San Leandro, CA). The image was imported into the National Institutes of Health Image Program (Rasband 1998), and the aggregate fluorescence of each band was determined using the

gel plotting macros distributed with the program. Background fluorescence was subtracted, and the contribution of each band to the total amount of DNA in each lane was calculated.

Cloning and sequencing: Standard cloning protocols were followed (Sambrook *et al.* 1989). PCR products were cloned into the pGEM-T vector system 1 (Promega, Madison, WI) as recommended by the manufacturer, except that half of the recommended amount of vector was used. In some cases the appropriate DNA fragment was purified from a low-melting agarose gel by binding the DNA to finely ground glass beads (Gene Clean, Bio 101, Carlsbad, CA). In other cases, unpurified PCR products were used. The ligation products were transformed into the XLI-Blue *Escherichia coli* strain (Stratagene, La Jolla, CA) and plated on MacConkey's lactose plates containing 40 μg/ml of ampicillin. Colonies were allowed to grow at 37° for 36 hr prior to scoring the lac phenotype. Lac$^-$ colonies were recognized by their white or light pink color and were patched onto Luria broth (LB) plates containing 40 μg/ml ampicillin. The size of each insertion in the pGEM vector was determined by a modified rapid cell cracking procedure (Chaconas *et al.* 1981). In brief, individual patches of single colonies of about 0.3 × 0.3 cm were picked off LB + ampicillin plates and suspended in 100 μl of 10 mm Tris, pH 8.0, 1 mm EDTA, 10 μg/ml RNAse A. An equal volume of 2× cracking buffer (28% sucrose, 100 mm Tris, pH 7.8, 4 mm EDTA, 2% SDS, 0.1 mg/ml bromophenol blue) was added. The tubes were inverted once and immediately centrifuged at 14,000 × *g* for 15 min. The nonviscous supernatant, usually about 20 μl, was fractionated on a 0.7% agarose gel.

Cloned products were sequenced using the ef+ oligonucleotide on an ABI Prizm sequencing apparatus at the Robarts Research Institute sequencing facility.

RESULTS

Lack of gene conversion from the sister chromatid in somatic cells: Excision of the *P* element from the w^{hd} allele in premeiotic germline cells results in most of the chromosomes retaining the *P* element because the nonexcised *P* element on the sister chromatid is copied into the excision site by gene conversion (Johnson-Schlitz and Engels 1993). We were interested to see if *P*-element excisions in head tissues were also repaired by gene conversion using the sister chromatid as a template. Gene conversion with a sister chromatid template would result in most of the chromosomes retaining the w^{hd} *P* element at its original insertion site. We reasoned that PCR amplification of the region of the *white* gene that contained the w^{hd} *P*-element insertion site could provide a crude estimate of the proportion of chromosomes in a DNA sample that carried the transposon. This would indicate the relative frequency of precise or imprecise *P*-element excision.

We induced *P*-element excision in female flies by crossing male flies of the genotype *w / Y*; *Sb* Δ2-3(99B) / *TM3, Ser* with female flies carrying the w^{hd} mutation. This cross segregated two classes of female progeny. The first class did not contain P transposase and had the genotype w^{hd} / *w*; *TM3, Ser* / +. The w^{hd} *P* element in this class was stable. The second class contained P transposase and had the genotype w^{hd} / *w*; *Sb* Δ2-3(99B) / +. The w^{hd} *P* element in this class was excised by the P transposase protein. DNA was prepared from heads of both classes of flies, and PCR amplification was used to estimate the proportion of X chromosomes that carried the w^{hd} *P* element. The primers ef+ and Hi− are complementary to *white* genomic DNA on the left and right sides of the w^{hd} *P*-element insertion site. Therefore, this primer pair amplifies a product of 607 bp in the absence of the w^{hd} *P* element or one of 1241 bp in its presence. Figure 1 shows the results of a representative PCR amplification. *P*-element excision is occurring in the samples that contain the *Sb* Δ2-3(99B) chromosome but not in those that lack it. The amplification products in the control samples contain two bands. The larger 1241-bp band represents amplification from the chromosome carrying the w^{hd} *P* element and the smaller 604-bp band represents amplification from the homologous chromosome. The amplification products in the *Sb* Δ2-3(99B)-containing samples contain only a ~604-bp band indicating the absence of the w^{hd} *P* element. This result indicates that the frequency of gene conversion using a sister chromatid template is very low in the soma and that the frequency of precise or imprecise excision is elevated compared to the frequency in the germline.

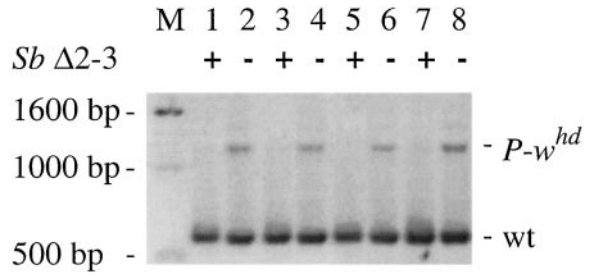

Figure 1.—Somatic excision products of the w^{hd} *P* element. The w^{hd} *P* element and closely flanking sequence in the *white* locus were amplified by PCR using the primers ef+ and Hi− using DNA prepared from individual fly heads with the genotypes described in the text. The PCR amplification was carried out for 20 cycles, which permits the major products of *P* element excision to be observed. The resulting products were fractionated on an agarose gel. The image shown is a negative of the ethidium bromide-stained gel. Excision of the w^{hd} *P* element is occurring in the samples containing *Sb* Δ2-3(99B) and is not occurring in the samples lacking it. The PCR amplifications are from female flies heterozygous for the *w* allele and the w^{hd} allele. The 604-bp band that represents the amplification product from the *w* chromosome and the 1241-bp band that represents the amplification product from the w^{hd} mutant chromosome are labeled as wt and *P*-w^{hd}, respectively.

The results in Figure 1 show that the 1241-bp band is less intense than the 607-bp band even when *P*-element excision was not occurring (the amplification products from the samples lacking transposase). Therefore, it was possible that the 1241-bp band is present in the *Sb* Δ2-3(99B)-containing samples but is undetectable with this PCR-based assay. This possibility was tested by mixing DNA prepared from male wild-type and w^{hd} flies in different proportions, followed by PCR amplification as be-

TABLE 1

Summary of w^{hd} *P*-element excision results

	Male		Female	
Size of PCR product (bp)	*Sb* Δ2-3(99B)	*TM3, Ser*	*Sb* Δ2-3(99B)	*TM3, Ser*
604	197	0	160	0
1241	2	58	5	51
Both bands	1	0	24	0

fore. Seven independent replicate experiments were performed. The results indicated that samples composed of 80% wild-type DNA and 20% w^{hd} DNA gave amplification products in which the 1241-bp band corresponding to the w^{hd} *P* element made up 20% of the amplified product (data not shown). Therefore, this assay was sufficiently sensitive to detect the 1241-bp band even if it made up 20% or less of the input DNA. Thus, at least 80% of the chromosomes that originally carried the w^{hd} *P* element had lost the transposon by precise or imprecise excision.

We next analyzed excision of the w^{hd} *P* element from the somatic tissue of flies that lacked a homologous *white* gene. In this instance the homologous *white* gene was not present and thus could not be used as a template for gene conversion. We reasoned that this should increase the proportion of sister chromatid gene conversions because there is no competition from the allelic *white* gene. We used either male flies with the genotypes w^{hd} / *Y* ; *Sb* Δ2-3(99B)/+ and w^{hd} / *Y* ; *TM3, Ser*/+ or female flies with the genotypes w^{hd} / w^{11E4} ; *Sb* Δ2-3(99B) /+ and w^{hd} / w^{11E4} ; *TM3, Ser* / +. [The w^{11E4} allele is a complete deletion of the *white* locus (Zachar and Bingham 1982).] DNA was prepared from the heads of 58 male and 51 female flies containing *TM3, Ser* and from 200 male and 189 female flies containing Δ2-3(99B). PCR amplification was performed as before with the ef+ and Hi− primers with the results shown in Table 1. The 1241 band, which results from amplification of the *white* gene in which the w^{hd} *P* element is intact, was the only band observed in the control samples. In contrast, the 1241-bp band was undetectable in 98.5% of the male flies and in 84% of the female flies in which excision was occurring because of the presence of the P transposase protein. We conclude that most of the double-strand breaks made by *P*-element excision in somatic cells results in precise or imprecise excision. This result implies that gene conversion in which the sister chromatid is used as a template is not a common pathway for double-strand break repair in the somatic tissues. In contrast, gene conversion is the predominant double-strand break repair product in the germline (Engels *et al.* 1990; Gloor *et al.* 1991; Johnson-Schlitz and Engels 1993). In this tissue precise or imprecise excision of the w^{hd} *P* element occurs infrequently (<30%; Johnson-Schlitz and Engels 1993). Thus, the double-strand break made following *P*-element excision in somatic cells is usually repaired by a different mechanism than that in premeiotic germ cells.

Imprecise excision products: We noted that the band amplified from the *white* gene when *P*-element excision was occurring in the soma was slightly larger and more diffuse than the band amplified from the wild-type *white* gene (Figure 1). This suggested that the majority of the *P*-element excision events were not precise and that a small footprint consisting of part of the *P*-element inverted terminal repeat might be found at the excision site. *P*-element footprints had been extensively characterized following *P*-element excision in the germline (excision from the *vestigial* gene and 3C on the X chromosome; Takasu-Ishikawa *et al.* 1992; Staveley *et al.* 1995) and in embryonic tissues (excision from a plasmid; O'Brochta *et al.* 1991), but not at the *white* locus. In addition, *P*-element footprints derived by excision of the same *P* element from germline and somatic tissues had not been characterized. Thus, we characterized the footprints following *P*-element excision from the *white* locus in both the germline and the soma.

We first needed to establish if imprecise excision of the w^{hd} *P* element in the germline produced footprints similar to those observed at other loci. Therefore, we isolated the footprints remaining at the *white* locus following excision of the w^{hd} *P* element in the germline and compared them to those reported at two other loci (Takasu-Ishikawa *et al.* 1992; Staveley *et al.* 1995).

P-element excisions in the male germline were generated by crossing 15 *y* w^{hd} / *Y*; *Sb* Δ2-3(99B) / + male flies individually with 3 *C(1)DX, y w f* female flies and collecting the male progeny. DNA was extracted from single flies, with between 4 and 10 progeny sampled per mating. The w^{hd} *P*-element insertion site was amplified using primers ef+ and wRR, and the size of the amplified products was determined by agarose gel electrophoresis. In total, 99 flies were examined, and 21 flies (21%, distributed among 13 families) showed single PCR products that were smaller than the 1.9-kbp band amplified from the w^{hd} *P*-element control. All but one of the smaller PCR products migrated near the 1.3-kbp band amplified from a wild-type *white* gene. The amplified PCR product from this single event migrated at about 1.4 kbp. The amplified PCR products from 13 indepen-

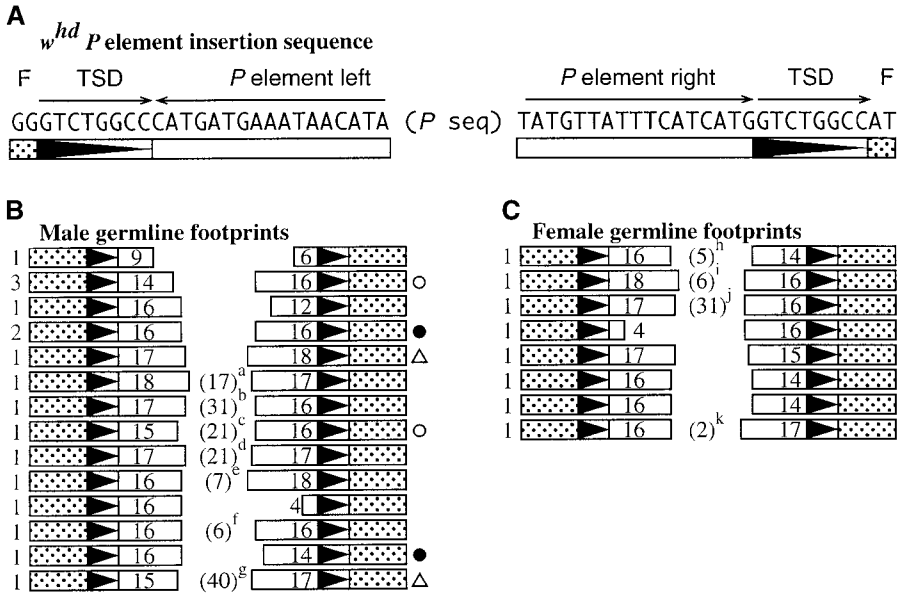

Figure 2.—Germline *P*-element footprints. *P*-element excisions in the male and female germlines were generated and sequenced as described in the text. A shows the sequence of the terminal 17 nucleotides of the *P*-element insertion in the w^{hd} allele (*P*-element left and *P*-element right), the 8-bp target site duplication (TSD), and 2 nucleotides of flanking sequence (F). The experimentally derived footprints that contain at least part of the 8-bp target site are shown in B and C. The number of bases of *P*-element end sequence is shown inside or adjacent to the open boxes. Numbers in parentheses indicate the sequence length of any extra sequence inserted between the *P*-element ends. The sequence of the non-*P*-element end insertions is given in Table 2; superscript letters indicate the source of the sequence. Symbols to the right of the footprints from the male germline indicate which footprints were derived from the same family. For example, the two different footprints adjacent to the small triangles were isolated from sibling males.

dent flies, and from four siblings, showing the 1.3-kbp band were sequenced.

The structure of the 17 germline-derived *P*-element excision products is shown in Figure 2B. Every excision product has two footprints consisting of between 6 and 18 nt of inverted repeat sequence derived from each *P*-element end. Several of the excision products have an extra sequence inserted between the *P*-element footprints. The most frequent footprint was 16 bp in length, and the average length was 15.1 bp ($n = 34$). These footprints were indistinguishable from those isolated from the *vestigial* locus (Staveley *et al.* 1995) and from cytological location 3C (Takasu-Ishikawa *et al.* 1992). We conclude that germline *P*-element footprints at all loci are equivalent.

We next examined the sequences remaining following somatic *P*-element excision events. These events were collected from DNA prepared from the heads of male flies with the genotype *y* w^{hd} /*Y* ; *Sb* Δ2-3(99B). PCR amplification of the w^{hd} *P*-element excision site was carried out with the primers ef+ and wRR. The DNA from each PCR amplification was cloned, and the insert size was determined (materials and methods). The clones that were approximately the size expected for an excision of the w^{hd} *P* element were sequenced using the ef+ primer. Every sequenced clone was independent, being derived from a different fly head. Twenty-eight of the clones retained one or both of the target site duplications. These clones contained imprecise or precise excisions of the w^{hd} *P* element and had the structures shown in Figure 3B. We also recovered several large deletions that were not characterized further.

The structures of these somatic excision events can be grouped into short and long footprint categories. The majority of the excision products (20/28) were short and contained seven nucleotides or fewer of the *P*-element inverted repeat sequence. Most of these short excision products (14/20) contained four or seven nucleotides derived from either *P*-element terminus. These excision products are defined as containing only one footprint. The majority of the remaining excision products (5/8) were long and were similar in size and structure to the germline footprints described above. Two of the remaining excision products had a small deletion, and one was a precise excision of the w^{hd} *P* element. Three of the long footprint clones contained extra sequences between the *P*-element footprints. The average length of the *P*-element footprint that remained following somatic *P*-element excision was 7.6 bp.

The effect of the Ku70 protein: We were interested to determine why the somatic and germline *P*-element excision products differed. Several reports in the literature determined the direction of the next experiments. First, Staveley *et al.* (1995) observed that germline *P*-element footprints corresponded to the number of nucleotides of *P*-element termini that are protected by binding of the IRBP. The IRBP is the Drosophila homologue of the mammalian Ku70 protein and is encoded by the *mus309* gene (Beall *et al.* 1994; Beall and Rio 1996). Second, P transposase cuts in the *P*-element inverted terminal repeat to generate a 17-nucleotide 3′ overhanging end during *P*-element excision (Beall and Rio 1997). Two groups proposed that the Ku70/80 heterodimer could bind to the single-strand end and facili-

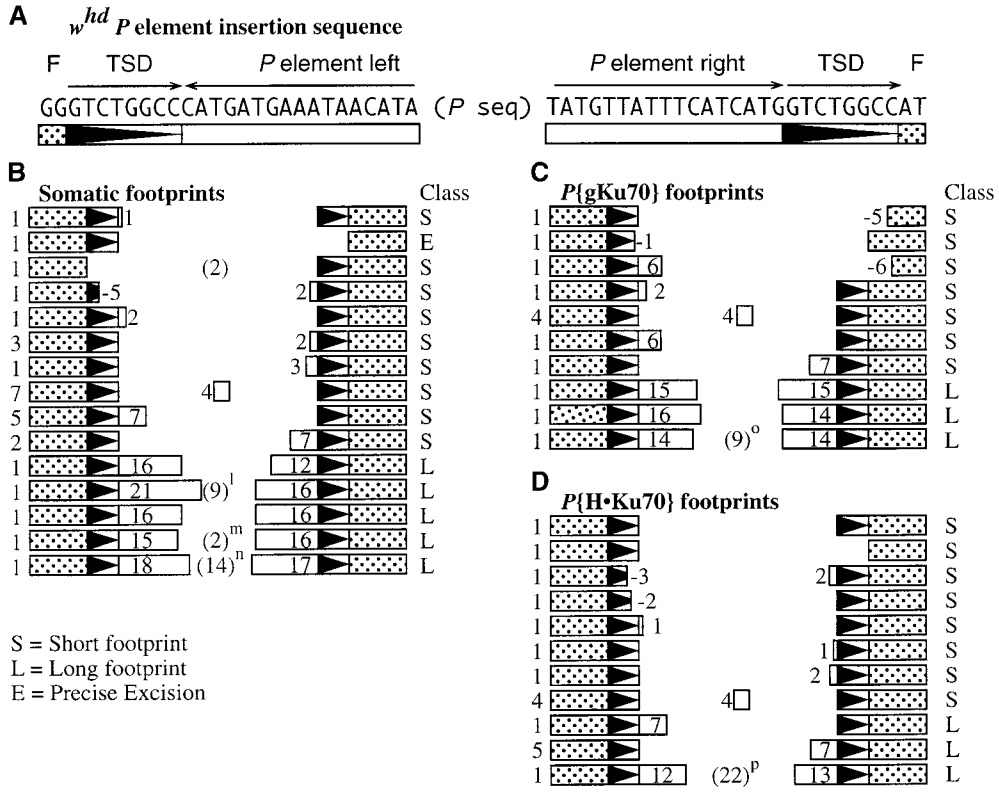

Figure 3.—Somatic *P*-element footprints. *P*-element excisions were isolated by PCR amplification of DNA extracted from fly heads and sequenced as described in the text. A shows the sequence of the terminal 17 nucleotides of the *P*-element insertion in the w^{hd} allele (*P* element left and *P* element right), the 8-bp target site duplication (TSD), and 2 nucleotides of flanking sequence (F). The experimentally derived footprints that contain at least part of the 8-bp target site are shown in B–D. The number of bases of *P*-element end sequence is shown inside or adjacent to the open boxes. Numbers in parentheses indicate the sequence length of each insertion that differs from the canonical *white* gene or *P*-element sequence. Negative numbers indicate that bases are deleted starting at the last recognizable junction. The sequence of the non-*P*-element end insertions is given in Table 2; superscript letters indicate the source of the sequence and are continuous with those in Figure 2.

tate DNA repair (Staveley *et al.* 1995; Beall and Rio 1997). Third, the Ku70 mRNA is abundant in the female germline and is found at lower levels in other tissues and developmental stages (Beall *et al.* 1994; Jacoby and Wensink 1994). Fourth, Johnson-Schlitz and Engels (1993) showed that the majority of germline *P*-element excision events were repaired without degradation of the ends at the breakage site, suggesting that a protein factor protects the broken ends. Taken together, these results suggested that one possible explanation for the discrepancy in structure between the excision products observed in the germline and somatic tissues was that higher levels of the Ku70 protein in the germline could protect the ends from degradation.

We tested the hypothesis that the *P*-element footprints observed in somatic tissues differed from the footprints observed in the germline because of different levels of the Ku70 protein. This question was addressed in two ways. First, we examined *P*-element excisions isolated from the female germline, where Ku70 is expressed at much higher levels than in the male germline (Jacoby and Wensink 1994). Second, we enhanced the expression of the Ku70 protein at the same time that excision of the w^{hd} *P* element was induced.

Imprecise excisions were obtained from the germline of female flies with the genotype *y* w^{hd} *spl* / + w^{1118} + ; *Sb* Δ2-3(99B) / +. These females were mated to males with the genotype *w* / *Y* ; *P*{CaSpeR}(35E). Progeny males containing the yellow and split mutant phenotypes (to ensure that the proper *white* locus was selected) and the *P*{CaSpeR}(35E) element were collected. Because the *P*{CaSpeR}(35E) insertion shows extreme mosaicism in the presence of Δ2-3(99B), males lacking the Δ2-3(99B) transposase source could be identified. This procedure was necessary because meiotic recombination could separate the Stubble bristle phenotypic marker from the otherwise undetectable Δ2-3(99B) locus. DNA was prepared from nonmosaic males and amplified with the primers ef+ and wRR. This primer pair will not amplify the *white* gene in the *P*{CaSpeR} element because a sequence complementary to the wRR primer is not found in the *P*{CaSpeR} element. The sizes of the amplified fragments were determined by agarose gel electrophoresis. Eight amplifications yielded fragments of approximately 1.3 kbp, indicating imprecise excision of the w^{hd} *P* element. These PCR products were sequenced with the results shown in Figure 2C. All eight of these excisions were similar to those recovered from the male germline. We conclude that the relative level of expression of the Ku70 protein in the germline does not affect the structure of the *P*-element excision products that are recovered.

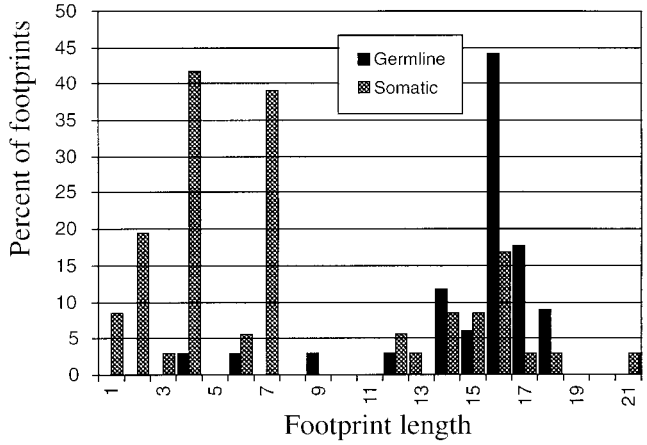

Figure 4.—Distribution of *P*-element footprints from somatic and germline excisions. The proportion of *P*-element footprints of each length is charted for the somatic and germline excision events. Germline footprints cluster around 16 nucleotides in length. Somatic footprints are either 4 or 7 nucleotides long, with a small minority being 16 ± 1 nucleotides in length.

Next, the effect of enhancing the expression of the Ku70 protein in somatic tissues was examined. Two transgenic Ku70-expressing fly lines were obtained from Dr. D. Rio; one expressed the Ku70 protein from its normal genomic promoter ($P\{w^+$ gKu70$\}$), and one expressed the Ku70 protein from an *hsP70* promoter ($P\{w^+$ H·Ku70$\}$). While the amount of expression of the Ku70 protein in either line is unknown, this expression is sufficient to rescue the *mus309* mutant phenotype. Therefore, biologically significant levels of the Ku70 protein must be made by these transgenic constructs. Both transgenes are located on the second chromosome. Male flies with the genotypes w^{hd} / Y ; $P\{w^+$ gKu70$\}$/+; *Sb* Δ2-3(99B)/+ and w^{hd} / Y ; $P\{w^+$ H·Ku70$\}$/+; *Sb* Δ2-3(99B)/+ were generated. The flies containing the $P\{w^+$ H·Ku70$\}$ transgene were heat-shocked at 35° for 1 hr each day for their entire life cycle. DNA was prepared from single fly heads as before, and the DNA sequence flanking the w^{hd} *P*-element excision site was amplified. The band expected for the w^{hd} *P* element was absent when samples of the PCR amplifications were fractionated on an agarose gel (data not shown). The PCR products from different heads were cloned and sequenced as before. A total of 31 clones that retained one or both of the 8-bp target site duplications were recovered. Thirteen independent clones were recovered from the flies with the Ku70 protein expressed from its normal promoter, and 18 independent clones were recovered from the flies with the Ku70 protein expressed from the *hsP70* promoter. The structure of the clones is shown in Figures 3C and 4D. As was seen for the somatic excision products that were obtained in the absence of enhanced expression of the Ku70 protein, the most common excision product retained the terminal four or seven nucleotides of the *P*-element inverted terminal repeat. We conclude that increasing the level of expression of the Ku70 protein has no effect on the length or structure of *P*-element footprints in either the germline or soma.

Figure 4 summarizes the size distribution of *P*-element footprints recovered in the germline and somatic tissues recovered from males and females. The predominant footprints from somatic tissues are 2, 4, and 7 bp long, whereas the predominant footprints from the germline are 16 ± 1 bp long.

It is striking that almost half (13/27) of the long footprint-containing excisions include an extra sequence. These extra sequences are shown in Table 2. Inspection of the extra sequences shows that they are rich in A and T residues and are similar in sequence to short segments of the *P*-element ends.

DISCUSSION

Lack of gene conversion from the sister chromatid template in somatic cell double-strand break repair: We examined the products that remained following *P*-element excision in somatic cells in the absence of a homologous template. On the basis of previous studies in the germline we expected that most *P*-element excision products would be repaired by gene conversion using the sister chromatid as a template and would be indistinguishable from the original chromosome (Johnson-Schlitz and Engels 1993). However, most of the somatic excision products had lost the w^{hd} *P* element. Thus the majority of the somatic excision events were repaired by nonhomologous end-joining.

There are three likely explanations for this observation. The first possibility is that the somatic tissues are incapable of gene conversion and use nonhomologous end-joining as the predominant double-strand break repair pathway. We do not favor this explanation because somatic gene conversion is readily observed as red patches of tissue in an otherwise white eye in female flies of the genotype w^{hd}/w^1 ; Δ2-3(99B)/+ (Engels *et al.* 1990). Similar eye-color mosaicism is observed in female flies containing w^{1118} or w^a in place of w^1 (G. B. Gloor, unpublished data), but not in females containing w^{hd} or w^{11E4}. Eye-color mosaicism is also not observed in males of the genotype w^{hd}/Y ; Δ2-3(99B)/+. These observations demonstrate that the somatic tissues can repair a double-strand break by gene conversion using a homologous template. However, our data showing the loss of the w^{hd} *P* element following excision indicates that the sister chromatid is not the preferred template for gene conversion in the soma. The second possibility is that different factors are needed to repair the double-strand break made by *P*-element excision using the sister chromatid template or the homologous chromosome template. This model proposes that the factor required for sister chromatid gene conversion is

TABLE 2

Extra sequences inserted in *P*-element footprints

Sequence	*P*-element 5′ end	Insertion and flanking sequence	*P*-element 3′ end
Reference sequence	CATGATGAAATAACATA	*P* element sequence	TATGTTATTTCATCATG
A[a]	CATGATGAAATAACATAA	TCATACATATTATAACA	TATGTTATTTCATCATG
B	CATGATGAAATAACATA	TTATTAATATATTATTAATATTATTAA TAAC	ATGTTATTTCATCATG
C	CATGATGAAATAACA	AATGTTTTTGTTATTATTAAC	ATGTTATTTCATCATG
D	CATGATGAAATAACATA	TGTTATATATGTTATAACATA	TATGTTATTTCATCATG
E	CATGATGAAATAACAT	TGTTATG	TTATGTTATTTCATCATG
F	CATGATGAAATAACAT	GTTAAC	ATGTTATTTCATCATG
G	CATGATGAAATAACATA	ATATGTTATAATAATAATATAATAATATG TTATATAATAA	TGTTATTTCATCATG
H	CATGATGAAATAACAT	GTTAT	TTATTTCATCATG
I	CATGATGAAATAACATAA	CATGAA	ATGTTATTTCATCATG
J	CATGATGAAATAACATA	TGTTATTATGATATTATATGATTATTA TATG	ATGTTATTTCATCATG
K	CATGATGAAATAACAT	GA	TATGTTATTTCATCATG
L	CATGATGAAATAACA TAAGGT	TATTTCATG	ATGTTATTTCATCATG
M	CATGATGAAATAACA	AC	ATGTTATTTCATCATG
N	CATGATGAAATAACATA	TGTTATATATGTTA	TATGTTATTTCATCATG
O	CATGATGAAATAA	TATATTCCC	GTTATTTCATCATG
P	CATGATGAAATA	TTTCATAATATTTCATAATATA	TTATTTCATCATG

[a] The sequences are lettered to correspond with those identified in Figures 2 and 3 as containing a sequence in addition to the *P*-element end.

absent in somatic cells. It is interesting to note that in *Saccharomyces cerevisiae* the RAD54 protein mediates gene conversion using a sister chromatid template while the TID1 protein (RDH54 protein) mediates gene conversion using a homologous chromosome template (Arbel *et al.* 1999). This possibility can be tested once the *RAD54* gene homologues are identified in Drosophila. The third possibility is that *P*-element excision in somatic cells occurs prior to replication of the *P* element. In this case the sister chromatid would not be available to serve as a template, and gene conversion using the homologous template or nonhomologous end-joining are the only pathways available. Cell-cycle control of transposition has been demonstrated for the *Ac*/*Ds* system in maize (Chen *et al.* 1987).

The source of large *P*-element footprints: Two research groups have previously characterized a large number of germline *P*-element excision products (Takasu-Ishikawa *et al.* 1992; Staveley *et al.* 1995). These groups showed that the predominant footprints recovered from germline excisions contained 16 ± 1 nucleotides of each *P*-element inverted terminal repeat. This footprint is independent of the excision site, because excisions at the *vg* locus on chromosome 2 and the 3C locus on the X chromosome and now the *w* locus all give the same types of footprints. P transposase makes a staggered cut that leaves a 17-nucleotide 3′ overhanging end at the site of *P*-element excision (Beall and Rio 1997). The terminal nucleotides at the excision site are complementary, leading to the hypothesis that 16-nucleotide footprints could arise by base-pairing of the terminal 2 nucleotides of the 3′ extended ends followed by DNA synthesis (Beall and Rio 1997). The germline-type pathway is at least partially active because 9 of 59 somatic excision products are germline-type *P*-element footprints.

The source of small *P*-element footprints: The short footprints observed in this experiment are similar to those observed with an injected plasmid-based *P*-element excision assay. The largest assay of this type was carried out by O'Brochta *et al.* (1991) who examined *P*-element footprints in embryos injected with plasmids carrying a *P* element and a Δ2-3-type P transposase source. Since the majority of the tissues in the embryo are somatic, the injected plasmid assay for *P*-element excision should give results similar to those observed in our somatic excision assay. O'Brochta *et al.* (1991) observed that 60 of 65 footprints were either four or seven nucleotides long, with the seven-nucleotide-long footprint being twice as common as the four-nucleotide footprint. We found that the four- and seven-nucleotide footprints were formed at equal frequencies in the somatic excision assay. O'Brochta *et al.* did not observe any of the long footprints. This may be because of the presence of multiple copies of the plasmid per cell in the embryonic excision assay but only one double-strand break per cell in our experiment. An excess of double-strand breaks could titrate one or more factors crucial to double-strand break repair in the injected embryos.

The final seven nucleotides of the *P*-element inverted

terminal repeat form an imperfect direct repeat. The sequence of the left and right junctions between the w^{hd} *P*-element insertion site and the *P*-element inverted terminal repeats is GTCTGGCCCATGATG and CAT CATGGTCTGGCC (the underlined nucleotides are the terminal seven nucleotides of the *P*-element inverted terminal repeat). Thus, the seven-nucleotide footprint could be formed by base-pairing of the entire imperfect repeat followed by mismatch repair or DNA synthesis to fix one or the other internal nucleotide in the repair product. The four-nucleotide repeat could be formed by base-pairing of the four terminal nucleotides of the repeat. Repeats of this type are commonly found when double-strand breaks are repaired by nonhomologous end-joining. They are formed *in vitro* in reactions that contain either the MRE11 exonuclease and DNA ligase (Paull and Gellert 1998) or DNA ligase, Ku70, Ku86, and DNA-PK (Baumann and West 1998).

We observed precise excisions of the w^{hd} *P* element in 2/59 somatic excisions. Precise excisions in the male germline occur with a frequency of about 1/1000 gametes sampled (Engels *et al.* 1990). The elevated somatic frequency may be a consequence of more extensive exonuclease degradation that would expose the target site duplications prior to end-joining.

The role of the Ku70 protein: The Drosophila homologue of the Ku70 protein (previously called IRBP) binds specifically to the terminal 17 nucleotides of the *P*-element inverted terminal repeat (Rio and Rubin 1988). Ku70 protein binding to the inverted terminal repeat sequences prior to transposition may help form an active transposase complex (Beall and Rio 1997); however, it seems to be dispensable for strand exchange (Beall and Rio 1998). Once *P*-element excision has occurred, the Ku70 protein is thought to remain bound to the single-strand ends at the excision site to facilitate DNA repair (Staveley *et al.* 1995; Beall and Rio 1996). This would be in accordance with the role of the Ku70 protein in other systems where it promotes nonhomologous end-joining (Baumann and West 1998; Featherstone and Jackson 1999).

We investigated the role of the Ku70 protein in *P*-element excision by enhancing its production and analyzing *P*-element excision products from these flies. We observed that increased production of the Ku70 protein in the soma did not result in an increase in the germline-type *P*-element footprints. This observation was contrary to that expected (Staveley *et al.* 1995; Beall and Rio 1997). Furthermore, the footprints recovered from the male and the female germlines were equivalent although there are vastly different levels of Ku70 in these tissues (Jacoby and Wensink 1994). Thus, our experiments indicate that the Ku70 protein is not the limiting factor for double-strand break repair in Drosophila germline or somatic cells. However, both the Ku70 and Ku80 proteins may need to be expressed at elevated levels for an effect on double-strand break repair to be observed because the heterodimer is the active form of the protein. Further studies will be required to determine the role of the Ku70 protein in *P*-element excision.

Control of *P* transposition: One of the major functions of double-strand break repair is to reconstitute the chromosome if the DNA replication fork proceeds through a single-strand nick in the DNA (Kuzminov 1995). An intimate association between double-strand break repair and DNA replication has been demonstrated in both *E. coli* and in *S. cerevisiae* (Kuzminov 1995; Haber 1999; Holmes and Haber 1999; Kuzminov and Stahl 1999). This suggests that double-strand break repair using a sister chromatid template may be most efficient if it occurs immediately following DNA replication.

P elements could take advantage of the association between double-strand break repair and DNA replication by excising after the passage of a DNA replication fork. Alternatively, *P*-element excision in the germline could be under cell-cycle control. Either regulatory method would ensure that a *P*-element-containing sister chromatid template was available for double-strand break repair. Therefore, the double-strand breaks generated by germline excisions would result in gene conversion from the sister chromatid template. In contrast, if a *P* element can excise in a somatic cell at any time in the cell cycle, a sister chromatid template might not be available to repair the resulting double-strand break by gene conversion. Instead, the double-strand breaks would be repaired by the nonhomologous end-joining pathway, resulting in short footprints. The unequivocal demonstration of such control of *P*-element excision would be a major advance in our understanding of both *P*-element biology and of double-strand break repair in Drosophila.

We thank Faye Males for technical assistance and Dave Haniford for stimulating discussions on the mechanism of transposition. Angela Coveny, Michael Simmons, and two anonymous reviewers made invaluable comments on the manuscript. Kathy Keeler was funded by a National Sciences and Engineering Research Council predoctoral fellowship. These experiments were funded by operating grant MT-11374 from the Medical Research Council of Canada.

LITERATURE CITED

Arbel, A., D. Zenvirth and G. Simchen, 1999 Sister chromatid-based DNA repair is mediated by RAD54, not by DMC1 or TID1. EMBO J. **18:** 2648–2658.

Baumann, P., and S. C. West, 1998 DNA end-joining catalyzed by human cell-free extracts. Proc. Natl. Acad. Sci. USA **95:** 14066–14070.

Beall, E. L., and D. C. Rio, 1996 Drosophila IRBP/Ku p70 corresponds to the mutagen-sensitive *mus309* gene and is involved in *P*-element excision in vivo. Genes Dev. **10:** 921–933.

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

Beall, E. L., and D. C. Rio, 1998 Transposase makes critical contacts with, and is stimulated by, single-stranded DNA at the *P* element termini in vitro. EMBO J. **17:** 2122–2136.

Beall, E. L., A. Admon and D. C. Rio, 1994 A Drosophila protein homologous to the human p70 Ku autoimmune antigen interacts

with the *P* transposable element inverted repeats. Proc. Natl. Acad. Sci. USA **91:** 12681–12685.

Capy, P., R. Vitalis, T. Langin, D. Higuet and C. Bazin, 1996 Relationships between transposable elements based upon the integrase-transposase domains: is there a common ancestor? J. Mol. Evol. **42:** 359–368.

Chaconas, G., R. M. Harshey, N. Sarvetnick and A. I. Bukhari, 1981 Predominant end-products of prophage Mu DNA transposition during the lytic cycle are replicon fusions. J. Mol. Biol. **150:** 341–359.

Chen, J., I. M. Greenblatt and S. L. Dellaporta, 1987 Transposition of *Ac* from the *P* locus of maize into unreplicated chromosomal sites. Genetics **117:** 109–116.

Clark, J. B., and M. G. Kidwell, 1997 A phylogenetic perspective on *P* transposable element evolution in Drosophila. Proc. Natl. Acad. Sci. USA **94:** 11428–11433.

Critchlow, S. E., and S. P. Jackson, 1998 DNA end-joining: from yeast to man. Trends Biochem. Sci. **23:** 394–398.

D'Aquila, R. T., L. J. Bechtel, J. A. Videler, J. J. Eron, P. Gorczyca *et al.*, 1991 Maximizing sensitivity and specificity of PCR by preamplification heating. Nucleic Acids Res. **19:** 3749.

Don, R., P. Cox, B. Wainwright, K. Baker and J. Mattick, 1991 "Touchdown" PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. **19:** 4008.

Engels, W. R., 1997 Invasions of *P* elements. Genetics **145:** 11–15.

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.

Featherstone, C., and S. P. Jackson, 1999 Ku, a DNA repair protein with multiple cellular functions? Mutat. Res. **434:** 3–15.

Geyer, P. K., K. L. Richardson, V. G. Corces and M. M. Green, 1988 Genetic instability in *Drosophila melanogaster*: *P*-element mutagenesis by gene conversion. Proc. Natl. Acad. Sci. USA **85:** 6455–6459.

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.

Gloor, G. B., C. R. Preston, D. M. Johnson-Schlitz, N. A. Nassif, R. W. Phillis *et al.*, 1993 Type I repressors of *P* element mobility. Genetics **135:** 81–95.

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

Holmes, A. M., and J. E. Haber, 1999 Double-strand break repair in yeast requires both leading and lagging strand DNA polymerases. Cell **96:** 415–424.

Jacoby, D. B., and P. C. Wensink, 1994 Yolk protein factor 1 is a Drosophila homolog of Ku, the DNA-binding subunit of a DNA-dependent protein kinase from humans. J. Biol. Chem. **269:** 11484–11491.

Johnson-Schlitz, D. M., and W. R. Engels, 1993 *P*-element-induced interallelic gene conversion of insertions and deletions in *Drosophila melanogaster*. Mol. Cell. Biol. **13:** 7006–7018.

Kaufman, P. D., and D. C. Rio, 1992 *P* element transposition in vitro proceeds by a cut-and-paste mechanism and uses GTP as a cofactor. Cell **69:** 27–39.

Kaufman, P. D., R. F. Doll and D. C. Rio, 1989 Drosophila *P* element transposase recognizes internal *P* element DNA sequences. Cell **59:** 359–371.

Kuzminov, A., 1995 Collapse and repair of replication forks in *Escherichia coli*. Mol. Microbiol. **16:** 373–384.

Kuzminov, A., and F. W. Stahl, 1999 Double-strand end repair via the RecBC pathway in *Escherichia coli* primes DNA replication. Genes Dev. **13:** 345–356.

Laski, F. A., D. C. Rio and G. M. Rubin, 1986 Tissue specificity of Drosophila *P* element transposition is regulated at the level of mRNA splicing. Cell **44:** 7–19.

Lindsley, D. L., and G. G. Zimm, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.

Nassif, N., and W. Engels, 1993 DNA homology requirements for mitotic gap repair in Drosophila. Proc. Natl. Acad. Sci. USA **90:** 1262–1266.

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. Trends Biochem. Sci. **14:** 1613–1625.

O'Brochta, D. A., S. P. Gomez and A. M. Handler, 1991 *P* element excision in *Drosophila melanogaster* and related drosophilids. Mol. Gen. Genet. **225:** 387–394.

O'Hare, K., and G. M. Rubin, 1983 Structures of *P* transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell **34:** 25–35.

O'Hare, K., C. Murphy, R. Levis and G. M. Rubin, 1984 DNA sequence of the *white* locus of *Drosophila melanogaster*. J. Mol. Biol. **180:** 437–455.

Paull, T. T., and M. Gellert, 1998 The 3′ to 5′ exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. Mol. Cell **1:** 969–979.

Rasband, W., 1998 NIH Image 1.62, http://rsb.info.nih.gov/nih-image/.

Rio, D. C., and G. M. Rubin, 1988 Identification and purification of a Drosophila protein that binds to the terminal 31-base-pair inverted repeats of the *P* transposable element. Proc. Natl. Acad. Sci. USA **85:** 8929–8933.

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.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Plainview, NY.

Siebel, C. W., and D. C. Rio, 1990 Regulated splicing of the Drosophila *P* transposable element third intron *in vitro*: somatic repression. Science **248:** 1200–1208.

Staveley, B. E., T. R. Heslip, R. B. Hodgetts and J. B. Bell, 1995 Protected *P*-element termini suggest a role for inverted-repeat-binding protein in transposase-induced gap repair in *Drosophila melanogaster*. Genetics **139:** 1321–1329.

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

Zachar, Z., and P. M. Bingham, 1982 Regulation of *white* locus expression: the structure of mutant alleles at the *white* locus of *Drosophila melanogaster*. Cell **30:** 529–541.

Communicating editor: M. J. Simmons