Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 3938–3942, June 1988
Genetics

Reversion of a *gypsy*-induced mutation at the yellow (*y*) locus of *Drosophila melanogaster* is associated with the insertion of a newly defined transposable element

(retrotransposon/mutagenesis/transcriptional control/suppression)

PAMELA K. GEYER*, M. M. GREEN†, AND VICTOR G. CORCES*‡

*Department of Biology, The Johns Hopkins University, 34th and Charles Streets, Baltimore, MD 21218; and †Department of Genetics, University of California, Davis, CA 95616

Contributed by M. M. Green, February 12, 1988

ABSTRACT **To understand the molecular basis of the phenotype of *gypsy*-induced mutations, we have analyzed the structure of phenotypic revertants of the y^2 allele, which is caused by the insertion of the *gypsy* element into the 5′ region of the yellow (*y*) locus. Seven spontaneous revertants examined fall into two different classes. Three of these revertants arose by homologous recombination between the two *gypsy* long terminal repeats (LTRs), leaving behind a solo LTR. Four additional revertants contain an intact 3′ LTR and half of the 5′ LTR, but the central portion of *gypsy* has been replaced by a different 6.5-kilobase transposable element that contains a poly(A) tail. These results suggest that the mutagenic effect of the *gypsy* element is not due to its insertion into sequences necessary for transcription or to the distancing between the yellow promoter and remote regulatory sequences but is a consequence of idiosyncratic properties of the element itself.**

Transposable elements can cause mutations by a variety of mechanisms. For example, the insertion of an element into the protein coding region of a gene may disrupt the reading frame and give rise to an aberrant product, or the element can affect initiation or termination of transcription, or splicing of the RNA, by inserting into sequences that control these processes (see refs. 1 and 2 for reviews). In addition, transposable elements can cause more subtle alterations in gene expression as a consequence of their insertion, bringing the mutated gene under the control of unrelated, unlinked loci. Such is the case with mutations induced by the *gypsy* transposable element whose phenotype can be reversed by second-site mutations at the suppressor of Hairy-wing [*su(Hw)*, 3–54.8] locus (3, 4).

The y^2 allele is caused by the insertion of *gypsy* at −700 base pairs (bp) from the transcription start site of the yellow (*y*, 1–0.0) gene (5). This locus encodes a single 1.9-kilobase (kb) RNA that is expressed in a temporal and tissue-specific fashion during *Drosophila* development (6–9). The yellow (*y*) gene is transcribed in the embryo and early larval stages to provide proper coloration of the mouth hooks and denticle belts of the larva and during pupal development, ensuring normal pigmentation of the adult cuticle and its derivatives such as hairs, bristles, and wings. The insertion of the *gypsy* element in the y^2 allele does not affect the early transcription of this gene but it disturbs the pupal expression, such that adult y^2 flies have normally colored bristles, whereas the wings and body cuticle are yellow (5, 8). We have recently reported that two enhancer-like elements, located in the 5′ region of the gene between −495 bp and −1868 bp and between −1868 bp and −2873 bp, are responsible for controlling the expression of the *y* locus in the body cuticle and wings, respectively (10, 11). The mutagenic effect of the *gypsy* element on *y* gene expression could be due to the insertion of *gypsy* into these regulatory sequences, or, alternatively, to the separation between these sequences and the "TATA box" of the *y* gene. Here we present evidence suggesting that this is not the case and that specific sequences in the *gypsy* element are responsible for its effect on the transcription of the *y* locus.

MATERIALS AND METHODS

Plasmid DNA isolation, DNA enzymology, and screening of λ libraries were carried out by standard procedures (12). *Drosophila melanogaster* DNA was prepared as described (8). DNA samples were digested with restriction enzymes and electrophoresed on a 1% agarose gel. Southern analysis was carried out as described (13). RNA was prepared by using the $NaDodSO_4$/phenol technique (14). RNA samples were electrophoresed on a 1.5% formaldehyde gel, blotted, and hybridized as reported (5). DNA sequencing was done according to Maxam and Gilbert (15).

RESULTS

Isolation of Spontaneous Revertant Lines. Phenotypic revertants of the *gypsy*-induced allele y^2 arose spontaneously in a stock of the genotype $y^2\ sc^1\ w^a\ sn^{83i7(1)}\ v^{36f}$. Crosses were set up either between 2 males and harems of 10 females in individual vials or between single females crossed to 2 or 3 males. A total of 96,371 progeny was scored. These revertants can be subdivided into two groups depending on the circumstances of their appearance. Four revertants arose as single female progeny, suggesting that the reversion was due to a postreplicative meiotic event that probably occurred in the parental male germ line, since no males were recovered in this category. We will refer to this group of revertants as class A. The second group of revertants (class B) arose as clusters of males and females among the progeny of a mating. Five independent reversion events were obtained in this class, and the size of the cluster varied, from 3 females and 1 male in the case of y^{2+} (no. 69) to 33 females and 25 males in the case of y^{2+} (no. 9), suggesting that the event leading to the reversion of the mutant phenotype was premeiotic and took place in the female germ line. Two revertants of the *gypsy*-induced allele sc^1 were obtained in the same experiment, but no revertants of the *copia*-induced mutation w^a were found.

Analysis of DNA from the *y* Locus of Revertant Flies. As a first step in determining the molecular basis of the reversion of the y^2 phenotype, genomic DNA was prepared, digested

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

Abbreviation: LTR, long terminal repeat.
‡To whom reprint requests should be addressed.

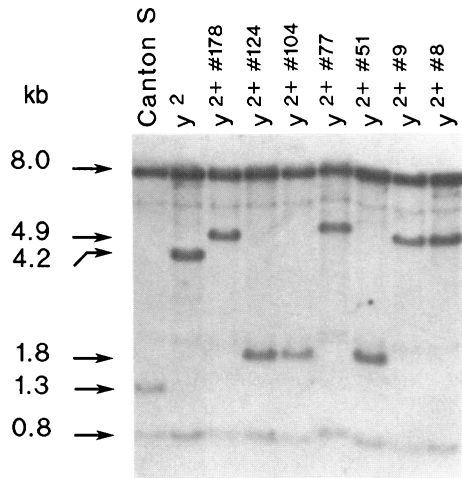

FIG. 1. Southern analysis of genomic DNA from revertant stocks. Five micrograms of total DNA from Canton S, the y^2 mutant allele, and each of the revertant stocks was digested with *Hin*dIII, electrophoresed on a 1% agarose gel, and blotted onto nitrocellulose paper. The filter was then hybridized with the *Sal* I-*Eco*RI fragment containing the *y* locus labeled with [^{32}P]dNTP (5). The numbers on the side correspond to the size in kb of the different restriction fragments indicated by the arrows. The revertant stocks are designated as y^{2+} followed by a number.

with *Hin*dIII, and subjected to Southern analysis. The result of this experiment is shown in Fig. 1. The size distribution of restriction fragments from these revertants follows two different patterns.

Some revertants (y^{2+} nos. 51, 104, and 124) contain the same size fragments as wild-type DNA, with the exception of a 1.8-kb band that replaces the normal 1.3-kb fragment. Sequence analysis of these fragments shows that the additional 0.5 kb is due to the presence of a solo *gypsy* long terminal repeat (LTR) inserted at the same nucleotide as the complete *gypsy* element in y^2 (Fig. 2*A*). Revertants showing this restriction pattern belong to class A as defined above, namely, they arose as single females, probably due to intrachromatid recombination between the two *gypsy* LTRs in the paternal germ line. The structure of the *y* locus in this class is the same as that described for a y^2 revertant obtained in a different set of experiments (5).

The second type of restriction pattern is displayed by all of the members of class B revertants (y^{2+} nos. 8, 9, 77, and 178). Southern analysis using several other restriction enzymes in addition to *Hin*dIII (Fig. 1 and data not shown) indicates that DNA sequences of equivalent size to *gypsy* are still present in the 5′ region of the *y* locus in this class of y^2 revertants, but the restriction sites, and thus the nature of the sequence, are different, indicating that these sequences do not correspond to the *gypsy* element (Fig. 2*B*).

Class B Revertants Contain a Newly Defined Transposable Element. Genomic DNA was isolated from class B revertant lines and used to prepare λ libraries. The *y* gene was then isolated and sequenced in order to understand the precise structure and nature of the DNA inserted at the *y* locus in these flies. All four revertants contained a 7.2-kb fragment of DNA inserted in the 5′ region of the *y* locus (Fig. 2*B*). The site of insertion remained the same as that in the y^2 parental stock. The inserted DNA is flanked by an intact *gypsy* 3′ LTR at the distal end and by approximately half of the 5′ LTR (from nucleotides 1 to 161) at the proximal end. The central part of the *gypsy* element is now replaced by a 6.5-kb piece of foreign DNA whose partial sequence is shown in Fig. 3.

Several characteristics of this sequence suggest that it might in fact correspond to a transposable element. The sequence inserted between the *gypsy* LTRs in class B revertants contains a poly(A) tract 43 bp long at one end, which is characteristic of some *Drosophila* transposable elements (16). In addition, this sequence is flanked by a 10-bp duplication of the nucleotides ACATTTCTTG, reminiscent of the sequence repetition that originates after the insertion of transposable elements (2). Another characteristic of transposable elements is their repetitive nature. To assess the number of copies of this sequence present in the *Drosophila* genome, total genomic DNA from Oregon R and Canton S was digested with several restriction endonucleases and subjected to Southern analysis. Fig. 4*A* shows the result of this experiment and indicates that there are around 40–50 copies of this sequence per haploid genome. The pattern of restriction fragments is different in both strains, suggesting that this sequence is mobile.

These results support the idea that class B revertants arose by a process in which the internal region of the *gypsy* element was deleted and substituted by a member of a different family of transposable elements. We have designated this newly defined element, which shows some general structural characteristics similar to the *F* elements (16), as *wallaby*. To determine if different members of this family of transposable elements show a strong structural conservation, the pattern of hybridization of internal restriction fragments in genomic DNA from Oregon R or Canton S flies was analyzed (Fig. 4*B*). Almost all of the hybridization is located in one (*Hin*dIII digest) or two (*Pst* I digest) specific bands of the size predicted from the restriction map of the *wallaby* element, suggesting that most members of the family have the same internal conserved structure.

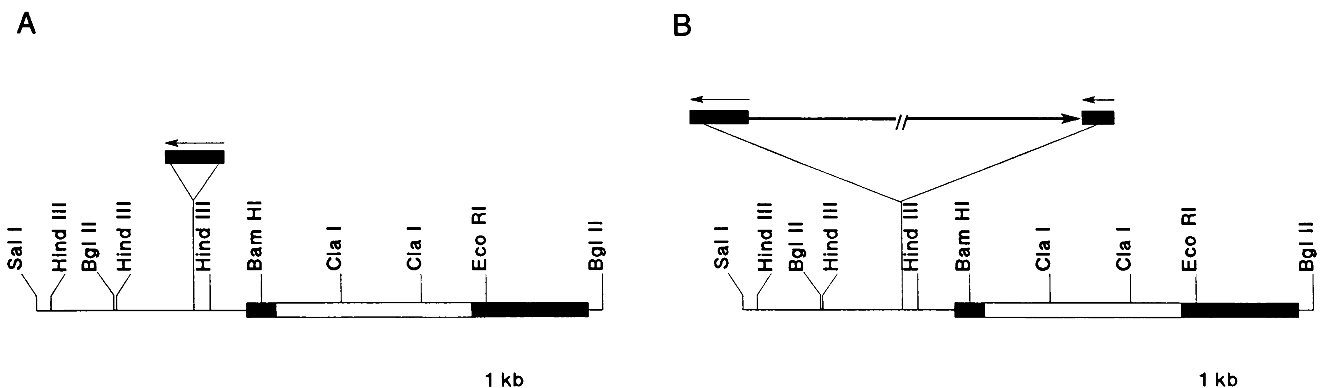

FIG. 2. Schematic representation of the structure of the *y* locus in both classes of revertant flies. (*A*) Class A revertants contain a complete *gypsy* LTR in the same place where the *gypsy* element was inserted in the y^2 allele. (*B*) Class B revertants contain the *gypsy* 3′ and half of the 5′ LTRs, but the central region of this element has been replaced by the *wallaby* element. This transposon is represented by an arrow to account for the location of the poly(A) tract at the 3′ end.

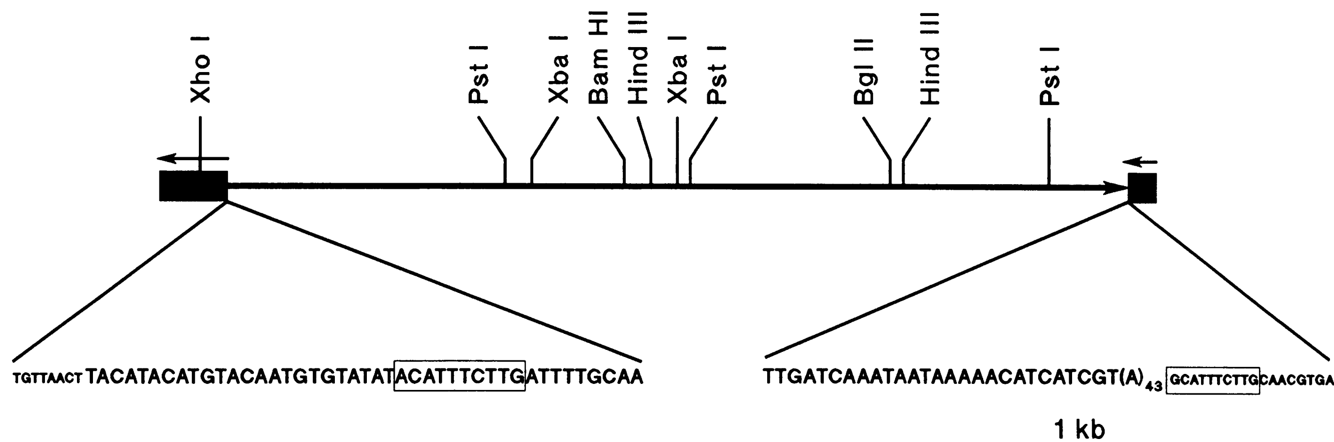

FIG. 3. Detailed structure of the *wallaby* element. The restriction map of *wallaby*, flanked by *gypsy* LTRs, is indicated. The lower part of the figure displays the sequence of the boundaries between both elements. Small capital letters indicate *gypsy* LTR sequences, whereas large capital letters refer to the *wallaby* element and adjacent sequences from the scute (*sc*) locus; boxed sequences designate the 10-bp duplication that originated when *wallaby* was inserted in the chromosome.

The *wallaby* element is different from *F* elements on the basis of three criteria. Approximately 500 bp of the sequence obtained for *wallaby* (data not shown) do not coincide with the published sequence of the *F* element (16). In addition, the characteristic internal restriction fragments conserved in *F* elements (17) are not present in *wallaby* (Fig. 4*B*). Finally, cloned DNA from both families of elements failed to cross-hybridize on Southern blots (data not shown). These results indicate that *wallaby* belongs to a different family of transposons that shares some structural characteristics but is distinct from the previously described *F* elements.

Reversion of the Mutant Phenotype Is Due to Normal Transcription of the *y* Gene. The newly defined transposable element located at the *y* locus in the class B revertants contains a poly(A) tail at the 3′ end, suggesting that, if expressed, transcription of the *wallaby* element would be from left to right in Fig. 2. This is the same transcriptional orientation as the adjacent *y* locus. It could be that the reversion of the mutant phenotype is due to the synthesis of *y* transcripts from a promoter located in the 5′ region of *wallaby*, giving rise to a functional hybrid RNA that contains sequences from this transposon as well as the *y* gene. To test this possibility, we prepared RNA from pupal stages of development of wild-type and class B revertant flies. These RNAs were subjected to transfer analysis and the result is shown in Fig. 5. The 1.6-kb transcript corresponds to the *ras2* gene, which is transcribed at constant levels during *Drosophila* development (18) and is shown here to compare the amount of RNA loaded in the different lanes. In Canton S flies, the amount of 1.9-kb *y* transcript increases from day 7 to day 8 and decreases again at day 9 of development. The revertant stock analyzed develops slower than wild type, with an 11-day period from embryo to adult at 22°C instead of the 10-day period elapsed by Canton S. In these revertants, the normal-size 1.9-kb transcript accumulates with the same temporal specificity as in wild type; its transcription is off in early pupal stages, it accumulates at normal levels in midpupae (day 9), and the steady-state levels of this RNA decrease again in day 10 (Fig.

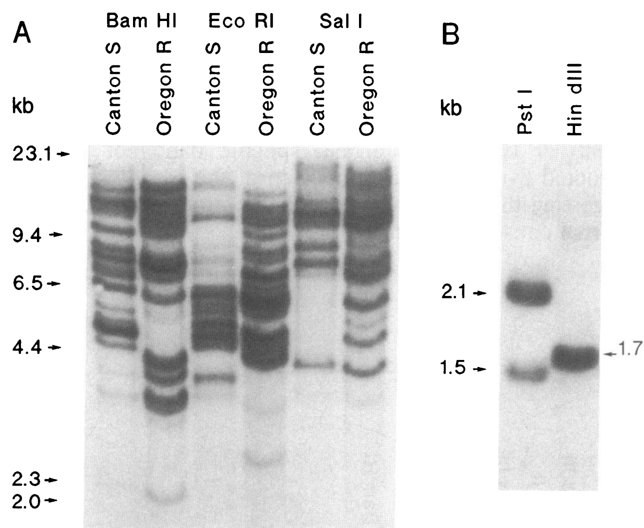

FIG. 4. Southern analysis of the *wallaby* element. (*A*) Genomic DNA from Canton S and Oregon R strains of *D. melanogaster* was digested with *Bam*HI, *Eco*RI, and *Sal* I, electrophoresed on a 1% agarose gel, blotted onto nitrocellulose, and probed with the internal 1.7-kb *Hin*dIII fragment of the *wallaby* element labeled with [^{32}P]dNTP (see Fig. 3). (*B*) Genomic DNA from Canton S flies was digested with *Pst* I or *Hin*dIII and subjected to Southern analysis as described above. The hybridization probe was the same 1.7-kb *Hin*dIII fragment.

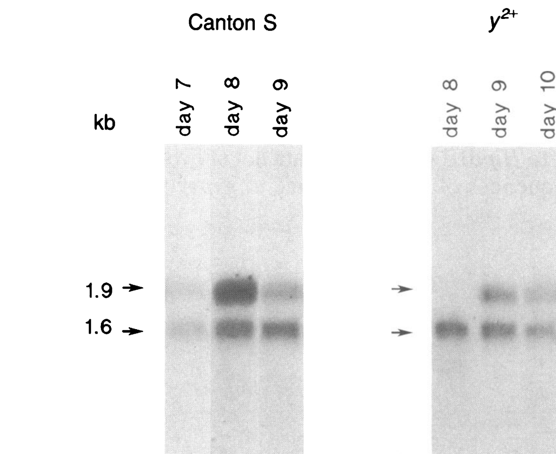

FIG. 5. Transfer analysis of RNA from wild-type and class B revertant stocks. Ten micrograms of poly(A)-containing RNA, isolated from three different days of pupal development of Canton S and y^{2+} (no. 8), was electrophoresed on a 1.5% agarose/formaldehyde gel, blotted onto a nylon membrane, and hybridized with ^{32}P-labeled DNA fragments containing the *y* (5) and *ras2* (18) *Drosophila* genes. The *y* locus hybridizes to a 1.9-kb RNA. *ras* gives rise to a 1.6-kb transcript that is expressed at approximately constant levels during *Drosophila* development and is shown as a marker for the amount of RNA.

5). Since these revertant flies express a *y* transcript of the same size as wild type and with the normal developmental transcription pattern, the reversion of the *y* phenotype must be due to the proper function of the *y* promoter in class B revertants.

DISCUSSION

We have shown that two classes of y^2 revertants arose by different mechanisms, although the cellular events underlying these phenomena might be very similar. Class A revertants contain a single *gypsy* LTR and originated by intrachromatid homologous recombination between the two *gypsy* terminal repeats in the paternal germ line during meiosis. Class B revertants have a more complicated structure and it is of interest to speculate on the mechanism by which they arose. The events leading from an intact *gypsy* element to a hybrid sequence containing the *wallaby* transposon and part of *gypsy* involved the deletion of the internal part of *gypsy* and half of its 5′ LTR and the insertion of the *wallaby* element. It is interesting to note that, in the sequences shown in Fig. 3, one of the 10-bp duplications that was created as a consequence of the insertion of *wallaby* is not immediately adjacent to the 3′ LTR of *gypsy* but there are 25 bp of DNA sequence between them. This suggests that these 25 bp are not part of the *wallaby* element proper but are adjacent to it. Therefore, these sequences must derive from the chromosomal location where *wallaby* was located before the reversion of y^2 took place. Inspection of the sequences immediately adjacent to the 3′ *gypsy* LTR shows that the 25-bp stretch begins with the sequence TACATA, which is homologous to the consensus sequence recognized by *gypsy* and is adjacent to the 5′ LTR upon its insertion (19, 20). This indicates that the 25-bp sequence and the *wallaby* element were located adjacent to the 5′ LTR of a *gypsy* element somewhere in the genome before the reversion process occurred. Examination of this sequence reveals its homology to the sequence immediately adjacent to the *gypsy* 5′ LTR in the *sc* locus in the sc^1 allele (ref. 19; and R. Freund, personal communication). These results suggest a mechanism that could account for the precision with which the reversion event happened in four independent cases and the excision and replacement of *gypsy* sequences in a single event. Class B revertants might have arisen by a double crossover event in which recombination took place between the 3′ LTR of the *gypsy* element located at the *y* locus in y^2 and the 5′ LTR of the *gypsy* located at the *sc* locus in sc^1; the second crossover event took place between the 10-bp duplication located immediately adjacent to the poly(A) tract of the *wallaby* element and a homologous sequence located in the 5′ LTR of the *gypsy* element at the *y* locus in y^2. Note that the 3′ copy of the 10-bp duplication starts with a guanosine instead of an adenosine. This is probably due to a point mutation in the parental stock that took place after this element inserted in the *sc* locus. The exact reason for the occurrence of this precise recombination event in four independent cases is not known at the moment, but it may be due to the presence of a mutator gene in the parental stock (21, 22). Additional experiments are now necessary to test this hypothesis. Phenotypic revertants of transposable element-induced mutations due to the insertion of a second transposon have also been reported by Mizrokhi *et al.* (23) and Mount *et al.* (24).

Analysis of these revertant stocks offers in addition important clues for understanding the molecular basis of *gypsy*-induced mutagenesis. The *gypsy* transposable element causes mutations whose phenotype is suppressed by second-site mutations at the *su*(*Hw*) locus. The exact mechanism whereby *gypsy* generates a mutant phenotype that can be reversed to wild type in the absence of the *su*(*Hw*) product is unknown. Based on the assumption that this mechanism should be the same for the different suppressible mutations, we have proposed that the *gypsy* element interferes with the proper transcription of the mutated gene, either at the level of initiation or elongation of the RNA. This interference is dependent on some idiosyncratic property of *gypsy*, such as its ability to be transcribed during the pupal stages of development, which, in turn, is contingent on the presence of *su*(*Hw*)-encoded product (8, 25). We present evidence in this study further supporting this hypothesis.

The expression of the *y* gene during the pupal stages of development is governed by three separate and independent enhancer elements that define the tissue-specific coloration of adult structures (10). Two of these enhancer elements are located in the 5′ region of the *y* locus, upstream from the insertion site of the *gypsy* element in the y^2 allele. It could then be argued that the mutant phenotype in these flies is a consequence of the distancing of these enhancer sequences from the *y* promoter by the insertion of a 7.4-kb *gypsy* element. The results presented above, showing that class B revertants still contain a 7.2-kb insert at −700 bp from the *y* transcription initiation site, argue against this possibility. These results state that it is not solely the insertion of DNA sequences but the nature of these sequences that is vital in originating the mutant phenotype. It is also apparent from these experiments that the properties of *gypsy* that cause suppressible mutations are dependent on sequences located in the internal region of the element, between both LTRs. The *su*(*Hw*)-encoded protein might modulate the mutagenic effect of *gypsy* by interacting with sequences in this internal region and disturbing the effect of the upstream enhancer elements on the transcription of the *y* locus.

We thank Dr. Robert Freund for communicating unpublished results on the sequence of the insertion point of the gypsy element in the sc^1 allele. This work was supported by Public Health Service Grants GM32036 and GM35463 from the National Institutes of Health. P.K.G. was supported by National Research Service Award GM11156 from the National Institutes of Health.

1. Rubin, G. M. (1983) in *Mobile Genetic Elements*, ed. Shapiro, J. (Academic, New York), pp. 329–361.
2. Finnegan, D. J. & Fawcett, D. H. (1986) in *Oxford Surveys in Eukaryotic Genes*, ed. Maclean, N. (Oxford Univ. Press, Oxford), Vol. 3, pp. 1–62.
3. Modolell, J., Bender, W. & Meselson, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1678–1682.
4. Parkhurst, S. M. & Corces, V. G. (1986) *BioEssays* **5**, 52–57.
5. Geyer, P. K., Spana, C. & Corces, V. G. (1986) *EMBO J.* **5**, 2657–2662.
6. Campuzano, S., Carramolino, L., Cabrera, C. V., Ruiz-Gomez, M., Villares, R., Boronat, A. & Modolell, J. (1985) *Cell* **40**, 327–338.
7. Biessman, H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7369–7373.
8. Parkhurst, S. M. & Corces, V. G. (1986) *Mol. Cell. Biol.* **6**, 47–53.
9. Chia, W., Howes, G., Martin, M., Meng, Y. B., Moses, K. & Tsubota, S. (1986) *EMBO J.* **5**, 3597–3605.
10. Geyer, P. K. & Corces, V. G. (1987) *Genes Dev.* **1**, 996–1004.
11. Geyer, P. K., Green, M. M. & Corces, V. G. (1988) in *Eukaryotic Transposable Elements as Mutagenic Agents*, eds. Lambert, M. E., McDonald, J. F. & Weinstein, I. B. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 123–130.
12. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
13. Geyer, P. K. & Fyrberg, E. A. (1986) *Mol. Cell. Biol.* **6**, 3388–3396.
14. Spradling, A. C. & Mahowald, A. P. (1979) *Cell* **16**, 589–598.
15. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
16. Di Nocera, P. P. & Casari, G. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5843–5847.
17. Di Nocera, P. P., Digan, M. E. & Dawid, I. (1983) *J. Mol. Biol.* **168**, 715–727.
18. Mozer, B., Marlor, R., Parkhurst, S. & Corces, V. (1985) *Mol. Cell. Biol.* **5**, 885–889.

19. Freund, R. & Meselson, M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4462–4464.
20. Marlor, R. L., Parkhurst, S. M. & Corces, V. G. (1986) *Mol. Cell. Biol.* **6,** 1129–1134.
21. Green, M. M. (1970) *Mutat. Res.* **10,** 353–363.
22. Green, M. M. & Lefevre, G., Jr. (1972) *Mutat. Res.* **16,** 59–64.
23. Mizrokhi, L. J., Obolenkova, L. A., Priimagi, A. F., Ilyin, Y. V., Gerasimova, T. I. & Georgiev, G. P. (1985) *EMBO J.* **4,** 3781–3787.
24. Mount, S. M., Green, M. M. & Rubin, G. M. (1988) *Genetics* **118,** 221–234.
25. Parkhurst, S. M. & Corces, V. G. (1985) *Cell* **41,** 429–437.