DOI: 10.1534/genetics.107.085498

Coincidence of *P*-Insertion Sites and Breakpoints of Deletions Induced by Activating *P* Elements in Drosophila

Jyotsna Sudi,[1] Sen Zhang,[2] Gino Intrieri, Ximing Hao and Ping Zhang[3]

Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269-2131

Manuscript received December 7, 2007
Accepted for publication February 26, 2008

ABSTRACT

We isolated a set of seven deletions in the 67B region by activating a nearby *P*-element insertion. The structures of the deletions were characterized by cloning and sequencing. The results showed that the *P*-induced deletions occurred nonrandomly in the genomic sites. One breakpoint of the deletions was located precisely at the end of the starting element, *i.e.*, at the end of the inverted terminal repeats. The other breakpoint was nearby the retained starting element and coincided with preferential *P*-element insertion sites that harbor transcription initiation activities. It is known that *P* elements induce male recombination near the starting elements, giving rise to deletions with one breakpoint precisely located at an inverted terminal repeat of the retained starting element. Database analyses further revealed that deletions generated in *P*-induced male recombination also contained the other breakpoint in genomic regions that coincided with preferential *P*-insertion sites. The results suggest that nonrandom distribution of the deletion breakpoints is characteristic of the mechanism by which *P* elements induce deletions near the starting elements.

TRANSPOSON insertional mutagenesis has been systematically employed to study genetic function in *Drosophila melanogaster* (Cooley *et al.* 1988; Spradling *et al.* 1999). Among the ~15,000 genes in the Drosophila genome, more than one-third have been mutated by *P* insertions alone (Bellen *et al.* 2004). A collection of ~38,000 molecularly mapped insertions of transposons including *P* elements are currently available (Crosby *et al.* 2007). The genomic target sites of *P* elements were diverse (Liao *et al.* 2000), but their distribution was apparently nonrandom. While many genes were hotspots for *P* insertions, the majority of the Drosophila genes have not been mutated by the insertions (Spradling *et al.* 1999; Bellen *et al.* 2004).

Nonrandom distribution of insertion sites not only was seen as hotspots and coldspots among genes, but also was characteristic of *P* insertions within genes. The insertion sites within a collection of 49 genes showed an overwhelming preference of *P* insertions in the 5′-noncoding regions (Spradling *et al.* 1995). Nonrandom distributions of *P* insertions, either among genes, or within genes, were also seen when *P* elements transposed locally (Timakov *et al.* 2002). Local insertions nearby a starting element in the polytene chromosome region of 67B2-3 were exclusively located in 5′-promoter regions of a selective set of genes. For example, among 32 local insertions in a ~12 kb segment, 17 insertions (53%) were located in the *Hsp26* gene, indicating that this was a hotspot. However, several other genes located between the hotspots were coldspots without hits.

P insertions have greatly facilitated the characterization of Drosophila gene functions, although their preferential locations in the 5′-regulatory regions often disrupted only the functions of a subset of the regulatory elements. To generate a null mutation, targeted mutagenesis by homologous recombination has widely been used in recent years (Rong *et al.* 2002). Null phenotypes could also be uncovered with small deficiencies that remove genomic DNA of a single gene. DrosDel is a system to construct deficiencies for specific genomic segments, including a single gene (Ryder *et al.* 2004). The method is based on the FLP site-specific recombinase activity on two closely located and genetically engineered *P* insertions that carry the *FRT* sites (Golic and Golic 1996). More than 500,000 *FRT*-containing insertions for DrosDel have been generated. Similar techniques have also been applied in a deletion project that produced many large deficiencies (Parks *et al.* 2004).

P elements transpose by a cut-and-paste mechanism (Engels 1989; Rio 2002). The excision of an element generates double-strand DNA breaks on the chromosome. Repair of the double-strand DNA breaks is erroneous, resulting in the production of partial *P* elements that appear as imprecisely excised transposons. The

[1]*Present address:* Department of Human Genetics, University of Chicago, Chicago, IL 60637.

[2]*Present address:* Department of Biology, Shanghai Normal University, Shanghai 200234, People's Republic of China.

[3]*Corresponding author:* Department of Molecular and Cell Biology, University of Connecticut, Beach Hall, Room 328, Storrs, CT 06269-2131. E-mail: ping.zhang@uconn.edu

imprecise excision is frequently accompanied by loss of genomic sequences immediately flanking the starting *P* element, which became a tool to produce null mutations for genes whose 5′ regulatory sequences were tagged with *P* insertions. In genetic screens to recover deficiencies from *P* imprecise excision, phenotypes associated with deletions extending into the flanking genes (Tsubota and Schedl 1986; Salz *et al.* 1987), molecular detections of flanking genomic rearrangements (Zhang and Spradling 1993), or a combination of phenotypic selection and molecular detection (Mellerick *et al.* 1992; Eanes *et al.* 2006; Xu *et al.* 2006; Hao *et al.* 2007), were applied. Here, we characterized seven deletions generated by activating a nearby *P* element. The results showed an intriguing correlation between the breakpoints of *P*-induced deletions and *P*-insertion sites. Our data suggest that preferential *P*-insertional activity at genomic sites with transcription initiation activities plays a significant role in inducing genomic deficiencies flanking *P* elements.

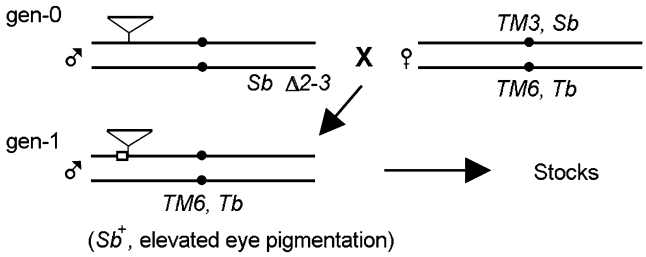

Figure 1.—A genetic screen to isolate candidate chromosomes that carried deletions nearby the starting element. *EP(3)3583* (left arm, triangle) was activated by a transposase source *Δ2-3* (99B on the right arm) in the male of the first generation (gen-0). A male progeny that was Sb^+ and displayed elevated eye pigmentation (gen-1) was used to establish a stock. No more than two gen-1 males from a given gen-0 were taken. An open box left of the original insertion site in the gen-1 male indicates a desired deletion to be molecularly determined. *TM6, Tb* and *TM3, Sb* are two balancers of chromosome 3.

MATERIALS AND METHODS

Drosophila strains: Flies were maintained on standard cornmeal/agar media at 25°. $hsp27^{-0.7}$ is a deletion induced by activating a *P*-element insertion in the promoter of the *Hsp27* gene (Timakov *et al.* 2002; Hao *et al.* 2007). It retained the *EP(3)3583* starting element along with the genomic sequences immediately flanking both ends of the element, but lost a 707-bp-genomic segment that contains the *Hsp27* ORF.

Genetic screen to isolate strains that carried deletions: The *EP(3)3583* element was activated in the male germ line with the *Δ2−3* transposase (Figure 1). Male progeny that displayed elevated eye color were selected to establish individual lines with a balancer chromosome (*TM6, Tb*).

Mapping with PCR: Drosophila genomic DNA was isolated using a standard protocol as described in (Ashburner 1989). The polymerase chain reactions were carried out using the GeneAmp system (PE Applied Biosystems, Foster City, CA) under standard conditions on a Robocycler (Stratagene, La Jolla, CA). PCR primers used to map *P*-element-induced breakpoints of genomic sequences were derived from either the *P* element (Rubin and Spradling 1983) or from the genomic sequences (Crosby *et al.* 2007).

Primers originated in the *P* element are as follows: Pp31, 5′ CGACGGGACCACCTTATGTTATTTCATCATG 3′; PE5′, 5′ AATTCGTCCGCACACAAC 3′.

Primers derived from the genomic sequences are as follows: R-a, 5′ ACATTGGGTGTGTTGTGG 3′; L-a, 5′ GAGCCAGAA GATGCGAGA 3′; L-b, 5′ CTTGGGAATACTGACGGT 3′; L-c, 5′ GATACAGCAAGTGAGTTT 3′; L1, 5′ TGCCCTTCTAT GAGCCCTAC 3′; L2, 5′ GCCTGCTGCTCTACCTCT 3′; L3, 5′ GGTCCCACTTGCTGAA 3′; L4, 5′ CCGTCAACAAG GATGGCTAC 3′; L5, 5′ CATTCTGGAAGAGATCCGATG 3′; L6, 5′ ATGTTCGCACTTCTTGCA 3′; R1, 5′ CCTCCTTG GGATTCTCCTTC 3′; R2, 5′ TAGCTGCACATTTGCTTG 3′; R3, 5′ GCGCGTACGACAACAACT 3′; R4, 5′ CTACTGAC TGGCGGCTTTGT 3′; R5, 5′ TGCCAGATATTCCCTTTGTCT 3′; R6, 5′ ACAGATGGCGCGAGAAAC 3′.

Inverse PCR analysis: To clone genomic sequences flanking *P* elements, inverse polymerase chain reaction was used by adopting the IPCR protocols in the Berkeley Drosophila Genome Project (Crosby *et al.* 2007). Two restriction enzymes each with a 4-bp recognition site were used in our experiments (*Msp*I and *Hha*I). Two IPCR primers complementary to the 5′ end of the *EP* element were Plac1, 5′ CACCCAAGGCTCTG CTCCCACAAT 3′; Pwht1, 5′ GTAACGCTAATCACTCCGAAC AGGTCACA 3′.

Sequencing: The PCR products were purified by using a PCR purification kit (QIAGEN, Valencia, CA) to remove primers and nucleotides and were then used as templates in subsequent sequencing reactions with nested primers.

RESULTS

Production of *P*-element-induced deficiencies: In *D. melanogaster*, a cluster of four small heat-shock genes, including *Hsp22*, *Hsp23*, *Hsp26*, and *Hsp27*, is located within a ~12-kb segment of the 67B region on chromosome 3. In an attempt to generate deficiencies for all of the small *Hsp* genes by *P*-imprecise excision, we took a combination of genetic and molecular approaches that were similar to the methods described in (Hao *et al.* 2007). Briefly, the *EP(3)3583* element retained in the $hsp27^{-0.7}$ deletion was used as the starting element (materials and methods). It was activated in males to produce progeny that displayed elevated eye pigmentation (Figure 1). Because the expression of the marker gene in the starting element, *mini-white*, is influenced by flanking genomic DNA, a rearrangement in the flanking sequences, such as a deletion, could modify the marker gene activity. Progeny with modified eye color were candidates for deficiencies that removed genomic sequences flanking the transposon. Since the *mini-white* marker gene of the starting element was expressed at a relatively low level, it was convenient to isolate progeny with elevated eye pigmentation, upon activation of the *P* element (Figure 1). A total of 382 strains were established from individual males with elevated eye pigmentation. The genomic DNA from each of these strains was isolated and subsequently analyzed to determine whether they contained novel deficiencies for the small *Hsp* genes.

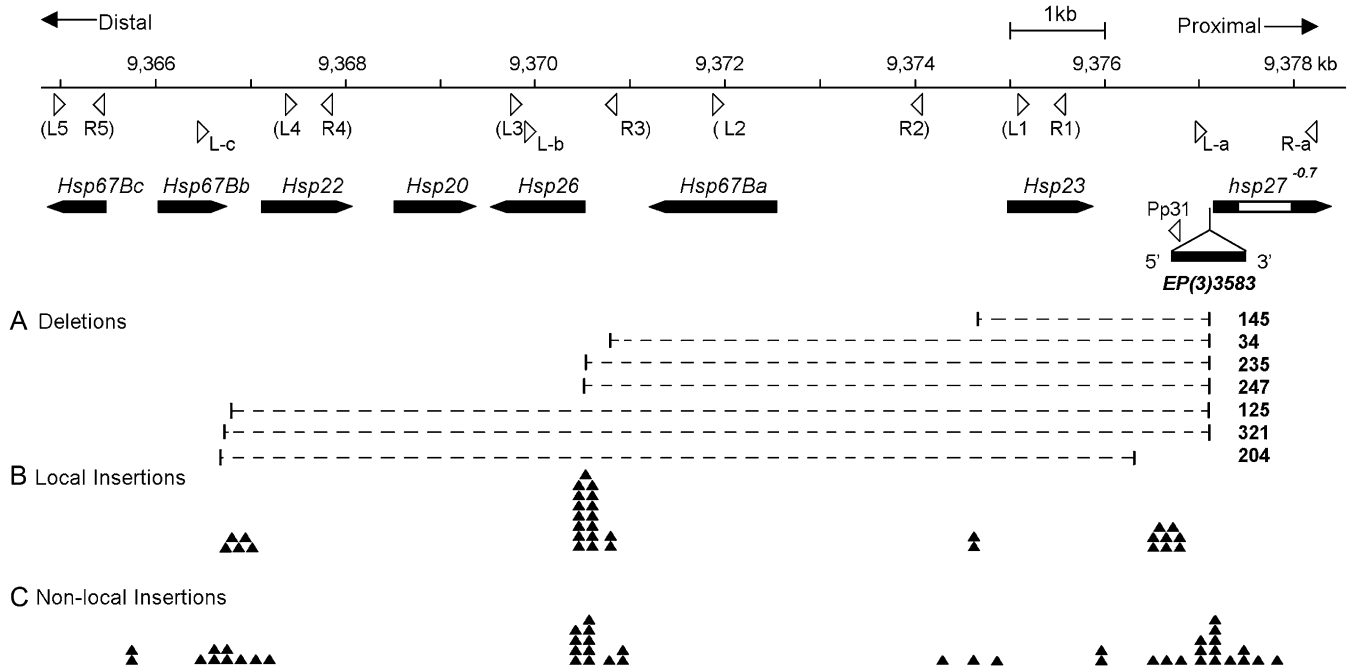

FIGURE 2.—The breakpoints of *P*-induced deletions in 67B coincided with the sites of *P* insertions. Below a genomic map with coordinates on chromosome 3, open arrowheads represent the 5′-to-3′ directions of primers in relation to the genomic sequences (including the L and R series). Pp31 (an open arrowhead) is a primer corresponding to the inverted terminal repeats of *P* elements in an outward direction. The positions of the genes in the region are shown as bars each with an arrow indicating the transcriptional direction. (A) Deletions. Dashed lines indicate seven deletions produced by activating *EP(3)3583* at 9,377,085 (large triangle, HAO *et al.* 2007). (B) Local induced insertions produced by activating *EP(3)3583* (TIMAKOV *et al.* 2002). (C) Nonlocal insertions in the FlyBase. Each triangle represents a local insertion or a nonlocal insertion. The $hsp27^{-0.7}$ allele carried a 707 bp deletion in the *Hsp27* gene (open box). It was produced by activating *EP(3)3583*.

Detection of *P*-element induced deficiencies: To detect genomic rearrangements 5′ to *EP(3)3583*, we used a simple PCR experiment with a pair of primers, Pp31 and L1. Pp31 was derived from the 31-bp-inverted-terminal repeats of the *P* element in an outward orientation, while L1 is located ~1.9 kp from *EP(3)3583* and is oriented toward the *P* element (Figure 2). Among the 382 genomic DNA samples, a total of 33 failed to produce PCR products with Pp31 and L1. To determine whether the candidate strains contained deficiencies for the *Hsp* genes, the 33 candidate strains were further examined by PCR with a pair of primers derived from the *Hsp23* gene, L1 and R1 (Figure 2). This PCR experiment was facilitated by the presence of flies homozygous for the candidate chromosome in all 33 strains. The results from PCR using genomic templates isolated from each of the homozygotes showed that 7 of the 33 strains failed to produced a PCR product with L1 and R1, indicating that this set of 7 strains contained mutations for the *Hsp23* gene (Table 1).

We then asked if the original *P* element was retained in this set of seven strains. A pair of primers, Pp31 and R-a, was used to amplify the genomic sequence flanking the 3′ end of *EP(3)3583* associated with the $hsp27^{-0.7}$ allele (Figure 2). Due to the 707-bp genomic deletion, the hsp27-0.7 allele was expected to produce a ~0.4-kb band, instead of the ~1.1-kb band that would have originated without the deletion adjacent to the starting element. We have determined that all of the seven strains in this set produced a ~0.4-kb band with the pair of Pp31 and R-a, indicating that at least the 3′ end of the initial *P* element was retained in each of the strains (Table 1).

The remaining 26 strains among the initial 33 strains each produced a product of the wild-type size with the L1 and R1 pair. Thus, these strains carried genomic arrangements in either the 31-bp-terminal-inverted repeat of the *P* element or in the genomic DNA immediately flanking the 5′ end of the *P* element. These strains were not further characterized.

Mapping seven deficiencies and determining their breakpoints: The breakpoints in seven deficiency candidates for the small *Hsp* genes were mapped by using PCR with primers mainly distributed in a ~12-kb genomic region distal to *EP(3)3583* (Figure 2). An additional pair, L6 and R6, was derived from the *CG4452* gene, which is located at a more distal segment on the genomic map (~6 kb to the left of *Hsp22* in Figure 2).

The results obtained from PCR mapping using genomic templates isolated from each of the homozygous deficiency candidates are summarized in Table 1. The data indicate that each of the seven strains contained breakpoints within the ~12-kb genomic region shown in Figure 2. The data also suggest that only the *204* strain

TABLE 1

Mapping the breakpoints of *P*-induced deletions

	Primer pair							
Strain	L6 R6	L5 R5	L4 R4	L3 R3	L2 R2	L1 R1	L-a Pp31	Pp31 R-a
145	+	+	+	+	+	−	−	+
34	+	+	+	+	−	−	−	+
235	+	+	+	+	−	−	−	+
247	+	+	+	+	−	−	−	+
321	+	+	−	−	−	−	−	+
125	+	+	−	−	−	−	−	+
204	+	+	−	−	−	−	+	+

PCR products are listed as either present (+) or absent (−) relative to the reference products of the control DNA that was isolated from the $hsp27^{-0.7}$ strain.

The L6 and R6 primers are derived from the *CG4452* gene that is located ~6 kb distal to *Hsp22*.

retained the genomic DNA immediately flanking the starting *P* element. In addition, two primer pairs distal to the *Hsp67Bb* gene (L5 and R5 and L6 and R6) were used to map the distal breakpoints of the deletions. With both of the primer pairs, all of the deficiencies produced PCR products of the sizes of the wild-type control, indicating that the deficiency breakpoints were restricted in the mapping region.

To determine the breakpoints for each of the deficiencies at the molecular level, genomic sequences containing the breakpoints were cloned and sequenced. Two complementary methods were used to clone the genomic sequences. One was to clone the breakpoints using inverse PCR (IPCR), while the other was to clone the sequences directly with PCR by using the mapping data described in Table 1. The IPCR method to clone genomic sequences flanking a *P* element was provided in detail by the FlyBase (Crosby *et al.* 2007). In the second method, the breakpoints were cloned directly from genomic DNA, which was facilitated from the mapping results in Table 1. With a primer derived from the 5′ end of the *P* element in an outward orientation (*e.g.*, Plac1 or Pp31), and another primer that was located near the breakpoints of a deficiency and orientated toward *EP(3)3583* (*e.g.*, L-b or L-c), a genomic fragment containing the breakpoints was amplified. The cloned fragments were then sequenced with nested primers.

The results obtained from the cloning and sequencing are summarized in Table 2 and schematically presented in Figure 2. Among the seven deficiencies, five contained deletions of genomic sequences beginning precisely at the 5′ end of the *EP(3)3583* transposon. Another deletion, the *321* deficiency, also had a precise deletion at the 5′ end of *EP(3)3583*, but contained an additional 26-bp sequence of GGACCACCTTATGTTA TTTCATCATG, a nearly complete direct copy of the 31-bp inverted terminal repeats. In contrast, the *204* deficiency retained the genomic sequence immediately flanking *EP(3)3583*, but contained a deletion beginning at 779 bp distal to the 5′ end of the transposon. In

TABLE 2

Deletion breakpoints and the relative positions to nearby genes

	Breakpoints					
Strain	Distal	Proximal	Deletion size (bp)	Initiation site[a]	Gene	Distance[b] (bp)
145	9,374,670	9,377,085	2,416	9,374,984	*Hsp23*	314
34	9,370,791	9,377,085	6,295	9,370,527	*Hsp26*	264
235	9,370,533	9,377,085	6,553	9,370,527	*Hsp26*	6
247	9,370,508	9,377,085	6,578	9,370,527	*Hsp26*	19
125	9,366,803	9,377,085	10,283	9,367,130	*Hsp22*	327
321	9,366,744	9,377,085[c]	10,342	9,367,130	*Hsp22*	386
204	9,366,674	9,376,306[d]	9,613	9,367,130	*Hsp22*	436

[a] Nearest transcription initiation sites and genes.

[b] Distance between the breakpoint and the nearest transcription initiation site.

[c] The deletion contained a 26-bp imperfect duplication of the 31-bp-*P*-terminal repeats at the 5′-*P* end, GGACCACCTTATGTTATTTCATCATG.

[d] The deletion contained an extra trinucleotide, ATG, at the junctions of the deletion breakpoints. The sequence at the junction is AAGTTCTTGACTACACCCATGATGAAATAATTACGTGCATGCACACAT, with the extra trinucleotides underlined.

addition, three extra nucleotides, ATG, were inserted at the junction of the breakpoints of the *204* deficiency. Interestingly, the ATG trinucleotides were also present on one side of the breakpoints that were apparently derived from the parental genomic sequence (Table 2).

The coincidence of the *P*-induced-deletion breakpoints with genomic sites of *P* insertions: Our mapping and sequencing data showed that, in all but one deficiency, one of the breakpoints was in a genomic site where a nearby small *Hsp* gene initiated its transcription, while the other was precisely located at the 5′ end of *EP(3)3583* (Table 2). The exception was the *204* deficiency, in which one of the breakpoints was different. Despite containing no breakpoint at the 5′ end of *EP(3)3583*, the distal breakpoints in the *204* deficiency were located in the 5′ region of the *Hsp22* genes (Table 2, Figure 2).

P elements are known to preferentially insert into 5′ ends of genes (Spradling *et al.* 1995). A collection of *P* insertions into the small *Hsp* genes at the FlyBase also showed that the insertions were located predominantly within the promoter regions of the target genes (Figure 2). When *EP(3)3583* was activated, local insertions into the nearby small *Hsp* genes were located exclusively within the promoter regions of the target genes (Figure 2). Thus, *P* insertions into the small *Hsp* genes, whether they were generated through local or nonlocal transpositions, were preferentially localized in the promoter regions of the small *Hsp* genes.

Compared with the insertion sites of *P* elements in the *Hsp22*, *Hsp23*, and *Hsp26* genes, the breakpoints of the deficiencies generated from activating *EP(3)3583* coincided with the genomic sites where the *P* elements inserted (Figure 2, A–C). For example, clusters of *P* insertions, local or nonlocal insertions, were found in two discrete genomic regions upstream of the transcription initiation site of the *Hsp26* gene. These two insertion sites coincided with three *EP(3)3583*-induced deficiencies. The *34* deficiency had a breakpoint coincide with one of the insertion sites, while the *235* and *247* deficiencies had breakpoints coincide with the other insertion site (Table 2 and Figure 2). Thus, the *EP(3)3583* induced deletions had nonrandom breakpoints that coincided with *P*-insertion sites, suggesting the involvement of *P*-insertional activities in the induction of these deletions.

The correlation between deletion breakpoints of *P*-induced male recombination and *P*-insertion sites in the 50C19-20 region: Of a total of seven deficiencies induced by activating *EP(3)3583*, six had a breakpoint located precisely at the 5′ terminus of the retained *P* element. The deletions with a breakpoint precisely at a *P*-element terminus are notably reminiscent of deletions associated with *P*-induced male recombination (Kidwell *et al.* 1977; Gray *et al.* 1996; Preston and Engels 1996). Although Drosophila meiotic recombination does not normally occur in males, recombination in the male germ line could be induced at low levels (~1%) by *P*-element transposition activities (Preston and Engels 1996). Male recombination was often accompanied by chromosome rearrangements, including deletions and duplications, which contained one of the breakpoints precisely located at one of the 31-bp-terminal-inverted repeats of a retained *P* element (Gray *et al.* 1996; Preston *et al.* 1996). A mechanism involving "hybrid element insertion" (HEI) was proposed to explain the occurrence of the chromosome rearrangements in male recombination. In the HEI model, the 5′ end of a *P* element on a sister chromatid interacts with the 3′ end of the *P* element copy on the other sister chromatid, to form an active hybrid element (Gray *et al.* 1996; Preston *et al.* 1996). The hybrid element composed of two *P*-element copies on the sister chromatids of a chromosome inserts in a nearby position on the homolog, giving rise to the deletions and duplications adjacent to the retained *P* element.

The similarity of the breakpoints at a terminus of the *P*-element inverted repeats between the deletions induced by activating *EP(3)3583* and by *P*-induced male recombination raised a possibility that the seven deletions might be the results of male recombination induced by activating *EP(3)3583*. Since the deletion breakpoints produced from activating *EP(3)3583* coincided with *P*-element preferential insertion sites, which were nearly exclusively located at the 5′ end of the small *Hsp* genes (Figure 2), it is possible that this coincidence is a common feature for deletions of male recombination induced by HEI.

To examine the relationship between *P*-induced deletions in male recombination and *P*-insertional activity, we carried out a database search for *P* insertions in a small chromosome 2 region, in which a large number of deletions of male recombination were previously investigated. Preston *et al.* (1996) activated a *P* insertion in the male germ line, *P{CaSpeR}Cp1^{50c}* (Preston *et al.* 1996, 2002) and isolated 21 deletions within a ~1.4-kb genomic segment in the polytene 50C19-20 region. The breakpoints of the deletions were disproportionately distributed, with the vast majority distal to the starting *P* element (17/21, or ~81%, Figure 3A). This pattern of an unequal distribution was well-correlated with that of *P* insertions found in the FlyBase, which were also located predominantly in genomic sequences distal to the *P{CaSpeR}Cp1^{50c}* starting site (11/13, or ~85%, Figure 3B).

It was intriguing to notice that the deletion breakpoints from both activating *EP(3)3583* and *P{CaSpeR}Cp1^{50c}* were nonrandom, displaying a strong correlation with *P*-insertional sites. However, frequent targeting of *P* elements in 50C19-20, which is in an intron of the Cp1 gene, was puzzling, because introns were rarely targeted by *P* elements (Spradling *et al.* 1995; Venken and Bellen 2005). Since *P* elements preferentially target transcription initiation regions, a database search was carried out to

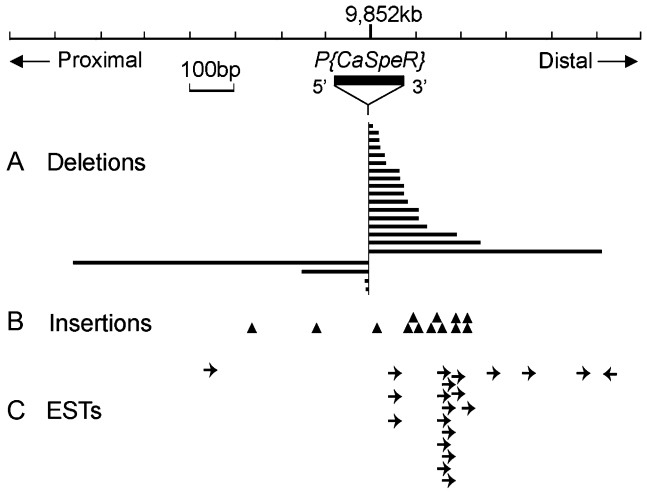

FIGURE 3.—The deletion breakpoints of *P*-induced male recombination in 50C19-20 coincided with the sites of *P* insertions. A ~1.4 kb genomic segment of the 50C19-20 region is shown with coordinates of chromosome 2. (A) Deletions of male recombination (horizontal bars) induced by activating *P{CaSpeR}Cp1*^{50c} at 9,851,974 (PRESTON *et al.* 1996, 2002). (B) *P* insertions in the region. Each triangle represents an insertion in the FlyBase. (C) Expression sequence tags (ESTs). Each arrow indicates the 5′ end of an EST in the FlyBase.

examine transcriptional activities in 50C19-20. As shown in Table 3, a total of 22 expression sequence tags (ESTs) with 5′-end sequences transcribed from 50C19-20 were identified in the FlyBase. Thus, despite being an intron, genomic sequences in the 50C19-20 region could serve as active transcription initiation sites, because of the identification of cDNA clones with 5′-end sequences localized in this intronic region. The transcriptional activities initiated in 50C19-20 could explain why *P* elements frequently targeted this genomic region.

In addition, the data also showed a nonrandom distribution of the ESTs. Similar to that of *P* insertions, all but one of these ESTs contained the 5′-end sequences localized to the distal side of the starting *P* element, indicating a correlation between *P* insertions and transcription initiation activities in 50C19-20 (Figure 3).

TABLE 3

Expression sequence tags with 5′-end sequences localized in 50C19-20

EST[a]	5′ Site[b]	Distance[c]
SD16867	9,851,634	−340
RE29616	9,852,034	+60
RE08377	9,852,034	+60
RE12916	9,852,034	+60
RH24812	9,852,158	+184
RH65853	9,852,158	+184
RE11004	9,852,158	+184
RH14017	9,852,158	+184
RH36876	9,852,158	+184
RH37208	9,852,158	+184
RH24883	9,852,158	+184
RH25710	9,852,158	+184
RH24779	9,852,158	+184
RH37626	9,852,158	+184
RH70448	9,852,159	+185
SD23574	9,852,173	+199
GM32440	9,852,173	+199
EN14717	9,852,184	+210
GH25862	9,852,259	+285
BP559375	9,852,337	+363
EN10012	9,852,461	+487
EK076238	9,852,535	+561[d]

The cDNA clones were identified in various tissue-specific libraries, including staged ovaries (GM), embryos (RE), adult head (GH, RH, and EN), embryonic gonads (BP), multiple tissues (EK), and tissue culture cells (SD).

[a] FlyBase ESTs (CROSBY *et al.* 2007) transcribed from the genomic region flanking the P{CaSpeR}Cp1^{50c} insertion in 50C19-20. Only ESTs with known 5′-end sequences are included.

[b] The position of the 5′end of an EST corresponding to the genomic map in Figure 4.

[c] The distance between an EST's 5′ end and the insertion site of P{CaSpeR}Cp1^{50c} at 9,851,974.

[d] EK076238 has a transcriptional direction opposite to the other ESTs.

DISCUSSION

***P*-induced deletions and imprecise excision:** Excision of *P* elements generates double-strand breaks, which are substrates of DNA repair machineries (ENGELS 1989; RIO 2002; PRESTON *et al.* 2006). Double-strand repairs are often aberrant, resulting in imprecise excision of *P* elements with deletions or additions of extra sequences (TAKASU-ISHIKAWA *et al.* 1992; JOHNSON-SCHLITZ and ENGELS 1993; STAVELEY *et al.* 1995). Deletions induced by imprecise excision could extend into the flanking genomic sequences (TSUBOTA and SCHEDL 1986; SALZ *et al.* 1987), which has been a widely used tool to generate null mutations near known *P* insertions (VENKEN and BELLEN 2005). In addition to genomic deletions extended from the imprecisely excised *P* elements, activating *P* insertions could also produce flanking genomic deletions while retaining the starting elements (ZHANG and SPRADLING 1993; HAO *et al.* 2007).

In this report, all of the seven deletions isolated by activating *EP(3)3583* apparently retained the starting element. First, the *mini-white* marker gene of the element was expressed in each deletion strain. Second, the inverted repeats on both ends of the transposon were present in each deletion strain (*e.g.*, Table 1). Third, the genomic sequence proximal to *EP(3)3583* was retained in each deletion strain. Finally, at least 1,000 bp sequences of the 5′ end of the starting *EP(3)3583* appear to be retained. This is because the restriction sites used in IPCR, as well as *P* sequences corresponding to the PE5′, Plac1 and Pwht1 primers, were retained. The presence of some of these 5′ sequences was confirmed in the sequencing experiments.

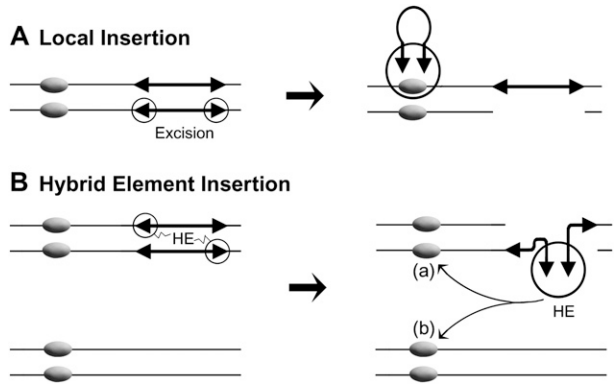

FIGURE 4.—Models of *P*-local insertion and hybrid element insertion (A) Local insertion. The ends of a *P* element (small circles) are excised. The excised *P* element is often tethered to the *P*-element copy on the sister chromatid (large circle). Local insertion occurs at a site with transcription initiation activity nearby the starting element on the sister chromatid (oval). (B) Hybrid element insertion (adapted from PRESTON *et al.* 1996). An excised 5′-end of an element (small circle) interacts with an excised 3′ end of the element copy on the sister chromatid (small circle) to form a hybrid element. The hybrid element (HE, large circle) inserts into a nearby site with transcription initiation activities (oval) on a sister chromatid (a), or on the homolog (b). The insertions lead to male recombination products including deletions.

In the genetic screen to isolate the deletions, a change of the marker gene expression could result from either a rearrangement of genomic sequences flanking the starting element or an insertion in another chromosomal location. The selection of the marker expression could thus be a main factor in the retentiveness of the starting transposon. However, this selection should not prevent the isolation of genomic rearrangements that contained small deletions or additions of *P*-element sequences, which were often associated with imprecise excision. To explain the isolation of the seven deletions with no obvious rearrangements of the starting element, we propose that a mechanism other than imprecise excision played a role in inducing the deletions.

***P*-induced deletions and the HEI model:** Mechanistically, the production of deletions from *P*-induced male recombination is best explained by the HEI model, in which a hybrid element inserts into a nearby site on the homolog, resulting in deletions and duplications (SVOBODA *et al.* 1995; GRAY *et al.* 1996; PRESTON *et al.* 1996). The hybrid element is composed of the 5′ end of one transposon on a chromatid and the 3′ end of the transposon copy on the sister chromatid (Figure 4).

The deletions induced by activating *EP(3)3583* also fit the HEI model, since they shared a number of molecular characteristics with deletions produced from *P*-induced male recombination. One of the breakpoints in the deletions produced from activating *EP(3)3583* was precisely located at the *P*-inverted terminal repeat (Table 2), resembling those of deletions isolated from *P*-induced male recombination (SVOBODA *et al.* 1995; GRAY *et al.* 1996; PRESTON *et al.* 1996). The similarities between the deletions described here and deletions of the male recombination extended to the genomic sites of the other breakpoint, which were nonrandomly distributed in targets of *P* insertions and harbored transcription initiation activities (Figures 2 and 3). Another shared feature is that these deletions involved only one end of the starting elements (Figures 2 and 3).

The *321* deficiency contained a 26-bp imperfect duplication of the 31-bp-inverted-terminal repeats (Table 2). Interestingly, a deletion from *P*-induced male recombination also had a 26-bp imperfect duplication of the terminal repeats at the deletion junction (PRESTON *et al.* 1996). These two duplications might result from *P*-element transposition into the starting element copy on the sister chromatid followed by an imprecise excision (PRESTON *et al.* 1996).

In addition to the deletions that were presumably induced by HEI, the *204* deficiency apparently resulted from a different mechanism. It did not contain a breakpoint precisely at the end of the starting element. This deletion could be explained by an insertion to a nearby genomic site followed by an imprecise excision of the newly inserted element. At its deletion junction, there was a trinucleotide insertion (ATG). The $hsp27^{-0.7}$ deletion contained similar features, including both of the breakpoints located within flanking genomic sequences and a trinucleotide insertion (GAA) at the deletion junction (HAO *et al.* 2007). In both deletions, the extra trinucleotide was identical to the junction sequence, as if it were a direct duplication of the genomic sequence at the junction (Table 2; and in (HAO *et al.* 2007).

Alternatively, the *204* deficiency could result from a more complicated rearrangement. A 13-nucleotide sequence at the junction of the deletion (underlined), AAGTTCTTGACTACACCCATGATGAAATAATTACGT GCATGCACACAT, is also homologous to the *P*-end inverted repeats. It is possible that this breakpoint is the result of a *P*-element insertion at the distal region, followed by base pairing to break-induced DNA replication from the original *P*-element excision site.

Local insertion and HEI: *P* elements preferentially insert into genomic sites near the starting elements (TOWER *et al.* 1993; ZHANG and SPRADLING 1993). The local insertions were nonrandomly distributed among genes, showing a pattern similar to that of nonlocal insertions with hotspots and coldspots (Figure 2). Local insertions seem to target only at promoter regions, since the insertion sites were mapped exclusively to nearby transcription initiation sites (TIMAKOV *et al.* 2002; SHILOVA *et al.* 2006).

Similar to local transposition, *P*-induced male recombination occurred mostly near the starting element (DUTTAROY *et al.* 1990; MCCARRON *et al.* 1994; PRESTON and ENGELS 1996). The HEI sites were also frequent *P*-insertion targets (Figure 3). The deletions induced by

activating *EP(3)3583*, which were presumably induced by HEI, showed similar coincidence among the deletion breakpoints, *P*-insertions sites, and transcription initiation sites (Figure 2). These similarities between local insertion sites and the sites of *P*-induced male recombination suggest that local insertion and HEI share some mechanistic aspects that mediate these two *P*-insertional activities.

A difference between local insertion and HEI seems to be the chromosomal apparatus on which *P* elements operate. In HEI, *P*-induced male recombination occurs between a pair of homologs (Figure 4). Although whether *P* elements transpose locally between a pair of homologs remains unclear, a study found that such nearby insertions on the homolog, if any, were rare (Tower and Kurapati 1994). Thus, preferential local insertions appear to be restricted to the donor chromosome. There is also evidence indicating that local insertions were further restricted on the two sister chromatids of the chromosome where the starting element was excised. If *P* elements transposed locally onto the chromatid on which the starting element was excised, the "cut-and-paste" model would predict loss of the starting element or imprecise excision due to aberrant double-strand DNA repair (Engels 1989). The vast majority of the local insertions were, on the contrary, isolated along with the retained starting elements (Tower *et al.* 1993; Zhang and Spradling 1993; Timakov *et al.* 2002; Shilova *et al.* 2006). The retaining of the starting elements indicates that the donor chromatid, on which a *P* element is excised, rarely serves as a target for local insertions (Figure 4).

The HEI model also predicts insertion of a hybrid element in a sister chromatid of the chromosome where the hybrid element was generated (Figure 4). Such sister-chromatid HEI insertions were implicated in a prior study using a ring *X* chromosome (Sved and Liang 2006). The results from the present study could not be used to investigate if sister-chromatid HEIs, or homolog HEIs, were involved in generating the deletions, since the chromosome region distal to the starting element was not marked. However, a recent study suggests that HEI on sister chromatids of the starting element might occur infrequently. In an effort to characterize the *Hsp27* function, Hao *et al.* (2007) used genetic and molecular methods, which are identical to those used here, to isolate genomic deletions proximal to *EP(3)3583* (Figure 2). Because the genomic region of interest is between *EP(3)3583* in 67B and the transposase source, Δ2-3 in 99B (linked with Sb), a male recombination in this region would produce a product together with the right arm carrying Sb and Δ2-3 (for the relative positions, see Figure 1). Since the resulting chromosome with Sb and Δ2-3 was selected against in subsequent genetic crosses, these experiments should have excluded the recovery of proximal deletions from HEI male recombination between a pair of homologs. Nonetheless, other proximal deletions derived from HEI on the *EP(3)3583* chromosome were recoverable, if sister-chromatid HEIs were to occur. Among a collection of 145 candidate strains, a deletion, $hsp27^{-0.7}$, was isolated (Figure 2). However, this deletion does not contain a breakpoint precisely located at the end of a 31-bp-inverted-terminal repeat, which is characteristic of HEI, although it retained the starting element. It also retained the genomic sequences immediately flanking both sides of the element. In addition, it carried three extra nucleotides inserted at the junction of the breakpoints. Thus, deletions proximal to the *EP(3)3583* element that might be produced from sister-chromatid HEIs were not recovered in the experiments (0/145, (Hao *et al.* 2007). The results suggest that HEI occurs infrequently on the sister chromatids, although it is also possible that the genetic screen, *i.e.*, eye color change after activating the starting element, played a role in the failure of isolating proximal deletions that were expected from sister-chromatid HEIs.

We propose that an excised *P* element preferentially inserts locally in a sister chromatid other than the one on which the starting element is excised, whereas an active HEI complex may frequently result in an insertion in the homologous chromosome (Figure 4). To explain frequent local insertion (Zhang and Spradling 1993), proteins bound to *P*-element termini (Lee *et al.* 1996; Beall and Rio 1998; Tang *et al.* 2005) are thought to tether the excised element to the unexcised copy on the sister chromatid, giving rise to a local insertion event near the starting element. An alternative explanation is that local transposition may occur shortly after DNA replication, during which the two copies of the *P* elements are still in a close association. The excised copy inserts locally due to its association with the unexcised copy on the sister chromatid (Zhang and Spradling 1993). HEI, on the other hand, may occur by and large at a different stage of the mitotic cell cycles of the spermatocytes, during which homologous sequences are paired in a preparation for meiosis. Perhaps, HEI not only requires a close proximity of insertion targets that are transcriptionally active, but also needs the presence of other factors, which are assembled only on the paired homolog complex at later stages of spermatogenesis, to resolve its products containing rearranged chromosomes.

We thank the Center for Applied Genomics and Technology at the University of Connecticut for sequencing. This work was supported in part by a grant from the National Science Foundation and a grant from the University of Connecticut Research Advisory Council.

LITERATURE CITED

Ashburner, M., 1989 *Drosophila: A Laboratory Handbook*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

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

Bellen, H. J., R. W. Levis, G. Liao, Y. He, J. W. Carlson *et al.*, 2004 The BDGP gene disruption project: single transposon insertions associated with 40% of Drosophila genes. Genetics **167:** 761–781.

Cooley, L., R. Kelley and A. Spradling, 1988 Insertional mutagenesis of the Drosophila genome with single *P* elements. Science **239:** 1121–1128.

Crosby, M. A., J. L. Goodman, V. B. Strelets, P. Zhang and W. M. Gelbart, 2007 FlyBase: genomes by the dozen. Nucleic Acids Res. **35:** D486–D491.

Duttaroy, A., M. McCarron, K. Sitaraman, G. Doughty and A. Chovnick, 1990 The relationship between *P* elements and male recombination in *Drosophila melanogaster*. Genetics **124:** 317–329.

Eanes, W. F., T. J. Merritt, J. M. Flowers, S. Kumagai, E. Sezgin *et al.*, 2006 Flux control and excess capacity in the enzymes of glycolysis and their relationship to flight metabolism in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **103:** 19413–19418.

Engels, W. R., 1989 *P* elements in *Drosophila melanogaster*, pp. 437–484 in *Mobile DNA*, edited by D. E. Berg and M. M. Howe. American Society for Microbiology, Washington, D.C.

Golic, K. G., and M. M. Golic, 1996 Engineering the Drosophila genome: chromosome rearrangements by design. Genetics **144:** 1693–1711.

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

Hao, X. M., S. Zhang, B. Timakov and P. Zhang, 2007 The *Hsp27* gene is not required for Drosophila development but its activity is associated with starvation resistance. Cell Stress Chaperones **12:** 364–372.

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.

Kidwell, M. G., J. F. Kidwell and J. A. Sved, 1977 Hybrid dysgenesis in *Drosophila melanogaster*: a syndrome of aberrant traits including mutation, sterility, and male recombination. Genetics **86:** 813–833.

Lee, C. C., Y. M. Mul and D. C. Rio, 1996 The Drosophila *P*-element KP repressor protein dimerizes and interacts with multiple sites on *P*-element DNA. Mol. Cell Biol. **16:** 5616–5622.

Liao, G. C., E. J. Rehm and G. M. Rubin, 2000 Insertion site preferences of the *P* transposable element in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **97:** 3347–3351.

McCarron, M., A. Duttaroy, G. Doughty and A. Chovnick, 1994 Drosophila *P* element transposase induces male recombination additively and without a requirement for *P* element excision or insertion. Genetics **136:** 1013–1023.

Mellerick, D. M., J. A. Kassis, S. D. Zhang and W. F. Odenwald, 1992 *castor* encodes a novel zinc finger protein required for the development of a subset of CNS neurons in Drosophila. Neuron **9:** 789–803.

Parks, A. L., K. R. Cook, M. Belvin, N. A. Dompe, R. Fawcett *et al.*, 2004 Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. Nat. Genet. **36:** 288–292.

Preston, C. R., W. Engels and C. Flores, 2002 Efficient repair of DNA breaks in Drosophila: evidence for single-strand annealing and competition with other repair pathways. Genetics **161:** 711–720.

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

Preston, C. R., C. C. Flores and W. R. Engels, 2006 Differential usage of alternative pathways of double-strand break repair in Drosophila. Genetics **172:** 1055–1068.

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

Rio, D. C., 2002 *P* transposable elements in *Drosophila melanogaster*, pp. 484–518 in *Mobile DNA II*, edited by E N. L. Craig, R. Craigie, M. Gellert and A. M. Lambowitz. American Society for Microbiology, Washington, DC.

Rong, Y. S., S. W. Titen, H. B. Xie, M. M. Golic, M. Bastiani *et al.*, 2002 Targeted mutagenesis by homologous recombination in *D. melanogaster*. Genes Dev. **16:** 1568–1581.

Rubin, G. M., and A. C. Spradling, 1983 Vectors for *P* element-mediated gene transfer in Drosophila. Nucleic Acids Res. **11:** 6341–6351.

Ryder, E., F. Blows, M. Ashburner, R. Bautista-Llacer, D. Coulson *et al.*, 2004 The DrosDel collection: a set of *P*-element insertions for generating custom chromosomal aberrations in *Drosophila melanogaster*. Genetics **167:** 797–813.

Salz, H. K., T. W. Cline and P. Schedl, 1987 Functional changes associated with structural alterations induced by mobilization of a *P* element inserted in the *Sex-lethal* gene of Drosophila. Genetics **117:** 221–231.

Shilova, V. Y., D. G. Garbuz, E. N. Myasyankina, B. Chen, M. B. Evgen'ev *et al.*, 2006 Remarkable site specificity of local transposition into the *Hsp70* promoter of *Drosophila melanogaster*. Genetics **173:** 809–820.

Spradling, A. C., D. M. Stern, I. Kiss, J. Roote, T. Laverty *et al.*, 1995 Gene disruptions using *P* transposable elements: an integral component of the Drosophila genome project. Proc. Natl. Acad. Sci. USA **92:** 10824–10830.

Spradling, A. C., D. Stern, A. Beaton, E. J. Rhem, T. Laverty *et al.*, 1999 The Berkeley Drosophila Genome Project gene disruption project: single *P*-element insertions mutating 25% of vital Drosophila genes. Genetics **153:** 135–177.

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

Sved, J. A., and X. Liang, 2006 Evidence of *P*-element-induced sister-chromatid exchange in a ring-*X* chromosome in Drosophila, with implication for a high rate of formation of hybrid elements. Genetics **172:** 975–979.

Svoboda, Y. H., M. K. Robson and J. A. Sved, 1995 *P*-element-induced male recombination can be produced in *Drosophila melanogaster* by combining end-deficient elements in *trans*. Genetics **139:** 1601–1610.

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.

Tang, M., C. Cecconi, H. Kim, C. Bustamante and D. C. Rio, 2005 Guanosine triphosphate acts as a cofactor to promote assembly of initial *P*-element transposase-DNA synaptic complexes. Genes Dev. **19:** 1422–1425.

Timakov, B., X. Liu, I. Turgut and P. Zhang, 2002 Timing and targeting of *P*-element local transposition in the male germline cells of *Drosophila melanogaster*. Genetics **160:** 1011–1022.

Tower, J., and R. Kurapati, 1994 Preferential transposition of a Drosophila *P* element to the corresponding region of the homologous chromosome. Mol. Gen. Genet. **244:** 484–490.

Tower, J., G. H. Karpen, N. Craig and A. C. Spradling, 1993 Preferential transposition of Drosophila *P* elements to nearby chromosomal sites. Genetics **133:** 347–359.

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.

Venken, K. J., and H. J. Bellen, 2005 Emerging technologies for gene manipulation in *Drosophila melanogaster*. Nat. Rev. Genet. **6:** 167–178.

Xu, Y., M. Condell, H. Plesken, I. Edelman-Novemsky, J. Ma *et al.*, 2006 A Drosophila model of Barth syndrome. Proc. Natl. Acad. Sci. USA **103:** 11584–11588.

Zhang, P., and A. C. Spradling, 1993 Efficient and dispersed local *P* element transposition from Drosophila females. Genetics **133:** 361–373.

Communicating editor: K. G. Golic