Cell, Vol. 44, 7–19, January 17, 1986, Copyright © 1986 by Cell Press

Tissue Specificity of Drosophila P Element Transposition Is Regulated at the Level of mRNA Splicing

Frank A. Laski, Donald C. Rio, and Gerald M. Rubin
Department of Biochemistry
University of California
Berkeley, California 94720

Summary

We show that the germline specificity of P element transposition is controlled at the level of mRNA splicing and not at the level of transcription. In the major P element RNA transcript, isolated from somatic cells, the first three open reading frames are joined by the removal of two introns. Using in vitro mutagenesis and genetic analysis we demonstrate the existence of a third intron whose removal is required for transposase production. We propose that this intron is only removed in the germline and that its removal is the sole basis for the germline restriction of P element transposition.

Introduction

P elements are a family of transposable elements found in Drosophila melanogaster. P elements are particularly interesting because their transposition is both genetically regulated and tissue-specific. They have been shown to be the causal agents of P–M hybrid dysgenesis, a syndrome whose traits include high rates of sterility, mutation, and chromosomal rearrangements (Kidwell et al., 1977; for review see Engels, 1983). The syndrome occurs in the progeny of a cross between a male of a P strain and a female of an M strain, but not in the reciprocal cross, a P × P cross or an M × M cross. The distinguishing characteristics of P strains are that their eggs have "P cytotype," a condition that results in repression of P element transposition, and that they carry autonomous 2.9 kb P elements that can encode transposase. M strains have "M cytotype," which allows P element transposition, but they lack autonomous P elements. Thus, transposition (and the resulting hybrid dysgenesis) only occurs when P elements are introduced into an M cytotype egg (Engels, 1983). The molecular mechanism for the P cytotype repression of transposition is not known. Furthermore, P element transposition is tissue-specific; transposition occurs at high levels in germline tissue but has not been detected in somatic tissues (Engels, 1983).

Autonomous P elements encode transposase, a *trans*-acting function that is necessary for P element transposition and excision (Spradling and Rubin, 1982; Engels, 1984). P elements have been analyzed at the molecular level and found to range in size from less than 500 bp up to the 2.9 kb autonomous P element (O'Hare and Rubin, 1983). A typical P strain carries 30–50 P elements, of which about one-third are 2.9 kb. The short nonautonomous P elements do not encode transposase activity but can be mobilized in *trans* when a source of transposase is supplied. All P elements have 31 bp inverted terminal repeats and create a direct 8 bp duplication of genomic target sequences upon insertion (O'Hare and Rubin, 1983). P element transcripts are unidirectional and contain sequences from four major open reading frames (O'Hare and Rubin, 1983; Karess and Rubin, 1984).

P elements can be used as vectors for the germline transformation of Drosophila (Spradling and Rubin, 1982; Rubin and Spradling, 1982). This method has facilitated the molecular analysis of P elements for two reasons. First, a single P element can be introduced into the genome of an M strain devoid of other P elements. Second, it enables one to mutate the P element in vitro, reintroduce it back into flies, and then genetically test the effect of the mutation on P element function. Using this strategy Karess and Rubin (1984) have shown that all of the four open reading frames are required to encode transposase. They placed frameshift mutations in each of the four open reading frames, re-introduced the mutated P elements into Drosophila, and found that all four mutants lack transposase activity. Furthermore, the four open reading frame mutants did not complement each other in any combination, suggesting that information from all four open reading frames was used to form one polypeptide.

In this paper we report the structure of P element poly(A)$^+$ RNA transcripts. We show that P element transcription in somatic tissues is not sufficient to allow somatic transposition. We also present genetic evidence demonstrating that the P element transcript has a germline-specific splice, that this splice is required for transposase activity, and that it is the sole basis for the germline restriction of P element transposition.

Results

The Germline Specificity of P Element Transposition Is Not Regulated at the Transcriptional Level

An hsp70-P Element Fusion Lacks Transposase Activity in Somatic Cells

The mechanism that restricts P element transposition to the germline might function at the transcriptional or post-transcriptional level. That is, P elements might be transcribed exclusively in the germline by virtue of a germline-specific promoter or enhancer element, thus limiting transposase production, and hence transposition, to the germline. Alternatively, P elements might be transcribed in all tissues, but functional transposase made only in the germline due to a block in other tissues at a subsequent step, such as RNA processing, translational control, or posttranslational covalent modifications. There is also the possibility that a germline-specific, non-P-element-encoded protein is required for transposition. As a first step toward distinguishing among these possibilities, we have placed the P element protein coding sequences under the transcriptional control of a promoter element known to be active in both somatic and germline tissue,

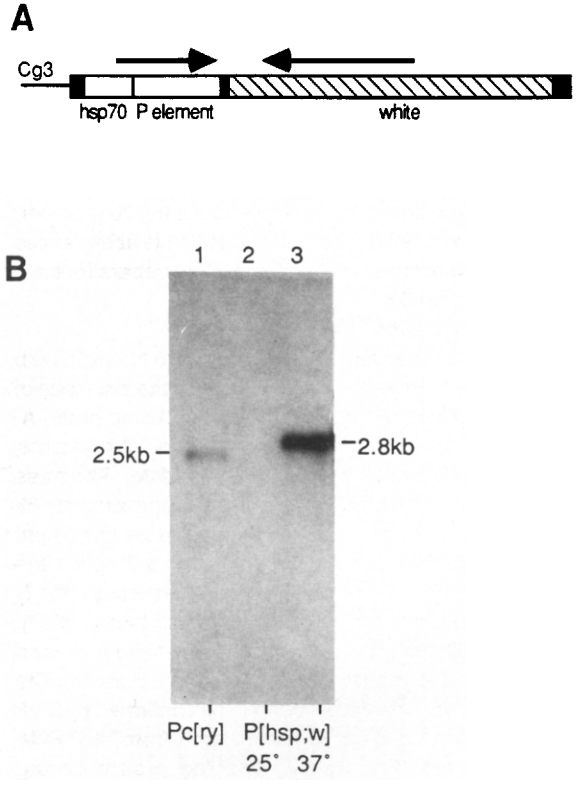

Figure 1. Structure and Expression of the Heat Inducible P Element Derivative P[hspP;w^+]

(A) Structure of P[hspP;w^+]. This vector carries a 5.2 kb Bam HI fragment containing the hsp70–P element fusion gene previously described by Rio and Rubin (1985) and indicated by the white (hsp70) and stippled (P element) boxes. It also carries a 10.7 kb Eco RI to Kpn I DNA fragment containing the white (w^+) gene (Levis et al., 1985). Both of these DNA fragments are inserted between the Eco RI and Kpn I sites of the Carnegie 3 P element transformation vector (Rubin and Spradling, 1983; see Experimental Procedures). The arrows indicate the direction of transcription of the hsp70–P element fusion gene and the *white* gene. The black boxes indicate the 31 bp P element inverted repeats. Note that only 22 bp of the inverted repeat is present between the hsp70–P element fusion gene and the white gene.

(B) RNA blot hybridization analysis. RNA was isolated from adult transformants carrying the P[hsp;w^+] element after a 1 hr heat shock (37°C; lane 3) or without a heat shock (25°C; lane 2). RNA was also isolated from embryos of the Pc[*ry*]2 line (Karess and Rubin, 1984) carrying a single 2.9 kb P element derivative (lane 1).

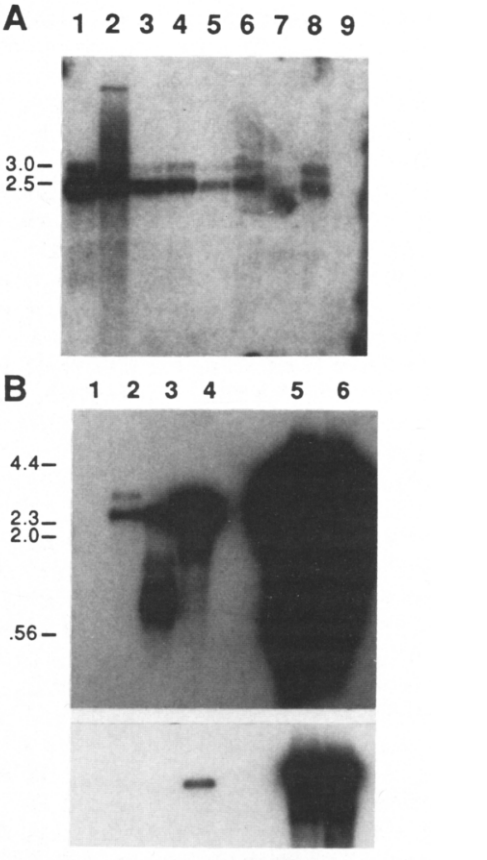

Figure 2. RNA Blot Hybridization to P Element Transcripts

(A) Lanes 1–7 have 1 μg of poly(A)$^+$ RNA isolated from a Pc[*ry*] strain at various stages of Drosophila development. Lane 1 has RNA isolated from 0–12 hr embryos; lane 2, 13–24 hr embryos; lane 3, first instar larvae; lane 4, second instar larvae; lane 5, third instar larvae; lane 6, pupae; lane 7, adults. Lane 8 contains 1 μg of poly(A)$^+$ RNA isolated from adult heads of a Pc[*ry*] strain. Lane 9 contains 1 μg of poly(A)$^+$ RNA isolated from 0–12 hr Canton S (M strain) embryos. The RNA blot was probed with P-element-specific probes. The 3.0 kb and 2.5 kb P element transcripts are indicated.

(B) Lanes 1–4 have 2 μg of poly(A)$^+$ RNA isolated from 0–12 hr embryos from Canton S (Lane 1), a Pc[ry] transformant(Lane 2), π_2(Lane 3), and a P[hspP;w^+] transformant (lane 4). Lane 5 and lane 6 have 1 μg of poly(A)$^+$ RNA isolated from NHP tissue culture cells. A shorter exposure is shown at the bottom of the figure. Marker sizes are in kilobases.

the hsp70 promoter (Lis, Simon, and Sutton, 1983; Bonner et al., 1984). If the P element promoter were germline-specific and if the germline restriction of transposition was due solely to this specificity, then expression of P element coding sequences from the hsp70 promoter should result in somatic P element transposition.

We have previously characterized expression of a gene fusion between the hsp70 promoter and P element coding sequences (Rio and Rubin, 1985). For the present experiments, this gene fusion was inserted next to the *white* eye color gene (w^+) in a P element vector to create P[hspP;w^+] (Figure 1). Upon microinjection of P[hspP;w^+] DNA into homozygous w^- embryos, two independent w^+ (red eyed) transformants were obtained. Genetic analysis indicated that each transformant line carried a single insertion of P[hspP;w^+]. Furthermore, heterozygous stocks of both transformants gave rise to spontaneous w^- (white eyed) revertants at high rates (1%–10%, even in the absence of heat shock) at each generation, indicating that the P[hspP;w^+] element was capable of catalyzing its own germline excision. If somatic excision events were also occurring, they would be detectable as clonal patches of w^- tissue on an otherwise w^+ (red eye color) background. However, no such somatic mosaics were observed in greater than 20,000 progeny from these two P[hspP;w^+] transformant lines, indicating a lack of somatic P element transposase activity. A more extensive genetic analysis, as well as the effect of heat shock on transposition of this element, will be published elsewhere (F. Spencer et al., unpublished).

We analyzed transcription from the hsp70–P element fusion gene in the P[hspP;w^+] transformed lines by blot hybridization of RNA from adult flies isolated after (Figure 1B, lane 3) or without (Figure 1B, lane 2) a 37°C heat shock. A dramatic increase in a 2.8 kb poly(A)$^+$ RNA is seen following heat shock. This RNA is the expected size for this construction (Rio and Rubin, 1985), and is approximately tenfold more abundant than the RNA from the single copy P element transformant Pc[*ry*]2 (Karess and Rubin, 1984; Pc[*ry*] RNA in Figure 1B, lane 1). The hsp70 promoter is active in both somatic and germline tissues, therefore we presume that a substantial fraction of the 2.8 kb mRNA is derived from adult somatic tissue and hence that P element coding sequences are being transcribed there. However, the absence of any somatic P element activity suggests that the germline restriction of P element transposition is not regulated solely at the level of P element transcription.

The P Element Promoter Is Active in Somatic Tissues

The role of the P element transcriptional promoter in tissue specificity was further studied by analyzing RNA isolated from various stages of Drosophila development by the RNA blot hybridization procedure. RNA isolated from adult heads, which are composed totally of somatic tissue, was also analyzed. Natural P strains contain in addition to the 2.9 kb P element, heterogeneous smaller P elements, which greatly complicate the molecular analysis of P transcripts. We therefore decided to use a strain, Pc[*ry*]2, that contains a single full length copy of Pc[*ry*] and no defective P elements (Karess and Rubin, 1984). Pc[*ry*] contains a 2.9 kb autonomous P element that is marked with the *rosy* eye color gene (see Figure 4A) and has wild-type transposase activity (Karess and Rubin, 1984). As shown by Karess and Rubin (1984), poly(A)$^+$ RNA isolated from 0–12 hr embryos of the Pc[*ry*]2 strain contains 2.5 and 3.0 kb P-element-specific transcripts (Figure 2A, lane 1). The analysis of RNA isolated from various developmental stages shows that these two transcripts are present throughout development (Figure 2A, lanes 1 through 7) and that they are both present in an RNA sample prepared exclusively from somatic tissue (head RNA, lane 8). We conclude that the P element promoter does not control the tissue specificity of P element transposition since P element transcripts are present in somatic tissues yet no transposase activity is observed there.

Structure of P Element Transcripts

The structure of P-element-encoded transcripts provided a clue to understanding the germline-restriction of transposition. Our strategy for mapping the transcripts was as follows: RNA from Pc[*ry*]2 was hybridized to ^{32}P-labeled complementary RNA made in vitro from a phage SP6-P element recombinant plasmid (see Figure 7 and Experimental Procedures). After hybridization, the RNA was treated with ribonuclease (under conditions where single-stranded but not double-stranded RNA is digested), denatured, and the sizes of the protected fragments determined. In addition to RNA isolated from the Pc[*ry*]2 strain, we analyzed embryonic RNA obtained from Canton S, a Drosophila M strain that lacks P elements, and π_2, a natural Drosophila P strain that carries both the 2.9 kb and heterogeneous smaller P elements. Because transcripts from the single P element in the Pc[*ry*]2 strain only constitute about 0.001% of the poly(A)$^+$ RNA in embryos (Karess and Rubin, 1984), we also analyzed RNA from NHP, a Drosophila tissue culture cell line which contains P element sequences transcribed from a hsp70 promoter (Rio and Rubin, 1985). The sizes and relative abundances of the P element transcripts in these RNA preparations are shown in Figure 2B.

ORF0–ORF1 Splice

Ribonuclease protection analysis of the 2.5 kb and 3.0 kb transcripts from Pc[*ry*]2 embryos detected the existence of two introns. The first intron was detected using probe A, a single-stranded RNA probe that spanned the junction between ORF0 and ORF1 (Figures 3a and 3e). The sizes of the ribonuclease protected products (approximately 85 and 100 nucleotides) in the Pc[*ry*]2 RNA were consistent with a 5′ splice site at nucleotide 442 and a 3′ splice site at nucleotide 501 (where consensus sequences for 5′ and 3′ splice sites are found; Mount and Steitz, 1984). Ribonuclease analysis showed that RNA from π_2 and NHP are also spliced at the same position (Figure 3a). The exact location of the splice junction was confirmed by DNA sequence analysis of cDNA clones made from NHP RNA (see Figure 3f). This splice joins the protein coding regions of ORF0 and ORF1 together into one continuous translational reading frame.

ORF1–ORF2 Splice

The second splicing event found in the P element transcripts joins the protein coding region of ORF1 to ORF2 in frame. Using probe B, an RNA probe that spanned the ORF1–ORF2 junction, ribonuclease protection analysis was consistent with the 5′ and 3′ splice sites of this intron being located at positions 1168 and 1222, where appropriate consensus sequences occur (Figures 3b and 3e). As with the first intron, this intron is also removed in π_2 and NHP RNA (Figure 3b). The precise location of the splice junction was again confirmed by DNA sequence analysis of a cDNA clone made from NHP RNA and also by direct sequencing of the NHP RNA using an oligonucleotide primer and reverse transcriptase (data not shown; see Figure 3g).

Alternate 5′ Splice Site for the ORF1–ORF2 Splice

In our analysis of cDNA clones made from NHP RNA, we discovered a cDNA clone that had an ORF1–ORF2 splice junction different from the one mapped above. This alternate splice used the same 3′ splice site as in the ORF1–ORF2 splice, at nucleotide 1222, but the 5′ splice site was located at 1156, 12 nucleotides 5′ of the junction mapped above (see Figure 3g). This splice puts ORF1 and ORF2 in frame but deletes four amino acids from the predicted polypeptide sequence. There is an acceptable 5′ splice site consensus sequence at this position so we believe this clone reflects an actual splicing event and is not an artifact created during cDNA cloning. However, there is no proof that this RNA is present in natural P strain flies and it may be the product of a cryptic splice resulting from the overexpression of the P element transcript in tissue culture cells.

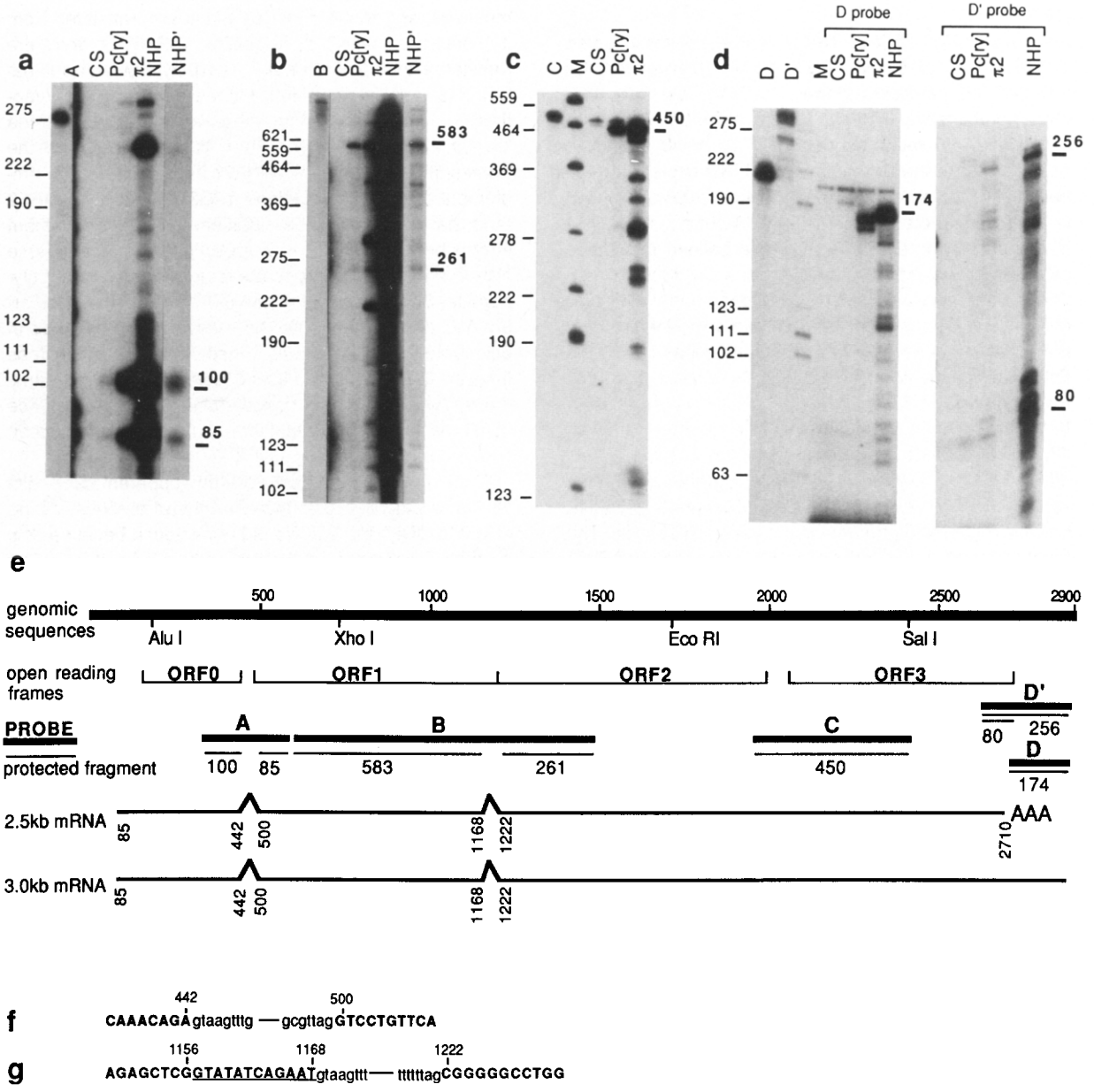

Figure 3. Ribonuclease Protection Analysis of P Element Transcripts

(a–d) Five ^{32}P-labeled single-stranded RNA probes (a, probe A; b, probe B; c, probe C; d, probe D and D′; the locations of these probes are shown in Figure 3e) were independently annealed to four different RNAs: 0–12 hr Canton S RNA (CS in a, b, c, d); 0–12 hr Pc[*ry*] transformant RNA(Pc[ry] in a, b, c, d); 0–12 hr π_2 RNA(π_2 in a, b, c, d); NHP tissue culture RNA (NHP in a, b, d). Samples were treated with ribonuclease and the protected products resolved on an 8% denaturing polyacrylamide gel. Lanes A, B, C, D and D′ have their respective untreated probe. Lane NHP′ in (a) and (b) is a shorter exposure of the NHP lane adjacent to it. Lane M is a Dde I digested ^{32}P-labeled polyoma DNA marker lane. Marker sizes to the left of each panel are in base pairs. The protected fragments and their approximate sizes are marked to the right of each panel. These sizes are not meant to represent the precise size (to the nucleotide) of the protected fragments but rather the exact sizes predicted for the protected fragments.

(e) Schematic presentation of the data shown in (a) through (d). The location of probes A, B, C, D, and D′ are shown in relationship to the ORFs in the P element. The locations of the ribonuclease protected fragments are also shown. The predicted structures for the 2.5 kb and 3.0 kb Pc[*ry*] transcripts are shown. The 2.5 kb mRNA initiates at about position 85 and polyadenylates at about position 2710. The sequences 443 to 499 and 1169 to 1221 are removed by mRNA splicing. The 3.0 kb mRNA is identical to the 2.5 kb mRNA except that it polyadenylates at a site outside of P element. The Alu I, Xho I, Eco RI, and Sal I restriction enzyme sites show the locations of the frameshift mutations constructed by Karess and Rubin (1984) that demonstrated that all four ORFs are required for the production of P element transposase.

(f) DNA sequence surrounding ORF0–ORF1 splice junctions. Intron sequences are in lower case.

(g) DNA sequence surrounding ORF1–ORF2 splice junctions. Intron sequences for the ORF1–ORF2 splice are in lower case. When the alternate 5′ splice site is used the 12 underlined nucleotides are also contained in the intron.

ORF3

As shown in Figures 3c and 3e, no splice site was detected in the 5′ half of ORF3. Probe C, which spanned the ORF2–ORF3 junction and the 5′ half of ORF3, was completely protected by the Pc[*ry*]2 RNA, except for the digestion of the SP6 vector sequences. This result raises the question of how the information in ORF3 is expressed (see below). It is interesting to note that about 20% of probe C is completely protected by the π_2 RNA but a majority of it is digested to smaller products. We believe that these products do not map splice junctions but instead map DNA deletions in the smaller nonautonomous P elements present in π_2. These RNAs are also detected with probe B (π_2 lane in Figure 3b), illustrating the necessity of using Pc[*ry*] transformants to map the P element transcripts.

Polyadenylation Site

Karess and Rubin (1984) showed that the 2.5 and 3.0 kb RNAs differed at their 3′ ends, and therefore most likely differed at their polyadenylation sites. Ribonuclease analysis of the Pc[*ry*]2, π_2 and NHP RNAs with 3′ end RNA probes (Figures 3d and 3e) indicates that this is true. The results detect two overlapping RNAs; a smaller RNA which terminates at about position 2710 and a larger RNA which spans the entire 3′ end of the P element. From this data and the RNA blot results of Karess and Rubin (1984) we conclude that the 2.5 kb RNA uses the polyadenylation signal (AATAAA) at position 2696 and polyadenylates 10 to 20 nucleotides downstream of it. The 3.0 kb RNA results from readthrough of this polyadenylation signal. The exact termination site of the 3.0 kb RNA has yet to be determined but it most likely spans the entire P element sequences in Pc[*ry*] and terminates in the adjoining *rosy* DNA sequences. The leakiness of the polyadenylation site in the P element is not limited to the Pc[*ry*] construct as readthrough products were also observed in π_2 RNA and NHP RNA (Figure 3d).

Structure of the P Element Transcript

The complete structure of the P element transcripts in the Pc[*ry*]2 strain are diagrammed in Figure 3e. The 2.5 kb transcript begins at nucleotide 85 (as mapped by Karess and Rubin, 1984) ends at about 2710, and is spliced twice at the positions shown. The removal of these introns connect ORF0, 1, and 2 into one continuous open reading frame. If translated from the first AUG in the transcript the protein product would initiate at the ATG at position 153, have a molecular weight of 66,000 kilodaltons, and terminate at the 3′ end of ORF2 (position 1991). We believe the 3.0 kb transcript is identical to the 2.5 kb transcript except that it has extended through the normal P element polyadenylation signal and is terminated at a site adjacent to the P element sequences.

Expression of ORF3 Information

The above results suggest that the protein made from P element transcripts contains information encoded from ORFs 0, 1, and 2 but not from ORF3. However, the genetic data of Karess and Rubin (1984) showed that sequences from all four ORFs are required for transposase activity. How is the information in ORF3 expressed? One possible answer is that the P element transcript is a polycistronic message, and that two polypeptides are translated from it. More specifically it is possible that ORFs 0, 1, and 2 are translated together to make one polypeptide and that ORF3 is translated separately to make a second polypeptide. However this is an unattractive hypothesis for three reasons. First, Karess and Rubin (1984) showed that the frameshift mutation at the Sal I site in ORF3 does not complement frameshift mutations in ORF0, ORF1, or ORF2 (and that ORF0, 1, and 2 mutations do not complement each other). Therefore, all four ORFs belong to the same cistron and most likely encode regions of the same polypeptide. Second, polycistronic messages that use different ATG translational initiation sites are extremely rare in eukaryotes (Baltimore, 1971). Finally, although ORF3 does have an ATG codon at its 5′ end (position 2059), the flanking sequences of this codon do not fit the sequence consensus for a typical eukaryotic translational initiation codon (Kozak, 1984).

The simplest way to reconcile this apparent contradiction is to postulate that there is indeed a splice joining ORF2 to ORF3 but that we did not detect it because only a small percentage of the transcripts have this splice. Given that P element transposition is restricted to the germline, it may be that the ORF2–ORF3 splice is germline-specific. The Pc[*ry*]2 and π_2 RNA studied were isolated from 0–12 hr old embryos which are composed of less than 2% germline tissue. Thus if the splice only occurs in germline cells we would not have detected it above the background of unspliced somatic RNAs. What is most attractive about this hypothesis is that it would also explain why P element transposition is limited to germline cells. Assuming that ORF3 is required for transposition, that a splice into ORF3 is required to express ORF3, and that this splice occurs only in germline cells, then it follows that transposition would only occur in the germline. This germline splicing hypothesis was tested genetically using P element mutants constructed in vitro.

An ORF2–ORF3 Splice Is Required for Transposase Production

Site-Directed Mutagenesis

Enough is known about the consensus sequences of mRNA splice junctions to predict potential splice sites with reasonable accuracy (Breathnach and Chambon, 1981; Sharp, 1981; Mount, 1982; Mount and Steitz, 1984). We therefore searched for a 5′ splice site located near the 3′ end of ORF2 and a 3′ splice site located near the 5′ end of ORF3, positioned such that removal of the putative intron would maintain the open translational reading frame. The best fit was found with a 5′ splice site at position 1947 and a 3′ splice site at 2138 (Figure 4B). The putative 5′ splice site has the mandatory consensus GT dinucleotide and has a 6 bp out of 9 bp homology with the consensus sequence proposed by Mount and Steitz (1984). The putative 3′ splice site has the mandatory AG dinucleotide and was preceded by the consensus polypyrimidine tract (Mount and Steitz, 1984).

If the above sites are actually used in splicing then a mutation at either site should abolish utilization of these splice sites (Montell et al., 1982; Treisman et al., 1982;

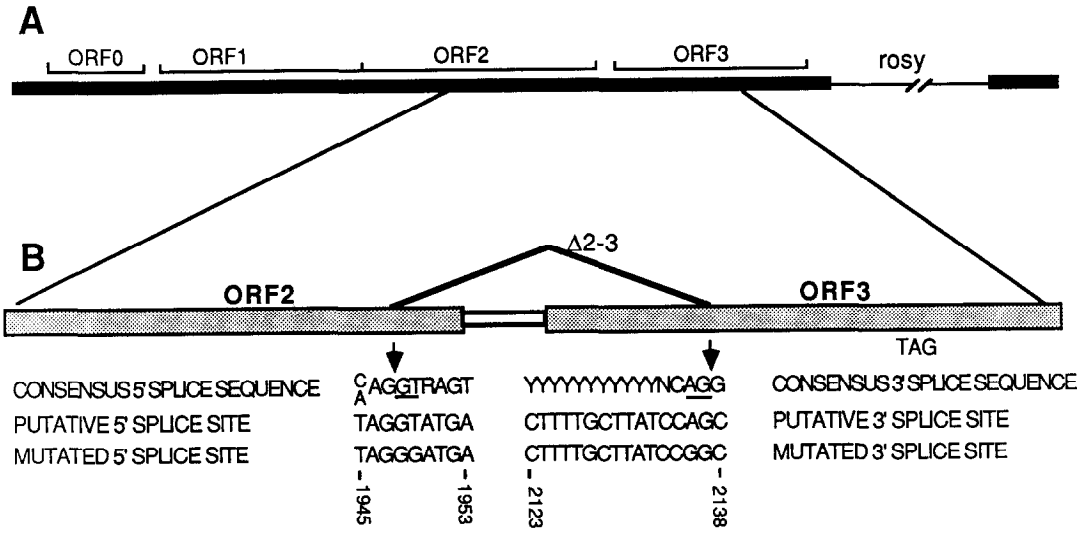

Figure 4. Site-Directed Mutagenesis of Pc[*ry*]

(A) Structure of Pc[*ry*]. The heavy line represents P element sequences. The 7.2 kb Hind III *rosy* fragment is represented by the thin line. The structure of Pc[*ry*] is described in detail in Karess and Rubin (1984).

(B) Site-directed mutagenesis of Pc[*ry*]. The region of the P element showing the ORF2–ORF3 junction is shown. Four site-directed mutations were made to test the germline-specific splicing theory. The putative 5′ splice site (which has 6 bp out of 9 bp homology to the consensus 5′ splice sequence) located between positions 1945 and 1953 is shown. The consensus GT at this site (underlined in the figure) was mutated to a GG (mutation 1949G) in order to test whether it is used as a splice junction. The putative 3′ splice site, which has good homology with the consensus 3′ splice sequence, is located between positions 2123 and 2138. The consensus AG was mutated to a GG (mutation 2136G) in order to test whether it is used as a splice junction. For both the 5′ and 3′ splice sites an arrow points to the predicted location of the splice junction. A third mutation (Δ2-3) which precisely deletes the proposed intron sequences (1948 to 2137) is shown. The location of a mutation that incorporates a nonsense mutation (TAG) into ORF3 (2340G) is also displayed.

Wieringa et al., 1983; Montell et al., 1984), and thereby abolish synthesis of functional transposase. Each site was mutated using primer-directed mutagenesis of an M13 phage recombinant (see Experimental Procedures). The consensus GT at the putative 5′ splice site was mutated to a GG and the consensus AG at the putative 3′ splice site was mutated to a GG (Figure 4B). Both of these point mutations (named 1949G and 2136G respectively) are in the third position of a codon and would not change the encoded amino acid if ORF2 and ORF3 were translated from unspliced RNA. In addition to the above two point mutations, we also constructed a deletion mutation that precisely removes the entire 190 base pairs (1948 through 2137) of the putative intron (Figure 4B). If the ORF2 to ORF3 splice is required to make transposase and uses the 5′ and 3′ splice sites proposed above, then this deletion mutation (named Δ2-3) should still retain transposase activity.

Site-directed mutagenesis was also used to confirm the conclusion of Karess and Rubin (1984) that translation of ORF3 is required for transposase activity, a conclusion that was based on a single mutation. At position 2340 in ORF3 the T was mutated to a G. This mutation (named 2340G) creates an amber stop codon (TAG) in the middle of ORF3 and should destroy transposase activity if the translation of ORF3 is required for transposase production (Figure 4B).

Transposase Assay

Our genetic assay for transposase activity is identical to that used by Karess and Rubin (1984) to assay Pc[*ry*] and Pc[*ry*] derivatives. Each of the four site-specific mutations described above were cloned into Pc[*ry*] after the removal of the equivalent wild-type sequences. Thus, these recombinant P elements were identical to Pc[*ry*] except that they now contained a site-directed mutation. The mutant P element derivatives were then introduced into the Drosophila melanogaster germline by P-element-mediated transformation. At least three independent transformants were isolated for each of the four mutant P elements. These transformant lines were then tested for transposase activity by their ability to destabilize the nonautonomous P elements of the *singed-weak* allele (sn^w) of the singed bristle locus. sn^w is hypermutable, is located on the X chromosome, and results from the insertion of two P elements in inverted orientation (Engels, 1984; Roiha et al., unpublished). In the absence of P element transposase the sn^w locus is very stable; however, when a source of P element transposase is provided in *trans*, sn^w is hypermutable to two new phenotypes (singed-plus,sn^+, and singed-extreme,sn^e; Engels, 1984). A single autonomous P element can result in rates of sn^w destabilization of approximately 10% (Spradling and Rubin, 1982; Karess and Rubin, 1984). The genetic crosses used to assay transposase activity among the transformants are diagrammed in Figure 5A.

As shown by Karess and Rubin (1984), a Pc[*ry*] element carrying wild-type P element sequences is able to destabilize the sn^w allele (Table 1). All of the 57 G1 progeny tested (from four independent transformants) had progeny that were either sn^+ or sn^e. sn^+ and sn^e are singed al-

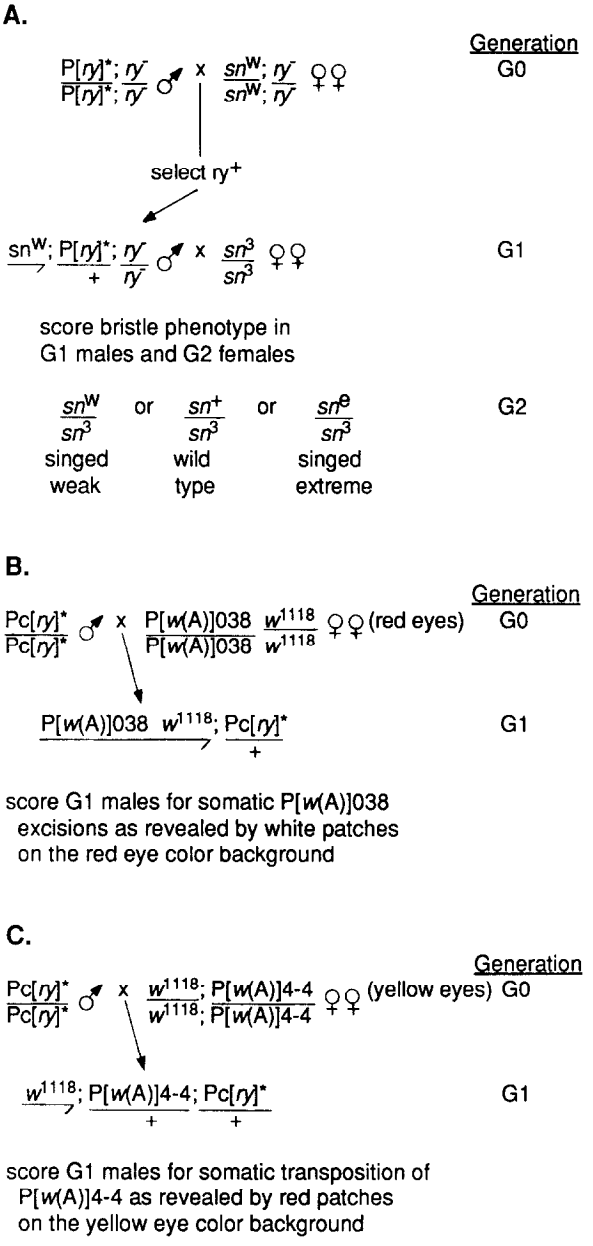

Figure 5. Protocol for Genetic Assays for P Element Transposase Activity

(A) Assay for *sn*^{w} destabilization. A male carrying the Pc[*ry*]* (Pc[*ry*] or its derivative) element is crossed to a *ry*$^-$ M strain female homozygous for the *sn*^{w} allele. The ability of the Pc[ry]* to destabilize the *sn*^{w} allele in the germline of the ry$^+$ G1 hybrids was assayed by crossing the G1 *sn*^{w} males to tester females homozygous for *sn*^{3}, a recessive severe allele of the singed locus. The phenotypes of the G2 female offspring are exclusively *sn*^{w} if no *sn*^{w} destabilization has occurred, but if the Pc[*ry*]* element is capable of providing transposase activity some are *sn*$^+$ or *sn*^{e}. The G1 *sn*^{w} males, which have one copy of *sn*^{w} and one copy of Pc[*ry*]*, were also scored for somatic excision events which would result in a mosaic bristle pattern. The Pc[*ry*]* stocks contained between one and several copies of the transposon and were a mixture of heterozygous and homozygous flies.

(B) Assay for P element somatic excision. A male carrying a Pc[*ry*]* element is crossed to a P[w(A)]038 female which is homozygous for a transposon carrying a *white*$^+$ allele on its X chromosome. The *white*$^+$ allele gives these females red eyes. G1 males have one X chromosome and the one copy of the P[*w*] transposon. Excision of this copy during the development of the eye will result in a mosaic pattern of white clonal patches on a red eye color background.

leles that result from the excision of a P element from the *sn*^{w} allele (Roiha et al., unpublished). However, of the 48 G1 progeny tested that carried the amber mutation in ORF3 (2340G), none showed any evidence of inducing *sn*^{w} instability. This confirms the earlier work of Karess and Rubin (1984) showing that ORF3 is required for transposase activity.

Neither the 5′ (1949G) nor the 3′ (2136G) putative splice site mutants had significant levels of transposase activity. In both cases *sn*^{w} destabilization decreased by at least 100-fold (Table 1). This indicates that both sites are indeed used as splice sites. The *sn*^{w} assay also showed that the Δ2-3 mutant had wild-type or slightly elevated levels of transposase activity (Table 1). Consistent with the results of the *sn*^{w} assay, P[*ry*(Δ2-3)] was found to be able to autonomously transpose in the Drosophila germline (data not shown). We conclude from these experiments that there is indeed a third splice in the P element transcript, that this splice joins ORF2 to ORF3 at the positions predicted in Figure 4B, and that this splicing event is required for transposase synthesis.

P[*ry*(Δ2-3)] Has Transposase Activity in Somatic Tissue

If the ORF2–ORF3 splice is the regulatory event that limits P element transposition to germline cells, then it follows that the Δ2-3 P element should have transposase activity not only in the germline but also in the soma. This prediction was tested using three different genetic assays. The first, an assay for P-element-mediated somatic excision events, uses the genetic cross diagrammed in Figure 5A. The G1 ry$^+$ males will have one copy of Pc[*ry*], or its mutant derivatives, and one copy of *sn*^{w}. When wild-type Pc[*ry*] is used in this cross, all the G1 ry$^+$ males have sn^w bristles because transposase is produced only in their germlines and not in the somatic cells that give rise to the bristles. However, when P[*ry*(Δ2-3)] was used, greater than 90% of the G1 ry$^+$ males were somatic mosaics, having a mixture of sn^w, sn$^+$, and sn^e bristles on the same individual.

A second assay for somatic excision (outlined in Figure 5B) uses the P[*w*] transformant P[*w*(A)]038. P[*w*(A)]038 is homozygous for the *white* null allele *w*^{1118} but has wild-type red eyes because it is transformed with a wild-type copy of the *white* locus. This *white* locus is carried within a nonautonomous P element vector, and thus can be mobilized if, and only if, transposase is provided in *trans* (Levis et al., 1985). The eyes of flies having one copy of P[*w*] and one copy of P[*ry*(Δ2-3)] were examined (G1 males in Figure 5B). Because the *white* locus acts in a cell

(C) Assay for P element somatic transposition. A male carrying the Pc[*ry*]* element is crossed to a P[*w*(A)]4-4 female which is homozygous for a transposon carrying a *white*$^+$ allele on its third chromosome. These G0 females have a yellow eye color because of the chromosomal location of the P[*w*] transposon. If this transposon somatically transposes to a new site during the development of the eye it will result in a mosaic pattern of red clonal patches on the background yellow eye color in G1 males.

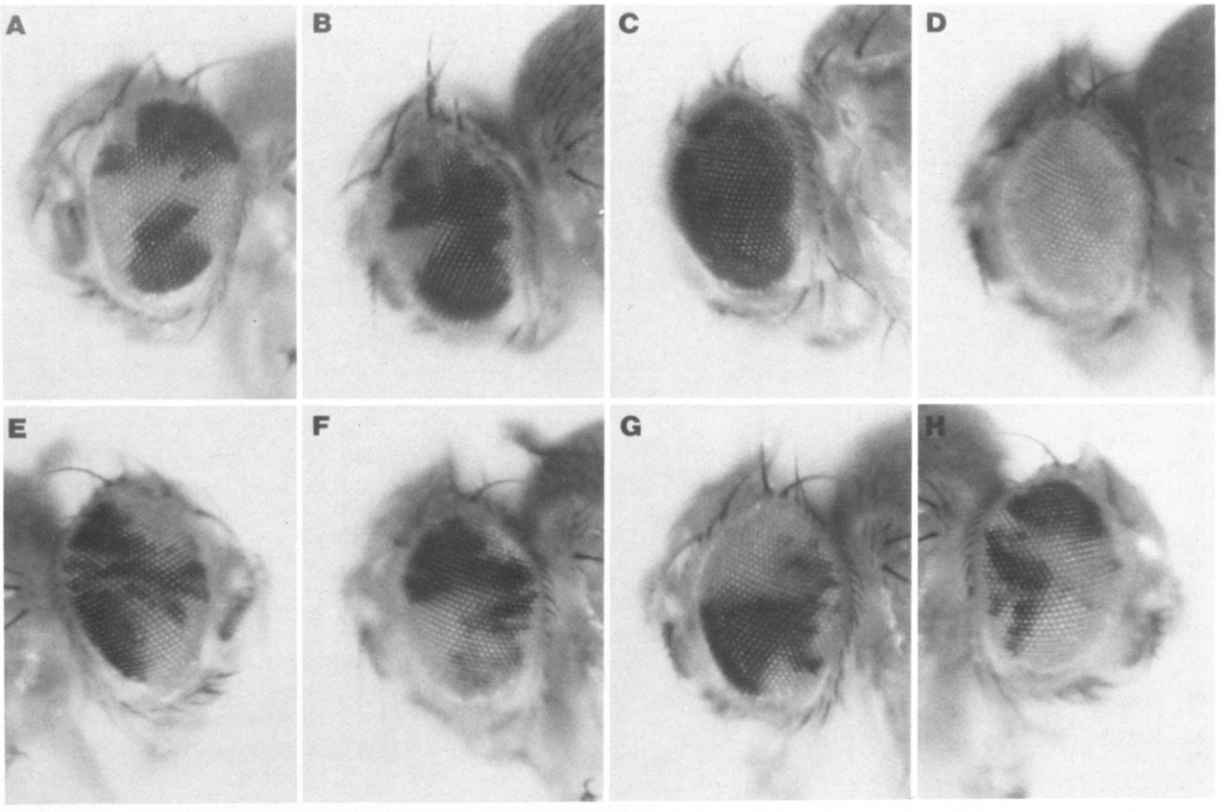

Figure 6. Assay for P Element Somatic Excision and Transposition in the Eye

(A, B, C) Eyes of male G1 progeny of P[*ry*(Δ2-3)] × P[*w*(A)]038 genetic cross.

(D) Eye of male G1 progeny of Pc[*ry*] × P[*w*(A)]4-4 genetic cross.

(E, F, G, H) Eyes of male G1 progeny of P[*ry*(Δ2-3)] × P[*w*(A)]4-4 genetic cross.

autonomous manner, the G1 males should have mottled eyes if *white* transposon excisions occur during growth of the cells destined to form the eye. As shown in Figures 6A, 6B, and 6C such excisions occur at high frequency. The eyes of G1 males contain multiple clonal white patches of w^- cells that apparently reveal the progeny of a cell that had an excision event of the P[*w*] transposon. Such mottled eyes occur in well over 50% of the G1 male progeny and no patches were ever observed in flies carrying the Pc[*ry*] control element.

The two assays for transposase activity described above scored somatic excision rather than forward transposition. A second *white* transformant line, P[*w*(A)]4-4, was used to test whether or not P[*ry*(Δ2-3)] could elicit a forward transposition event in somatic tissue. P[*w*(A)]4-4 is transformed with a wild-type allele of *white* but flies of this transformed line have a yellow eye color because the P[*w*] transposon is inserted in a region of the chromosome that inhibits full expression of the *white* gene (Levis et al., 1985). The assay for transposition asks whether or not G1 males of the genetic cross diagrammed in 5C will have mosaic eyes due either to transposition of the P[*w*] allele to a new chromosomal location or due to other local rearrangements (see Levis et al., 1985). Figure 6D shows a typical eye of a G1 male whose father carried Pc[*ry*]. The eye is pale yellow and not mosaic indicating that the transformed *white* allele could not transpose to a new chromosomal position. However, when the father was P[*ry*(Δ2-3)] greater than 50% of the male progeny had mosaic eyes such as those shown in Figures 6E–6H. We interpret these results as showing that during the development of the eye the *white* transposon moved to a new location where its expression is restored to wild-type.

Due to the high transposition and excision rates of P[*ry*(Δ2-3)], we were unable to isolate and maintain pure homozygous stocks of these transformants. Therefore all the assays for transposase activity used inbred stocks of P[*ry*(Δ2-3)] flies and these stocks contained a mixture of homozygous and heterozygous flies. We believe this is why some of the flies did not have mosaics eyes in the assays for somatic excision and transposition using P[*w*(A)]038 and P[*w*(A)]4-4. In the genetic crosses diagrammed in Figures 5B and 5C we were unable to use the rosy marker on P[*ry*(Δ2-3)] to tell which G1 males carried the transposon because these strains carry a wild-type *rosy* gene. We believe that the G1 males that did not have mosaic eyes most likely did not contain a copy of P[*ry*(Δ2-3)]. For the sn^w analysis we could use the *rosy* marker to detect which of the G1 males carried P[*ry*(Δ2-3)]. Of these males greater than 90% had a mosaic bristle pattern. We

Table 1. *Singed-weak* Destabilization by Pc[*rosy*] and Its Derivatives

Strain	G1 Progeny		G2 Progeny		
	Total No. Tested	No. Showing Germline sn^w Instability	Total	sn^+	sn^e
Pc[*rosy*]					
Pc[*ry*]7-2	18	18	1148	230 (20%)	205 (18%)
Pc[*ry*]9-1	15	15	1224	205 (17%)	166 (14%)
Pc[*ry*]4	14	14	734	56 (8%)	31 (4%)
Pc[*ry*]10	10	10	451	37 (8%)	41 (9%)
ORF3 Amber Mutant					
P[*ry*(2340G)]2	14	0	1252	0	0
P[*ry*(2340G)]4-1	19	0	1610	0	0
P[*ry*(2340G)]6	15	0	1424	0	0
5′ Splice Mutant					
P[*ry*(1949G)]2	15	0	1040	0	0
P[*ry*(1949G)]6	14	0	926	0	0
P[*ry*(1949G)]14-1	19	0	1027	0	0
3′ Splice Mutant					
P[*ry*(2136G)]5	15	0	1219	0	0
P[*ry*(2136G)]22-1	20	0	1215	0	0
P[*ry*(2136G)]23	15	1	1077	0	1 (.09%)
Δ2-3 Mutant					
P[*ry*(Δ2-3)]1	24	23	721	73 (10%)	80 (11%)
P[*ry*(Δ2-3)]3-2	16	15	489	182 (37%)	214 (44%)
P[*ry*(Δ2-3)]10-2	18	18	985	167 (17%)	206 (21%)
Δ1-2 Mutant					
P[*ry*(Δ1-2)]19	14	14	522	55 (11%)	56 (11%)
Δ123 Mutant					
P[*ry*(Δ123)]17	3	3	216	42 (19%)	17 (8%)
Δ1-2 Alternate 5′ Splice Mutant					
P[*ry*(Δ1-2A5′)]3	15	0	1267	0	0

believe the nonmosaic individuals can be explained by P element mediated somatic excisions of P[*ry*(Δ2-3)] early in the lineage of the cells that eventually develop into bristles.

Analysis of ORF1–ORF2 Introns

Even though it was not feasible to directly examine the structure of P element transcripts in germline cells, our genetic evidence indicates that a splice connects ORF2 to ORF3 in at least some germline P element transcripts. It is therefore possible that there are other splices or alternate splice junctions that also occur at levels too low to biochemically detect in RNA isolated from whole organisms. As a first step in addressing this question we have constructed Pc[*ry*] derivatives that lack the ORF1–ORF2 intron, the ORF1 alternate 5′ splice site–ORF2 intron, and a combination of the ORF1–ORF2 and ORF2–ORF3 introns. In these constructs a particular splicing pattern is locked into the P element genome thereby preventing the 5′ and 3′ splice sites from using other partners. Our results show that a P element that lacks the ORF1–ORF2 intron still retains germline-specific transposase activity (Table 1, Δ1-2 mutants). P elements lacking both the ORF1–ORF2 and ORF2–ORF3 introns have transposase activity in both the germline (Table 1, Δ123 mutant) and the soma (data not shown). We conclude that neither the 5′ nor the 3′ splice junction of the ORF1–ORF2 splice is used in an alternate splice in a transcript that is required for transposase activity. We also find that a Pc[*ry*] derivative carrying the 5′ alternative splice junction of the ORF1–ORF2 splice can not encode transposase activity (Table 1, Δ1-2 alternate 5′ splice mutant); however it is possible that a protein encoded by this alternatively spliced RNA may serve some other function.

Discussion

P element transposition can occur at high levels in the germline, but not in somatic tissue. We show here that this tissue specificity is controlled at the level of mRNA splicing. The most compelling evidence supporting our conclusion is that point mutations in the 5′ and 3′ sites of the putative ORF2–ORF3 intron abolish transposase activity and that a P element carrying a deletion that precisely removes this putative intron has transposase activity in the soma in addition to the germline.

Two results led us to carry out the experiments that indicated germline-specific splicing. First, the tissue specificity of P element transposition was shown not to be controlled at the level of transcription. Using the P[hspP;w^+] recombinant, which fuses a Drosophila heat shock promoter (hsp70) to the P element coding sequences, we

showed that transposition is still limited to germline cells even though the hsp70 promoter is known to be active in somatic cells. (Similar experiments reaching identical conclusions were done by Steller and Pirrotta, submitted.) We have also directly shown that the P element promoter is active in somatic tissue; P element transcripts were found in adult head RNA. These data argue strongly that the P element promoter is not involved in regulating the tissue specificity of transposition.

Second, our RNA mapping data suggested that alternative RNA splicing was involved in controlling transposase synthesis. The P element transcripts found in 0–12 hr embryos contained two introns that translationally connect the first three open reading frames, but did not have a splice joining ORF2 to ORF3. This result was unexpected because a frameshift mutation in ORF3 eliminated transposase activity and could not complement frameshift mutations in any of the three other ORFs (Karess and Rubin, 1984), suggesting that each of the four ORFs contributes to a single polypeptide having transposase activity. Our RNA mapping data, however, suggested that only ORFs 0, 1, and 2 were translated together into a single polypeptide. The most plausible resolution of this apparent paradox would be a germline-specific splicing event joining ORF2 to ORF3. This hypothesis is consistent with our failure to detect the splice in embryonic RNA which is primarily composed of somatic cells. It is also consistent with the results of Karess and Rubin (1984) because it suggests all four ORFs contribute to a single polypeptide in the germline. Furthermore, it can explain how the germline restriction of P element transposition is regulated.

Our approach to testing the germline splicing theory was to identify the splice sites by sequence criteria, specifically mutant them in vitro, reintroduce them into flies, and use a genetic assay to ask whether the mutations reduced transposase activity. Because we had never obtained any biochemical evidence for the splice it was not obvious where the splice junctions would be, if they existed at all. The genetic data predicted one polypeptide covering all four ORFs which suggested that the splice should connect ORF2 and ORF3 in frame. From data of Karess and Rubin (1984) we knew that the transposase must contain the information encoded at the Eco RI site in ORF2 and the Sal I site in ORF3; therefore, any intron must be between these two restriction enzyme sites. By analysis of the P element DNA sequence we selected the consensus 5′ and 3′ splice sites that best met these criteria (see Figure 4). A mutation at either of these sites was shown to dramatically reduce the transposase activity (Table 1), providing strong evidence that these sites are in fact used in an ORF2–ORF3 splice. Evidence for the germline specificity of this splice came from a deletion in Pc[*ry*] that precisely removes the entire 190 bp ORF2–ORF3 intron. This deletion mutant retains transposase activity in germline cells but now also has transposase activity in somatic tissue as well. We conclude that not only is there a germline-specific splice between ORF2 and ORF3 but that this splice is the sole factor governing the germline restriction of P element transposition.

There are now many examples of alternative mRNA splicing (for example, Montell et al., 1984; Breitbart et al., 1985; Falkenthal et al., 1985). However the P element case is different from most because it is an all or nothing event; that is, the intron either is removed or remains. For most cases of alternative splicing there is always a splicing event but different junctions are used. Whether this distinction has biological significance will not be known until the molecular mechanism (or mechanisms) of alternative splicing is elucidated. In the case of the P element ORF2–ORF3 splice, the possible mechanisms can be divided into two distinct classes. The first is that there is a factor present only in germline tissues which specifically recognizes the ORF2–ORF3 splice sites and catalyzes the splicing event. The most likely possibility for such a factor would be a new or modified small nuclear ribonucleoprotein (snRNP). snRNPs have been shown to be required during in vitro splicing reactions (Padgett et al., 1983; Krämer et al., 1984) and it has been shown that the snRNPs U1 and U2 bind to splice junctions in vitro (Mount et al., 1983; Black et al., 1985). It is possible that in germline tissues, U1, U2, or another part of the splicing machinery is specifically modified so that it can identify a new class of splice junctions, a class to which the ORF2–ORF3 splice belongs. The second possibility is that the tertiary structure of the P element RNA is different in somatic and germline tissues and this is what regulates the ORF2–ORF3 splice. The P element RNA transcript might normally be folded into a tertiary structure that prevents access of the normal splicing apparatus to the ORF2–ORF3 splice, but in germline tissue the structure of the RNA is altered allowing the normal splicing machinery to act. What can alter the structure of the RNA? There are many possibilities including a difference in ionic concentrations, another RNA which has a small region of complementarity to the P element transcript, or germline-specific RNA binding proteins. We wish to emphasize that in the first hypothetical mechanism, the primary difference between germline and somatic tissue is a difference in the splicing machinery, whereas in the second class of mechanisms, the primary difference is not directly involved with splicing. However, no matter what the mechanism, P elements would be under strong selective pressure to develop a way to limit their transposition to the germline in order to favor their rapid spread through a population without the deleterious effects to the host that would result from high levels of somatic transposition.

The difference between the two models (germline-specific splicing machinery vs. germline-specific alteration of the tertiary structure of P element transcripts) has implications as to whether or not other germline-specific genes are likely to be regulated by mRNA splicing. If the splicing machinery is specifically modified in the germline then these modifications must have evolved not for the benefit of P elements but rather for the regulation of other genes required for proper Drosophila development. However, if the P element simply takes advantage of specific conditions in the germline to alter its RNA structure to allow splicing, then there is no reason a priori to believe that other genes will be regulated in this way.

Our RNA blot analysis showed that P elements are tran-

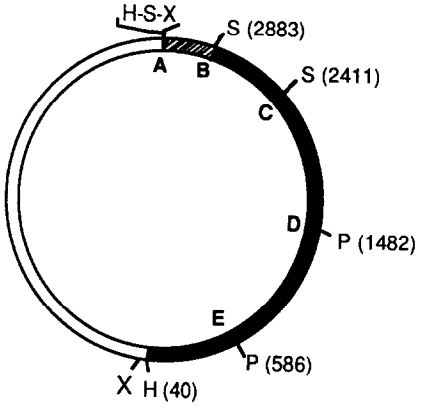

Figure 7. Structure of pPele[-0]

The white area represents pSP64 sequences, the cross hatched area represents pBR322 sequences, and the black area represents P element sequences. Restriction sites: H, Hind III; S, Sal I; P, Pvu II; X, Xba I. Sites A, B, C, D, and E are explained in Experimental Procedures.

scribed throughout development and in all tissues, raising the question of whether the P element promoter has any regulatory role at all. Our data suggest that it does not, but are inconclusive because we have never been able to study germline RNA directly due to the difficulty of obtaining large amounts of germline tissue.

Karess and Rubin (1984) identified the 2.5 kb and 3.0 kb transcripts from Pc[*ry*] transformants and suggested that the differences between the two may have a regulatory role. From the mapping data presented here we now think that this is unlikely. The two transcripts are identical through the protein coding regions up to the polyadenylation site. At this point the 2.5 kb transcript is polyadenylated, but the 3.0 kb message appears to be polyadenylated 300 to 400 nucleotides downstream in the adjacent *rosy* DNA sequences. We therefore believe that the presence of a discrete transcript of 3.0 kb is an artifact of the Pc[*ry*] construct. However, the leakiness of the P element polyadenylation site is real; it has been observed before in transcripts through an individual P element (Levis et al., 1984), and in total P element transcripts from the π_2 P strain (Figure 3d; Karess and Rubin, 1984). Whether this leakiness has any function is unknown.

From our mRNA mapping data we can predict the amino acid sequence of the P element transposase and of the predicted polypeptide product of the somatic cell P element transcript (which contains information from ORF0, 1, and 2 but not ORF3). These proteins and their activities are studied in the accompanying paper (Rio et al., 1986).

Experimental Procedures

Drosophila Strains

The Drosophila strains ry^{506}, *y* sn^3v/y^+Y, π_2, *y* sn^w are described in Karess and Rubin (1984). w^{1118}, and P[(*w,ry*)A^R]4-4 are described in Hazelrigg et al. (1984) and Levis et al. (1985). P[(*w,ry*)A^R]038 is the white transformant located at 10D in Figure 1 in Levis et al. (1985). Two Pc[*ry*] lines were used for RNA analysis; Pc[*ry*]2 (from Roger Karess) was used as a source of RNA for the ribonuclease protection analysis and Pc[ry^+]81.1/CyO (from Forest Spencer) was used for the RNA blot in Figure 2A.

Plasmid Construction

All nucleotide positions listed are in reference to the published sequence of the 2.9 kb autonomous P element (O'Hare and Rubin, 1983).

The plasmid pπ25.7wc (Karess and Rubin, 1984) was partially digested with both Hind III and Sal I and the P element fragment from 40 (Hind III site) to 2883 (where a Sal I site had been inserted at the Ava II site) was cloned into Hind III, Sal I cut pBR322. This recombinant was digested with Cla I and Nru I, the Cla I site was made blunt using the Klenow fragment of DNA-polymerase, and Xba I linkers were added. After Xba I digestion the fragment was ligated into the Xba I site of pSP64 (Melton et al., 1984) forming the plasmid pPele(-0) (see Figure 7). Subclones of pPele(-0) were then constructed to make the plasmids used for the ribonuclease protection analysis.

pPele(-1) was constructed by partial digestion of pPele(-0) with Sal I, removing the sequences between sites A and B (see Figure 7). pPele(-2) was constructed by complete digestion of pPele(-0) with Sal I, removing the sequences between sites A and C. pPele(-4) was constructed by digestion of pPele(-0) with Sal I, treating with Klenow DNA-polymerase, partial digestion with Pvu II, and ligating, thereby removing the sequences between sites A and D. pPele(-5) was constructed by digestion of pPele(-0) with Sal I, treatment with Klenow DNA-polymerase, digestion with Pvu II, and ligation, thereby removing the sequences between sites A and E.

The plasmid pP[hsp;w^+] was constructed from Cg3*w* (C. Zuker, unpublished). Cg3*w* carries the 10.7 kb Eco RI to Kpn I fragment from the white locus (O'Hare et al., 1984), inserted into the Eco RI and Kpn I sites of the Cg3 P element transformation vector (Rubin and Spradling, 1983). The plasmid pP[hspP;w^+] was made by cleaving Cg3*w* with Kpn I and treating with T4 DNA polymerase to create flush ends. Into this site, a 5.2 kb Bam HI fragment, from the plasmid pNHP (Rio and Rubin, 1985), that was treated with the Klenow fragment of DNA polymerase I to create flush ends, was inserted creating pP[hsp;w^+].

M13-R1-Sal was constructed by cloning the 669 bp Eco RI–Sal I fragment from Pc[*ry*] (Karess and Rubin, 1984) into Eco RI–Sal I digested M13mp18 (Yanisch-Perron et al., 1985). The four site-specific mutations [1949G, 2136G, 2340G, and Δ2-3] were cloned back into Pc[*ry*] using a three-way ligation involving the Eco RI–Sal I fragment from the M13-RI-Sal mutants, the 531 bp Kpn I (1070)–Eco RI (1601) P element fragment, and the 12 kb Sal I (2110)–Kpn I (1070) fragment from Pc[*ry*]. These plasmids are named pP[*ry*(1949G)], pP[*ry*(2136G)], pP[*ry*(2340G)], and pP[*ry*(Δ2-3)].

The plasmids pP[*ry*(Δ123)] and pP[*ry*(Δ1-2)] are derived from Pc[*ry*] (Karess and Rubin, 1984) as follows: Pc[*ry*] was first digested with Kpn I and Sal I and the 12 kb fragment was isolated. This Kpn I–Sal I fragment was ligated to a Kpn I to Eco RI cDNA fragment (which spanned the ORF1–ORF2 P element intron) and the Eco RI to Sal I fragment from either pP[*ry*(Δ2-3)] or Pc[*ry*]. This creates derivatives of Pc[*ry*] that lack either both the second and third P element introns (pP[*ry*(Δ123)] or only the second P element intron (pP[*ry*(Δ1-2)].

The plasmid pP[*ry*(1-2A5′)] was derived from Pc[*ry*] (Karess and Rubin, 1984) as follows. After cleavage with Kpn I and Sal I the 12 kb fragment was isolated and ligated to a mixture of a Kpn I to Eco RI cDNA segment carrying the alternative 5′ ORF1–ORF2 splice and the genomic Eco RI to Sal I P element fragment. This creates a derivative of Pc[*ry*] lacking the second P element intron that utilizes the alternative 5′ splice site.

Nucleic Acids

RNA preparations were as described in O'Hare and Rubin (1983). For RNA blot analysis poly(A)$^+$ RNA was dissolved in 10 μl of 50% formamide, 6% formaldehyde, 1× E buffer(18 mM Na_2HPO_4, 2 mM NaH_2PO_4), incubated at 60°C for 3 min, and 2 μl of dye mix (50% glycerol, 50% formamide, 1× E buffer, 0.1% bromophenol blue) was added. The RNA samples were loaded onto a 1.0% agarose gel that was 6% formaldehyde and 1× MOPs buffer (40 mM morpholinopropanesulfonic acid, pH 7.0, 10 mM sodium acetate 1 mM EDTA) and run for 13 hr at 25 volts in 1× MOPs buffer. The gel was soaked 20 min in 10× SSC and the RNA blotted onto Genescreen 2 (New England Nuclear). After baking for 2 hr the blots were probed with a pool of three different P-element-specific M13 single-strand probes (Karess and Rubin, 1984).

M13 single-strand probes were made by mixing 1 μg of M13 single-strand DNA with 1 μl 10× annealing buffer (100 mM Tris, pH 7.5, 500

mM NaCl, 10 mM β-mercaptoethanol, 50 mM $MgCl_2$) and 2 pmol of universal primer in 10 μl total volume. After incubation at 65°C for 10 min and then room temperature for 30 min the following were added: 2 μl 10× dNTP-buffer (1 mM dGTP, 1 mM dTTP, 250 mM Tris, pH 7.5, 150 mM $MgCl_2$, 10 mM β-mercaptoethanol), 10 μl of ^{32}P-dNTP mix (5 μl ^{32}P-dCTP, 10 mCi/ml, 5 μl of ^{32}P-dATP 10 mCi/ml, 0.5 μl 5 mM dGTP, 0.5 μl 5 mM dTTP; the ^{32}P-dNTPs are 800 Ci/mmol from Amersham), and 0.5 μl of DNA-polymerase-Klenow. After incubation at 37°C for 1 hr, 2 μl of 100 μM dATP-dCTP was added and the incubation was continued at 37°C for another 30 min. Salts were adjusted for endonuclease restriction digests and tenfold excess restriction enzyme added. After digestion an equal volume of loading buffer was added (60% DMSO in 10 mM Tris, pH 8.0, with bromophenol blue), the mixture placed into a boiling water bath for 3 min, on ice for 1 min, and loaded onto a 1.4% neutral agarose gel. After running the bromophenol blue to the bottom of the gel the DNA was located by autoradiography (1 min), cut out of the gel, and electroeluted in a dialysis bag.

Single-stranded RNA probes were made according to the procedures of Melton et al. (1984). Probe D was made using Dra I digested pPele(-1). Probe D has 27 bp of the pSP64 promoter and polylinker sequences, followed by P element sequences from the Ava II site (2883) to the Dra I site (2708). Probe D′ was made using Sca I digested pPele(-1). Probe D′ has 27 bp of the pSP64 promoter and polylinker sequences, followed by P element sequences from the Ava II site (2883) to the Sca I site (2630). Probe C was made using Dra I digested pPele(-2). Probe C has 27 bp of the pSP64 promoter and polylinker sequences, followed by P element sequences from the Sal I site (2411) to the Dra I site (1961). Probe B was made using Pvu II digested pPele(-4). Probe B has 24 bp of the pSP64 promoter and polylinker sequences, followed by P element sequences from the Pvu II site (1482) to the Pvu II site (586). Probe A was made using Dra I digested pPele(-5). Probe A has 24 bp of the pSP64 promoter and polylinker sequences, followed by P element sequences from the Pvu II site (586) to the Dra I site (342).

For ribonuclease analysis 1 μl of poly(A)$^+$ RNA (2 mg/ml in H_2O) was mixed with 2 μl of the single-stranded RNA probe (about 2 × 10^6 cpm in 80% formamide solution) and 17 μl of 80% formamide solution (80% formamide, 40 mM Pipes, pH 6.6, 400 mM NaCl, 1 mM EDTA). Samples were placed at 80°C for 5 min and then incubated 12–18 hr at 42.5°C. Ten microliters of the samples were then added to 200 μl of RNAase digestion buffer (10 mM Tris, pH 7.5, 5 mM EDTA, 300 mM NaCl) containing 20 units ribonuclease T1 or 20 units ribonuclease T1 and 0.1 μl RNAase A (1 mg/ml). After incubation for 1 hr at room temperature the samples were extracted with an equal volume of phenol/chloroform and ethanol precipitated. The pellets were resuspended in loading buffer, placed in a boiling water bath for 5 min, and loaded.

For ribonuclease protection analysis, probes that were identical to Probes A, B, and C, except that they had different 3′ ends, were used to determine the orientation of the protected fragments.

cDNA Cloning

Double-stranded cDNA was synthesized starting with 15 μg of poly(A)$^+$ mRNA from the heat-shocked NHP cell line as described (Huynh et al., 1985). Following S1 nuclease treatment and Eco RI methylase treatment, Eco RI linkers (8-mer, Collaborative Research) were added and the cDNA was cloned into Eco RI ant-λgt10 (Huynh et al., 1985). Following in vitro packaging in λ particles (Maniatis et al., 1982), the phage were plated on E. coli (600) and screened directly without amplification by plaque hybridization (Maniatis et al., 1982) using single-stranded M13 probes from various P element regions (O'Hare and Rubin, 1983). After plaque purification, phage DNA was prepared (Maniatis et al., 1982) and Eco RI fragments were subcloned into the pEMBL (Dente et al., 1983) plasmid vectors for DNA sequence analysis (Sanger et al., 1977) using short deoxyribonucleotides located 3′ to each P element intervening sequence as primers.

Site-Specific Mutagenesis

Mutagenesis was carried out using a minor modification of the procedure described by Hutchinson et al. (1978), Razin et al. (1978), and Kudo et al. (1981). Oligonucleotides used for mutagenesis were: AATTTGTCATCCCTATTATATA for 1949G; TATTCCTGGCCGGATAAGCAAA for 2136G; GTATTCCTGGCTATTATATATTTTC for Δ2-3; GCCCGCGATCTATTCCATCGC for 2340G.

P Element Transformation

Transformation of ry^{506} M-strain embryos was as described by Karess and Rubin (1984). Independent rosy$^+$ transformants were inbred to establish stocks. Transformation of w^{1118} embryos was as described by Hazelrigg et al. (1984).

Transposase Assays

The *singed-weak* destabilization assay was as described by Karess and Rubin (1984). For the somatic excision assay a male carrying either a Pc[*ry*] or P[*ry*(Δ2-3)] was crossed to P[(*w,ry*)A^R]038 (abbreviated P[*w*(A)]038) females and the G1 male progeny scored. For the somatic transposition assay a male carrying either a Pc[*ry*] or P[*ry*(Δ2-3)] transposon was crossed to P[*w,ry*)A^R]4-4 (abbreviated P[*w*(A)]4-4) females and the G1 male progeny scored.

Acknowledgments

We are grateful to Steve Mount for his advice on the location of possible splice sites in P element. We thank Forrest Spencer, Bruce Baker, Danielle Thierry-Mieg, Ulrike Heberlein, Michael Levine, Bruce Malcolm, Karen Ronan, and members of the Rubin lab for their help and advice. We extend our special thanks to Spyros Artavanis-Tsakonas and Barry Yedbovnik for their help and hospitality during the preparation of cDNA libraries at Yale. We appreciate the expert technical assistance of Todd Laverty, Patricia Zuker, and Katie Rendahl. We thank Herman Steller for discussing results prior to publication and Kevin Jones, Daniel Kalderon, Larry Kauvar, Todd Laverty, Robert Levis, Frank McNally, Dzrisilav Mismer, Mary Mullins, and Herman Steller for critical reading of the manuscript. F. A. L. is an Exxon Education Foundation Fellow of the Life Sciences Research Foundation. D. C. R. is a fellow of the Jane Coffin Childs Memorial Fund for Medical Research. This work was supported by grants from the National Institutes of Health to G. M. R.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 18, 1985

References

Baltimore, D. (1971). Polio is not dead. In From Molecules to Man, Perspectives in Virology, Volume 7. (New York: Academic Press).

Black, D. L., Chabot, B., and Steitz, J. A. (1985). U2 as well as U1 small nuclear ribonucleoproteins are involved in pre-mRNA splicing. Cell *42*, 737–750.

Bonner, J. J., Parks, C., Parker-Thornburg, J., Mortin, M. A., and Pelham, H. R. B. (1984). The use of promoter fusions in Drosophila genetics: isolation of mutations affecting heat shock response. Cell *37*, 979–991.

Breathnach, R., and Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. Ann. Rev. Biochem. *50*, 349–383.

Breitbart, R. E., Nguyen, H. T., Medford, R. M., Destree, A. T., Mahdavi, V., Nadal-Ginard, B. (1985). Intricate combinatorial patterns of exon splicing generate multiple regulated troponin T isoforms from a single gene. Cell *41*, 67–82.

Dente, L. G., Cesareni, G., and Cortese, R. (1983). pEMBL: a new family of single-stranded plasmids. Nucl. Acids Res. *11*, 1645–1655.

Engels, W. R. (1983). The P family of transposable elements in Drosophila. Ann. Rev. Genet. *17*, 315–344.

Engels, W. R. (1984). A trans-acting product needed for P factor transposition in Drosophila. Science *226*, 1194–1196.

Falkenthal, S., Parker, V. P., and Davidson, N. (1985). Developmental variation in the splicing pattern of transcripts from the Drosophila gene encoding myosin alkali light chain result in different carboxy-terminal amino acid sequences. Proc. Natl. Acad. Sci. USA *82*, 449–453.

Hazelrigg, T., Levis, R., and Rubin, G. M. (1984). Transformation of *white* locus DNA in Drosophila: dosage compensation, *zeste* interaction, and position effects. Cell *36*, 469–481.

Hutchison, C. A., Phillips, S., Edgell, M. H., Gillam, S., Jahnke, P., and Smith, M. (1978). Mutagenesis at a specific position in a DNA sequence. J. Biol. Chem. *253*, 6551–6560.

Huynh, T. V., Young, R. A., and Davis, R. W. (1985). Constructing and screening cDNA libraries in λgt10 and λgt11. In DNA Cloning Techniques: A Practical Approach, David Glover, ed. (Oxford: IRL Press).

Karess, R. E., and Rubin, G. M. (1984). Analysis of P transposable element functions in Drosophila. Cell *38*, 135–146.

Kidwell, M. G., Kidwell, J. F., and Sved, J. A. (1977). Hybrid dysgenesis in *Drosophila melanogaster*. A syndrome of aberrant traits including mutation, sterility, and male recombination. Genetics *36*, 813–833.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucl. Acids Res. *12*, 857–872.

Krämer, A., Keller, W., Appel, B., and Lührmann, R. (1984). The 5′ terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. Cell *38*, 299–307.

Kudo, I., Leineweber, M., and RajBhandary, U. L. (1981). Site-specific mutagenesis on cloned DNAs: generation of a mutant of Escherichia coli tyrosine suppressor tRNA in which the sequence G-T-T-C corresponding to the universal G-T-ψ-C sequence of tRNAs is changed to G-A-T-C. Proc. Natl. Acad. Sci. USA *78*, 4753–4757.

Levis, R., O'Hare, K., and Rubin, G. M. (1984). Effects of transposable element insertions of RNA encoded by the *white* gene of Drosophila. Cell *38*, 471–481.

Levis, R., Hazelrigg, T., and Rubin, G. M. (1985). Effects of genomic position on the expression of transduced copies of the *white* gene of *Drosophila*. Science *229*, 558–560.

Lis, J. T., Simon, J. A., and Sutton, C. A. (1983). New heat shock puffs and β-galactosidase activity resulting from transformation of Drosophila with an *hsp70-lacZ* hybrid gene. Cell *35*, 403–410.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucl. Acids Res. *12*, 7035–7055.

Montell, C., Fisher, E. F., Caruthers, M. H., and Berk, A. J. (1982). Resolving the functions of overlapping viral genes by site-specific mutagenesis at an mRNA splice site. Nature *295*, 380–384.

Montell, C., Fisher, E. F., Caruthers, M. H., and Berk, A. (1984). Control of adenovirus E1B mRNA synthesis by a shift in the activities of RNA splice sites. Mol. Cell. Biol. *4*, 966–972.

Mount, S. M. (1982). A catalogue of splice junction sequences. Nucl. Acids Res. *10*, 459–472.

Mount, S. M., and Steitz, J. A. (1984). RNA splicing and the involvement of small ribonucleoproteins. Modern Cell Biol. *3*, 249–297.

Mount, S. M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J. A. (1983). The U1 small nuclear RNA–protein complex selectively binds a 5′ splice site in vitro. Cell *33*, 509–518.

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

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

Padgett, R. A., Mount, S. M., Steitz, J. A., and Sharp, P. A. (1983). Splicing of messenger RNA precursors is inhibited by antisera to small nuclear ribonucleoprotein. Cell *35*, 101–107.

Razin, A., Hirose, T., Itakura, K., and Riggs, A. D. (1978). Efficient correction of a mutation by use of chemically synthesized DNA. Proc. Natl. Acad. Sci. USA *75*, 4268–4270.

Rio, D. C., and Rubin, G. M. (1985). Transformation of cultured Drosophila melanogaster cells with a dominant selectable marker. Mol. Cell. Biol. *5*, 1833–1838.

Rio, D. C., Laski, F. A., and Rubin, G. M. (1986). Identification and immunochemical analysis of biologically active Drosophila P element Transposase. Cell, in press.

Rubin, G. M., and Spradling, A. C. (1982). Genetic transformation of Drosophila with transposable element vectors. Science *218*, 348–353.

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

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Sharp, P. A. (1981). Speculations on RNA splicing. Cell *23*, 643–646.

Spradling, A. C., and Rubin, G. M. (1982). Transposition of cloned P elements into Drosophila germ line chromosomes. Science *218*, 341–347.

Treisman, R., Proudfoot, N. J., Shander, M., and Maniatis, T. (1982). A single-base change at a splice site in a β^0-thalassemic gene causes abnormal RNA splicing. Cell *29*, 903–911.

Wieringa, B. J., Meyer, F., Reiser, J., and Weissmann, C. (1983). Unusual splice sites revealed by mutagenic inactivation of the rabbit beta-globin gene. Nature *301*, 38–43.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene *33*, 103–119.

Note Added in Proof

In Figures 3e and 3f the position of the ORF0–ORF1 3′ splice site should be labeled 501 instead of 500.

The arrow in Figure 1A indicating the direction of *white* gene transcription is incorrect as shown and should be in the same orientation as the hsp70–P element transcript.