MOLECULAR AND CELLULAR BIOLOGY, Mar. 1994, p. 1613–1625
0270-7306/94/$04.00+0

Vol. 14, No. 3

Efficient Copying of Nonhomologous Sequences from Ectopic Sites via P-Element-Induced Gap Repair†

NADINE NASSIF,[1] JANICE PENNEY,[2]‡ SEBELY PAL,[2]§ WILLIAM R. ENGELS,[1] AND GREGORY B. GLOOR[2]*

Laboratory of Genetics, University of Wisconsin—Madison, Madison, Wisconsin 53706,[1] and Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland A1B 3V6, Canada[2]

Received 23 August 1993/Accepted 2 December 1993

P-element-induced gap repair was used to copy nonhomologous DNA into the *Drosophila white* locus. We found that nearly 8,000 bp of nonhomologous sequence could be copied from an ectopic template at essentially the same rate as a single-base substitution at the same location. An in vitro-constructed deletion was also copied into *white* at high frequencies. This procedure can be applied to the study of gene expression in *Drosophila melanogaster*, especially for genes too large to be manipulated in other ways. We also observed several types of more complex events in which the copied template sequences were rearranged such that the breakpoints occurred at direct duplications. Most of these can be explained by a model of double strand break repair in which each terminus of the break invades a template independently and serves as a primer for DNA synthesis from it, yielding two overlapping single-stranded sequences. These single strands then pair, and synthesis is completed by each using the other as a template. This synthesis-dependent strand annealing (SDSA) model as a possible general mechanism in complex organisms is discussed.

We used a transposable P-element insertion allele of the *white* gene to study the repair of a double-stranded DNA break in *Drosophila melanogaster*. Unrepaired double-stranded DNA breaks can be cell lethal after mitosis. Consequently, organisms have developed efficient methods for their repair. Double strand break repair is thought to occur by a process in which the broken ends search for a homologous sequence, invade it, and serve as primers for DNA synthesis to reconstitute the broken chromosome (29, 32, 39).

P elements are thought to transpose by a cut-and-paste process in which excision of the element breaks both DNA strands of the chromosome (6, 12, 19). Experiments have shown that the break is usually repaired by copying the corresponding sequences from the sister strand (6), thus restoring a copy of the P element to the break site. Alternatively, sequence may be copied from a template on the homolog (18, 25) or one inserted in the genome at an ectopic site (12). The DNA sequence flanking the P excision site is usually converted as a continuous block with an average tract length of about 1,400 nucleotides. Recent results have also shown that small insertions and deletions from the homolog can be converted into the excision site (18).

These observations suggested to us that the gap repair process could be adapted for use as an efficient gene targeting method in *D. melanogaster* (Fig. 1). We tested this possibility by using a P-element insertion in the *white* gene, w^{hd} ($w^{hd80K17}$), which is known to excise at high frequencies in the presence of P transposase. We then recovered gap repair products that occurred in the presence of one of five ectopically located templates. In four of these templates, the *white* gene had been altered by the addition of unrelated sequences of lengths up to approximately 7,970 bp. We found that each of the nonhomologous sequences could be copied into the *white* locus almost as efficiently as a single-base alteration. Similarly, we tested a fifth template with a 136-bp deletion relative to the wild-type *white* sequence and found that it could also be efficiently copied.

We also recovered several classes of aberrant conversion events in which large portions of the template sequence had been duplicated. The endpoints of these duplications corresponded to sites of directly repeated sequences present in the template. Such events can be explained if each of the two broken ends conducts an independent genome search for a template. According to this model, DNA synthesis occurs from each end to produce overlapping single strands which rejoin at a point of direct homology to complete the synthesis. As will be discussed below, this kind of mechanism differs from models involving a double Holliday structure (Fig. 1) in several testable ways and may prove to be a common pathway for double strand gap repair in metazoans.

MATERIALS AND METHODS

Construction of insertion and deletion templates. DNA cloning was done by standard techniques (23, 35). The plasmid pP[*w*alter] was produced from pP[CaSpeR] (31) as previously described (12). The P[*w*alter] transposon carries a *white* gene with 12 single-base changes, each of which adds or removes a restriction site relative to the canonical *white* gene sequence (27). The mutated *white* gene still encodes a functional product and extends the 3′ end of the *white* gene 162 nucleotides beyond that contained in P[CaSpeR].

The plasmid pP[*w*alL], shown in Fig. 2, was derived from pP[*w*alter] as follows: the polylinker in pP[*w*alter] at the 5′ end of the *white* gene was removed between the *Eco*RI and *Pst*I sites by a double digest. The linear plasmid was digested with S1 nuclease to blunt the ends, and the plasmid was recircularized. A new polylinker was introduced into the plasmid at

* Corresponding author. Present address: Department of Biochemistry, University of Western Ontario, London, Ontario N6A 5C1, Canada. Phone: (519) 661-3526. Fax: (519) 661-3175. Electronic mail address: ggloor@julian.uwo.ca.

† Paper 3380 from the Laboratory of Genetics.

‡ Present address: Department of Biochemistry, University of Western Ontario, London, Ontario N6A 5C1, Canada.

§ Present address: Department of Animal Sciences, University of Adelaide, Glen Osmond 5064, South Australia, Australia.

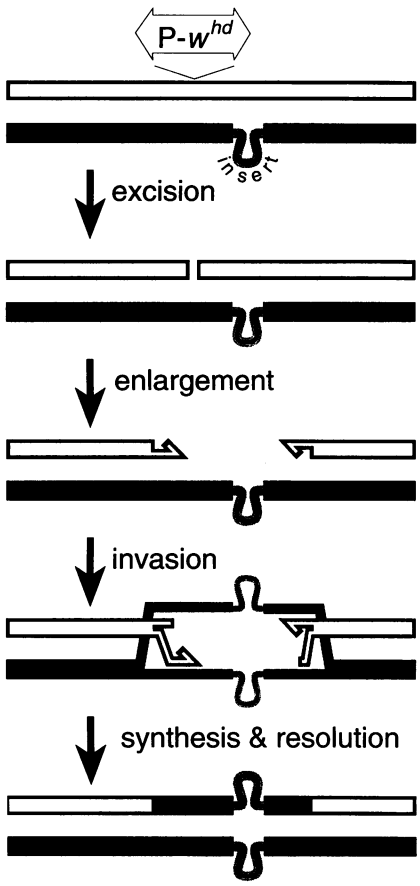

FIG. 1. Gap repair process for gene targeting. Excision of the P-element insertion (P-w^{hd}) in the *white* gene (white box) leaves behind a double-stranded DNA break. The homologous template sequence (black line) has a nonhomologous sequence inserted into it at some distance from the site corresponding to the P-w^{hd} insertion site. The nonhomologous insertion is invisible during the homology search if the broken ends are sufficiently degraded prior to their finding the template. The invading ends serve as primers for DNA synthesis to copy the template, and the nonhomologous sequence is copied into the *white* locus.

nucleotide position −2368 in the O'Hare numbering scheme (27) at the 3′ end of the *white* gene by recombinant PCR (17). The polylinker sequence is GTTAACTGCAGGATCCTCG AGAATTCTAGA and contains unique sites for *Hpa*I, *Pst*I, *Bam*HI, *Xho*I-, *Eco*RI-, and *Xba*I. The first two and last three bases correspond to the original *white* gene sequence. This polylinker site is 34 nucleotides downstream of the 3′ end of the *white* gene cDNA (27, 30).

The plasmid pP[*w*alLKP], shown in Fig. 2, carries an insertion derived from KP, an internally deleted P element common in natural populations, as described previously (3). A fragment consisting of all of the KP element except the first and last 15 nucleotides was cloned into the *Bam*HI site of the pP[*w*alL] polylinker. To avoid confusion in later discussions, the complete P-element ends derived from P[*w*alL] have been named P_T 5′ and P_T 3′, and the incomplete KP-derived ends are called KP5′ and KP3′.

The plasmid pP[*w*alL*y*] (Fig. 2) contains a wild-type allele of the *yellow* gene inserted into the *Xho*I site of the P[*w*alL] polylinker. The insert consists of the *Sal*I fragment from the plasmid D-2873 (11). This *Sal*I fragment is about 243 bp larger than reported and is closer to 7,970 bp in length. The extra sequence lies near the 3′ end of the gene.

A fourth insertion template, P[dbl*w*alL*y*], was produced spontaneously from P[*w*alL*y*] during the germ line transformation process. As shown in Fig. 2, it has one copy of P[*w*alL*y*] inserted into another in direct orientation.

Finally, we made a deletion template starting with a vector derived from the P-element-bearing plasmid pπ25.1 (28). This vector, pπPvu, has 188 bp of P 5′ and 173 bp of P 3′ end sequence with a unique *Pvu*II site located between the P ends. It lacks any *white* gene sequences flanking the P element, thus facilitating modification of *white* sequences placed internally. The mini-*white* gene was excised from pP[*w*alL] by double digestion with the restriction enzymes *Hin*dIII and *Ppu*MI. The ends were blunted by filling them in with the large fragment of *Escherichia coli* DNA polymerase I, and this *white* gene fragment was ligated into the *Pvu*II site of pπPvu, regenerating the *Ppu*MI site. One clone in which the orientation of the *white* gene and the P ends was the same as in pP[*w*alL] was chosen. A *Not*I site was introduced into this plasmid by recombinant PCR (17). This site changed 6 bases of *white* sequence with the result that the last two amino acids of the *white* gene sequence were mutated from Lys Glu to Gly Arg. A 136-bp deletion was introduced between a *Taq*I site and a *Bss*HII site by double digestion, filling in with the Klenow fragment, and religation. The resulting construct was named pP[*w*al*Not*Δ] (Fig. 2) and has a deletion between nucleotide positions 223 and 368 relative to the P-w^{hd} insertion site.

The templates described above were introduced into the *Drosophila* genome by P-element transformation as previously described (33, 34, 36).

Genetic screen for reversions. Genetic symbols not defined here are defined elsewhere (22). *Drosophila* stocks were maintained on a standard cornmeal-molasses-agar formulation. Crosses were performed at 22 or 25°C, brooded at 5 to 7 days, and scored until 1 to 2 days before the F_2 generation was expected to eclose (24 days at 22°C and 18 days at 25°C).

The scheme for recovering template-dependent reversions of w^{hd} was similar to one used previously (12) and is shown in Fig. 3. The P-element excision and gap repair events occur in the germ line of the parental male which carries an immobile P-element transposase source, Δ2-3(99B) (33), the w^{hd} target site, and one of the five template constructs described above and in the legend to Fig. 2. Revertant progeny were identified by their fully wild-type eye color. Only revertant sons which did not receive Δ2-3(99B), as recognized by the Sb^+ phenotype, were kept for further study.

Molecular analysis of revertants. DNA for PCR was prepared from individual flies as previously described (13). Oligonucleotide priming sites for PCR are shown in Fig. 4. Each DNA sample was amplified with primer pairs p1-p2 and p3-p4 to determine whether the left end and right end of the template were present. Flies positive for both ends were presumed to carry a copy of the template transposon. The rest were analyzed by PCR and restriction mapping, as described below, to determine the structure of the revertant allele and the extent of the conversion tract. Additional PCR amplifications were used to determine the structure of the complex aberrant events. The usual primer pairs used, their amplification product, and the site(s) analyzed (restriction enzyme used) were as follows: p5-p6, 298 bp, −1906 (*Nde*II); p7-p8, 682 bp, −1025 (*Nde*II); p9-p10, 537 bp, −393 (*Nde*II); p7-p10, 1,421 bp, −1025 (*Nde*II), −393 (*Nde*II); p11-p12, 509 bp, −135 (*Nde*II), −24 (*Nde*II), 1 (*Hae*III), 28 (*Rsa*I), 72 (*Hae*III), 82 (*Nde*II), 238 (*Nde*II); p13-p14, 348 bp, 238, 311 (*Rsa*I); and

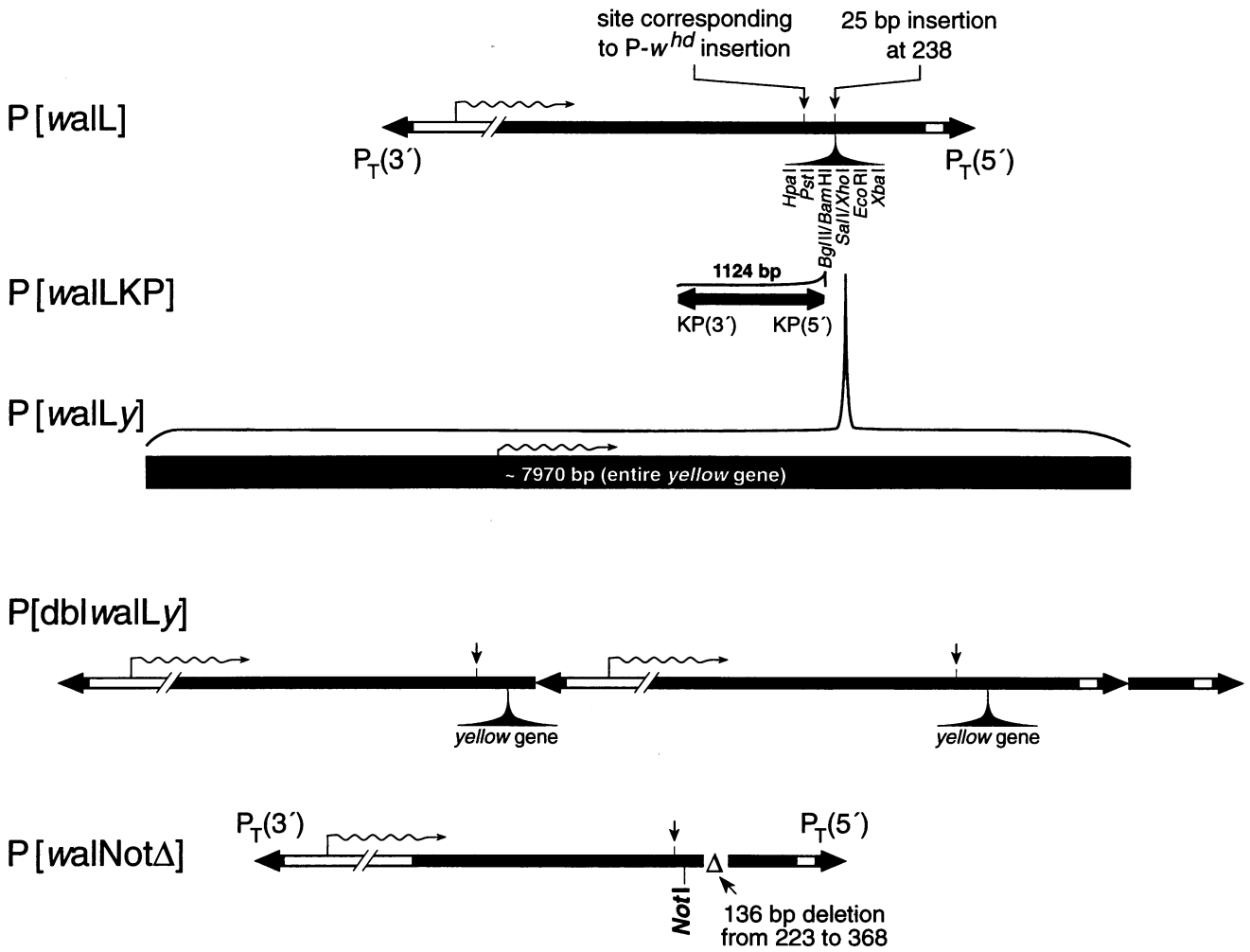

FIG. 2. Gene targeting templates. The starting construct for the insertion templates is P[*w*alL], which was derived from P[*w*alter] (12), and has the same 12 single-base alterations that introduce or remove restriction endonuclease cleavage sites from the canonical *white* gene sequence (27). The 2,790 bp of *white* gene sequence which contains these alterations is shown as a black line, and the other portions of the *white* gene are shown as a white box. The cloning sites and insert sizes of the nonhomologous sequences are indicated. The truncated, inverted terminal repeat sequences of the KP ends are represented as blunt arrows. The P[dbl*w*al*Ly*] template is a tandem duplication of P[*w*al*Ly*], and it was produced during the transformation process. The construct P[*w*al*Not*Δ] has the *white* gene from P[*w*alter] cloned in between P-element ends derived from pπ25.1 (28). It has nucleotides 65 to 70 changed to introduce a *Not*I restriction cleavage site and a deletion of 136 bp between nucleotides 223 to 368. The small arrows represent the site in the templates which correspond to the P-w^{hd} insertion site. In this and the figures to follow, the nucleotide positions are given with reference to the single base alteration in the P-w^{hd} target site duplication. They may be converted to the coordinates of O'Hare et al. (27) by subtracting positive numbers from −2028 and subtracting negative numbers from −2029.

p15-p16, 355 bp, 827 (*Nde*II), 884 (*Nde*II). When the insertion from P[*w*alL] was copied in, the sizes of the amplified product from p11-p12 and p13-p14 were 534 and 373 bp, respectively. When the P[*w*alLKP] insertion was present, we amplified each end of the insert independently by using p11-p17, 606 bp, KP 3′ end; and p14′-p18, 286 bp, KP 5′ end. For the P[*w*al*Ly*] insert, we used p11-p19, 449 bp, *yellow* gene 5′ end; and p14′-p20, ≈480 bp, *yellow* gene 3′ end. The primer sequences were p1, GAGTGTCGTATTGAGTCTGAG; p2, AAGAGA TAGCGGACGCAGCG; p3, GGTTGGCGCGATCTCGCG CTCT; p4, GGCTATACCAGTGGGAGTAC; p5, CGCAGT CGGCTGATCTGTGT; p6, AGGTGTTCCCTGGCCGTT AG; p7, AGCCCACCTCCGGACTGGAC; p8, AGAAG GGTGTGGAATCAGGC; p9, CTCTTCCTGACCAACATG AC; p10, CCATTGGTTAATCAGCAGAC; p11, GGTTG TCGTACCTCTCATGG; p12, GTGTTTTATGTAACGATA AACGAG; p13, GCCGACATAAATCCGAAATAACTG; p14, ACAGCGAAAGAGCAACTACG; p14′, CGAAAGAG CAACTACGAAACG; p15, TTTGTGACCTGTTCGGAGT GA; p16, AGGACCTATTTCCGCTGCAC; p17, CCGTCAC AGCTGAGTTAATTCAAACCCCACG; p18, CGTCCGCA CACAACCTTTCC; p19, TAATAGTCGAGGATCCTGC; p20, TACGCAATTTTCTTGAGCGG; p21, GTGCAGCGG AAATAGTTAATAAC; p22, CTCCCTCAGCTGCTGCTC TA;p23, CAGTTATTTCGGATCTATGTCGGC; p24, TCGC TGTCTCACTCAGACTC; p25, CGAGATCGCGCCAACC AAGTAGCGATAGAT; p26, GTGGGTATCACAAATTTG GG; p27, TGCTGCAAAGCTGTGACTGG; p28, GCGACT GTCGTTAGTTCCGC; p29, ACGTGGGTCTAGCCATTCT CATCGTGAGCTTCCGGGTACTCGCATATCT; p30, GGT CATCCTGGAGACGC; p31, AGGTATGCAGGTGTGTAA GTC; p32, CTCGTTTATCGGTACATAAAACAC; p33,

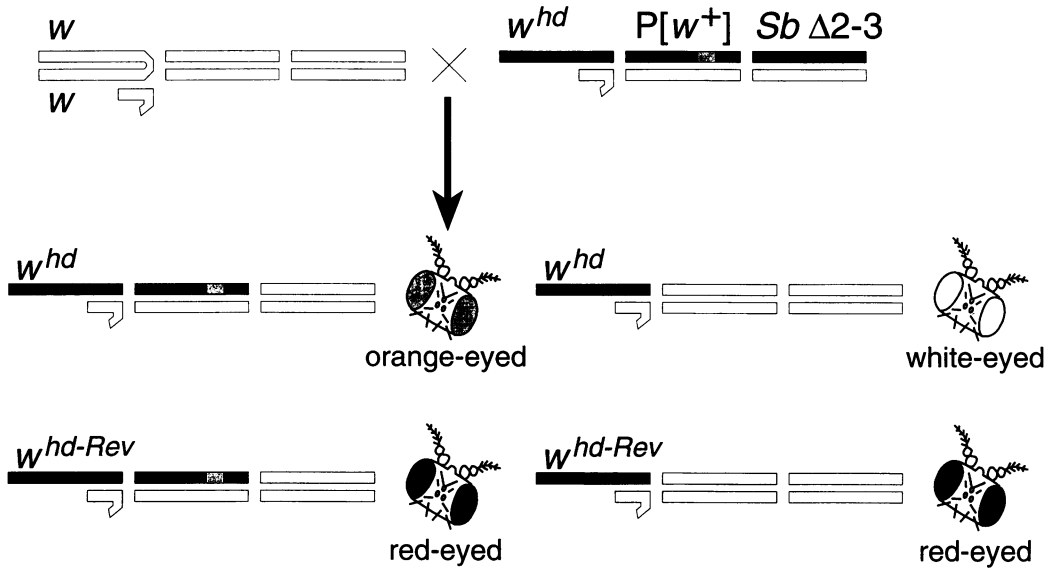

FIG. 3. Mating scheme for generation of w^{hd} revertant flies. Single male flies carrying the three P elements necessary for the gene conversion reaction were end mated to three to five compound-X virgin females, and the sons were scored for their eye phenotype. If the w^{hd} allele was not reverted, the progeny were either orange or white eyed, depending on whether they inherited the template (P[w^+]) or not. Sons with a wild-type eye color, denoted $w^{hd\text{-}Rev}$, were kept and mated to several more compound-X virgin females to establish a line. The P[w^+] template could not be phenotypically detected in these flies. Autosomal insertions of the template could be detected in the next generation as female progeny with pigmented eyes. The *Sb* chromosome was selected out to eliminate the Δ2-3(99B) transposase source. Third chromosomal insertions of the templates were used in an analogous mating scheme.

CGATCGCATCATTAGTTGCG; p34, ATGCATTCTATGC ACGAGCC; p35, CTTTACATGAGGGTGCTCTC; p36, CG ACAGTCGCTCCAAGGCAT; p37, CATCTGCCGAGCAT CTGAAC;p38, GATTAGCCAGGCTGGGCTAG; and p39, ATATGTTCTTAAGCAGTCCG). The 3′ end of p36 has 3 bases of homology to the P element 5′ end, and the primer's 5′ end has 17 bases of homology to the *white* gene sequence. It was used to amplify specifically the internal P_T 5′ end of P[dbl*w*alL*y*].

DNA sequencing. Single-stranded DNA for some sequencing reactions was prepared by using the asymmetric PCR approach (14) as described previously (13). Others were prepared by digesting one strand of the PCR product with an exonuclease prior to sequencing (2). Standard dideoxy sequencing procedures were then applied.

RESULTS

Reversion rates. We used the method described in the legend to Fig. 3 to determine the reversion rate of the w^{hd} allele in the presence of P transposase and each of the five ectopic templates described above and in the legend to Fig. 2. Table 1 shows the number of flies scored and the reversion rates associated with each template insertion. The rates ranged from 0.20% for one insertion of P[*w*alL] to 1.56% for one insertion of P[*w*alL*y*]. There was no apparent correlation between reversion rate and the size of the inserted sequence.

The average reversion rates were somewhat lower than that reported previously for the P[*w*alter] template (i.e., lacking any inserted sequence) (12). For example, the 11 insertion sites for P[*w*alL] had an average reversion rate of 0.7% as opposed to 1.2% observed previously for 26 autosomal sites of P[*w*alter]. We used a permutation test to assess the significance of this difference without relying on any assumptions of the underlying distribution (8). We computed the average rate for each of the 11-member subsets of the 36 total insertion sites. Only 17,804,841 of the total 854,992,152 subsets had a lower average than the observed set of P[*w*alL] insertions, giving us a significance level of 0.02.

Conversion tracts. We selected 293 independent revertant lines for molecular analysis as described in the legend to Fig. 4. Of these, 67 were found to have both termini of the template transposon. These were assumed to represent transposition events and were not analyzed further. There were also 29 cases of apparent precise excision of the P insert in w^{hd} (hereafter denoted P-w^{hd}), including one copy of the flanking 8-bp host duplication. These events were assumed to be template independent, since the same frequency, approximately 0.1%, was observed previously in the absence of a template (6).

The remaining 197 lines represent true conversion events and are the subject of the rest of this report. Most of the events were simple conversion tracts (Fig. 5), defined as complete loss of the P-w^{hd} element and substitution of some of the flanking sequence for the corresponding template DNA, possibly including the insertion or deletion carried by the template. Such events constituted the majority for each of the five templates we studied: 88.3% (68 of 77) for P[*w*alL], 69.8% (31 of 43) for P[*w*alLKP], 76.2% (32 of 42) for P[*w*alL*y*], 51.7% (15 of 29) for P[dbl*w*alL*y*], and 83.3% (5 of 6) for P[*w*al*Not*Δ].

Figure 5 also shows six cases of discontinuous conversion tracts. This is approximately the same proportion of discontinuous events as was reported from previous experiments (12, 25). Such events can be interpreted in terms of heteroduplex formation followed by mismatch repair.

The simple conversions can be further subdivided into two categories, those which copied only *white* gene sequence from the template and those which also included the nonhomologous site (insertion or deletion). The proportion of simple conversions which copied the nonhomologous sequence was

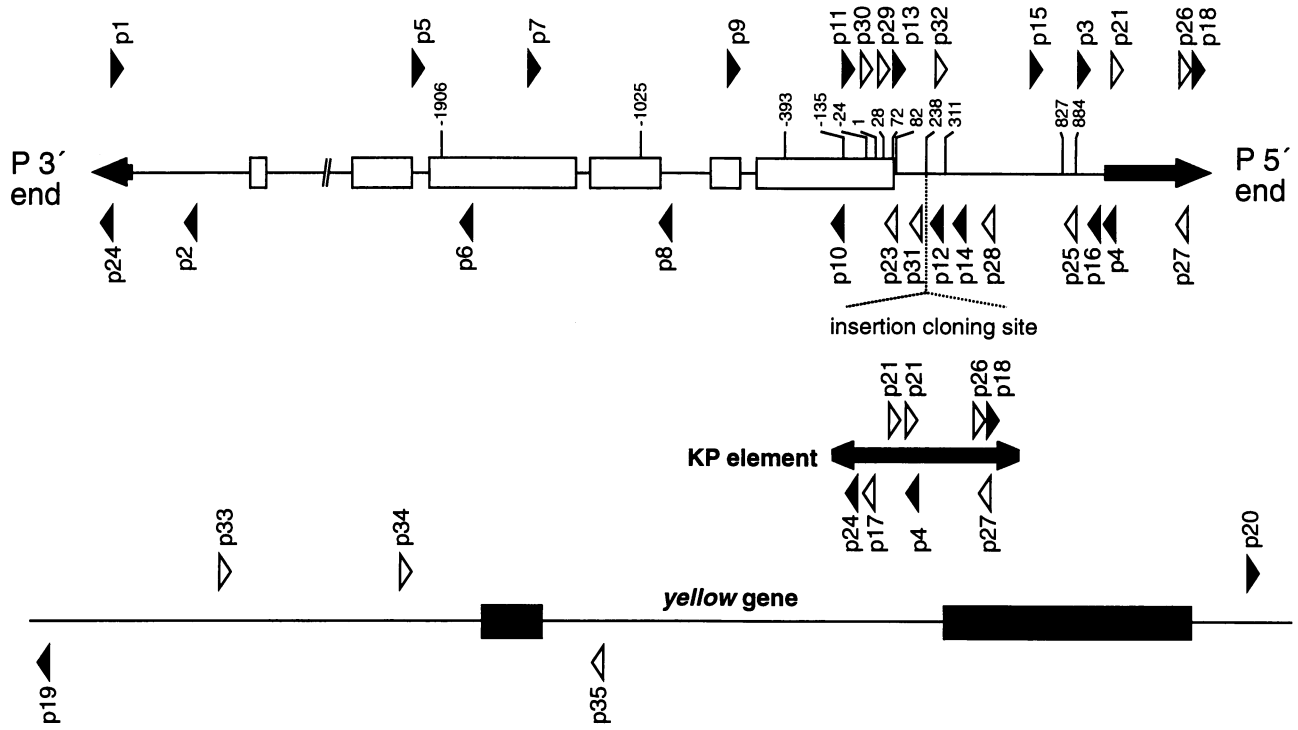

FIG. 4. Oligonucleotide primers for PCR. Black arrows represent primers that were used to analyze all the reversions from at least one of the templates, and white arrows indicate additional primers used to analyze some of the revertants—usually the complex ones. The *white* and *yellow* gene exons are shown as boxes. Primers p14 and p14′ produced amplification products that differed by only 4 bp.

67.6% (46 of 68) for P[*w*alL], 51.6% (16 of 31) for P[*w*alLKP], 100% (5 of 5) for P[*w*al*Not*Δ], 61% for P[*w*alL*y*], and 57% for P[dbl*w*alL*y*]. The last two estimates come from the proportion of w^+ revertants that were phenotypically yellow$^+$, thus avoiding possible bias from preferentially selecting w^+ y^+ revertants for molecular analysis.

Figure 6 shows the conversion frequency of each marker site plus the nonhomologous insert. The present data are compared with previous results (12, 25) obtained by using either the homolog as a template or an ectopic template, P[*w*alter], with only homologous DNA. The P[dbl*w*alL*y*] and P[*w*al*Not*Δ] were excluded because of small sample sizes. As expected, the most frequently converted marker sites were those closest to the P-w^{hd} insertion point, with the frequency falling off in both directions.

The distribution in Fig. 6 can be interpreted as representing widening of the double strand gap prior to repair. A good fit to the data can be obtained from a model in which each exonucleolytic cleavage is an independent trial (12, 25). If nonhomologous DNA gets copied into the gap as efficiently as homologous sequences, then we expect that any gap extending more than 238 bp rightward would include the nonhomologous sequence. The similarity between the insertion-bearing templates and P[*w*alter] in their conversion tract distributions (Fig. 6) is in good agreement with this interpretation. In particular, the conversion frequencies for the insertions themselves, indicated at 238 bp on the abscissa of Fig. 6, appear to fit the same curve as the base substitutions. This observation suggests that copying of even large insertions occurs at approximately the same frequency as single-base substitutions. However, the insertion templates do show slightly lower conversion frequencies than P[*w*alter] for the

TABLE 1. Reversion and conversion rates

Template line[a]	Total no. of males scored	Reversion rate (%)	No. of revertants		Conversion rate (%)
			Tested	Converted	
P[*w*alL](II)-2	182	0.55	0		
P[*w*alL](II)-3	1,048	0.29	3	2	0.19
P[*w*alL](95D)-5	8,515	0.73	27	13	0.35
P[*w*alL](II)-6	710	0.56	3	1	0.19
P[*w*alL](TM6B)-7	136	0.66	0		
P[*w*alL](45D)	7,039	1.13	19	11	0.65
P[*w*alL](43D)	5,032	0.68	13	8	0.42
P[*w*alL](61D)	9,273	1.17	24	17	0.83
P[*w*alL](II)-10	457	0.20	1	1	0.20
P[*w*alL](TM6B)-13	3,607	0.98	9	7	0.70
P[*w*alL](95D)-21	6,501	0.78	22	17	0.60
P[*w*al*Not*](45D)	7,915	0.38	15	6	0.15
P[*w*alLKP](94F/71C)[b]	15,819	1.03	15	9	0.62
P[*w*alLKP](70D)	31,783	0.42	43	30	0.29
P[*w*alLKP][c]	2,596	0.35	6	4	0.23
P[*w*alL*y*](70C)	17,412	1.56	53	42	1.24
P[dbl*w*alL*y*](86C)	12,818	1.12	40	29	0.81
Total	129,497		293	197	

[a] Template insertion sites are shown in brackets; a unique identifying number is given for those templates requiring it.
[b] This insertion is on TM3 and hybridizes to one of the breakpoints.
[c] Data for this row are pooled from several insertions of undetermined location.

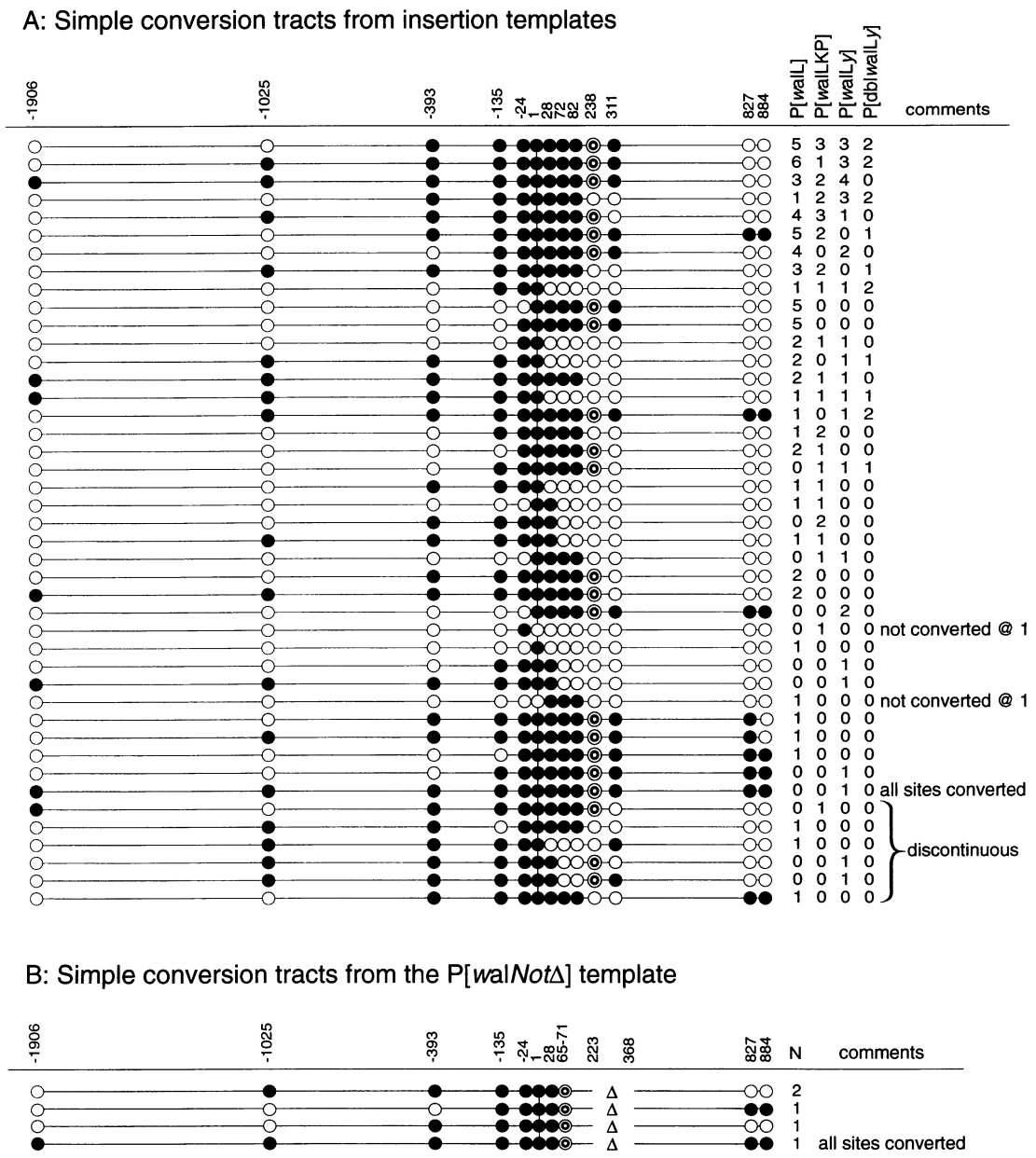

FIG. 5. Simple conversion tracts. The black circles represent the single base marker sites that were converted to the template sequence, and the target symbols indicate conversion of nonhomologous DNA sequence. (A) Target symbols represent 25, 1,124, or about 7,970 bp of DNA; (B) target symbol represents conversion of 6 bases in P[*walNot*Δ]. Δ, conversion of the deletion. The columns on the righthand side show the number of independent events recovered when each template was used. Those conversions with noteworthy characteristics are indicated.

marker site at position 311 and a markedly lower conversion rate for the rightmost two sites (positions 827 and 884 on the Fig. 6 abscissa). This suggests that there is a small but detectable decrease in repair efficiency when a nonhomologous insertion is copied in.

Complex conversion and repair events. The remaining 46 of the revertant lines included a variety of complex structures as determined by DNA sequencing and PCR analysis with various combinations of the primers in Fig. 4. These structures, shown in Fig. 7 to 10, provide further information about the mechanism of gap repair.

The most common type of complex event is shown in Fig. 7. These structures, which we call conversion-duplications, represent 12.2% of all conversions and occurred with each of the five templates. Similar events were seen in our previous experiments (12) and constituted approximately the same proportion of conversions. All conversion-duplications carry one terminus of the template transposon and have a duplication of *white* sequence from the P-w^{hd} insertion site extending rightward. The extent of their conversion tracts vary in the left end. These structures are expected as bitemplate convertants according to the synthesis-dependent strand annealing (SDSA) model, which we will discuss below.

There were also several classes of complex events that were

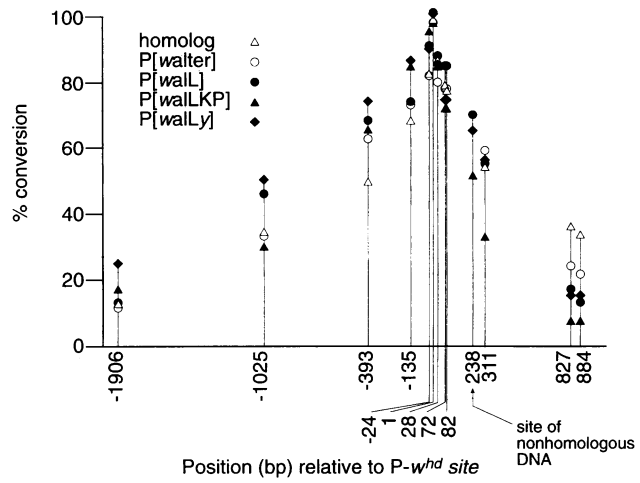

FIG. 6. Conversion frequencies. The proportions of each of the sites which have been copied into the revertant *white* genes in each of the simple conversion tracts are shown. Included here are data from previous experiments in which the template did not have any nonhomologous sequence (12, 25).

specific to the P[*w*alLKP] template (Fig. 8) and the P[dbl*w*alL*y*] template (Fig. 9). All of the structures in Fig. 8 are also predicted on the basis of the SDSA model (see below).

Finally, there were seven structures (Fig. 10) which did not fall into any of these categories. Only one case of each of these structures was observed. Several of them were similar to the more common categories except for the presence of a deletion (Fig. 10A to C), an insertion (Fig. 10D and E), or an uncharacterized alteration (Fig. 10F). One of them (Fig. 10G) had an apparently template-independent base substitution, similar to a case observed previously (12).

DISCUSSION

Insertions and deletions can be copied with high frequency. These results show that even large stretches of nonhomologous DNA can be copied into an arbitrary genomic site near a P-element insertion. We have targeted the insertion of four different nonhomologous sequences into a site more than 200 bp from the P-w^{hd} site: a 25-bp insertion, a 1,124-bp insertion, an ~7,970-bp insertion, and a 136-bp deletion in conjunction with a 6-base alteration. These events are accompanied by loss of the P element and replacement of various lengths of the flanking sequence with information from the ectopic template. In each case, the frequency was approximately the same as would be expected for a single-base substitution at the corresponding position. Our results indicate that most conversions copying even the largest insertion are copied completely. This suggests that insertions significantly larger than those used here may also be efficiently converted, although the upper limit is not known.

The ability to copy arbitrary sequences into genomic sites has several potential applications. Approximately 1% of all progeny receiving the targeted chromosome were found to carry the insertion at the expected site. This frequency is sufficiently great that a PCR-based screen, as opposed to the phenotypic one used here, would be adequate to recover the events. Such a screen has already been successfully completed. McCall and Bender have used PCR to screen excisions of a P[*rosy*$^+$] element to target an insertion to Ubx (24). One potential use for this technique could be the replacement or modification of sequences controlling gene expression in *D. melanogaster*. Another would be the placement of a gene into the control region of another, resulting in ectopic expression of the inserted gene in the tissue and time of development specified by nearby regulatory sequences. The tendency of P elements to insert in or near the upstream regulatory regions of genes (20) is likely to provide many opportunities for experiments of this kind.

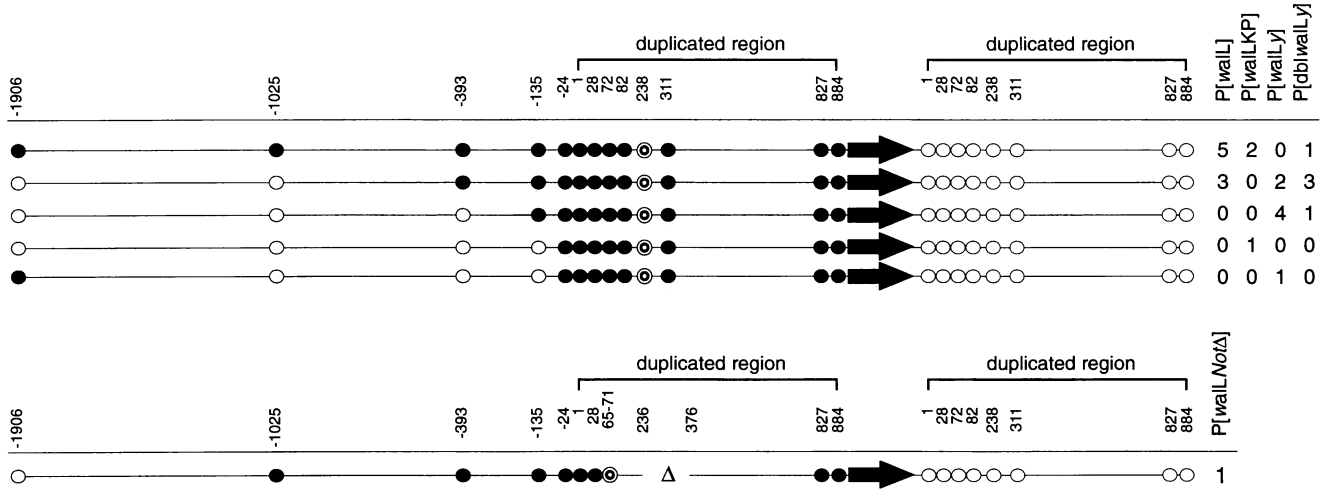

FIG. 7. Conversion-duplications. Events of this type had the P_T 5′ end inserted at the site of the 3′ end of the P-w^{hd} element. This resulted in a duplication of all the *white* gene sequence as shown. Each of these structures was confirmed by amplification with PCR primers p3 and p12 or an equivalent combination (p23-p26 or p23-p32). We determined the conversion tracts at the upstream nucleotide locations 238, 311, 827, and 884 relative to the P_T 5′ end by amplification with p4 in combination with p15 or p13. The extent of the conversion tracts at the downstream nucleotide sites 1, 28, 72, 82, 238, and 311 was determined by amplification with primer p14 or p14′ in combination with p18 or p26. These combinations promoted amplification out from the P-element end. The conversion status of sites that could not be directly analyzed was inferred from amplifications expected to yield both the converted and the unconverted patterns. The P_T 5′ end of the template is shown as the heavy line and arrow; other symbols are the same as for Fig. 5.

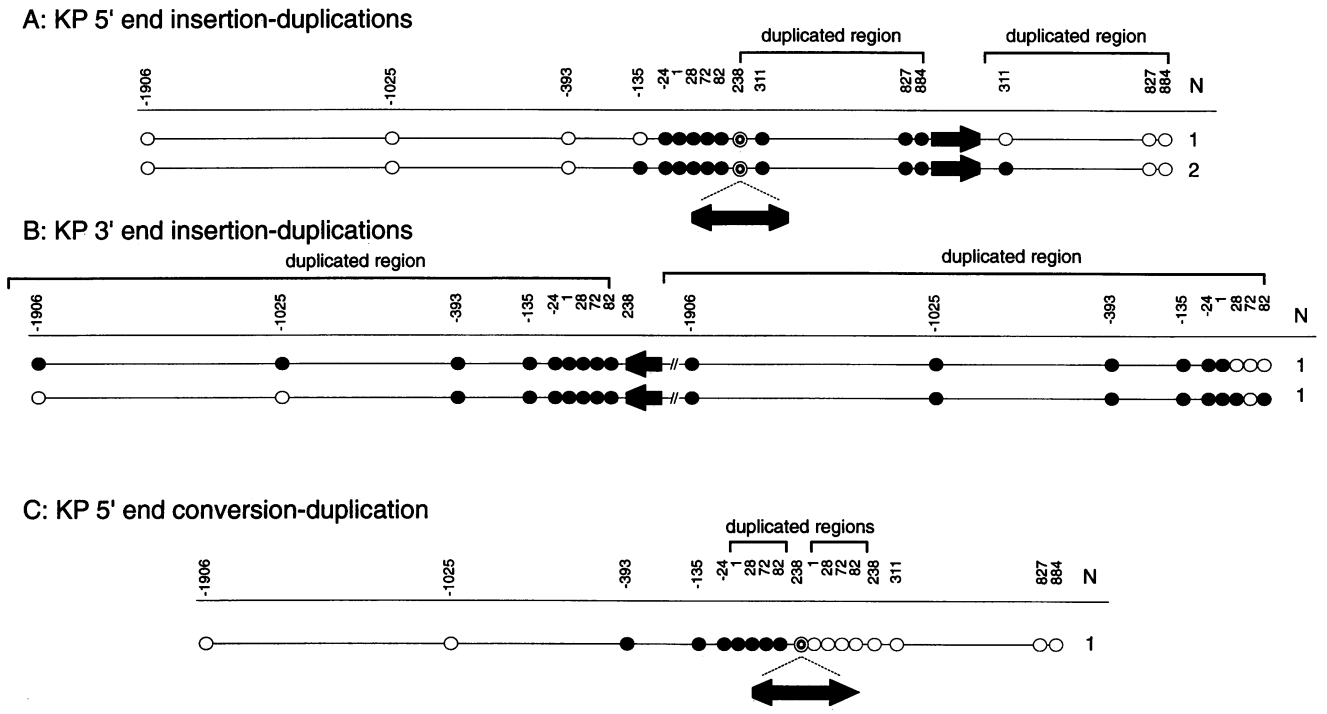

FIG. 8. Duplications involving the KP ends. The conversions in A and B had the P_T 5′ or P_T 3′ ends inserted into the site occupied by the KP end in the template. However, as indicated by the blunt arrowheads, the P-element terminus was truncated to resemble that of the KP insert. The conversion in C had the KP 5′ end inserted into the P-w^{hd} site, similar to a conversion-duplication. In this instance, the KP 5′ end was restored to a complete P terminus, as indicated by a pointed arrowhead. Primer combinations used to determine the P 5′ end structures and the associated conversion tracts were p21-p25, p3-p12, p28-p14′, p21-p12, p3-p14′, and p18-p27. For the P 3′ end structures and their conversion tracts, we used p13-p2, p29-p2, and p17-p13. Primers used to determine the structure of the single KP insertion into the P-w^{hd} excision site were p30-p31 and p23-p26. Presence of the polylinker sequence adjacent to the KP ends in A was ascertained by amplification with p18-p14′ (amplifies both P 5′ ends), p22-p14′ (amplifies only KP 5′ end), and p21-p14′ (amplifies P_T 5′ end only). The P 5′ end in C was amplified with p18-p23. The amplified fragments were digested with either *Eco*RI or *Pst*I, which cut within the polylinker. We distinguished between intact P ends (pointed arrowheads) and truncated P ends (blunt arrowheads, presumably derived from the KP fragment) by amplifying across each end sequence and digesting the product fragment with *Nla*III, which recognizes the terminal 4 bp of the P sequence. The amplifications were performed with primers p22-p14′ or p21-p14′ for the 5′ end in structure A, p18-p23 for the 5′ end in structure C, and p13-p24 for the 3′ end in each case. The resulting structures of the KP elements represented by the target symbol are shown below the symbols. All sites not shown to the right of the site at 82 in B were unconverted.

Fidelity of conversion. Several lines of evidence suggest that P-induced gap repair occurs without frequent misincorporation. In previous work (12), we sequenced more than 3,500 bp in and near the conversion tracts without finding any template-independent changes. The present data provide further evidence on this score. There were 44 full-length *yellow* genes copied into the *white* locus, and each retained its yellow$^+$ expression. This observation argues for a high fidelity of repair, since our selection was for the white$^+$ phenotype and not for *yellow* gene expression. These 44 *yellow* genes represent more than 350,000 bp, of which approximately 87,500 bp are coding sequence. Although some misincorporations would be tolerated without loss of *yellow* expression, this result is not consistent with a high rate of misincorporations.

The high fidelity of conversion we observe contrasts with the results of Banga and Boyd (1), who found that one of only two conversion tracts they examined had a misincorporation very close to the P-w^{hd} site. Both conversion events utilized oligonucleotide templates rather than a chromosomal sequence. It is possible that the repair process is more error prone when an oligonucleotide serves as the template.

Gap repair models. Double strand gap repair is thought to occur in yeast cells by a mechanism involving an intermediate structure with two Holliday junctions, as shown in Fig. 11A (29, 39). From our previous work, we suggested that the mechanism for P-induced gap repair in *D. melanogaster* might be analogous, except that the resolution of the double Holliday structure would have to be such that crossing over is greatly reduced, and transfer of information from the broken duplex to the template occurs rarely, if at all. Hastings (15) suggests one such resolution mechanism. However, the present data, especially the complex conversion events, are more consistent with an entirely different kind of model in which no double Holliday structure is formed.

To better explain the data, we propose an SDSA model in which the mechanism for DNA synthesis is similar to the one studied by Formosa and Alberts (10) for end-dependent DNA replication in late stages of bacteriophage T4-infected *E. coli* cells. The SDSA model also shares features with the single strand annealing model proposed by Sugawara and Haber to explain how deletions are formed when a double strand break is introduced between two directly repeat sequences in the yeast *Saccharomyces cerevisiae* (38), with a model proposed by Lin et al. (21) to explain how DNA transfected into mouse L cells is concatenated, and with a model proposed by Strathern to explain the lack of recombination associated with mating

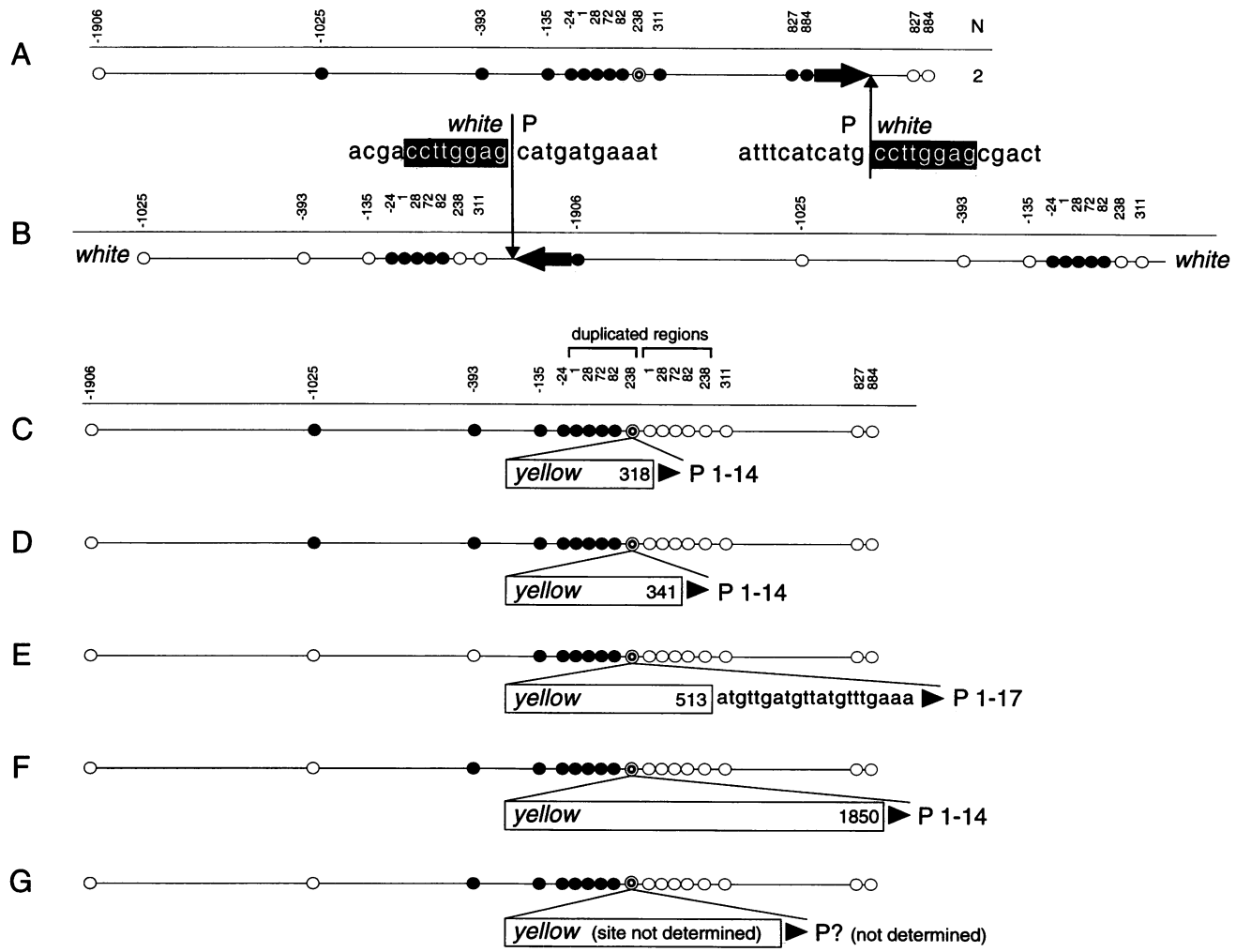

FIG. 9. Conversions from the P[dblwalLy] template. Two conversion tracts (A) had the P_T 5′ end inserted 465 bp 3′ to the P-w^{hd} excision site, and one (B) had the P_T 3′ end inserted at the same site. The structures of the two P_T 5′ insertions were confirmed by DNA sequencing, and the P-*white* junction in each was found to be the same as in the template. The P_T 3′ sequence was also confirmed by sequencing. The deletion in intron 1, which originated with the P[*CaSpeR*] construct, was present in the right copy of *white* in structure B but not the left copy, as indicated by amplifications with primers p37-p6 and p38-p6. Primers p11 and p17 amplify the *white* gene in the left copy but not the right copy in structure B, allowing us to determine its conversion tract. We inferred the same pattern for the right copy, because primer pair p11-p12, which is expected to amplify both copies, yielded the same pattern. The five conversions shown in panels C to G had part of the 3′ end of the *yellow* gene deleted. The deletion endpoints were determined by DNA sequencing for four of them (C to F) and are indicated according to the published numbering scheme (4). Each was truncated at a different site of the *yellow* gene and contained at least 14 bp of the P-terminal inverted repeat between the *yellow* and *white* sequences. In some cases the T nucleotide at the junction could have come from either the *yellow* or P sequence. Primers used to amplify the novel junctions were p23-p33 for structures C, D, and E and p23-p39 for structure F. These and other combinations failed to yield amplifications across the right end of *yellow* in structure G.

type switching in yeasts (37). While the majority of our events can be explained by existing models for double strand break repair, none of them can account for the bitemplate conversions as readily as the SDSA model.

In the SDSA model, repair is initiated by each terminus finding and invading a homologous sequence. The ends may have been degraded prior to this invasion. In some cases, the two ends will locate different homologous sites, as shown in Fig. 11B. DNA synthesis proceeds independently from each 3′ end, with the newly synthesized DNA being immediately displaced from its template. Complete displacement of the newly synthesized DNA must occur for the next step in the reaction. This process eventually produces single-stranded "tails" on the strands with overlapping sequences. The two single strands combine in a region of overlap, and any nonhomologous sequences beyond the overlap are removed. Analogous removal of nonhomologous overhang in *S. cerevisiae* has been found (9). DNA synthesis then continues, with each newly made strand serving as the template for the other until only double-stranded DNA remains.

The primary advantage of this model is that it predicts all the complex conversion events shown in Fig. 7 and 8. Each such event comes about by an appropriate choice for the templates (one for each terminus) and the region of overlap, as shown in Table 2. Previously, we attempted to explain the conversion-duplications (Fig. 7) in terms of either P element internal breakage or a two-step process (12), neither of which is likely to result in the high frequencies observed for this class.

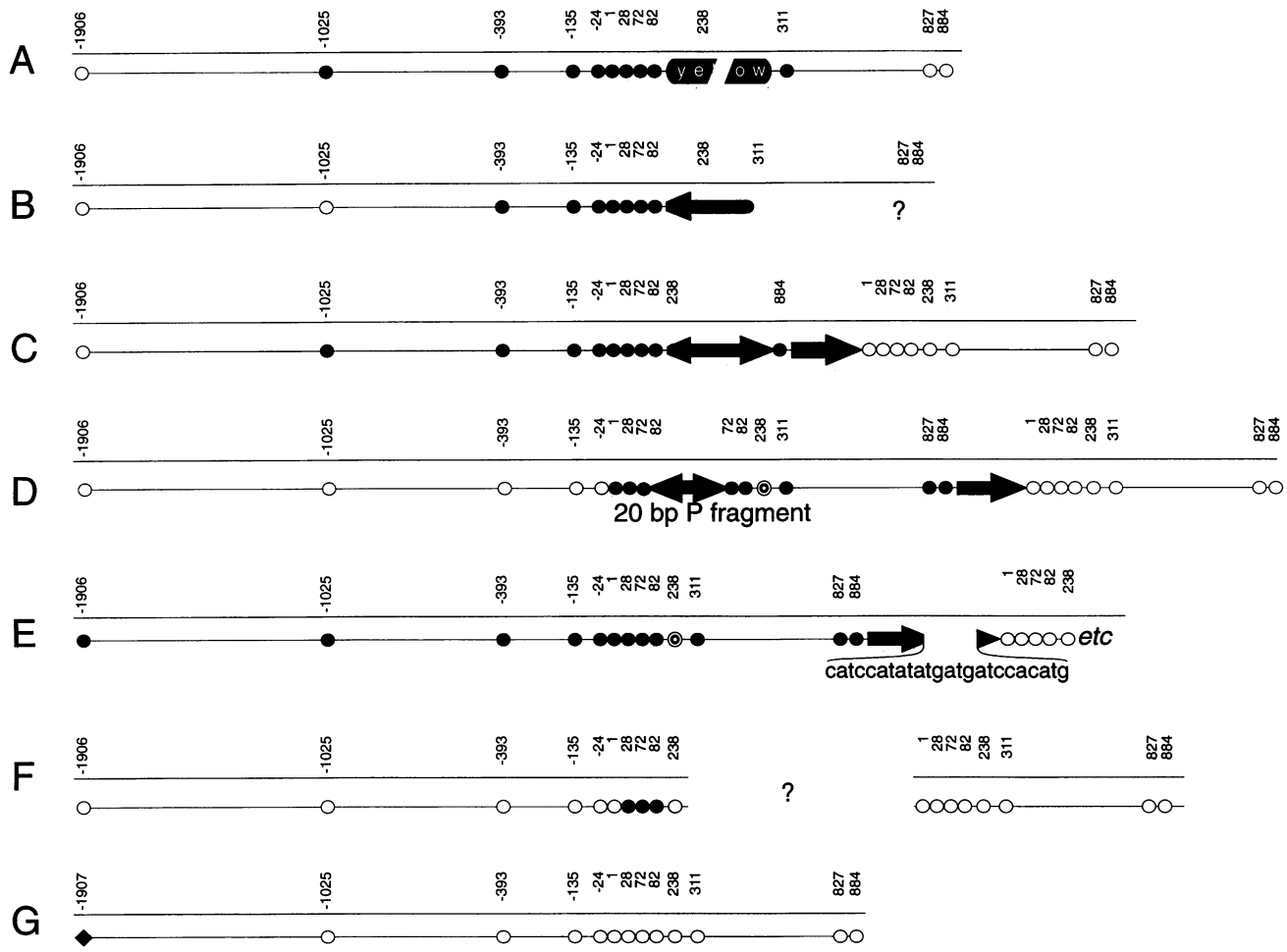

FIG. 10. Exceptional conversion tracts. Seven conversion tracts did not fit into any of the previous categories. (A) From template P[*w*alL*y*]. The structure was identical to the class of simple conversions which included the nonhomologous insertions, except for a deletion of approximately 1,035 bp within the *yellow* gene, as indicated by amplification with primers p33 and p35 but not p34 and p35. (B) From P[*w*alLKP]. It did not amplify with any of the primer combinations used to amplify sequences downstream of the insertion, and the status of the KP 5′ end is unknown. (C) From P[*w*alLKP]. This structure differs from a conversion-duplication by having a complete KP 5′ end and an extra copy of P_T 5′, as was confirmed by DNA sequencing. (D) From P[*w*alL*y*]. Amplification with primers p11-p12 gave a larger fragment than usual. DNA sequencing revealed the presence of an internally deleted P element which had an 8-bp target site duplication and 13 bp of one P end and 7 bp of the other. This structure could occur by a transposition of a P element and then by an internal deletion. (E) From template P[*w*alL*y*]. Sequencing revealed 24 bp of non-P and non-*white* sequence between two partial P-element ends. The lefthand P end was missing the last nucleotide of the 31-bp repeat, and the other end had the final 16 nucleotides of the repeat. (F) From P[*w*alLKP]. This conversion had a duplication as shown. PCR with various primer pairs failed to amplify from one *white* sequence to another, leading us to conclude that there was a large insertion or an inversion between the two *white* sequences. However, cytological examination showed no gross rearrangement, and in situ hybridization to polytene chromosomes (7) revealed hybridization to the *white* sequence only at cytological position 3C, the normal *white* locus. (G) From P[*w*alL]. DNA sequencing revealed a single G-T transversion at site −1907 that was not derived from the template or the w^{hd} allele.

However, under the SDSA model, such structures should occur normally from bitemplate conversion in which one terminus invades the sister strand and the other invades the ectopic template (Table 2). Moreover, the events shown in Fig. 8A and B cannot be explained by P element internal breakage, but they should occur readily, according to the SDSA model. The structures derived from the P[dbl*w*alL*y*] template (Fig. 9) are not easily explained by the SDSA model. However, this is not a major disadvantage, since the complex nature of the P[dbl*w*alL*y*] template makes predictions difficult under any model. Similarly, the events in Fig. 10 are each unique and could have come about by a secondary step occurring prior to, or following, the repair, such as a deletion or a P-element transposition.

Two additional structures predicted to be common according to the SDSA model can be obtained by reversing the sister strand and ectopic template in the second and last rows of Table 2. We would not expect to observe these structures in our screen because neither results in a fully wild-type *white* gene.

According to the SDSA model, crossing over of flanking genetic markers is not normally associated with gap repair. Therefore, this model explains the lack of correlation observed between P-induced gap repair and genetic recombination (6, 18, 25). In addition, the SDSA model accounts for the lack of "backwards conversion," defined as the transfer of sequence information from the broken strand to the template. Finally, our observations of discontinuous conversion tracts can be

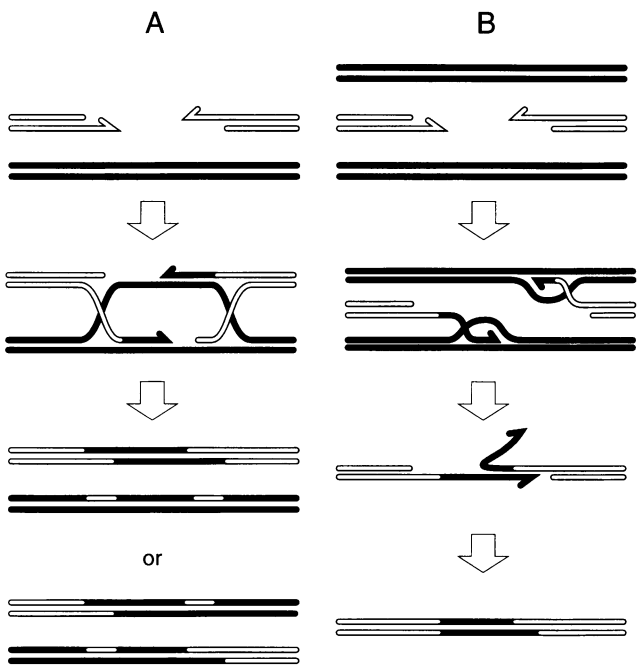

FIG. 11. Two models for double strand break repair. In both these models the P element makes a double-stranded DNA break which is enlarged (12). The models here have the 5′ ends degraded more quickly than the 3′ ends (arrows). (A) model for double strand break repair in yeasts (29, 39). The broken DNA ends invade a template and serve as primers for DNA synthesis, and the newly synthesized DNA is ligated to the parental strands. This results in a double Holliday junction. Resolution of this structure by cleavage results in a gene conversion either with or without recombination as shown. (B) The SDSA model proposes that the ends invade independently and displace only a local loop or bubble of DNA. DNA synthesis primed from the 3′ ends copies sequence from the template. In this model, as has been found for phage T4 (10), the bubble is collapsed behind the DNA polymerase by rapid displacement of the newly synthesized DNA. If the rate of displacement of the DNA is slower than or keeps pace with new synthesis, then most of the gap can be filled in before the strands are completely displaced. These strands could then anneal (38), and the gap could be completely filled by extension of the annealed ends. If synthesis of one strand on the template proceeds as far as the opposite gapped end, annealing will result in a heteroduplex. Pairing at regions of homology which result in nonpaired sequences will result in a free single strand tail. The unpaired sequence may be removed by the *Drosophila* homolog of the yeast RAD1 gene product (9).

explained if we assume that the 5′ end of the break is degraded more and the newly synthesized 3′ end is long enough to anneal with the unconverted strand (Fig. 11B). The result is a stretch of heteroduplex DNA, which can result in a discontinuous tract after mismatch repair.

Another advantage of the SDSA model is that it provides a plausible explanation for the frequent occurrence of P-element internal deletions in the presence of transposase. Most such deletions occur at internal sites of short fortuitous direct repeats of 4 or fewer bp (5, 26, 28, 41) or are found to have several base pairs of AT-rich sequence at the breakpoint (18, 40). We suggest that these deletions occur when both termini invade the sister strand but DNA synthesis from the template does not proceed far enough to produce an overlap. For example, in the Formosa and Alberts bubble migration model, the extent of synthesis is limited when active displacement of the newly synthesized DNA strand overtakes the DNA polymerase (10). The resulting short stretches of single-stranded DNA could then pair at a random site of short homology or in AT-rich bases that might have been added to the 3′ ends by cellular terminal transferase activity. The resulting partial replacement of the excised P element would appear to be an internal deletion with breakpoints similar to those observed.

Chimeric P elements revisited. A similar process can also provide an alternative explanation for the two reported chimeric P elements (16, 42). These authors described the insertion of P elements which had apparently "captured" a long stretch of flanking DNA from another part of the genome. We suggest that the chimeric elements did not actually transpose but were rather the result of aberrant gap repair events, as follows. The first step is similar to the process proposed above for the formation of internal deletions, resulting in short, nonoverlapping DNA sequences synthesized from the sister strand and matching the P-element termini. However, instead of joining with each other, one of them invades a P element lying at another site in the genome. This P element, along with a substantial amount of flanking DNA, is copied entirely. When the newly synthesized DNA is displaced, it anneals at a site that happens by chance to be homologous to a few bases of the P terminus. Trimming of non-base-paired tails and DNA synthesis to complete the gap-filling reaction would result in the production of a chimeric P element in the genome. Both chimeric P elements were seen to excise the adjacent DNA from the new insertion site (16, 42), which was interpreted to mean that the chimeric element was mobile. However, such events could also result from simple excision of just the P element and then gap enlargement and repair from the homolog as a template, thus accounting for the precise loss of

TABLE 2. Formation of conversion structures according to SDSA model

Figure	*white* sequence invaded by broken end		Homologous overlap between single strands	
	Left end	Right end	Left	Right
5 (simple conversions)	Ectopic template	Ectopic template	*white* sequence	*white* sequence
7 (conversion-duplications)	Ectopic template	Sister strand containing P-w^{hd}	P_T 5′ end (at terminal inverted repeat)	P-w^{hd} 3′ end (at terminal inverted repeat)
8A	Ectopic template	Ectopic template	P_T 5′ end (at terminal inverted repeat)	KP 5′ end (at truncated terminal repeat)
8B	Ectopic template	Ectopic template	KP 3′ end (at truncated terminal repeat)	P_T 3′ end (at terminal inverted repeat)
8C	Ectopic template	Sister strand containing P-w^{hd}	KP 5′ end (at truncated terminal repeat)	P-w^{hd} 3′ end (at terminal inverted repeat)

the captured DNA without postulating mobility of the chimeric P element.

Tests and implications of the SDSA model. Models involving the double Holliday junction intermediate differ in several testable ways from the SDSA model. One is that bitemplate conversion events, such as those we postulated for the structures in Fig. 7 and 8C, should be common under the SDSA model but not the alternatives. Such events should be detectable with suitable templates. Another prediction of the SDSA model is that the process should be independent of topoisomerase I activity (10). Such independence is unlikely if a double Holliday junction structure is required, because a large D loop must be formed.

A repair mechanism such as the SDSA model might be more advantageous than mechanisms involving a double Holliday junction for organisms with a complex genome and much dispersed repetitive DNA. It provides a robust repair process that does not lead to unequal crossing over if the template is ectopic, even if the two termini invade different templates. An important open question is whether the repair process we observe for P-induced DNA breakage is the same as that of more general double strand break repair in *Drosophila* cells.

ACKNOWLEDGMENTS

We thank Chantal Gloor for analyzing the P[*walNot*Δ] revertants, Clarann Flynn for making P[*walNot*Δ], and Chris Preston for determining the transgene insertion sites by in situ hybridization.

G.B.G. was supported by grant MT11374 from the MRC of Canada, and the work with P[*walNot*Δ] was supported by a grant to G.B.G. from the Cancer Research Society of Canada, Inc. W.R.E. was supported by NIH grant GM30948. This is paper number 3380 from the Laboratory of Genetics.

REFERENCES

1. **Banga, S. S., and J. B. Boyd.** 1992. Oligonucleotide-directed site-specific mutagenesis of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **89:**1735–1739.
2. **Bastia, D., J. Germino, J. Grosa, and J. Ram.** 1981. The nucleotide sequence surrounding the replication terminus of R6K. Proc. Natl. Acad. Sci. USA **78:**2095–2099.
3. **Black, D. M., M. S. Jackson, M. G. Kidwell, and G. A. Dover.** 1987. KP elements repress P-induced hybrid dysgenesis in *D. melanogaster*. EMBO J. **6:**4125–4135.
4. **Chia, W., G. Howes, M. Martin, Y. Meng, K. Moses, and S. Tsubota.** 1986. Molecular analysis of the *yellow* locus in Drosophila. EMBO J. **13:**3597–3605.
5. **Engels, W. R.** 1989. P elements in Drosophila, p. 437–484. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
6. **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.
7. **Engels, W. R., C. R. Preston, P. Thompson, and W. B. Eggleston.** 1986. In situ hybridization to Drosophila salivary chromosomes with biotinylated probes and alkaline phosphatase. Focus **8:**6–8.
8. **Fisher, R. A.** 1935. Design of experiments, 5th ed. Oliver and Boyd, Edinburgh.
9. **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.
10. **Formosa, T., and B. M. Alberts.** 1986. DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. Cell **47:**793–806.
11. **Geyer, P. K., and V. G. Corces.** 1987. Separate regulatory elements are responsible for the complex pattern of tissue-specific and developmental transcription of the *yellow* locus in *Drosophila melanogaster*. Genes Dev. **1:**996–1004.
12. **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.
13. **Gloor, G. B., C. R. Preston, D. M. Johnson-Schlitz, N. A. Nassif, R. W. Phillis, W. K. Benz, H. M. Robertson, and W. R. Engels.** 1993. Type I repressors of P element mobility. Genetics **135:**81–95.
14. **Gyllensten, U., and H. Erlich.** 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the *HLA-DQA* locus. Proc. Natl. Acad. Sci. USA **85:**7652–7656.
15. **Hastings, P. J.** 1988. Recombination in the eukaryotic nucleus. Bioessays **9:**61–64.
16. **Heslip, T. R., J. A. Williams, J. B. Bell, and R. B. Hodgetts.** 1992. A P element chimera containing captured genomic sequences was recovered at the *vestigial* locus in Drosophila following targeted transposition. Genetics **131:**917–927.
17. **Higuchi, R.** 1990. Recombinant PCR, p. 177–183. *In* M. A. Innis, D. H. Gelfand, J. J. Sinsky, and T. J. White (ed.), PCR protocols: a guide to methods and applications. Academic Press, Inc., New York.
18. **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.
19. **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.
20. **Kelley, M. R., S. Kidd, R. L. Berg, and M. W. Young.** 1987. Restriction of P element insertions at the *Notch* locus of *Drosophila melanogaster*. Mol. Cell. Biol. **7:**1545–1548.
21. **Lin, F., 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.
22. **Lindsley, D. L., and G. G. Zimm.** 1992. The genome of *Drosophila melanogaster*. Academic Press, Inc., New York.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **McCall, K., and W. Bender.** 1993. Personal communication.
25. **Nassif, N. A., and W. R. Engels.** 1993. DNA homology requirements for mitotic gap repair in Drosophila. Proc. Natl. Acad. Sci. USA **90:**1262–1266.
26. **O'Hare, K., A. Driver, S. McGrath, and D. M. Johnson-Schlitz.** 1992. Distribution and structure of cloned P elements from the *Drosophila melanogaster* P strain π_2. Genet. Res. **60:**33–41.
27. **O'Hare, K., C. Murphy, R. Levis, and G. Rubin.** 1984. DNA sequence of the *white* locus of *Drosophila melanogaster*. J. Mol. Biol. **180:**437–455.
28. **O'Hare, K., and G. M. Rubin.** 1983. Structure of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. Cell **34:**25–35.
29. **Orr-Weaver, T. L., and J. W. Szostak.** 1985. Fungal recombination. Microbiol. Rev. **49:**33–58.
30. **Pepling, M., and S. M. Mount.** 1990. Sequence of a cDNA from the *Drosophila melanogaster white* gene. Nucleic Acids Res. **18:**1633.
31. **Pirrotta, V.** 1988. Vectors for P-mediated transformation in Drosophila. Bio/Technology **10:**437–456.
32. **Resnick, M. A.** 1976. The repair of double-strand breaks in DNA: a model involving recombination. J. Theor. Biol. **59:**97–106.
33. **Robertson, H. M., C. R. Preston, R. W. Phillis, D. Johnson-Schlitz, W. K. Benz, and W. R. Engels.** 1988. A stable genomic source of P element transposase in *Drosophila melanogaster*. Genetics **118:**461–470.
34. **Rubin, G. M., and A. C. Spradling.** 1982. Genetic transformation of Drosophila with transposable element vectors. Science **218:**348–353.
35. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
36. **Spradling, A. C., and G. M. Rubin.** 1982. Transposition of cloned

P elements into Drosophila germ line chromosomes. Science **218:**341–347.

37. **Strathern, J. N.** 1988. Control and execution of homothallic switching in *Saccharomyces cerevisiae*, p. 445–464. *In* R. Kucherlapati and G. R. Smith (ed.), Genetic recombination. American Society for Microbiology, Washington, D.C.
38. **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.
39. **Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl.** 1983. The double-strand-break repair model for recombination. Cell **33:**25–35.
40. **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.
41. **Tsubota, S., and P. Schedl.** 1986. Hybrid dysgenesis-induced revertants of insertions at the 5′ end of the *rudimentary* gene in *Drosophila melanogaster*: transposon-induced control mutations. Genetics **114:**165–182.
42. **Tsubota, S. I., and H. Dang-Vu.** 1991. Capture of flanking DNA by a P element in *Drosophila melanogaster*: creation of a transposable element. Proc. Natl. Acad. Sci. USA **88:**693.