Cell, Vol. 38, 471–481, September 1984, Copyright © 1984 by MIT 0092-8674/84/090471-11 $02.00/0

Effects of Transposable Element Insertions on RNA Encoded by the *white* Gene of Drosophila

Robert Levis,* Kevin O'Hare,† and Gerald M. Rubin*
Carnegie Institution of Washington
Department of Embryology
115 West University Parkway
Baltimore, Maryland 21210

Summary

We have examined the manner in which transposable element insertions affect the expression of the *white* gene of Drosophila by analyzing polyadenylated RNA of flies with each of nine insertions in or near the gene. In five mutants having insertions in the transcribed sequences of *white*, transcripts initiating at the *white* promoter are truncated within the insertions. Two insertions in the 3 kb intron of *white* alter neither the amount nor the structure of the mature *white* RNA. An insertion near the 5′ end of the gene blocks the accumulation of any *white* transcripts. Another insertion, located 1.2 kb upstream from the transcribed region of the gene, causes a mutant phenotype yet surprisingly has no obvious effect on the structure or abundance of the major *white* RNA. We also show that a mutation at each of two other loci that modulate the phenotype of the *white–apricot* insertion mutant are correlated with small but significant changes in the pattern of *white* transcripts.

Introduction

Insertional mutagenesis by transposable elements is a major cause of spontaneous mutations in D. melanogaster. This generalization is based on surveys of mutations in genes as diverse as *white* (reviewed by Rubin, 1983), *bithorax* (Bender et al., 1983), *rosy* (Bender, Cote, McCarron and Chovnick, unpublished results cited in Bender et al., 1983), *scute* (Carramolino et al., 1982), and *Notch* (Artavanis-Tsakonas et al., 1983; Kidd et al., 1983). A substantial proportion of spontaneous mutations in these genes were shown to be caused by, or at least correlated with, the insertion of transposable elements.

Little is known about the way in which transposable element insertions in eucaryotes affect gene expression and, ultimately, the phenotype. The position of an insertion within a gene is certain to be an important factor in determining its genetic consequences. A large insertion in protein-coding sequences would most likely severely alter the structure of the protein, while an insertion in the promoter may interfere with transcription, as appears to be the case for two Drosophila mutants (Snyder et al., 1982; McGinnis et al., 1983). It is less clear how an insertion located far upstream of the start of transcription or in the intron of a gene would affect its expression (Hawley et al., 1982; Harbers et al., 1984). The introduction of the transcriptional signals of a transposable element into an intron may cause RNA transcripts of the gene to be truncated or abnormally spliced. An insertion anywhere in a gene might provide a new promoter, such that transcription of all or part of the gene would originate in the transposable element (Hayward et al., 1981). Many transposable elements are actively transcribed at specific developmental times (reviewed for Drosophila elements by Rubin, 1983). Therefore, if an insertion provides a new promoter or if the transcription of the transposable element itself acts in *cis* to alter the transcription of the gene into which it is inserted, the element may change the tissue or temporal specificity of the gene's expression. In addition, cases are known in both Drosophila (Modelell et al., 1983) and yeast (reviewed by Roeder and Fink, 1983) in which the insertion of a transposable element causes the expression of a gene to be modulated by an unlinked gene.

To explore the mechanisms by which the insertion of a transposable element can alter gene expression, we have undertaken a study of the RNA transcribed from alleles of the *white* gene containing such insertions. The *white* gene product serves an unknown but vital role in the pigmentation of the adult eyes, ocelli, and testes sheath and of the larval Malpighian tubules (reviewed by Judd, 1976; Phillips and Forrest, 1980). Null mutants of *white* lack pigment in these tissues and have white eyes, while other mutants have eye colors intermediate between white and wild-type red. A 2.6 kb polyadenylated RNA has been identified as the major transcript of the *white* gene in embryos, larvae, pupae, and adults (Pirrotta et al., 1983; O'Hare et al., 1983; Pirrotta and Brockl, 1984). The extremely low abundance of *white* mRNA, which is estimated to be 0.0005% of the poly(A)$^+$ RNA of the fly, has made precise characterization of its structure difficult. The available data indicate that the sequences coding for the 2.6 kb *white* RNA are interrupted by a 3 kb intron located about 300 bp from the 5′ end of the RNA as well as by three or four small introns (O'Hare et al., 1983; Pirrotta and Brockl, 1984; O'Hare et al., unpublished data). Insertions associated with *white* mutations have been localized throughout most of the transcribed portion of the gene, as well as up to 6 kb upstream from the putative 5′ end of the RNA (Zachar and Bingham, 1982; O'Hare et al., 1983; Pirrotta and Brockl, 1984). Some of the insertions near the 5′ end of the gene appear to alter the regulation of the gene. Genetic or physical mapping of more than 20 *white* mutations has suggested, however, that few if any lie within the 3 kb intron. The only transposable element insertion previously shown to be within this intron does not cause a mutant phenotype (O'Hare et al., 1983).

Because of the variety of inserted elements, insertion sites and mutant phenotypes available among *white* alleles, we have chosen to survey the effects of these insertions on *white* transcription despite the technical dif-

* Present address: Biochemistry Department, University of California at Berkeley, Berkeley, CA 94720.
† Present address: Biochemistry Department, Imperial College of Science and Technology, London SW7 2AZ, England.

ficulties inherent in this analysis of rare transcripts. We report here our analysis for nine *white* alleles. The insertions associated with these alleles include representatives of copia-like, P, and F elements, three of the four structurally distinct classes of transposable elements in Drosophila (reviewed by Rubin, 1983; DiNocera et al., 1983). In addition, we have also tested the effect on *white* transcripts of two mutations in other genes that alter specifically the phenotype of the *white–apricot* allele.

Results

white Alleles Studied

We have studied the transcripts from the nine *white* insertion alleles listed in Table 1. Figure 1 shows the exon–intron organization of *white* that we have proposed (O'Hare et al., unpublished data) and the positions of these insertions. The eye colors of flies with these alleles, and the size and type of element inserted are described in Table 1. The eyes of flies with the $w^{hd80k17}$ mutation are bleach white. The w^{bf}, w^{a}, w^{a4}, w^{sp55}, w^{sp}, and w^{e} alleles (the full names are given in Table 1) confer eye colors intermediate between white and wild-type red. The w^{i+A} and w^{zm} alleles, which both contain insertions within the 3 kb intron, confer a wild-type red eye color. The w^{zm} allele, in contrast to the wild-type w^{+} allele, shows an interaction in males with the z^{1} mutation of the *zeste* locus (see footnote to Table 1 for details). However, not all insertions into the 3 kb insertion affect the *zeste–white* interaction; we have found that the w^{i+A} allele interacts with z^{1} in a manner indistinguishable from that of a wild-type allele (see Experimental Procedures).

Initial Analysis of Transcripts From *white* Alleles Containing Insertions

We initially analyzed the RNA of eight *white* alleles in parallel. We prepared poly(A)$^{+}$ RNA from flies carrying each of these eight alleles and from flies carrying a w^{+} (wild-type) *white* allele. RNA blots of equivalent amounts of these samples that had been fractionated by gel electrophoresis were hybridized separately with each of five ^{32}P-labeled, single-stranded genomic DNA probes (indicated in Figure 1). Probe 1 includes the small 5′ exon, probe 2 is part of the 3 kb intron, while probes 3, 4, and 5 encompass the other exons and introns that form the 3′ portion of the gene.

Figures 2A, 2B and 2C, show the autoradiographs of hybridization with probes 1, 2, and 4. The w^{+}, w^{sp}, and w^{i+A} samples showed hybridization of similar intensity to a 2.6 kb RNA with probes 1, 3, 4, and 5. We conclude that the w^{sp} and w^{i+A} insertions do not grossly alter the structure or amount of the major *white* transcript. The w^{sp55} RNA showed little or no hybridization with any of the probes. Transcripts of aberrant sizes accumulate in $w^{hd80k17}$, w^{bf}, w^{a}, w^{a4}, and w^{e} flies. These aberrant transcripts hybridized to those probes that are upstream or span the position of the particular insertion and that hybridized to the w^{+} RNA. This suggests that these transcripts have the same 5′ end

Table 1. *white* RNA from Alleles Containing Insertions

Allele		Eye Color[a]	Inserted Element	Insertion Size (kb)	RNAs[b] (kb)	Probes[c]
w^{+}	(wild-type)	red	—	—	2.6 5.5 (weak)	1,3,4,5 2
w^{sp}	(*white–spotted*)	tiny dark spots on orange background, males browner	B104	8.7	2.6 5.5 (weak)	1,3,4,5 2
w^{sp55}	(*white–spotted* 55)	♂ yellow ♀ pale yellow	?	5.8	None detected	
w^{e}	(*white–eosin*)	♂ red-orange ♀ red-brown	F-like[d]	5.7	5.0;2.6–2.7 (weak)	1
w^{i+A}	(*white–ivory revertant A*)	red	F	3.0	2.6	1,3,4,5
w^{a}	(*white–apricot*)	yellow-orange	copia	5.0	5.7;1.25	1,3
w^{a4}	(*white–apricot* 4)	yellow-orange	BEL	7–10	1.0	1,3
w^{bf}	(*white–buff*)	pale yellow	B104	8.7	2.1	1,3,4
$w^{hd80k17}$	(*white–hd80k17*)	white	P	0.63	3.2 (weak);2.7	1,3,4,5
w^{zm}	(*white–zeste–mottled*)	red[e]	?	6.0	2.6	mixture 1–5

[a] The eye colors described are for flies grown at 25°C, which was the growth temperature for all flies used for RNA analysis.

[b] Only the main bands of hybridization are listed.

[c] Figure 1 shows the genomic *white* DNA sequences included in each probe. The coordinates of these probes are listed in Experimental Procedures.

[d] Partial sequencing of the w^{h} insertion has shown that it has the oligo(A) terminus characteristic of an F element, but that its internal sequences share only scattered homology to those of a prototypic F element (O'Hare et al., unpublished).

[e] The eye color of w^{zm} flies is indistinguishable from wild type. However, males of the genotype $z^{1}w^{zm}$ have a red–brown mottled eye color at 25°C, in contrast to the wild-type red eye color of $z^{1}w^{+}$ males. Females that are $z^{1}w^{zm}$ or $z^{1}w^{+}$ both have a yellow eye color at 25°C, but differ in eye color at 14°C (Becker, 1960).

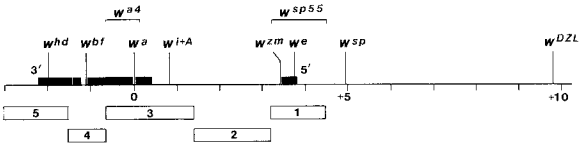

Figure 1. The *white* Transcription Unit and Positions of Insertions

A physical map of *white* locus DNA is diagrammed, with the exons of the 2.6 kb w^+ RNA represented as solid bars and hybridization probes as numbered open bars. The coordinate scale is that of Levis et al. (1982) with the copia insertion in w^a defined as 0.0 and 1 unit equalling 1 kb. The orientation of this map follows the genetic convention for the chromosomal orientation of the gene with transcription being from right to left. This map is based on the nucleotide sequence of w^+ DNA, determined by O'Hare et al. (unpublished). Some of the distances differ slightly from those of previously published restriction enzyme cleavage maps in which distances were calculated from the relative electrophoretic mobilities of restriction enzyme digest fragments. The positions of exons are those proposed by O'Hare et al. (unpublished) and are consistent with the available data on the hybridization of the 2.6 kb RNA with genomic DNA probe (O'Hare et al., 1983; Pirrotta and Brockl, 1984) and with S1 nuclease mapping of this RNA (Pirrotta and Brockl, 1984). This exon structure also provides a continuous open reading frame spliced at consensus junction sequences (Breathnach and Chambon, 1981; Mount, 1982). The position of the 3′ end of the transcript is known from the DNA sequence of a partial cDNA copy of the transcript (O'Hare et al., unpublished). The sites of insertion of transposable elements are indicated for insertions associated with the alleles discussed in the text. For seven of these alleles, the exact position of the insertion has been located by DNA sequencing (O'Hare et al., unpublished). These are the $w^{hd80k17}$, w^{bf}, w^a, w^{i+A}, w^{zm}, w^{sp}, and w^{DZL} alleles. We presume that the w^e insertion is at the same position as the w^h insertion, which has also been determined by DNA sequencing. The w^e allele and the w^h allele are both derivatives of the w^1 allele and have been reported to differ by rearrangements internal to the insertion (Zachar and Bingham, 1982). The precise location of the w^{a4} and w^{sp55} insertions is not known; the brackets in Figure 1 span the segments within which these insertions occurred (Zachar and Bingham, 1982; Goldberg et al., 1983).

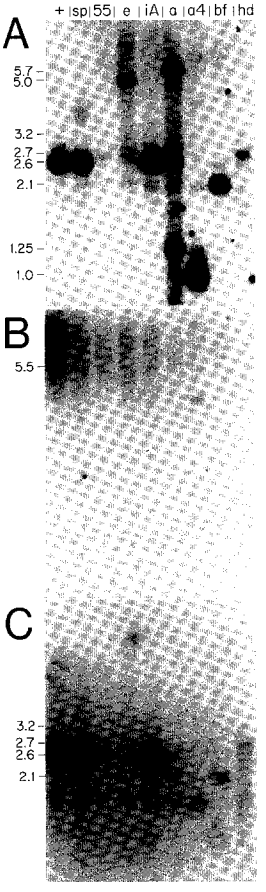

Figure 2. *white* RNA Species Present in Flies with *white* Insertion Mutations

RNA filter blots were hybridized with the following probes containing the DNA segments indicated by the open bars in Figure 1: (A) probe 1; (B) probe 2; (C) probe 4. The lane order of *white* allele RNAs is the same for each panel. The alleles have been abbreviated above each lane as follows: + = w^+ (wild type), sp = w^{sp}, 55 = w^{sp55}, e = w^e, iA = w^{i+A}, a = w^a, a4 = w^{a4}, bf = w^{bf}, hd = $w^{hd80k17}$. Approximately 2–5 μg of poly(A)$^+$ RNA was loaded per lane. The estimated nucleotide length in kb of each band of hybridization is indicated on the left. These estimates were made based on comparisons with the mobility of the Hind III digest fragments of λ DNA and/or actin and copia RNA markers. Only the portions of the autoradiographs showing hybridizing bands are included.

as the w^+ transcript, but have their 3′ ends within the insertions. Further analysis of the structures of these aberrant RNAs is described for each allele below. Probe 2 from the 3 kb intron hybridized faintly to a band of about 5.5 kb in the w^+ and w^{sp} samples (Figure 2B). Hybridization with probe 2 may be to an unspliced primary precursor (predicted to be 6 kb) or a partially spliced intermediate. There was also hybridization of probe 2 to a 4.1 kb RNA in the w^{a4} sample. The sizes of RNA of each allele that hybridized to each probe are given in Table 1. No hybridization was seen when a mixture of five probes prepared from the opposite strand of the same segments was used (not shown).

In addition to the bands listed in Table 1, hybridization was occasionally seen with one or two somewhat diffuse bands with apparent lengths of 1.8–2.1 kb (for example, see Figure 2C). The intensity of hybridization to these bands varied from experiment to experiment, although it seemed to be reproducibly stronger with certain probes. When hybridization to these bands occurred, it was seen in all or most samples. It is unlikely that this represents hybridization to specific artifactual breakdown products of the 2.6 kb *white* transcript since it appeared equally intense in samples lacking substantial amounts of any *white* transcript larger than 2.1 kb (for example, compare w^+ and w^{sp55} in Figure 2C). It is possible that this hybridization is to an abundant RNA, such as ribosomal RNA, that shares a weak sequence complementarity to certain *white* probes. This is probably not a complete explanation for the phenomenon, since one of these minor w^+ bands is generally more intense in male flies than female flies (O'Hare et al., 1983; Pirrotta and Brockl, 1984). Pirrotta and Brockl have proposed that these minor bands may arise from transcripts initiating downstream from the normal *white* start site. We will not consider hybridization to these bands further.

The RNA of w^{zm} was analyzed separately from the others. Because of our interest in the basis for the difference between the phenotypes of z^1w^+ and z^1w^{zm} males,

poly(A)$^+$ RNA was prepared from male and female flies of the genotypes w^+, z^1w^+, w^{zm}, and z^1w^{zm}. Gel-fractionated RNAs were transferred to a nitrocellulose filter and the filter was hybridized with a mixture of the same five ^{32}P-labeled single-stranded genomic DNA probes (Figure 1) described above. A predominant 2.6 kb band was observed in all eight samples (not shown). The intensity of this band in the w^+ males and females and the w^{zm} males appeared to be roughly equivalent, whereas this band was several times less intense in the w^{zm} female sample. We conclude that the w^{zm} insertion does not detectably alter the size of the major *white* gene transcript. A more careful examination will be required before any conclusions can be drawn concerning a possible sex difference in the abundance of the transcript in w^{zm} flies. To a first approximation, the z^1 mutation had no effect on the size or abundance of the 2.6 kb w^{zm} RNA in either sex. The z^1 mutation has previously been shown to have little or no effect on the size or abundance of the 2.6 kb w^+ RNA (O'Hare et al., 1983; Pirrotta and Brockl, 1984).

Further Analysis of the Structure of Aberrant Transcripts

white–eosin (w^e): The insertion associated with w^e is at coordinate +3.71 (see legend to Figure 1 for an explanation of the coordinate system). We propose that this insertion lies within the transcribed but untranslated 5′ region of the 2.6 kb *white* transcript (O'Hare et al., unpublished data). A 5.0 kb transcript in the w^e sample hybridized with probe 1 (Figure 2A) that contains the DNA segment between coordinates +3.17 and +4.44 spanning the insertion site. Weak hybridization can also be seen to a 2.6–2.7 kb band in w^e RNA. Probes 2–5, located 3′ to the insertion site, did not hybridize to these transcripts (Figures 2B and 2C; Table 1). Figure 3 shows the results of hybridizing filter strips of gel-fractionated w^+ and w^e poly (A)$^+$ RNA with smaller probes from close to the point of insertion. Probe A in Figure 3 (coordinates +3.38 to +3.60), located downstream from the insertion point, hybridized to the 2.6 kb w^+ RNA, but not to any w^e RNA. Probe B in Figure 3 (coordinates +3.61 to +3.84) spans the site of insertion at +3.71 and hybridized to the 2.6 kb w^+ RNA and the 5.0 kb and 2.6–2.7 kb w^e RNAs. Probes C (+3.76 to +3.95) and D (+3.86 to +4.16) in the diagram in Figure 3 failed to detect either the w^+ or w^e transcripts (data not shown). We therefore propose that the 2.6 kb w^+ RNA and the 5.0 kb and 2.6–2.7 kb w^e RNAs have the same 5′ terminus, but the 3′ termini of the w^e RNAs are within the insertion. In their analysis of w^e, Pirrotta and Brockl (1984) observed only a 2.6 kb w^e RNA, much reduced in abundance compared to w^+, and concluded that this insertion proably lies upstream of the 5′ terminus of the w^+ mRNA.

white–buff (w^{bf}): The w^{bf} mutation is associated with an insertion of a B104 element at coordinate −1.13. Our model for the *white* transcription unit places this insertion within the intron between the fourth and fifth exons. The

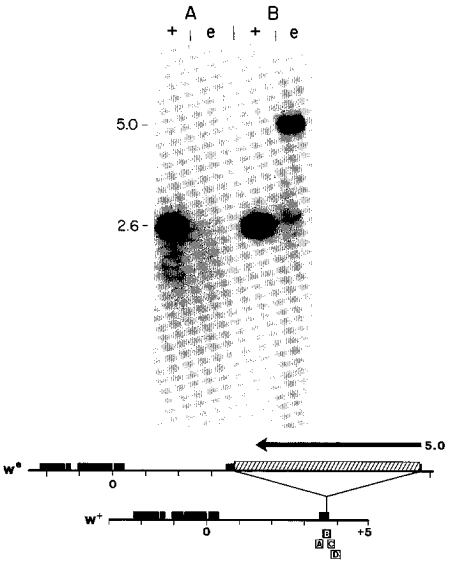

Figure 3. Hybridizaton of w^e RNA with Probes from the Vicinity of the Insertion

Filter blot strips, each containing both a lane of w^+ (abbreviated +) and a lane of w^e (abbreviated e) RNA were each hybridized with probes A, B, C, or D. The 5.0 kb w^e and 2.6 kb w^+ bands are marked. The diagram below the autoradiographs provides a map of the w^+ and w^e alleles with exons shown as solid bars and the insertion in w^e as a striped bar. The arrow above the w^e map represents our interpretation of the structure of the 5.0 kb w^e RNA. The lettered boxes below the w^+ map show the map positions of four probes tested. Only the autoradiographs of strips hybridized with probes A and B are shown.

results presented in Figure 2 and Table 1 show that probes 1, 3, and 4 hybridized to a 2.1 kb w^{bf} RNA. Probes 1 and 3 are upstream of the insertion and probe 4 spans the insertion point. In addition to hybridizing very strongly to the 2.1 kb RNA, probe 1 also hybridized very weakly to a 2.6 kb RNA that comigrates with the w^+ RNA (Figure 2A). Probe 5, which is downstream from the point of insertion, did not detect these RNAs. Additional hybridizations with shorter probes on either side of the point of insertion are shown in Figure 4. The 2.1 kb w^{bf} RNA was detected by the upstream probe B (−1.01 to −0.61), but not by the downstream probe A (−1.42 to −1.15). It therefore appears that the 2.1 kb w^{bf} RNA has the same 5′ terminus as the w^+ RNA and has its 3′ terminus within the B104 element. The failure of the 3 kb intron probe 2 in Figure 2B to hybridize to the w^{bf} RNA indicates that exons 1 and 2 are spliced together in the 2.1 kb w^{bf} transcript; we presume that the splices between exons 2 and 3 and between exons 3 and 4 are also made in the w^{bf} RNAs. From the length of the 2.1 kb w^{bf} transcript and that of the *white* exons 1–4 that it contains, we calculate that the 3′ end of the 2.1 kb w^{bf} RNA probably maps to a region close to the upstream boundary of the B104 element. The B104 element (also known as roo; Meyerowitz and Hogness, 1982) contains 429 bp direct terminal repeats (Scherer et al., 1982). Scherer et al. (1982) describe two transcripts of B104 elements, both of which have their 5′

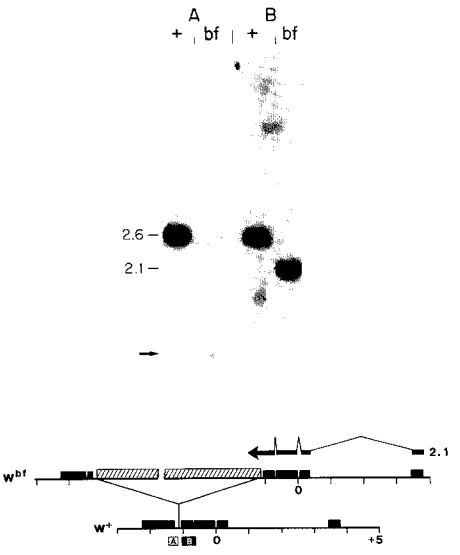

Figure 4. Hybridization of w^{bf} RNA with Probes from the Vicinity of the Insertion

Two filter strips, each containing adjacent lanes of w^+ (abbreviated +) and w^{bf} (abbreviated bf) RNA, were each hybridized with probe A or B. The 2.6 kb w^+ RNA and the 2.1 kb w^{bf} RNA bands are marked to the left of the autoradiographs. The arrow to the left of the autoradiographs points to an apparent band of hybridization in the w^{bf} RNA with probe A. This may represent transcription downstream of the insertion or may be an experimental artifact; there was no hybridization in this region with the downstream probe 5 of Figure 1. The diagram below the autoradiographs provides a map of the w^+ and w^{bf} alleles with exons shown at solid bars and the insertion in w^{bf} as a striped bar. The arrow above the w^{bf} map represents our interpretation of the structure of the 2.1 kb w^{bf} RNA with splices between exons indicated. The segments included in probes A and B are shown below the w^+ map.

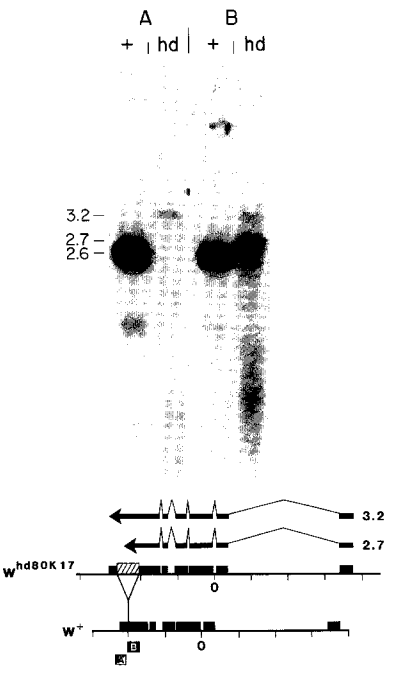

Figure 5. Hybridization of $w^{hd80k17}$ RNA with Probes from the Vicinity of the Insertion

Two filter strips, each containing adjacent lanes of w^+ (abbreviated +) and $w^{hd80k17}$ (abbreviated hd) RNA, were each hybridized with probe A or B. The 3.2 and 2.7 kb $w^{hd80k17}$ bands and the 2.6 kb w^+ bands are marked to the left of the autoradiographs. The diagram below the autoradiographs provides a map of the $w^{hd80k17}$ and w^+ alleles with exons shown as solid bars and the insertion as a striped bar. The arrows above the $w^{hd80k17}$ map represent our interpretation of the structures of the 3.2 and 2.7 kb $w^{hd80k17}$ RNAs with splices between exons indicated. The lettered boxes below the w^+ map show the map positions of probes with which strips A and B were hybridized.

ends within one of the terminal repeats. The shorter RNA has its 3′ end about 1 kb downstream, while the longer one probably terminates in the other terminal repeat. The B104 transcription unit in w^{bf} is in the same 5′–3′ orientation as that of *white*. The length of the w^{bf} transcript is consistent with its having its 3′ end within the upstream terminal repeat. We propose that sequence elements of the terminal repeat, which normally cause 3′-end formation of the B104 transcripts in its downstream terminal repeat, cause truncation of the *white* RNA at an analogous position in the upstream terminal repeat.

$w^{hd80k17}$: This null mutation is caused by the insertion of a 629 bp P element at coordinate −2.03 (Rubin et al., 1982; O'Hare and Rubin, 1983). This insertion is within the putative protein coding sequences of the last exon of *white*, 0.21 kb from the 3′ end of the transcript. Probes 1, 3, 4, and 5 of Figure 1 hybridized to two bands of 3.2 and 2.7 kb in $w^{hd80k17}$ RNA (Figures 2A and 2C; Table 1). Probes 1, 3, and 4 are upstream of the insertion while probe 5 spans the insertion point. The 2.7 kb RNA hybridized with an intensity about three times that of the 3.2 kb RNA. Figure 5 shows that probe B (−2.03 to −1.70), upstream of the point of insertion, hybridized to both RNAs while probe A (−2.34 to −2.08), downstream from the insertion site, hybridized to only the larger RNA. On the basis of its size and these hybridization data, we hypothesize that the 3.2 kb RNA consists of a properly spliced complete transcript of the *white* gene plus the P-element insertion. The shorter 2.7 kb RNA apparently has its 3′ end within the P-element insertion. The *white* and P-element transcription units (Karess and Rubin, 1984) in $w^{hd80k17}$ are oriented in the same direction. The 3′ ends of the major transcripts of the 2.9 kb prototypic P element are at a site that is present in the partially deleted 0.63 kb element inserted in $w^{hd80k17}$, at a position that is consistent with that of the 3′ end of the 2.7 kb $w^{hd80k17}$ RNA. There is a translational stop codon occurring six triplets inside the 5′ end of the P element that is in frame with what we believe to be the reading frame of *white* (O'Hare et al., unpublished). Therefore, both the 3.2 kb and 2.7 kb $w^{hd80k17}$ RNAs would each be expected to be translated into a fusion polypeptide containing amino-terminal *white* amino acids, and a few amino acids coded by the P-element terminal repeat sequence.

white-apricot 4 (w^{a4}): The w^{a4} mutation is associated with an insertion in the interval between coordinates −0.66 and +0.12 (Zachar and Bingham, 1982; Goldberg et al.,

1983). In the initial series of hybridizations (Figure 2 and Table 1) probe 1 and probe 3 both hybridized strongly to a 1.0 kb w^{a4} RNA. Probe 1 includes the 5′ exon of *white* and probe 3 spans the insertion site. Two weakly hyridizing bands of about 6.4 and 4.0 kb can be seen in the w^{a4} sample in Figure 2A as well. Probes 4 and 5, which are 3′ to the w^{a4} insertion, did not detect any of these RNA species. The hybridization with three shorter probes from the vicinity of the insertion was tested and the results are shown in Figure 6. Probes A (−0.63 to −0.44) and B (−0.34 to −0.09) hybridized to the 2.6 kb w^+ RNA but not to any w^{a4} RNA. The 1.0 kb w^{a4} RNA is complementary to probe C (+0.18 to +0.43). We propose the following: that the 5′ end of the 1.0 kb w^{a4} RNA coincides with that of the 2.6 kb w^+ RNA; that the sequences from the 3 kb intron are correctly spliced from the w^{a4} 1.0 kb RNA; and that the 3′ end of this RNA is within the insertion, near its upstream edge. Given the combined lengths of *white* exons 1 and 2, the length of the w^{a4} RNA and the results of Zachar and Bingham (1982), we would predict that the w^{a4} insertion lies between coordinates −0.25 and +0.12. This region includes the 3′ most 0.1 kb of exon 2, the 5′ most 0.2 kb of exon 3, and the intron between these two exons. Because w^{a4} flies have a residual eye color that is very similar to that of w^a flies, we believe that it is most likely that the w^{a4} insertion is within the intron between exons 2 and 3, as is the w^a insertion. Our preliminary DNA sequence data support this localization.

white–apricot (w^a): The w^a mutation results from a copia-element insertion at a site defined to be our coordinate 0.0 (Bingham and Judd, 1981; Bingham et al., 1981; Levis et al., 1982; Goldberg et al., 1982). The gene organization we propose places the w^a insertion in the intron between exons 2 and 3. In w^a RNA the 5′ exon probe 1 (Figure 1), hybridized strongly to two bands of 5.7 and 1.25 kb and more weakly to a cluster of bands of 2.6–3 kb (Figure 2A). Probe 4, located downstream from the insertion site, did not detect either the 5.7 or 1.25 kb bands (Figures 2C). Figure 6 shows that probe C from upstream of the insertion hybridized to the 5.7 and 1.25 kb w^a RNAs, while probe B, from just downstream of the insertion, did not. These data lead us to conclude that: the 5′ ends of the 5.7 and 1.25 kb w^a RNAs are the same as that of the 2.6 kb w^+ RNA; the 3 kb intron sequences are spliced from both; and both RNAs have their 3′ end within the copia element. The transcription unit of copia in w^a is oriented in the same direction as that of *white*. Copia elements produce a 5 and a 2 kb RNA, which are 5′-coterminal and differ in the position of their 3′ ends (Carlson and Brutlag, 1978; Young and Schwartz, 1981; Flavell et al., 1981; Schwartz et al., 1982). The 3′ end of the 5 kb copia RNA is within the downstream direct terminal repeat of the element (Schwartz et al., 1982). The lengths of the 5.7 and 1.25 kb w^a transcripts are consistent with these RNAs having their 3′ ends within the downstream and upstream copia-terminal repeats, respectively. The 2.1 kb w^a RNA that hybridized weakly with probe 4 in Figure 2C may result

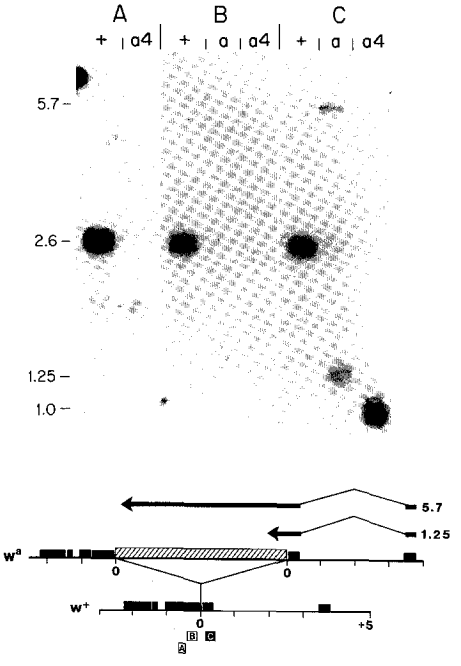

Figure 6. Hybridization of w^a and w^{a4} RNA with Probes from the Vicinity of the Insertions

Filter strip A contains a lane of w^+ (abbreviated +) and a lane of w^{a4} (abbreviated a4) RNA, while strips B and C contain lanes of w^+, w^a (abbreviated a) and w^{a4} RNA. The three strips were each hybridized with a different probe. The 5.7 and 1.25 kb w^a bands, the 2.6 kb w^+ band, and the 1.0 kb w^{a4} band are marked to the left of the autoradiographs. The diagram below the autoradiographs provides a map of the w^+ and w^a alleles with exons shown as solid bars and the insertions in w^a as a striped bar. See Figure 1 and the text for the approximate position of the w^{a4} insertion. The arrows above the w^a map represent our interpretation of the structures of the 5.7 and 1.25 kb w^a RNAs with splices between exons indicated. The lettered boxes below the w^+ map show the map positions of the probes with which strips A, B, and C were hybridized.

from a transcript having its 5′ end within the downstream terminal repeat of copia and its 3′ end formed at the site of 3′-end formation of the w^+ transcript. The 2.6 kb w^a band that was also detected by probe 4 may be wild-type in structure and arise from a spliced transcript of the entire *white* gene with its copia insertion. Pirrotta and Brockl (1984) have presented additional data supporting these interpretations for the structures of the 2.6 and 2.1 kb w^a RNAs.

The Effect on w^a RNAs of Mutations at Other Loci That Modulate the w^a Phenotype

The w^a phenotype can be altered by mutations of other genes that have no effect on the phenotype of other *white* alleles. Flies with both the w^a and *suppressor of white–apricot* (*su*(w^a)) mutations have a more nearly wild-type eye color than w^a flies (Green, 1959a). Conversely, the mutation *suppressor of forked* (*su*(*f*)) gives a nearly white eye color in combination with w^a, in addition to decreasing the severity of the phenotypic effects of certain *forked* bristle mutations (Green, 1959a). The triple mutant, *su*(w^a)

w^a *su(f)* has an eye color only slightly lighter than w^a. Figure 7 demonstrates that *su*(w^a) and *su(f)* have little or no effect on the accumulation of the major 5.7 and 1.25 w^a RNAs. However, several of the weakly hybridizing w^a bands are significantly increased or decreased in the various strains. In particular, note that the 2.6 kb w^a band is increased in abundance in the presence of *su*(w^a), absent in the presence of *su(f)*, and unchanged when both are present. A 3.0 kb band shows an opposite behavior in that it appears only when *su(f)* is present, regardless of whether *su*(w^a) is present.

The Abundance of the *white* Transcript Is Not Affected by Blocking Pigment Formation

A functional *white* gene is required for the deposition of both the drosopterin (red) and ommochrome (brown) pigment groups in all tissues where they are normally present. We have tested the premise that the abundance of the *white* transcript might be decreased by some sort of feedback mechanism if other genetic blocks were imposed on the accumulation of these pigments. The *brown* mutation eliminates drosopterins, while the *scarlet* mutation eliminates the ommochromes (Lindsley and Grell, 1968). A filter blot of gel-fractionated poly(A)$^+$ RNA of wild type and double mutant *brown*$^-$, *scarlet*$^-$ flies was hybridized with a mixture of the *white* probes 1–5 of Figure 1. Hybridization to the 2.6 kb *white* RNA was equally intense in both samples; the filter was later rehybridized with a cloned actin DNA probe to assure that equal amounts of RNA of each sample were present on the filter (results not shown). We conclude that the *white* transcript accumulates to normal levels even in these white-eyed flies.

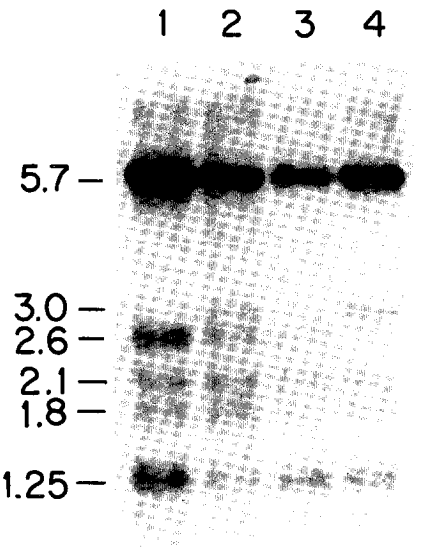

Figure 7. Effects of *su*(w^a) and *su(f)* Second Site Modifiers on the *white* RNA of w^a Flies

A filter blot of gel-fractionated RNAs was hybridized with a mixture of probes 1–5 of Figure 1. The estimated sizes in kb of the RNA bands referred to in the text are marked to the left of the autoradiograph. The lanes were loaded with poly(A)$^+$ RNA from flies of the following genotypes: 1) *su*(w^a) w^a; 2) *su*(w^a) w^a su(*f*); 3) w^a; 4) w^a *su(f)*.

Discussion

Insertions That Truncate *white* Transcripts

Of the nine *white* alleles we studied, five are associated with transposable element insertions in the transcribed region of *white* and result in *white* mutant phenotypes. In all five cases transcripts of discrete novel sizes accumulate that contain the sequences of the wild-type *white* transcript 5′ to the site of insertion. These transcripts initiated at the *white* promoter appear to be truncated within the insertion by termination or cleavage, and then polyadenylated. We hypothesize that these hybrid transcripts are truncated at the same sites in the inserted elements as the elements' own transcripts.

Hybrid transcripts initiated from within insertions are not as common in the mutants we studied as those initiated at the *white* promoter. The copia and B104 elements associated with w^a and w^{bf}, respectively, are similar in their overall structures to vertebrate retroviruses (Scherer et al., 1982; reviewed by Rubin, 1983, and Varmus, 1983), and transcription of the host c-*myc* gene from the 3′ terminal repeat of the avian leukosis provirus has been reported (Hayward et al., 1981; Payne et al., 1982). However, the provirus rearrangements seen in these cases may be prerequisite for efficient transcription from the 3′-terminal repeat (Cullen et al., 1984). We confirm the results of Pirrotta and Brockl (1984) that the steady-state level of transcripts of *white* sequences downstream from the copia insertion in w^a is extremely low. In w^{bf} flies, we find that little if any RNA accumulates from *white* sequences downstream of the B104 insertion.

Insertions within Introns

One might expect that insertions within the intron of a gene would have less severe effects on the phenotypic expression of that gene than insertions into a protein-coding region. In the nuclear genes of metazoans, the sequences that are necessary for proper intramolecular processing of the transcripts of these genes are restricted to short regions adjacent to the exon–intron junctions (reviewed by Elder et al., 1981). Therefore, if the transcript is not truncated or improperly spliced within the insertion, the sequences of the insertion within an intron may be spliced out of the transcript along with flanking intron sequences. It is in this way that we presume that functional *white* mRNA with a wild-type structure is made in w^{i+A} and w^{zm} flies.

Moreover, we believe that this same process accounts for the partial phenotypic expression of the w^{bf}, w^a, and w^{a4} alleles. We hypothesize that all three of these insertions are within small introns and that their residual gene expression results from the few transcripts that evade truncation

within the insertion and are processed to form mature transcripts with a wild-type structure. Consistent with this, in w^a and w^{bf} flies we see faint hybridization to a band comigrating with the w^+ 2.6 kb RNA (Figure 2A). Pirrotta and Brockl (1984) have also obtained evidence from RNA blot hybridizations and S1 nuclease analysis of RNA that there is a low level of RNA with a wild-type structure in w^a flies. Since the eyes of w^a flies have less than 10% of the pigment of eyes of w^+ flies (Rasmuson et al., 1960; Smith and Lucchesi, 1969), it is not surprising that the functional mRNA is so difficult to detect in these mutants.

Why do some insertions within introns interfere with transcript production while others do not? Two important factors may be the type of element inserted and its orientation. For example, the insertions in w^h (*white–honey*) and w^{i+A} both contain at one end the eucaryotic RNA cleavage signal AATAAA (Montell et al., 1983) followed by a string of As (DiNocera et al., 1983; O'Hare et al., unpublished data). It may be significant that the w^h insertion, which causes a mutant phenotype, is oriented such that the AATAAA sequence is inserted in the sense strand of *white,* whereas in the w^{i+A} insertion, which does not cause a mutant phenotype, this RNA-processing signal is inserted in the anti-sense strand.

We speculate that the position of an intron within the transcription unit may also influence whether an insertion within the intron causes a mutant phenotype. We are led to this proposal by the uneven distribution of intron insertion mutations in *white.* Phenotypic mutations caused by insertions in the small introns of *white* are more numerous than those in the 3 kb intron of *white,* despite the much smaller combined target sizes of the small introns. While it is conceivable that insertions into the 3 kb intron are less frequent, we prefer the alternative explanation, which states that insertions that do occur in the 3 kb intron are less likely to alter *white* expression. Evidence in support of this view is that the two known insertions in the 3 kb intron, in the w^{i+A} and w^{zm} alleles, do not appear to alter *white* expression in an otherwise wild-type genetic background. Transposable element insertions in introns that do not cause a mutant phenotype are also known for the Drosophila *Antennapedia* gene (Garber et al., 1983; Scott et al., 1983). It is possible that the frequency of 3′-end formation of a gene transcript within an insertion may vary with the position of the insertion in the gene, due to the structure of the nascent transcript or of the chromatin in the transcription complex. It does not appear that the relatively large size of the 3 kb *white* intron is the crucial factor; for example, Kidd et al. (1983) have identified three transposable element insertions into much larger introns of the *Notch* locus that are associated with mutant phenotypes.

Insertion Mutations and the Regulation of *white* Expression

It is sometimes possible to deduce the function of a DNA sequence element from the phenotype of mutants in which this sequence is changed. However, such inferences must be made with particular caution for mutations caused by transposable element insertions. The phenotype of an insertion mutant may reflect the juxtaposition of the inserted sequences with those of the mutated gene, rather than simply the disruption of the seqeunces at the site of the insertion. With this in mind, we next consider the issue of the location of the *cis*-acting regulatory sequences of *white*.

Mutations that appear to alter the regulation, as well as the level of *white* expression, cluster at one end of the gene (Green, 1959a; Judd, 1976) and are all associated with insertions or deletions (Zacher and Bingham, 1982) near to or upstream from the small 5′ exon (O'Hare et al., 1983). The w^e mutation is one example; in addition to having a reduced level of pigmentation, w^e does not dosage-compensate and fails to show the interaction with *zeste* that is chracteristic of the w^+ gene and mutants mapping downstream from the 3 kb intron. We propose that the w^e insertion occurred in the *white* sequences that are part of the untranslated 5′ leader of the mRNA (O'Hare et al., unpublished) and that most w^e transcripts have their 3′ ends within the insertion. Although the sites of initiation of transcription and translation of the *white* mRNA have not been precisely mapped, the hypothesis that the w^e insertion is within the transcribed sequences of the gene is supported by our analysis of w^e RNA. If the insertion in w^e is within the transcribed region of *white* as we propose, this might indicate a role for transcribed sequences in the regulation of the expression of the gene. However, an alternative interpretation, which we favor, is that the sequences at the site of the w^e insertion are not normally involved in the regulation of *white,* but an insertion at this position can nevertheless affect regulation (a more definitive example of such an effect is the w^{DZL} insertion, as discussed below). Since the w^e RNAs we have detected lack most of the w^+ protein-coding sequences, it is unlikely that they are translated into proteins that function in the partial pigmentation of the eyes of w^e. It is possible that the w^e insertion is within a yet undetected small intron near the 5′ end of the w^+ mRNA and that some transcripts from the *white* promoter are processed to form an undetectably small amount of an mRNA of the wild-type structure. There are several other hypothetical ways in which a functional w^e mRNA might be formed; it is impossible to infer what mechanism operates on the basis of the available data.

The sexually dimorphic eye color of w^{sp55} suggests that this insertion also alters the regulation of *white.* This insertion has not been mapped precisely, but is known to be in the region of the small 5′ exon of *white.* We speculate that this insertion lies upstream from the 5′ end of the 2.6 kb *white* transcript and that it disrupts the *cis*-acting control region of the gene, reducing the rate of transcription rather than altering the stucture of the transcript. We suspect that a small amount of RNA that is below our level of detection is responsible for the slight pigmentation of the eyes of w^{sp55} flies. Transposable element insertions 32 bp and 33

bp upstream from the mRNA coding sequences inactivate a larval cuticle protein gene (Snyder et al., 1982) and a salivary protein gene (McGinnis et al., 1983). Our experiments to not rule out other possibilities including a very short, unstable, or poly(A)$^-$ RNA transcribed at normal rates from w^{sp55}. Furthermore, even if the initiation of transcription is blocked by this insertion, it is possible that this is a consequence of the properties of the insertion rather than disruption of the promoter.

The w^{sp} insertion occurs further upstream from the putative site of initiation of *white* transcription than does the w^{sp55} insertion. The w^{sp} mutation not only reduces the overall pigment of the eye, but also causes a slightly variegated pattern of eye pigmentation, an abnormal dosage compensation, and abnormal interaction with the *zeste* mutation. The w^{sp} phenotype is reproduced in three other alleles: w^{sp2}, w^{sp3}, and w^{sp4}. Each of these four mutants contains an insertion into or a deletion breakpoint within the same 0.9 kb DNA region (Zachar and Bingham, 1982) 1.0 to 1.9 kb upstream from the 5′ end of the *white* RNA (O'Hare et al., 1983). This provides strong evidence that the sequences within this 0.9 kb w^{sp} region are part of the control region of *white*.

Despite the several-fold reduction of eye pigment in w^{sp} flies, we confirm the results of Pirrotta and Brockl (1984) that the size, structure, and abundance of the 2.6 kb transcript of w^{sp} and w^+ flies are not detectably different. We estimate that we would have detected a two-fold reduction in the amount, or a 0.1 kb change in the size of this RNA. We are left to speculate that while the overall level of the 2.6 kb *white* RNA is unchanged in w^{sp} flies, the spatial or temporal control of transcription has been affected or the structure of the mRNA has been subtly changed. It is conceivable that the w^{sp} insertion affects a separate transcript that is necessary for a wild-type eye color. In this context it is interesting to note that w^{sp} partially complements many other *white* alleles; females heterozygous for w^{sp} and any null or lightly pigmented allele not deleted for *white* have an eye color that is more nearly wild-type than either of their parents (Green, 1959b).

From other genetic evidence, Jack and Judd (1979) have postulated the existence of such a separate transcript involved in the *zeste–white* interaction, but it has not been detected biochemically. The z^1 mutant interacts with *white* in *trans* to reduce severely eye pigmentation (Gans, 1953); as in w^{sp} flies, this decrease in pigmentation is not accompanied by any drastic changes in the 2.6 kb *white* transcript (O'Hare et al., 1983; Pirrotta and Brockl, 1984). The 5′ control region of *white* is thought to mediate this interaction with *zeste* (Green, 1959a; Goldberg et al., 1983; reviewed by Judd, 1976). The w^{DZL} insertion, 6 kb upstream from the putative initiation of transcription of the 2.6 kb RNA (O'Hare et al., 1983), mimics the z^1 phenotype (Bingham, 1980) and also has no detectable effect on the abundance of the 2.6 kb *white* transcript (O'Hare et al., 1983). Several lines of evidence have shown that the wild-type sequences near the site of w^{DZL} insertion are not necessary for a wild-type eye color (Zachar and Bingham, 1982; Levis and Rubin, 1982; Hazelrigg et al., 1984), indicating that this insertion causes a *white* mutation even though it lies outside the boundaries of the gene.

Clearly, the apparent paradox between the phenotypic changes caused by w^{sp}, w^{DZL}, and z^1 and the lack of change in the putative *white* mRNA indicates that we remain ignorant of some fundamental aspects of the control of *white* expression. It is intriguing that each of these mutations may be interfering with this same control mechanism.

The w^{zm} allele confers a wild-type eye color in a z^+ background but shows an abnormal interaction with z^1 (see footnote to Table 1). The localization (O'Hare et al., unpublished) of an insertion associated with the w^{zm} allele to a site within the 3 kb intron, only 16 bp from the donor splice site, suggests that a change in splicing may be involved in the interaction of *white* and *zeste*. We did not observe any difference between z^+w^{zm} and z^1w^{zm} flies in the pattern of mature *white* RNA, but we would not have expected to detect any small change in the splicing pattern with the methods we used.

Modulation of the w^a Phenotype

We observed changes in the abundance of two minor RNA bands when w^a flies with and without the second site modifiers *su*(w^a) and *su*(*f*) were compared (Figure 7). A minor 2.6 kb band was most abundant in the *su*(w^a) w^a strain that has the most nearly wild-type phenotype, of intermediate abundance in w^a and triple mutant flies that have an intermediate phenotype, and undetectable in w^a *su*(*f*) flies that have the most mutant phenotype. This 2.6 kb minor band may be RNA with a wild-type structure. However, because there are several such minor bands in the same size range, further analysis will be required to prove this identification. The small increase or decrease in the abundance of the 2.6 kb w^a transcript may be sufficient to account for the modulation of the phenotypes observed. Although the effects of *su*(w^a) and *su*(*f*) on the eye color of w^a flies are visually clear-cut, each second site mutant has only a two to three fold effect on the eye pigment levels (Rasmuson and Rasmuson, 1961). The *su*(*f*) mutation was correlated with the appearance of a 3.0 kb RNA, which we detected only in the presence of *su*(*f*).

We have postulated that the residual pigment in w^a flies results from a low level of transcripts with the wild-type structure, which are produced by splicing of rare *white* transcripts that are not truncated in the copia insertion. A corollary to this hypothesis is that the second-site modifiers may act by modulating the frequency of truncation of transcripts within copia that are initiated at the *white* promoter, or may act by changing the efficiency of proper splicing of those rare transcripts from the *white* promoter that have the normal *white* 3′ terminus. Some partial revertants of w^a continue to interact with these second site-modifiers while others do not (Rasmuson and Rasmuson, 1961). An analysis of the structures and transcrip-

tion of these partial revertants may provide clues as to the mode of action of these modifying loci.

Experimental Procedures

Drosophila Strains

The genotypes and sources of the fly strains from which RNA was extracted are given in Table 2. In the text we have not included in the strain designations the irrelevant marker alleles, which are described by Lindsley and Grell (1968). We verified by DNA blot anlaysis the presence of insertions in the *white* genes of the w^{zm} and w^{i+A} flies. The w^a *su(f)* and *su*(w^a) w^a *su(f)* strains were constructed by recombining *su*(w^a)w^a and w^a with f^5 *su(f)* (obtained from the Mid-America Drosophila Stock Center). Recombinant males were identified by their eye color and individually mated to *C(1)DX,yf* females to establish stocks. These chromosomes were made homozygous using the FM6 balancer chromosome. The w^{i+A} allele was combined with z^1 by recombination between *gt* w^{i+A} *sp1* and y^2 z^1 w^{11E4} f^5 (the source of the z^1 allele of y^2 z^1 w^{11E4} f^5 was a z^1 w^{11E4} chromosome obtained from W. Gelbart and C. T. Wu). Recombinants between *y* and *w* were recovered, approximately half of which should be z^1w^{i+A} recombinants. Stocks were established from ten single *y w* males. These X chromosomes were made homozygous using the FM7c balancer and several lines were found in which the females had yellow eyes indistinguishable from those of z^1w^+ females. We presume that these lines are genotypically $y^2z^1w^{i+A}$*spl*. Males of these lines have wild-type red eyes.

RNA Blot Analysis

O'Hare et al. (1983) have described the procedures used for purification of polyadenylated RNA from adult flies, electrophoresis of RNA on agarose/formaldehyde gels, transfer to nitrocellulose, and hybridization with single-strand probes. The probes indicated in Figure 1, used for the hybridization in Figures 2 and 7, were prepared from restriction fragments of wild-type (Canton S) strain genomic DNA from the lambda phage clones isolated by Levis et al. (1982), which were subcloned into the M13 vectors mp7, mp8, or mp9 (Messing and Vieira, 1982). The subcloned fragments with coordinates of their termini are: probe 1: Bam HI (+4.44)–Hind III (+3.17); probe 2: Hind III (+3.17)–Bam HI (+1.38); probe 3: Bam HI (+1.38)–Sal I (−0.67); probe 4: Sal I (−0.67)–Sal I (−1.53); probe 5: Sal I (−1.53)–Sal I (−3.05). Probe B of Figure 5 was prepared from an Hae III restriction fragment (coordinates −1.69 to −2.03) subcloned in mp8. All of the other probes used in Figures 3–6 were prepared from sonicated DNA fragments of cloned Canton S genomic DNA, which were subcloned into M13 vectors as part of the DNA sequence determination of O'Hare et al. (unpublished). The endpoints of these cloned segments are known from DNA sequencing.

Table 2. Drosophila Strains

Genotype	Source	Ref.
wild type (P2)	A. Spradling	—
y sn^w;*bw;st*	Spradling and Rubin (1982)	
w^a	Mid-America Drosophila Stock Center	Lindsley and Grell (1968)
su(w^a) w^a	Mid-America Drosophila Stock Center	Green (1959a)
w^{a4}	Mid-America Drosophila Stock Center	Lindsley and Grell (1968)
$w^{bf}f^5$	Mid-America Drosophila Stock Center	Lindsley and Grell (1968)
w^e	Mid-America Drosophila Stock Center	Lindsley and Grell (1968)
y w^{sp}	B. H. Judd	Lindsley and Grell (1968)
w^{sp55}	CalTech Drosophila Stock Center	Zachar and Bingham (1982)
$w^{hd80k17}$	M. G. Kidwell	Rubin et al. (1982)
gt w^{i+A}*spl*	M. M. Green	Karess and Rubin (1982)
w^{zm}	B. H. Judd	Lindsley and Grell (1968)

Acknowledgments

We thank Dr. Roger Karess for his instructions in the preparation of single-stranded labeled probes and Dr. Stephen Mount for his help in the analysis of *white* DNA sequence data. We thank Dr. Vincent Pirrotta for communication of results prior to publication. We are grateful to Mr. Vincent Young and Dr. B. H. Judd for providing an unpublished clone containing the insertion in the w^{zm} allele. We thank the stock centers and investigators listed in Table 2 for providing fly stocks. This work was supported by grants from the National Institutes of Health and the American Cancer Society 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 June 6, 1984

References

Artavanis-Tsakonas, S., Muskavitch, M. A. T., and Yedvobnick, B. (1983). Molecular cloning of *Notch*, a locus affecting neurogenesis in *Drosophila melanogaster*. Proc. Nat. Acad. Sci. USA *80*, 1977–1981.

Becker, H. J. (1960). Variegation in the *zeste* eye color alleles and its bearing on gene action during the development of the eye of Drosophila melanogaster. Genetics *45*, 519–534.

Bender, W., Akam, M., Karch, F., Beachy, P. A., Peifer, M., Spierer, P., Lewis, E. B., and Hogness, D. S. (1983). Molecular genetics of the bithorax complex in Drosophila melanogaster. Science *221*, 23–29.

Bingham, P. M. (1980). The regulation of *white* locus expression: a dominant mutant allele at the *white* locus of Drosophila melanogaster. Genetics *95*, 341–353.

Bingham, P. M., and Judd, B. H. (1981). A copy of the copia transposable element is very tightly linked to a w^a allele at the *white* locus of D. melanogaster. Cell *25*, 705–711.

Bingham, P. M., Levis, R., and Rubin, G. M. (1981). Cloning of DNA sequences from the *white* locus of D. melanogaster by a novel and general method. Cell *25*, 693–704.

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

Carlson, M., and Brutlag, D. L. (1978). One of the copia genes is adjacent to satellite DNA in *Drosophila melanogaster*. Cell *15*, 733–742.

Carramolino, L., Ruiz-Gomez, M., Guerrero, M., Campuzano, S., and Modelell, J. (1982). DNA map of mutations at the *scute* locus of *Drosophila melanogaster*. EMBO J. *1*, 1185–1191.

Cullen, B. R., Lomedico, P. T., and Ju, G. (1984). Transcriptional interference in avian retroviruses: implications for promoter insertion model of leukaemogenesis. Nature *307*, 241–245.

DiNocera, P. P., Digan, M. E., and Dawid, I. B. (1983). A family of oligoadenylate-terminated transposable sequences in Drosophila melanogaster. J. Mol. Biol. *168*, 715–727.

Elder, J. T., Spritz, R. A., and Weissman, S. M. (1981). Simian virus 40 as a eukaryotic cloning vehicle. Ann. Rev. Genet. *15*, 295–340.

Flavell, A. J., Levis, R., Simon, M., and Rubin, G. M. (1981). The 5′ termini of RNAs encoded by the transposable element copia. Nucl. Acids Res. *9*, 6279–6291.

Gans, M. (1953). Étude génétique et physiologique du mutant *z* de

Drosophila melanogaster. Bulletin Biologique de France et de Belgique *38,* Suppl., 1–90.

Garber, R. L., Kuroiwa, A., and Gehring, W. J. (1983). Genomic and cDNA clones of the homeotic locus *Antennapedia* in *Drosophila.* EMBO J. *2,* 2027–2036.

Goldberg, M. L., Paro, R., and Gehring, W. J. (1982). Molecular cloning of the *white* locus region of Drosophila melanogaster using a large transposable element. EMBO J. *1,* 93–98.

Goldberg, M. L., Sheen, J.-Y., Gehring, W. J., and Green, M. M. (1983). Unequal crossing-over associated with asymmetrical synapsis between nomadic elements in the *Drosophila melanogaster* genome. Proc. Nat. Acad. Sci. USA *80,* 5017–5021.

Green, M. M. (1959a). Spatial and functional properties of pseudoalleles at the *white* locus in *Drosophila melanogaster.* Heredity *13,* 303–315.

Green, M. M. (1959b). Putative non-reciprocal crossing over in Drosophila melanogaster. Zeitschrift fur Vererbungslehre *90,* 375–384.

Harbers, K., Kuehn, M., Delius, H., and Jaenisch, R. (1984). Insertion of retrovirus into the first intron of α1 (I) collagen gene leads to embryonic lethal mutation in mice. Proc. Nat. Acad. Sci. USA *81,* 1504–1508.

Hawley, R. G., Shulman, M. J., Murialdo, H., Gibson, D. M., and Hozumi, N. (1982). Mutant immunoglobulin genes have repetitive DNA elements inserted into their intervening sequences. Proc. Nat. Acad. Sci. USA *79,* 7425–7429.

Hayward, W. S., Neel, B. G., and Astrin, S. M. (1981). Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. Nature *290,* 475–480.

Hazelrigg, T. H., 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.

Jack, J. W., and Judd, B. H. (1979). Allelic pairing and gene regulation: a model for the *zeste–white* interaction in *Drosophila melanogaster.* Proc. Nat. Acad. Sci. USA *76,* 1368–1372.

Judd, B. H. (1976). Complex loci. In The Genetics and Biology of Drosophila, 1b, M. Ashburner and E. Novitski, eds. (London: Academic Press), pp. 767–799.

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

Kidd, S., Lockett, T. J., and Young, M. W. (1983). The *Notch* locus of Drosophila melanogaster. Cell *34,* 421–433.

Levis, R., and Rubin, G. M. (1982). The unstable w^{DZL} mutation of Drosophila is caused by a 13 kilobase insertion that is imprecisely excised in phenotypic revertants. Cell *30,* 543–550.

Levis, R., Bingham, P. M., and Rubin, G. M. (1982). Physical map of the *white* locus of Drosophila melanogaster. Proc. Nat. Acad. Sci. USA *79,* 564–568.

Lindsley, D., and Grell, E. (1968). Genetic variations of *Drosophila melanogaster.* Carnegie Inst. Washington Pub. 627.

McGinnis, W., Shermoen, A. W., and Beckendorf, S. K. (1983). A transposable element inserted just 5′ to a Drosophila glue protein gene alters gene expression and chromatin structure. Cell *34,* 75–84.

Messing, J., and Vieira, J. (1982). A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene *19,* 269–276.

Meyerowitz, E. M., and Hogness, D. S. (1982). DNA sequence organization at the 68C Drosophila polytene chromosome puff. Cell *28,* 165–176.

Modolell, J., Bender, W., and Meselson, M. (1983). *Drosophila melanogaster* mutations suppressible by the suppressor of Hairy-wing are insertions of a 7.3-kilobase mobile element. Proc. Nat. Acad. Sci. USA *80,* 1678–1682.

Montell, C., Fisher, E. F., Caruthers, M. H., and Berk, A. J. (1983). Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3′ consensus sequence AAUAAA. Nature *305,* 600–605.

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

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., Levis, R., and Rubin, G. M. (1983). Transcription of the *white* locus in *Drosophila melanogaster.* Proc. Nat. Acad. Sci. USA *80,* 6917–6921.

Payne, G. S., Bishop, J. M., and Varmus, H. E. (1982). Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. Nature *295,* 209–214.

Phillips, J. P., and Forrest, H. S. (1980). Ommochromes and Pteridines. In The Genetics and Biology of Drosophila. 2d, M. Ashburner and T. R. F. Wright, eds. (London: Academic Press), pp. 541–623.

Pirrotta, V., and Brockl, C. (1984). Transcription of the Drosophila *white* locus and some of its mutants. EMBO J. *3,* 563–568.

Pirrotta, V., Hadfield, C., and Pretorius, G. H. J. (1983). Microdissection and cloning of the *white* and the 3B1–3C2 region of the Drosophila X chromosome. EMBO J. *2,* 927–934.

Rasmuson, B., and Rasmuson, M. (1961). Diversity of interaction properties in eight back-mutations from the mutant w^a in Drosophila melanogaster. Hereditas *47,* 619–630.

Rasmuson, B., Green, M. M., and Ewertson, G. (1960). Qualitative and quantitative analyses of eye pigments and pteridines in back-mutations of the mutant w^a in Drosophila melanogaster. Hereditas *46,* 635–650.

Roeder, G. S., and Fink, G. R. (1983). Transposable elements in yeast. In Mobile Genetics Elements, J. A. Shapiro, ed. (New York: Academic Press), pp. 299–328.

Rubin, G. M. (1983). Dispersed repetitive DNAs in *Drosophila.* In Mobile Genetic Elements, J. A. Shapiro, ed. (New York: Academic Press), pp. 329–361.

Rubin, G. M., Kidwell, M. G., and Bingham, P. M. (1982). The molecular basis of P-M hybrid dysgenesis: the nature of induced mutation. Cell *29,* 987–994.

Scherer, G., Tschudi, C., Perera, J., Delius, H., and Pirrotta, V. (1982). B104, a new dispersed repeated gene family in *Drosophila melanogaster* and its analogies with retroviruses. J. Mol. Biol. *157,* 435–451.

Schwartz, H. E., Lockett, T. J., Young, M. W. (1982). Analysis of transcripts from two families of nomadic DNA. J. Mol. Biol. *157,* 49–68.

Scott, M. P., Weiner, A. J., Hazelrigg, T. I., Polisky, B. A., Pirrotta, V., Scalenghe, F., and Kaufman, T. C. (1983). The molecular organization of the *Antennapedia* locus of Drosophila. Cell *35,* 763–776.

Smith, P. D., and Lucchesi, J. C. (1969). The role of sexuality in dosage compensation in Drosophila. Genetics *61,* 607–618.

Snyder, M. P., Kimbrell, D., Hunkapillar, M., Hill, R., Fristrom, J., and Davidson, N. (1982). A transposable element that splits the promoter region inactivates a *Drosophila* cuticle protein gene. Proc. Nat. Acad. Sci. USA *79,* 7430–7434.

Varmus, H. E. (1983). Retroviruses. In Mobile Genetic Elements, J. A. Shapiro, ed. (New York: Academic Press), pp. 411–503.

Young, M. W., and Schwartz, H. E. (1981). Nomadic gene families in *Drosophila.* Cold Spring Harbor Symp. Quant. Biol. *45,* 629–640.

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

Note Added in Proof

Material referred to in this paper as O'Hare et al. (unpublished data) is now in press: O'Hare, K., Murphy, C., Levis, R., and Rubin, G. M. (1984). The DNA sequence of the *white* locus of Drosophila melanogaster. J. Mol. Biol.