Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 8593–8597, November 1988
Genetics

Mutant gene phenotypes mediated by a *Drosophila melanogaster* retrotransposon require sequences homologous to mammalian enhancers

(transposable element/mutagenesis/transcriptional enhancer/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, August 16, 1988

ABSTRACT **We have analyzed the molecular structure of phenotypic revertants of *gypsy*-induced mutations to understand the molecular mechanisms by which this retrotransposon causes mutant phenotypes in *Drosophila melanogaster*. The independent partial revertants analyzed are caused by the insertion of different transposons into the same region of *gypsy*. One partial revertant of the yellow allele y^2 arose as a consequence of the insertion of the *jockey* mobile element into *gypsy* sequences, whereas a second incomplete revertant is due to the insertion of the *hobo* transposon. In addition, a previously isolated partial revertant of the Hairy-wing allele Hw^1 resulted from the integration of the *BS* transposable element into the same *gypsy* sequences. The region affected by the insertion of the three transposons contains 12 copies of a repeated motif that shows striking homology to mammalian transcriptional enhancers. Our results suggest that these sequences, which might be involved in the transcriptional control of the *gypsy* element, are also responsible for the induction of mutant phenotypes by this retrotransposon.**

The *gypsy* mobile element of *Drosophila melanogaster* is a retrovirus-like transposon composed of (*i*) two 482-base-pair (bp)-long terminal repeats (LTRs) containing transcription initiation and termination signals and (*ii*) a central region with three different open reading frames that encode products homologous to retroviral proteins (1). The insertion of this element at different gene loci generates mutant alleles whose phenotype can be modified by nonallelic mutations of suppressor and enhancer genes (2–4).

The molecular basis of *gypsy*-induced phenotypes has been studied by using various genes as model systems. The recessive, X-chromosome yellow [*y*, 1-0.0 (location 0.0 on chromosome 1)] mutation y^2 is caused by the insertion of *gypsy* 700 bp upstream of the *y* mRNA cap site, resulting in a temporal and tissue-specific phenotype manifested in adult flies with wild-type pigmentation in the bristles but mutant wing blade and body cuticle coloration (5–8). This phenotype can be reversed by mutations at the unlinked suppressor of Hairy-wing [*su(Hw)*, 3-54.8] locus and is thought to be a consequence of the interaction of the *su(Hw)*-encoded protein with tissue-specific transcriptional enhancers of the yellow gene (9) mediated by sequences located in the central region of the *gypsy* element (10). The interaction of specific *gypsy* sequences with regulatory regions of the affected gene might also be responsible for the mutant phenotype in the forked (*f*, 1-56.7), bithorax (*bx*, 3-58.8), and bithoraxoid (*bxd*, 3-58.8) mutations that result from the insertion of the *gypsy* element into an intron or in the 5′ region of these genes, although other types of effects were not completely ruled out (11, 12). An apparently different situation is that of the Hairy-wing (*Hw*, 1-0.0) Hw^1 mutation caused by the insertion of *gypsy* into an exon of one of the transcription units of the achaete-scute locus. This insertion results in high levels of accumulation of a truncated RNA whose transcription terminates at the *gypsy* 5′ LTR (13). In this case, the mutant phenotype might not be due to premature termination of transcription but rather to the high levels and/or improper spatial expression of the mutant gene, again due to the inadequate functioning of regulatory sequences as a consequence of the insertion of *gypsy* DNA.

In spite of the wide variety of insertion sites of the *gypsy* element within the mutated genes, the molecular mechanisms by which each mutant phenotype arises seems invariably to involve the interaction of sequences located in the central region of *gypsy*, probably towards the 5′ end (10, 12), with regulatory signals of the affected genes. To identify the *gypsy* sequences involved in this phenomenon, we have characterized two partial revertants of y^2 and a previously isolated partial revertant of Hw^1 (13). Here we present evidence that a specific region of *gypsy*, which contains 12 copies of a sequence homologous to mammalian enhancers, is responsible for the generation of mutant phenotypes by this retrotransposon.

MATERIALS AND METHODS

Fly stocks were maintained at 22°C and 65% relative humidity. Plasmid DNA isolation, DNA enzymology, and screening of phage λ libraries were carried out by standard procedures (14). *D. melanogaster* DNA was prepared as described (7). DNA samples were digested with restriction enzymes, electrophoresed on a 1% agarose gel, and subjected to Southern analysis (10). Phage λ libraries were constructed by inserting an *Mbo* I partial digest of genomic DNA into the *Bam*HI site of the EMBL-3 λ vector (15). DNA sequencing was done with Sequenase (United States Biochemical, Cleveland) by following the protocol provided by the manufacturer.

RESULTS

Origin of the Partial Revertants of y^2. Two independent partial revertants of the y^2 mutation, designated y^{2PR1} and y^{2PR2}, were isolated. Each was identified objectively in males by the change in the pigmentation phenotype of the caudal abdominal sclerites from the brown of y^2 to black. In females, neither y^{2PR1} nor y^{2PR2} homozygotes could be distinguished from homozygous y^2. The y^{2PR1} males could be separated

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.

from y^{2PR2} males on the basis of the darker wing coloration of the former.

The y^{2PR1} mutation arose at a frequency of *ca.* 1 in 5000 chromosomes in females homozygous for both y^2 and a third chromosome mutation designated *mu-1* (16). (The *mu-1* mutation is in all probability a DNA repair-defective mutation.) Once recovered, y^{2PR1} mutations were stable.

The y^{2PR2} mutation was recovered as two males from a female homozygous for y^2 sn^3 lz^{50e} ras^2 *v m* (see ref. 17 for a description of these mutations) crossed to a male with a wild-type X chromosome. Since y^{2PR2} could not be identified in females, it is probable that it occurred as a premeiotic mutation whose cluster was greater than two. Presumably the mutation occurred spontaneously. In contrast to y^{2PR1}, y^{2PR2} was not mutationally stable. Thus far seven independent mutations of y^{2PR2} to y^2 have been identified as well as two independent deletion mutations, which include the *y* and *ac* gene loci and which we designate y^- ac^-. Mutations of y^{2PR2} to y^2 occurred in either sex at a crudely estimated frequency of 1 in 10,000 X chromosomes.

An attempt was made to mobilize the *jockey* and *hobo* elements in y^{2PR1} and y^{2PR2} (see below) by using an *MR* element, which is known to excise *P* elements. Accordingly, males carrying the patroclinously inherited *MR* plus y^{2PR1} or y^{2PR2} were crossed to double-X-chromosome females (17), and their progeny were scored for *y* mutations. In the y^{2PR1} experiment, no *y* locus mutations occurred among 6261 male progeny scored. In the y^{2PR2} experiment, 1 mutation to y^2 was recovered among 7072 male progeny scored. The latter frequency does not differ from that of controls. Thus, the *MR* element failed to generate excision of either *jockey* or *hobo*.

The partial revertant of Hw^1 designated Hw^{BS} had a slightly weaker phenotype than its parental stock, and its characterization has been described elsewhere (13).

Molecular Analysis of yellow and Hairy-Wing Partial Revertants. As a first step in determining the molecular basis of the partial reversion of y^2, genomic DNA was obtained from y^{2PR1} and y^{2PR2} adult flies, digested with various restriction enzymes, and subjected to Southern analysis (data not shown). Results obtained from this experiment suggested that both reversion events were due to the insertion of DNA sequences into the *gypsy* element present in y^2. Genomic phage λ libraries were then constructed and screened with a ^{32}P-labeled *Sal* I–*Eco*RI fragment containing sequences from the yellow gene (7). A restriction map of this gene in each of the revertants was deduced from the restriction analysis of clones obtained from the λ libraries and is shown in Fig. 1.

DNA sequence analysis was then performed to determine the nature and precise position of the insert. Partial revertant y^{2PR1} was caused by the insertion of 2.6 kilobases (kb) of DNA in the 5′ region of the *gypsy* element, downstream from the 5′ LTR, at 823 bp from the beginning of the element (Fig. 2) (1). The inserted DNA contains a poly(A) tail, and its restriction map (Fig. 1) and sequence (Fig. 2) coincide with those of the *jockey* element (18). Partial revertant y^{2PR2} resulted from the insertion of 1.8 kb of DNA in the same region of the *gypsy* element at nucleotide 813 (Fig. 2). The sequence and restriction map of this DNA correspond to a copy of a deleted *hobo* element (19).

In addition, we determined the precise location of the insertion of the *BS* element into *gypsy* in the Hw^{BS} partial revertant of Hw^1 (13). The *BS* element contains a poly(A) tail, but it is structurally different, both in its restriction map and DNA sequence, from the *jockey* transposon (Fig. 2). It is inserted at nucleotide 933, in the same region of *gypsy* where *hobo* and *jockey* are inserted in the yellow partial revertants (Fig. 3*A*).

Molecular Analysis of Mutations from y^{2PR2} to y^2 and to y^- ac^-. As noted, y^{2PR2} is unstable and mutates spontaneously to a mutation whose phenotype is inseparable from that of y^2 as well as to a more extreme y^- ac^- mutation, the latter presumably a deletion of both genes. To understand the molecular basis of these events, DNA was prepared from adult flies homozygous for each of four new y^2 and one y^- ac^- mutations. The DNAs were digested with different restriction enzymes and subjected to Southern analysis (data not shown). The results of this experiment suggested that all four new mutations to y^2 that originated from y^{2PR2} are structurally the same and that most if not all of the *hobo* element inserted into *gypsy* in y^{2PR2} had been excised in these new mutants. The coding region of the yellow locus, located to the right of the *hobo* insertion site on the restriction map shown in Fig. 1, was deleted in the y^- ac^- flies, although the exact limits of the deletion were not determined. To more precisely analyze the events that restored the parental phenotype of the new y^2 mutations, we prepared genomic phage

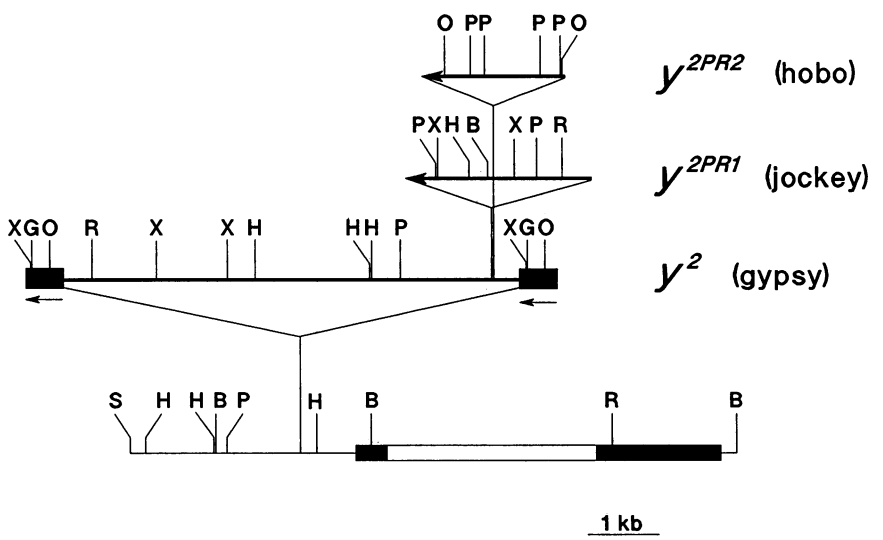

FIG. 1. Restriction map of the yellow locus in the gypsy-induced y^2 allele and partial revertants. The transcription unit of the yellow locus is indicated by boxes; filled-in boxes indicate exons, and the empty one indicates an intron. The *gypsy* element in the y^2 allele is inserted in the 5′ region of the yellow gene; filled-in boxes indicate the two LTRs. The *jockey* and *hobo* elements inserted into *gypsy* in the y^{2PR1} and y^{2PR2} partial revertants are designated by arrows that indicate their direction of transcription. Restriction enzyme symbols are as follows: B, *Bam*HI; G, *Bgl* II; H, *Hin*dIII; O, *Xho* I; P, *Pst* I; S, *Sal* I; X, *Xba* I.

y^{2PR1} 802
TACATTGCATACCCTCTTTTAAgattggcatgggagattagcaatcaagtggacgtgttcacaggactcgcggataa...

...tgggtaacagtgcgttgatgaaataataaaaacatcatcatcatgaatag(a)$_{10}$ ataccctctttaaTAAAAAATATTGC

y^{2PR2} 792
TTTAATAAAATACATTGCATACcagagaactgcagcccgccactcgcactctacgtccacccgataaacactcggta...

... aaaattgtagggtgtgagcgagtggaaaaaagtgccacccttgcagttctctgttgcatacCCTCTTTTAATAAAAAAT

Hw^{BS} 912
ACATTGCATACCCTCTTTTAATgacctgcagttttagcgctat.....gaggtttggttttatcttttatatgttaattgcgctgtt

atgttactgttattgcattgtattgattcatcgcttctaaataaataaat(a)$_{12}$ accctcttttaatAAAAAATATTGCATACGTT

FIG. 2. Sequence analysis of revertant stocks. The DNA sequence of the region surrounding the insertion point of the *jockey*, *hobo*, and *BS* elements into the *gypsy* retrotransposon was determined for the three partial revertants shown. Uppercase letters correspond to the *gypsy* element, whereas lowercase letters indicate the sequence of the transposon inserted into *gypsy*. The base pair duplications generated as a consequence of the insertion of these elements are boxed.

λ libraries and isolated clones containing the yellow locus from two stocks carrying these mutations, using as hybridization probe a *Sal* I–*Eco*RI fragment containing the wild-type yellow gene (7). Sequence analysis of the *gypsy* elements isolated from these stocks shows that the two new y^2 mutations examined were caused by the precise excision of the *hobo* element, including the 8-bp duplication created upon its insertion, leaving an intact *gypsy* element with exactly the same sequence as in the original parental y^2 mutation.

DISCUSSION

The results presented here offer new insights into the mechanisms by which the *gypsy* element causes mutant phenotypes. The two partial revertants of y^2 plus the previously isolated partial revertant of Hw^1 contain insertions of different transposable elements into the same region of *gypsy* (Fig. 3*A*). Although the effect of the insertion of these elements on *gypsy* function could be due to distinctive properties of these transposons, the fact that three structurally different ele-

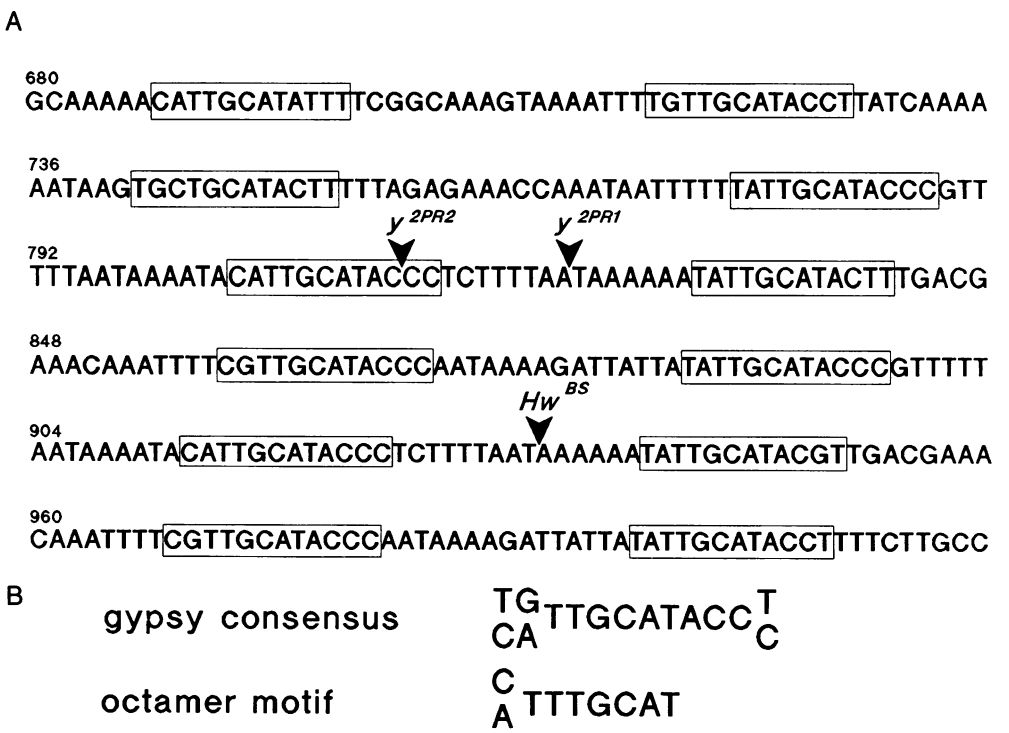

FIG. 3. Analysis of *gypsy* sequences involved in the induction of mutant phenotypes. (*A*) Sequence structure of the *gypsy* region affected in the partial revertants analyzed. Each of the 12 copies of the consensus sequence homologous to mammalian enhancer elements is boxed. The insertion points of the various transposable elements found in the different revertants analyzed are indicated by solid arrows. (*B*) Consensus sequence of the *gypsy* repeats and the octameric motif present in mammalian transcriptional enhancers.

ments have the same phenotypic effect suggests that the result is due to the alteration of *gypsy* sequences rather than to specific properties of the elements themselves and indicates that the region of the *gypsy* element where the *jockey*, *hobo*, and *BS* elements insert is directly involved in the generation of the mutant phenotype. In support of this conclusion, a partial revertant of the cut (*ct*, 1-20.0) allele ct^6 has been found to result from the insertion of *jockey* into the same region, although at a different place from the insertion of this element in y^{2PR2} (18). In addition, two partial revertants of the bithoraxoid allele bxd^1 are caused by the deletion of 109 bp from this same section (20).

Since the same region of the *gypsy* element is involved in the generation of the mutant phenotype in four different genes, the mechanisms underlying this phenomenon must be the same at all four loci. What is the role of this domain in normal *gypsy* function and in its mutagenicity? The sequence of this particular region of *gypsy* is displayed in Fig. 3*A*. Its most dramatic structural characteristic is the existence of 12 copies of the consensus sequence 5′-YRTTGCATACCY-3′ (R = unspecified purine nucleoside and Y = unspecified pyrimidine nucleoside), which shows striking homology to the octameric motif 5′-$^{C}_{A}$TTTGCAT-3′ found in transcriptional enhancers of viruses such as simian virus 40 and the mammalian immunoglobulin heavy chain, histone H2B, and U2 snRNA genes (Fig. 3*B*) (21, 22). These structural similarities suggest that this region might be involved in the transcriptional regulation of the *gypsy* element. This is supported by recent evidence indicating that the protein product of suppressor of Hairy-wing, a gene shown to be necessary for proper *gypsy* transcription *in vivo* (6, 11), also interacts *in vitro* with the specific region of *gypsy* that contains the 12 copies of the enhancer-like repeat (23, 24). The particular details of this interaction are not well understood at present, but the analysis of mutation events from y^{2PR2} to y^2 offers some insights into this problem. Although the mechanisms of reversion of *hobo*-induced mutations have not been studied in detail, its structural similarity with the *P* element suggests some possible outcomes. Revertants of *P*-induced mutations can occur by several mechanisms. The most frequent type of reversion results from imprecise excision of *P* elements, either by an excision of the element leaving the 8-bp duplication or by an imprecise excision involving the deletion of some additional adjacent sequences or leaving part of the 31-bp inverted repeats (25). Complete excisions of the *P* element, including the 8-bp duplication, have been obtained in cases in which the *P* element has inserted into the protein coding region of a gene, and therefore a perfect excision is required to restore function but at a frequency lower than usual by a factor of 10 (26). The fact that two independent y^{2PR2} to y^2 mutation events examined involved the clean excision of the *hobo* element including the additional 8-bp duplication suggests either that the *hobo* element excises precisely with high frequency or that a return to the y^2 phenotype requires this particular region of the *gypsy* element to be intact and suggests that even the insertion of 8 bp would cause a partial yellow phenotype similar to that of y^{2PR2}. Our results indicate that the severity of the phenotype is independent of the size of the inserted sequences because the phenotype resulting from the insertion of the 1.8-kb *hobo* element is similar to that induced by the 2.6-kb *jockey* element, and the insertion of the 6.5-kb *BS* element in Hw^{BS} also results in only partial reversion of the Hw^1 phenotype.

The results presented here suggest that the *gypsy*-induced phenotype is mediated by the interaction of the *su(Hw)*-encoded protein with the specific sequences of the *gypsy* element that are homologous to mammalian enhancers. This interaction could have two types of effects on the mutated genes. It could interfere with the proper interplay between the promoter of the mutant gene and transcriptional enhancers located distal to the *gypsy* insertion point; the consequence from such interference would be a loss of function phenotype such as it occurs in the y^2 allele (9). On the other hand, the insertion of the *gypsy* element could bring new transcriptional regulatory sequences in the proximity of a gene. These sequences could take over the normal regulation of the mutated gene and result in an altered gain of function mutation as is the case in Hw^1. The molecular mechanisms underlying both situations are the same, and it is only the particular characteristics of regulatory elements of the mutant gene and their arrangement with respect to the *gypsy* insertion point that determine the outcome.

Our molecular analysis of the partial revertants of y^2 provides some insights on the questions of integration of mobile elements and the origin of spontaneous mutations. As noted earlier, the y^{2PR1} mutation arose frequently under the influence of the *mu-1* mutation. The cytogenetics of *mu-1*-induced X-chromosome recessive lethal mutations showed most to be chromosome deletions (27). Thus, *mu-1* is in some way defective in repairing chromosome or chromatid breaks. Presumably the defect allows chromosome or chromatid breaks to remain "open" longer than normal, thereby allowing *jockey* to integrate into the *gypsy* element. This situation is not unlike that found for the integration of the *P* mobile element, which is mediated by the *MR* element. The *MR* element also generates inordinate numbers of chromosome (chromatid) breaks as evidenced by the increased frequency of mitotic crossing over in males (28). Thus, chromosome (chromatid) breakage is a concomitant of mobile-element integration. What is not explained by these observations is why specific mobile elements are integrated at specific sites and why *jockey* has an affinity for *gypsy*.

It should be noted that the *gypsy* element is not alone in being prone to secondary insertions of mobile elements. Partial reversions of the white-apricot mutation of *D. melanogaster* are found to be associated with secondary insertions into the *copia* mobile element (29), and the reversion of the original *w* mutation to w^e is associated with the insertion of the *pogo* element into *doc* (30). The reversions mediated by *mu-1* and mutations induced by *MR* emphasize the need for caution when concluding that mutational events are spontaneous. The discovery of the *mu-1* mutation was a wholly serendipitous event, and its cytogenetic identification is at best laborious. Therefore, there is good reason to believe that *mu-1* is segregating undetected in numerous *D. melanogaster* genetic stocks. Since *mu-1* mediates the integration of *jockey* into *gypsy*, it could be responsible for the "burst" of partial reversions of a *gypsy*-associated cut (*ct*) wing mutation of *D. melanogaster* (31, 32). Thus, before concluding that a specific mutational event is spontaneous by implication of an autonomous event, it is necessary to be certain that no mutator genes are involved. In all likelihood, mutator mutations analogous to *mu-1* are prevalent in *D. melanogaster* laboratory stocks and wild populations.

We thank J. Modolell and S. Campuzano for generously providing the clone containing the *BS* element and S. Mount for critical reading of the manuscript and valuable suggestions. This work was supported by Public Health Service Grants GM32036 and GM35463 from the National Institutes of Health.

1. Marlor, R. L., Parkhurst, S. M. & Corces, V. G. (1986) *Mol. Cell. Biol.* **6**, 1129–1134.
2. Modolell, J., Bender, W. & Meselson, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1678–1682.
3. Parkhurst, S. M. & Corces, V. G. (1986) *BioEssays* **5**, 52–57.
4. Rutledge, B. J., Mortin, M. A., Schwarz, E., Thierry-Mieg, D. & Meselson, M. (1988) *Genetics* **119**, 391–397.
5. Biessman, H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7369–7373.
6. Parkhurst, S. M. & Corces, V. G. (1986) *Mol. Cell. Biol.* **6**, 47–53.

7. Geyer, P. K., Spana, C. & Corces, V. G. (1986) *EMBO J.* **5,** 2657–2662.
8. Chia, W., Howes, G., Martin, M., Meng, Y. B., Moses, K. & Tsubota, S. (1986) *EMBO J.* **5,** 3597–3605.
9. Geyer, P. K. & Corces, V. G. (1987) *Genes Dev.* **1,** 996–1004.
10. Geyer, P. K., Green, M. M. & Corces, V. G. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3938–3942.
11. Parkhurst, S. M. & Corces, V. G. (1985) *Cell* **41,** 429–437.
12. Peifer, M. & Bender, W. (1986) *EMBO J.* **5,** 2293–2303.
13. Campuzano, S., Balcells, L., Villares, R., Carramolino, L., Garcia-Alonso, L. & Modolell, J. (1986) *Cell* **44,** 303–312.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
15. Frischauf, A.-M., Lehrach, H., Poustka, A. & Murray, N. (1983) *J. Mol. Biol.* **170,** 827–842.
16. Green, M. M. (1970) *Mutat. Res.* **10,** 353–363.
17. Lindsley, D. & Grell, E. (1968) *Carnegie Inst. Washington Publ.* **627.**
18. Mizrokhi, L. J., Georgieva, S. G. & Ilyin, Y. V. (1988) *Cell* **5,** 685–691.
19. Streck, R. D., MacGaffey, J. E. & Beckendorf, S. K. (1986) *EMBO J.* **5,** 3615–3623.
20. Peifer, M. & Bender, W. (1988) *Proc. Natl. Acad. Sci. USA*, in press.
21. Bohmann, D., Keller, W., Dale, T., Scholer, H. R., Tebb, G. & Mattaj, I. W. (1987) *Nature (London)* **325,** 268–272.
22. Rosales, R., Vigneron, M., Macchi, M., Davidson, I., Xiao, J. H. & Chambon, P. (1987) *EMBO J.* **6,** 3015–3025.
23. Parkhurst, S. M., Harrison, D. A., Remington, M. P., Spana, C., Kelley, R. L., Coyne, R. S. & Corces, V. C. (1988) *Genes Dev.* **2,** 1205–1215.
24. Spana, C., Harrison, D. A. & Corces, V. G. (1988) *Genes Dev.*, in press.
25. Geyer, P. K., Richardson, K. L., Corces, V. G. & Green, M. M. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6455–6459.
26. O'Hare, K. & Rubin, G. M. (1983) *Cell* **34,** 25–35.
27. Green, M. M. & Lefevre, G., Jr. (1972) *Mutat. Res.* **16,** 59–64.
28. Hiraizumi, Y. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 268–270.
29. Mount, S. M., Green, M. M. & Rubin, G. M. (1988) *Genetics* **118,** 221–234.
30. Finnegan, D. J. & Fawcett, D. H. (1986) in *Oxford Surveys on Eukaryotic Genes*, ed. Maclean, N. (Oxford Univ. Press, Oxford), Vol. 3, pp. 1–62.
31. Gerasimova, T. I., Matjunina, L. V., Mizrokhi, L. J. & Georgiev, G. P. (1985) *EMBO J.* **4,** 3773–3779.
32. Mizrokhi, L. J., Obolenkova, L. A., Priimagi, A. F., Ilyin, Y. V., Gerasimova, T. I. & Georgiev, G. P. (1985) *EMBO J.* **4,** 3781–3787.