P-Element-Induced Male Recombination and Gene Conversion in Drosophila

Christine R. Preston and William R. Engels

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

Manuscript received June 14, 1996
Accepted for publication September 11, 1996

ABSTRACT

A P-element insertion flanked by 13 restriction fragment length polymorphism (RFLP) marker sites was used to examine male recombination and gene conversion at an autosomal site. The great majority of crossovers on chromosome arm 2R occurred within the 4-kb region containing the P element and RFLP sites. Of the 128 recombinants analyzed, approximately two-thirds carried duplications or deletions flanking the P element. These rearrangements are described in more detail in the accompanying report. In a parallel experiment, we examined 91 gene conversion tracts resulting from excision of the same autosomal P element. We found the average tract length was 1463 bp, which is essentially the same as found previously at the white locus. The distribution of conversion tract endpoints was indistinguishable from the distribution of crossover points among the nonrearranged male recombinants. Most recombination events can be explained by the "hybrid element insertion" model, but, for those lacking a duplication or deletion, a second step involving double-strand gap repair must be postulated to explain the distribution of crossover points.

THE discovery of male recombination in certain hybrids between wild-caught male and laboratory strain female Drosophila was the first indication of the existence of the Pfamily of transposable elements (HIRAIZUMI 1971). In more than 20 years following this observation, P elements have been studied in great detail and have been used as ubiquitous tools for Drosophila molecular genetics (ENGELS 1989, 1996). Much is now known about how these elements transpose and how their activity is regulated. However, the mechanism by which they induce male recombination has never been resolved. When P elements are mobilized, recombination can be detected in the progeny of either sex. However, the phenomenon is usually called "male recombination" because it is most conspicuous in males where meiotic recombination is absent.

Many experiments employing classical genetic techniques and strains bearing multiple P elements in scattered locations revealed several features about the process (HIRAIZUMI et al. 1973; KIDWELL and KIDWELL 1976; KIDWELL et al. 1977; WOODRUFF and THOMPSON 1977; SVED 1978; WOODRUFF et al. 1978; ENGELS 1979; YANNOPOULOS 1979; ISACKSON et al. 1981; SINCLAIR and GRIGLIAATTI 1985; MCCARRON et al. 1989; DUTTAXOV et al. 1990): From these studies, the average frequency of male recombination in "dysgenic" hybrids was found to be ~1% per chromosome arm but varied depending on which strains were used. The distribution of crossover points within the genome was also a function of the P strain. Most P-induced male recombination was found to occur in mitotic germ cells before meiosis. There was little or no interchromosomal effect or positive interference, as is seen with normal meiotic recombination. More recently, it has been possible to study male recombination in better detail with the use of stocks in which P-element transposase is supplied from a single, nonmobilizable source (ROBERTSON et al. 1988), and a single mobile P element is present at a known site (SVED et al. 1990, 1991, 1995; MCCARRON et al. 1994; SVOBODA et al. 1995). These studies showed that the frequency of male recombination was greatly increased when a mobile P element was made homozygous, and that mobilizing P elements in somatic cells leads to somatic recombination.

A key question in these studies was whether the recombination events were occurring at the sites of mobile P elements, as opposed to random sites that can be attacked by P transposase. The experiments favoring random sites (McCarron et al. 1989, 1994; Duttaroy et al. 1990) were limited by the use of highly selective screens for recombination, raising the question of whether the recombinations recovered are actually typical of the majority of events. The work favoring localization to P sites (SVED et al. 1990, 1991, 1995; SVOBODA et al. 1995) relied on large genetic intervals and therefore could not place the recombination point with precision.

Another mystery was the possible relationship between male recombination and double-strand break repair leading to gene conversion. Strong evidence showed that P-element transposition leaves behind a stretch of homologous template sequence (ENGELS et al. 1990; GLOOR et al. 1991; JOHNSON-SCHLITZ and ENGELS 1993; NASSIF and ENGELS 1993; ENGELS et al. 1994; NAS-
sif et al. 1994). The template is often the homologue, which would seem to provide a clear opportunity for male recombination, especially through the resolution of Holliday junction intermediates. However, several observations argued against this possibility. (1) There was only a weak correlation between conversion and recombination (Engels et al. 1990; Johnson-Schlitz and Engels 1993). (2) Many recombinant chromosomes retained a copy of the original P element (Sved et al. 1995), which would have been removed had the break been repaired off the homologue. (3) The array of conversion events observed suggested an underlying gap repair mechanism that did not involve Holliday junctions (Nassif et al. 1994).

In the present work, we resolve these questions through the use of restriction site polymorphisms flanking an autosomal Pelement site. This strategy allowed us to analyze male recombinant chromosomes with molecular precision without the necessity of using strong selective measures that tend to complicate the interpretation. In addition, the conversion tract distribution was obtained based on the same restriction site markers and mobile P element. The results show that the great majority of Pinduced recombination sites do occur within 2 kb on either side of the mobile P element. We conclude that most male recombination events are not due to the random transposase cuts hypothesized by Chovnick and coworkers. Approximately two-thirds of the recombinants carried a flanking duplication or deletion, and the remaining third were simple crossovers. Among the simple recombinants, the distribution of crossover points matched the distribution of conversion endpoints associated with double-strand gap repair. The duplications and deletions are described in detail in the accompanying paper (Preston et al. 1996).

Overall, the results suggest that most male recombination events occur by an aberrant transposition process (Gray et al. 1996; Preston et al. 1996) as opposed to double-strand break repair. A minority of events might also be explained as a secondary step following double-strand break repair initiated by unligated single-strand nicks at the conversion tract endpoints.

### MATERIALS AND METHODS

**Drosophila crosses:** Flies were raised on standard cornmeal-molasses-agar medium and grown at room temperature (24° ± 1) unless otherwise noted. All genetic symbols not described in the text are in the Drosophila reference works (Lindsley and Zimm 1992; FlyBase 1996).

**DNA sequencing:** All sequencing of PCR products from recombinants was done by the dye termination method (Lee et al. 1992; Rosenthal and Charnock-Jones 1992) using ABI Prism model 373 and 377 automated sequencers according to the manufacturers instructions. Primers for the 50C region were based on the initial determination of portions of the sequence surrounding the insertion P(CaSpeR)(50C), which was done by the dideoxy method (Sambrook et al. 1989) using plasmid clones described previously (Sved et al. 1991). Sequence information in these areas was kindly provided by J. Sved, L. Blackman and C. Flores.

**PCR and restriction mapping:** PCR was performed in 14 µl volumes in Stratagene RoboCyclers on Drosophila genomic DNA samples extracted by the simple single-fly procedure (Gloor and Engels 1992). Annealing and denaturing were at 61° and 94° for ~1 min each. Extension was at 72° for between 1 and 10 min, depending on the expected length of the amplicon. The number of cycles was 29 in most cases. All amplifications used Taq DNA polymerase, but those requiring amplification of fragments longer than 3 kb also included a small amount of another thermostable DNA polymerase that has 3 exonuclease, thus enabling "long PCR" (Barnes 1994).

Amplified DNA was cut with restriction enzymes without further purification. Each PCR tube received 5–8 units of enzyme and sufficient 10X buffer to bring the final concentration to 1X. These components were added under the oil to the aqueous part, and the mixture was incubated for 90 min at the optimum reaction temperature for each enzyme. Samples were then electrophoresed on gels of agarose, Metaphor (FMC), or a mixture of the two, at a concentration determined by the expected size of the DNA fragments.

Most reactions utilized the primers shown in Figure 1A, whose sequences are as follows: D5 CTAACGAGCAAAATACGGC; G4 CTATTGCACAGTGATTTAGC; D4 CTTGCTGCTGAGAATT; G3 TCTTATCCGCCACAGTGAC; D5 GTCAGCAGCAACTTTTC; G0 GCCATCTCAGAATTTGTT; 2223 CTGCCGGAACACACCTCTTCC; 2231 TCAGCTGCTACCTGACGTC; G2 ATTCGATTTGCGGCGGATACG; D0 GCAACCGAACTGATCGC; D2 AGCTGCATCCTCGTGC; G1 TCAATTTGCTGCAAGTGTCG; D1 AATCAGACGTCCTCCATGTCG. For most amplifications, primer G4d GCAGTGATTGTTGCTGTTAG, which is located very close to G4, was used in place of G4. Some PCRs also included a second pair of primers to serve as a positive control for integrity of the genomic target DNA and PCR conditions and reagents. These primers were 8494 CGTCCGAGTAGAGATGGTCGC and 1422 GCCGAACTTGCGGTCG, which amplify a 1.0-kb fragment from the singed locus.

### RESULTS

**Development of the marker system:** We used a P(CaSpeR) insertion at cytological position 50C to study both male recombination and conversion tracts. This insert was chosen because it has been used previously in studies of male recombination (Sved et al. 1990, 1991, 1995; Svolboda et al. 1995), and because the non-autonomous P element, P(CaSpeR) carries a mini-white gene (Pirrotta 1988).

The first step was to obtain a set of closely linked markers on both sides of P(CaSpeR)(50C). These markers would serve to localize the crossover point in male recombination events and also to delineate the conversion tracts in Pinduced gene conversion. Our approach was to screen a series of distantly related Drosophila lines to find one that differed from the line carrying P(CaSpeR)(50C) at several restriction sites near the insertion point.

Primer pairs G4d-D4, G3-D3, G2-D2 and G1-D1 can be used to amplify four fragments of ~1 kb each in the vicinity of the P(CaSpeR)(50C) insert, as shown in Figure 1B. We amplified these four segments from the P(CaSpeR)(50C)-bearing stock plus 21 other Drosophila.
Recombination and Conversion

Figure 1.—Maps of the P(CaSpeR)(50C) insertion region. All map elements are drawn to scale with the exception of the CaSpeR transposon, which is 4982 bp, and the arrowheads indicating primer sites. Orientation is according to the standard Drosophila second chromosome map, with the centromere to the left and the 2R tip to the right. The orientation of the mapped 50C region with respect to the centromere was inferred from the recombinants described below. (A) Positions of primers. See MATERIALS AND METHODS for primer sequences. (B) Amplified segments used to screen for RFLPs. Fragments 4, 3 and 1 were used to characterize conversion tracts and recombinants. In most cases, fragment 2 was not used for this purpose, since the much smaller amplicons, D0-G0 or D0-2231, were used to test for the Ed site. (C) RFLP differences between chromosomes A and C. Cutting sites on chromosome A are shown above the line, and those on chromosome C are below. The map does not include sites on both A and C chromosomes. For some of the tests, we used CiaI (ATCGAT) to test for the TaqI (TCGA) difference. The feature indicated as a 300-bp insert on chromosome C could also be considered as a 300-bp deletion on the A chromosome.

Gene conversion tracts: Previous work at the white locus suggested that in the presence of transposase, P elements excise leaving behind a double-strand break. Gap expansion and repair result in sequence being copied into the excision site from a homologous template (Engels et al. 1990; Gloor et al. 1991; Johnson-Schlitz and Engels 1993; Nassif et al. 1994). The most likely templates are the sister chromatid and the homologous chromosome. Excision of P(CaSpeR)(50C) and use of the A chromosome homologue as template would result in clean loss of the CaSpeR element concomitantly with some of the RFLP markers on the C chromosome being replaced with the corresponding A markers. In the previous work, clean loss of the P insert in w^td was phenotypically detectable. However, the 50C insertion does not mutate any gene, and its excision is therefore not phenotypically detectable. In addition, the loss of the internal w^t gene contained in P(CaSpeR)(50C) is not a reliable indicator of clean loss of the element because internal deletion will yield the same result. Therefore, we used a PCR screen, as shown in Figure 2, to collect conversion events. These events are expected to occur in the premeiotic germline of the gen-0 males and recovered in the gen-1 sons. From each gen-0 cross, up to three w^a1 b Dr^+ sons were selected and allowed to mate individually with balancer females as shown in
There were 391 iso-8 gen-0 crosses that were fertile, of which 191 yielded at least one son with the w a1 b Dr+ phenotype. No amplicon provided a more precise indication of the clean loss of CuSpeR, which was eventually confirmed by loss of the NluIII site recovered in the gen-1 sons selected for w a1 b Dr+ phenotype. Transposase was eliminated in the gen-1 by using only in a previous experiment. This fragment retains the NluIII site (CATG) on both sides, but lacks the necessary sequences for mobility. Thus, it serves as a nonmobile marker to ensure that no wild-type-size DO-GO amplicon comes from the maternal allele. We then cut the DO-GO amplicon with NZuIII, whose recognition sequence, CATG, begins and ends the P-fragment of -430 bp derived from an internal excision of P{CuSpeR/(5OC) in generation R2 if no cn bw progeny appeared. Lethal chromosomes were eliminated in R3. RFLP tests to map the conversion candidates, and their gen-3 progeny were used for further tests.

We performed this screen on 369 gen-1 white-eyed al b sons and identified 176 with approximately the wild-type-size D0-G0 fragment. The rest had either a larger-than-wild-type D0-G0 amplicon, indicative of an internal CaSpeR deletion, or else no amplicon at all, as would be expected if CaSpeR remained intact. All but eight of the 176 lines with a wild-type D0-G0 amplicon had lost the NluIII cut site, indicating full loss of the CaSpeR element, and presumably conversion from the homologue. Since 36.9% of the gen-1 al b sons were white eyed, we estimate the overall frequency of conversion at 0.369(168/369) = 16.8%.

Figure 2. The same gen-1 males were then used for DNA extraction and PCR.

Those gen-1 males that yielded an approximately wild-type-size amplicon with primers D0 and G0 (145 bp, see Figure 1A) were candidates for conversion. To confirm that a clean loss of CaSpeR had taken place, we then cut the D0-G0 amplicon with NlaIII, whose recognition sequence, CATG, begins and ends the P-element sequence. Loss of the NlaIII site is therefore a strong indication of a clean CaSpeR excision and rules out the possibility of a near precise excision in which a small part of one or both P ends is retained. Such near-precise losses are known to be the most common type of P-element excision when no wild-type template is present (TSUBOTA and SCHbddel 1986; JOHNSON-SCHLITZ and ENGELS 1993). Those lines in which the NlaIII site had been removed were retained as conversion candidates, and their gen-3 progeny were used for further tests.

Figure 2.—Scheme for collecting P-induced gene conversions. Excision of P(CaSpeR)(50C) in the gen-0 male germline was recovered in the gen-1 sons selected for w al b Dr+ phenotype. Transposase was eliminated in the gen-1 by using only Dr+ sons. Selection for w phenotype provided a crude initial screen for full or partial loss of CaSpeR. Screening for the wild-type DO-GO amplicon provided a more precise indication of the clean loss of CaSpeR, which was eventually confirmed by loss of the NluIII site. The maternal second chromosome carries a P-fragment of -430 bp derived from an internal excision of P(CaSpeR)(50C) in a previous experiment. This fragment retains the NluIII site (CATG) on both sides, but lacks the necessary sequences for mobility. Thus, it serves as a nonmobile marker to ensure that no wild-type-size DO-GO amplicon comes from the maternal allele. We then cut the DO-GO amplicon with NZuIII, whose recognition sequence, CATG, begins and ends the P-fragment of -430 bp derived from an internal excision of P{CuSpeR/(5OC) in generation R2 if no cn bw progeny appeared. Lethal chromosomes were eliminated in R3. RFLP tests to map the conversion candidates, and their gen-3 progeny were used for further tests.
(50C) $x = 0.99864$
(white) $x = 0.99855$

**FIGURE 3.**—Distribution of conversion frequencies. The proportion of the 91 conversion tracts that include each of the 13 RFLP markers is indicated. Coordinates are given as distances from the insertion site of P(CaSpeR)(50C). Five of the lines were lost before the right endpoints could be tested for $Ea$. Since 73 of the other 82 lines had their right endpoints to the right of the $Ea$ site, we assumed the same proportion ($4/5$) of the untested ones would extend beyond $Ea$. The theoretical curves are $x^n$, where $n$ is the positive distance to the CaSpeR insertion point. The value of $x$ was estimated by maximum likelihood assuming each endpoint is an independent trial, as described previously (GLOOR et al. 1991).

be continuous, meaning the run of A markers was not interrupted by C markers. In addition, all conversion tracts overlapped the P(CaSpeR)(50C) insertion site. These properties are similar to those of conversion tracts obtained previously at the white locus (GLOOR et al. 1991; NASSIF and ENGELS 1993). The frequency at which each marker site was converted is shown in Figure 3.

For a quantitative comparison with previous work at the white locus, we used a model in which conversion tract endpoints are assumed to be determined by a random gap-enlargement process, as described in footnote 23 of GLOOR et al. (1991). According to this model, the distribution of conversion tracts is determined by a single parameter, $x$, defined as the probability of an exonuclease removing one base. Thus, the proportion of conversion tracts that include a site $n$ nucleotides away from the $P$ insertion site is $x^n$. Using a maximum likelihood procedure as described (GLOOR et al. 1991), we estimate $x = 0.998635 \pm 0.00011$ from the present data. The corresponding distribution is indicated by the solid curve in Figure 3. This value corresponds to an average conversion tract length of $2x/(1 - x) = 1463$ bp. This result is in remarkably good agreement with the analogous work at the white locus, where $x$ was estimated to be 0.99855 (GLOOR et al. 1991), and the average tract length was 1379. The dashed curve in Figure 3 shows the theoretical distribution for the white locus data.

**Male recombination screen:** The scheme in Figure 4A was used to generate a collection of male recombination lines in which the A and C RFLP markers could be used to locate the crossover point. Note that recombination occurred in the germline of the gen-0 males, which are identical to the gen-0 males of Figure 2 that were used to obtain conversion events. The only phenotypic selection was for recombination between cinnabar and brown, which are on opposite ends of chromosome arm 2R (FLYBASE 1996). The frequency of male recombination was found to be 1.05% among 23,188 flies scored. These were equally distributed among the two reciprocal classes, $cn$ $bw^+$ and $cn^+$ $bw$. The frequency of male recombination decreases when the rearing temperature of the gen-0 males is reduced to 19° (PRESTON et al. 1996).

Some of the 243 recombinant males from generation gen-1 were members of clusters of recombinants from single gen-0 crosses. By taking no more than one of each recombinant class from each cross, we obtained 128 lines for further study. Of these, 26 were homozygous lethal and 17 were sterile. These lines were maintained as heterozygous stocks over the balancer chromosome, CyO. Complementation tests reported elsewhere (PRESTON et al. 1996) showed that most of these were allelic to one or a combination of three essential loci, one female sterility locus and two male sterility loci in the 50CD region. PCR analysis showed that most of the lethal or sterile effects
Select recombinants

Select Cy cn bw+ DrC or gen-2 Cy cn+ bw Dr+ males and females. Cross them to test for lethality and make homozygous stocks.

B

FIGURE 4.—P-induced male recombination. (A) Mating scheme. Recombination events occur in the germline of the gen-0 males and are recovered among gen-1 progeny. All 677 fertile gen-0 crosses were iso-δ. Wavy lines indicate balancer chromosomes. White and hatched boxes represent the A and C RFLP markers, as in Figure 2. Question marks on the recombinant chromosomes signify the possibility that all or part of CaSpeR remains at the recombination site. The balancer Cy0 carries amorphic alleles of both cn and bw (Craymer 1980) to facilitate recombinant screening. The gen-0 males were derived as in Figure 2. (B) Flow chart showing the classification of gen-1 progeny and recombinant chromosomes.

can be attributed to structural changes in the vicinity of the crossover point (Preston et al. 1996). The 85 recombinants that were not lethal or sterile were kept as homozygous stocks. Figure 4B shows how the recombinants were classified.

**Crossover points usually lie in the vicinity of CaSpeR:**
Crossover points were determined using the RFLP markers in Figure 1C. Amplified fragments 4, 3 and 1 (Figure 1B) were cut with the restriction enzymes of Figure 1C to determine the pattern, A or C, of each. This was done with homozygotes for the nonlethal lines. For the lethals, it was done in heterozygotes in two ways: recombinant/A and recombinant/(=. We could then deduce the pattern of the lethal by subtracting the bands of the homologue. In addition, those whose crossover points lay between the MspI marker and the 300-bp size difference were also tested for the Earl RFLP using amplicon D0-G0 or D0-2231, depending on whether CaSpeR was present.

The crossover point was considered to be within the 4-kb interval surrounding P(CaSpeR(50C) if the RFLP markers were C on the left side of the interval and A on the right for cn bw+ recombinant, and the reverse for cn bw recombinants. Of the 128 recombinants, we found that 110 followed this rule (81 + 29 in Figure 4B), implying that 86% of the recombinants had crossover points within the 4-kb interval. Since this interval is only 0.02% of the total distance between cn and bw,
these results show that P-induced male recombination occurs primarily at the site of a mobile P element.

The remaining 14% of the recombinants that occurred outside the 4-kb interval can be explained as second-step events following a transposition of CaSpeR to a site elsewhere on 2R. All but two of the 18 recombinants outside the 4-kb interval were in the structurally simple category (see below), although they might have had structural changes at the crossover point outside our analyzed interval. They were approximately equally divided between the proximal and distal directions relative to 50C, with 10 being proximal and eight distal. They were also about equally divided between cn bk+ and cn+ bk- recombinant types (eight vs. 10).

There were 12 cases in which a single gen-0 cross produced both cn bk+ and cn+ bk- recombinants among the gen-1 sons. Only one of each type was kept for analysis, thus accounting for 24 of the 128 recombinant lines. Eight of the pairs included deletions and/or duplications, and are discussed in detail elsewhere (Preston et al. 1996). In the other four we found two pairs in which the crossover points were in different RFLP intervals, suggesting that they came from independent events. The remaining two pairs had crossover points in the MspI-EarI interval, which also contains the CaSpeR insertion site, and is the most frequent site of recombination (Figure 5A). These two cases could be explained either as reciprocal products of a single event, or from two independent recombination events in the MspI-Ead interval.

Structural classification of recombinants: Each of the 128 recombinant chromosomes was classified as either "simple" or rearranged according to a series of tests including PCR, restriction mapping, and DNA sequencing. The simple classification implies that no deletion, duplication, or other rearrangement was detected. Specifically, the following tests were done in various combinations depending on the structure of each recombinant.

(1) PCR fragments 4, 3 and 1 were amplified and in some cases fragment 2 as well (see Figure 1B). Classification as simple required the amplicon to be present and of the normal size for each fragment.

(2) Each RFLP site analyzed as described above must show either the A or C pattern for the recombinant to be categorized as simple. In some cases, both patterns appeared indicating a duplication. For the homozygous lethal chromosomes, where it was necessary to examine
both kinds of heterozygotes as mentioned above, there
were some cases in which we observed only A bands in
the recombinant/A heterozygote and only C bands in
recombinant/C. In those cases we assumed that the
fragment from the recombinant chromosome failed to
amplify, suggesting a deletion or other rearrangement.
(3) PCR using primer combinations G0/2223 and
D0/2231 was done to determine whether each CaSpeR
boundary was still present and flanked by the same
genomic sequences. The simple classification required
that both CaSpeR ends be present, or that both be ab-
sent. These two results correspond to the presence or
complete loss of CaSpeR.
(4) PCR with primer combination D0/G0 was done
to determine whether the wild-type CaSpeR-free frag-
ment was present and of the expected size. If the two
boundary fragments described in the previous test were
missing, then the simple classification required the D0/
G0 fragment to be amplified and of the normal size
(145 bp).
(5) DNA sequencing of the boundary fragments
(G0/2223 and D0/2231) and/or the empty fragment
(D0/G0) was carried out for a subset of the recombi-
nants. Only those with the expected A or C sequence
were classified as simple.
By these criteria we classified 45 of the chromosomes
as simple and 83 as structurally rearranged. Twelve of
the simple recombinants were confirmed by sequencing
as described in step 5 above, and the rest were inferred
from PCR fragment sizes. In the remainder of this re-
port, we will focus on the simple recombinants. The
rearranged cases will be described in detail in the ac-
companying paper (Preston et al. 1996).
We found that 10 of the simple recombinants in the
4-kb interval had both P ends, as indicated by PCR am-
plication of both boundary fragments and lack of the
expected size empty fragment (primers D0/G0). There
were two additional cases in which a CaSpeRinternal
deletion prevented amplification of one of the boundary
fragments. In one of these, the D0/G0 combination
produced a larger-than-normal fragment. Further tests,
especially digestion with NalII, showed that both ends
of the P element were still present. Therefore, this re-
combinant was placed in the structurally simple cate-
gory, and the other was assumed to be similar. The
remaining 17 simple recombinants failed to amplify either
boundary segment but did yield the wild-type size for
the CaSpeR-empty fragment. Thus CaSpeR was cleanly
lost from the 50C site of those chromosomes. For those
with crossover points outside the CaSpeR RFLP interval,
the presence/absence of CaSpeR was mostly as expected.
That is, recombinants in which the MspI–EarI interval
carried the C markers also had the CaSpeR element (Fig-
ure 5A). The one exception to this rule can be explained
as a subsequent excision of CaSpeR.

Comparison of crossover points and conversion end-
points: The crossover point for each simple recombi-

<table>
<thead>
<tr>
<th>Interval</th>
<th>Crossover</th>
<th>Conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>NalII–ThaI</td>
<td>1 (3.3)</td>
<td>6 (3.6)</td>
</tr>
<tr>
<td>ThaI–TaqI</td>
<td>1 (3.3)</td>
<td>5 (3.0)</td>
</tr>
<tr>
<td>TaqI–MspI</td>
<td>7 (23.3)</td>
<td>22 (13.0)</td>
</tr>
<tr>
<td>MspI–EarI</td>
<td>15 (43.3)</td>
<td>70 (41.4)</td>
</tr>
<tr>
<td>EarI–gap</td>
<td>4 (14.3)</td>
<td>60 (35.5)</td>
</tr>
<tr>
<td>gap–Sau3AI</td>
<td>4 (13.3)</td>
<td>6 (3.6)</td>
</tr>
</tbody>
</table>

Data pooled from Figure 5. The number (and percentage
in parentheses) of conversion points and conversion tract
endpoints in each interval is indicated. These numbers were
treated as a 2 × 6 contingency table in the test for inde-
pendence.

nant is mapped in Figure 5A. Note that many of the
crossover points lie within the 4-kb region but are not
within the MspI–EarI interval that includes the CaSpeR
insertion site. The density of crossover points is greatest
in the CaSpeR-containing interval and decreases as one
moves away in either direction. Indeed, the distribution
is similar to that of the conversion tract endpoints, as
shown in Figure 5B.

To test this similarity, we divided the 50C region into
six intervals and totalled the number of crossover points
and conversion breakpoints in each interval. Table 1
shows the counts. The chromosomes not tested with
EarI (Figure 5) were excluded. The 2 × N generaliza-
tion of Fisher’s Exact Test for independence yielded
P = 0.08. This test was performed by computing the
probability and likelihood ratio statistic of all 129,051
tables with the same marginal sums as the observed.
We conclude that the crossover points of the simple
recombinants follow a distribution similar to that of
conversion tract endpoints, with no statistically detect-
able difference between the two distributions.

**DISCUSSION**

**Male recombination occurs at P-element sites:** Is P-
induced male recombination the result of direct action
of P transposase on the Drosophila genome, or is it a
consequence of the transposition process itself and/or
subsequent repair events? McCarron et al. (1994) ar-
gue for the former model based on their results show-
ing that recombination can occur in the absence of
any mobile P elements provided a transposase source
is available. Their experimental design for detecting
recombination under these conditions was such that
only recombinant progeny survived. Therefore, the fre-
cuency of such events could not be measured with accu-
racily, but it was clearly several orders of magnitude lower
than what is typically seen when even a single mobile
P element is present. In addition, studies have shown
that when a mobile P element is present, most of the
recombination events occur in the same genetic interval.
as the element (Sved et al. 1991, 1995; McCarron et al. 1994; Svoboda et al. 1995), which would seem to indicate that the events are the result of transposase action on the mobile P-element rather than on random genomic sites. An alternative interpretation advocated by Chovnick and coworkers (Dutartre et al. 1990; McCarron et al. 1994) is that the mobile P element merely tends to concentrate the P transposase to the general vicinity where it is more likely to interact directly with genomic sequences. As evidence for this view, they point out that the presence of a mobile P element also elevates recombination rates in neighboring genetic intervals, albeit to a lesser extent.

Use of the RFLP markers surrounding P(CaSpeR(50C)) allowed us to examine this issue with much higher resolution. We find that 86% of the recombinant chromosomes had crossover points within 2 kb on either side of a single mobile P element. This means that the concentration of crossover points in the immediate vicinity of the P-element is ~30,000-fold greater than in the rest of the chromosome arm. Such a high concentration strongly supports the idea that male recombination results from the action of transposase on mobile P elements rather than on the genome directly.

If one assumes that male recombination occurs only at P-element sites, the lower frequency crossover events in neighboring regions can then be explained by a two-step process in which the mobile P-element transposes to a new location, followed by crossing over at the new site in a future germ cell generation. The phenomenon of “short hopping” observed for some P-element insertions (Tower et al. 1993; Zhang and Spradling 1993) can be invoked to account for any excess of crossover events on the same chromosome arm as the mobile P element. Finally, to explain the rarer-still events in which transposase induces crossing over in a genome lacking any mobile P elements, we postulate that the Drosophila genome contains occasional sequences with sufficient similarity to P-element ends to allow transposase attack, at least with reduced affinity.

Conversion tract distributions are the same at white and 50C: A useful side result of the present work is the conversion tract distribution in Figure 3, which can be compared to analogous data generated previously at the white locus. Specifically, Figure 3 of the present work is comparable to Figure 7 of Gloor et al. (1991), Figure 4 of Nassif et al. (1993) and Figure 6 of Nassif et al. (1994). From these comparisons, it is clear that the distribution of conversion tracts is almost identical between the white locus and 50C. The average conversion tract length at 50C was 1463 bp, essentially the same as the estimated average of 1379 bp for conversion tracts in white (Gloor et al. 1991). For both loci, the distribution of conversion tracts was highly symmetrical with most tracts being bidirectional. This contrasts with one study in Saccharomyces cerevisiae where strong polarity was observed and where most tracts were unidirectional (Sweed et al. 1994). However, conversion tracts can be bidirectional in yeast when the site of the double-strand break is flanked by homologous sequences (Nelson et al. 1996).

The mechanism of conversion tract generation is not known, but the data provide a good fit to a simple model in which a double-strand break is expanded by exonuclease activity such that the extent of the expansion determines the conversion tract endpoints. In this model, each base cleavage event occurs with some probability \( x \), and the gap expansion process stops after a given base with probability \( 1 - x \). Estimating \( x \) by maximum likelihood yields almost the same value whether the data come from 50C or the white locus (0.99864 vs. 0.99855). This near identity suggests that gap expansion occurs through a cell-wide system that is independent of genomic position.

P-element-induced gap repair has been proposed for use as a gene replacement technique for target sites in the vicinity of P insertions (Gloor et al. 1991). The distance between the target site and the P insertion is crucial for this technique, since the fraction of conversion tracts that include a target site \( n \) bases distant from the P-insertion will be \( x^n \). Planning of such experiments will be facilitated if one can assume that \( x \) is essentially invariant between loci, as the present data suggest.

The mechanism of simple recombination events: Any proposed mechanism for the simple recombinants must explain the retention of CaSpeR in a large fraction of recombinant chromosomes, and the distribution of crossover points that resembles the distribution of conversion tract endpoints (Figure 5). The relationship between conversion tracts and crossing over has been investigated for Drosophila meiosis (Curtis et al. 1989; Curtis and Bender 1991), but the underlying mechanisms are likely to be different for P-induced recombination, as discussed below.

Approximately two-thirds of the male recombination events we analyzed contained structural rearrangements, usually duplications or deletions. The true proportion might be even greater, since some of those with crossover points outside the 50C region might also carry undetected rearrangements. In the accompanying paper (Preston et al. 1996) we conclude that duplication- and deletion-bearing recombinants are most easily explained by the hybrid element insertion (HEI) model (Svoboda et al. 1995; Gray et al. 1996).

In the HEI model, recombination comes about via an aberrant transposition event resulting in a recombinant chromosome carrying an intact P element and a flanking duplication or deletion (Figure 6A). Evidence for the HEI model comes from studies in which each homologue carries only one functional P-element end (Svoboda et al. 1995; Gray et al. 1996). Studies of the structure of transposase-induced recombinant chromosomes from males of this kind provided strong evidence for the HEI model (Gray et al. 1996). We apply the model...
FIGURE 6.—Two step model for formation of simple crossover lacking CaSpeR. The RFLP markers are represented by C1, C2, . . . , or A1, A2, . . . for the C and A chromosomes. The 8-bp insertion site of P(CaSpeR)50C is indicated by ▼. (A) The four most common recombinant structures produced by a hybrid element insertion event from CaSpeR elements on the two sister chromatids to the homologue. The steps involved in this kind of HEI event are described in detail elsewhere (PRESTON et al. 1996). For each structure, a novel junction connects the CaSpeR end with an arbitrary site in the flanking sequence of the A chromosome. The duplication or deletion is indicated by a dashed box. (B) CaSpeR excises from one of the recombinant chromosomes at some time after the HEI event. This might occur in a subsequent mitotic germ cell division. Gap expansion and repair from the unmodified A chromosome results in a simple recombinant lacking CaSpeR, but with an apparent crossover point offset from the original CaSpeR insertion site. The process is shown starting with a cn+ bw deletion, but would be similar for the other three structures.

In addition to the structural rearrangements, some of the 29 simple recombinants with crossover points within the 50C marked interval (Figure 5) can also be explained by the HEI model. If the hybrid element happens to insert into the homologue at precisely the same point as the original element, the result would appear as a simple crossover in the MspI–EarI interval with retention of the CaSpeR element. From Figure 5, we see that there are nine recombinants in this category: four cn bw+ and five cn+ bw, five of which were confirmed by DNA sequencing (PRESTON et al. 1996). This interpretation requires that nine independent HEI events, or 6% of the total, occurred at the same target site. Such an extreme hotspot would be unlikely for ordinary P-element transposition. However, insertion site selection for HEI events is probably different from that of ordinary P-element transposition. Physical constraints on the movement of the hybrid element are expected since it remains covalently linked to both sister strands of the homologue. Indeed, five potential cases HEI insertion into this site were reported in another experiment (PRESTON et al. 1996, see type V in Table 2). Therefore, an extreme excess of insertion events into the original site cannot be ruled out.

A two-step process was suggested by J. SVED and M. TANAKA (personal communication) to account for the simple recombinants in the 50C interval that have lost the CaSpeR element. The first step is an HEI event, leaving a recombinant chromosome with an intact CaSpeR and a flanking duplication or deletion (Figure 6A). The second step is ordinary transposition of the CaSpeR element followed by double-strand gap repair (Figure 6B). The template for this repair would be the nonrecombinant A chromosome, which lacks CaSpeR. Following repair, the product is a recombinant chromosome in which the CaSpeR element and the flanking duplication or deletion have been lost. In their place is sequence derived from the A chromosome. We observed 16 simple recombinants in this category. Those with crossover points within the MspI–EarI interval are interpreted as cases in which the second step occurred by a short conversion tract so as not to replace any of the RFLP markers, yet large enough to eliminate any dupli-
cation or deletion. Those with apparent crossover points outside the \( MspI-EaI \) interval are assumed to have longer conversion tracts. The two cases of an apparent double crossovers would, according to this interpretation, be due to discontinuous conversion events have longer conversion tracts. The two cases of an application or deletion. Those with apparent crossover

There remain four cases of simple recombinants that cannot be explained by the HEI model, even with multiple steps. These are the chromosomes with crossover points outside the \( MspI-EaI \) interval but which retain the \( CaSpeR \) element (Figure 5). The occurrence of such chromosomes suggests that not all male recombination comes about via HEI events. One way to account for these structures is to assume that ordinary transposition of the \( CaSpeR \) element followed by gap repair creates a recombinogenic intermediate structure that interacts with the homologue. According to various models of double-strand gap repair (Szostak et al. 1983; Formosa and Alberts 1986; Nassif et al. 1994), a likely intermediate is a repaired DNA duplex with single-stranded nicks at each end of the filled-in gap. Such nicks can be recombinogenic, invading another DNA duplex and initiating recombination (Meelson and Radding 1975). If the initial repair event used the sister chromatid as the template, the copy of \( CaSpeR \) would be restored, but nicks would be left at some distance from the element. Subsequent recombination at these nicks sites would result in the observed structures. This kind of mechanism also offers an alternative way to explain the nine \( CaSpeR \)-bearing crossovers within the \( MspI-EaI \) interval without requiring a high frequency of insertion into the original \( CaSpeR \) site.

To summarize, the HEI model can explain 25 of the 29 simple recombinants whose crossover points lie within the RFLP region. However, other mechanisms must be invoked to account for the remaining cases. The correspondence between the conversion tract distribution and that of the simple crossover points (Figure 5) is expected from the two-step mechanism in Figure 6, since the apparent crossover point corresponds to an end of the conversion tract. In addition, the model of recombinogenic nicks at the boundaries of conversion tracts would also account for the similarity in distributions because conversion tract endpoints determine the recombination point.

Helpful comments were provided by Dena Johnson-Schlitz, Carlos Flores, Koji Kusano, James Crow and Jing Li. We especially thank John Sved and Mark Tanaka for the suggestion that HEI events might be involved. This is paper 3465 from the University of Wisconsin Laboratory of Genetics (http://www.wisc.edu/genetics) supported by National Institutes of Health grant GM-30948.

**LITERATURE CITED**


Communicating editor: R. S. HAWLEY