Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 1751–1755, March 1985
Genetics

Molecular genetics of a transposon-induced dominant mutation in the *Drosophila* locus Glued

(premature termination of gene transcription)

ANAND SWAROOP, M. LUISA PACO-LARSON, AND ALAN GAREN

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511

Contributed by Alan Garen, October 15, 1984

ABSTRACT **The organization of the *Drosophila* locus Glued containing the dominant allele *Gl* was shown to differ from that of the normal locus by an insertion of a 9-kilobase-pair DNA segment near the 3′ end of a transcribed region. The insertion causes the formation of a truncated polyadenylylated transcript of 5.1 kilobases instead of the normal 6.0 kilobases. The inserted DNA segment has the properties of a transposon and was identified by its corresponding restriction map as *B104*, which is a retrovirus-like transposon with direct terminal repeats. *B104* appears to be oriented in *Gl* with the same polarity of transcription as *Gl*. The truncated *Gl* transcript terminates prematurely inside the 5′ terminal repeat of *B104*, in the region of a putative polyadenylylation signal. We discuss the general implications of this finding for transposon- and retrovirus-induced mutagenesis and for the origin of dominant mutations.**

Although dominant mutations are generally rarer events than recessive mutations, their impact on a population is disproportionately great because of the relatively high frequency of heterozygous individuals that express a dominant phenotype. Dominant mutations have been identified and genetically characterized in various organisms, notably in *Drosophila* (1), but the molecular basis of dominant effects is not well understood. In *Drosophila*, most genes appear to be haplo-sufficient (2); that is, viable adults are still produced when one copy of a diploid gene is deleted from one of the homologs. Accordingly, a dominant allele of a haplo-sufficient gene would have to act in *trans* to interfere with the function controlled by the normal gene in heterozygous individuals, presumably by specifying either an abnormal product or a product that is regulated abnormally. Since such effects are likely to be highly specific, this requirement could account for the rarity of dominant mutations.

The dominant allele *Gl* of the haplo-sufficient *Drosophila* locus Glued was chosen for the molecular analysis of a dominant effect, because there is extensive information about the genetic and developmental properties of the locus (3–6), and genomic clones of the Glued region were isolated recently from wild-type strains (unpublished results). *Gl* causes pleiotropic nonlethal developmental abnormalities in heterozygous Gl/Gl^+ flies, most strikingly evident in the severe structural and functional defects of the visual system, including the eye and its neural connections to the optic lobe of the brain (3–5). *Gl* also has a recessive lethal effect during early development (5, 6). In this report, we compare the organization and expression of the normal Glued locus with the locus containing the dominant allele *Gl*, and demonstrate that the two differ by a transposon insertion in the dominant allele, which results in the formation of a truncated transcript.*

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.

MATERIALS AND METHODS

The wild type carrying the normal Glued locus, located at 70C2 on chromosome 3L (6), is an inbred strain of *Drosophila melanogaster* Oregon R maintained in this laboratory for about 12 yr by mass transfers; a derivative with the chromosome 3 markers *st red ca* was used for certain experiments as indicated. The mutant strain carrying the dominant Glued allele *Gl* and the chromosome 3 marker *st* was maintained over the chromosome 3 balancer *TM3 Ser*, as described previously (6). The heterozygous flies used for the preparation of genomic DNA were *Gl st*/Gl^+ *st red ca* progeny of a cross between *Gl st/TM3 Ser* and *st red ca/st red ca* parents. The heterozygous third-instar larvae used for the preparation of poly(A)$^+$ RNA were collected from the *Gl st/TM3 Ser* stock.

The methods used to prepare genomic DNA (7), cloned DNA from λ phage and pBR322 plasmids (8), total RNA (9), and poly(A)$^+$ RNA (8) are described in the references cited. For preparing and screening libraries of *Drosophila* DNA clones, mapping the restriction sites of the cloned DNA, and preparing and hybridizing Southern blots of DNA and blots of electrophoretically resolved poly(A)$^+$ RNA, the methods were as described (8). DNA probes for hybridization with DNA and RNA blots were prepared by nick-translation (10).

RESULTS

In a parallel study (to be published elsewhere) of the molecular organization of the genomic DNA from the normal Glued locus, it was shown that the region specifies two coordinately expressed polyadenylylated transcripts encoded on opposite DNA strands (Fig. 1). The clones required for a molecular analysis of the dominant Glued allele (*Gl*) had to be isolated from a heterozygous Gl/Gl^+ genomic library, because homozygous *Gl/Gl* individuals die too early in development to be a source of homozygous DNA. Before attempting to screen the heterozygous DNA library with a cloned DNA fragment from the normal locus, we first tested several of the fragments to determine whether it would be possible to distinguish between the DNA from *Gl* and Gl^+ clones. The tests involved comparing the patterns of hybridization bands produced when the fragments were hybridized with Southern blots of *Eco*RI digests of homozygous Gl^+ and heterozygous Gl/Gl^+ genomic DNA (Fig. 2). The results show that fragment 2 hybridizes with two additional fragments in the heterozygous DNA, which are not detected in the homozygous DNA and therefore should be specific for *Gl*. Genomic clones containing the *Gl*-specific fragments and flanking DNA then were isolated from the heterozygous DNA li-

Abbreviations: kb, kilobase(s); kbp, kilobase pair(s).

*Our results demonstrating the formation of a truncated *Gl* transcript as a result of a transposon insertion were presented at the 25th Annual Drosophila Research Conference in Chicago, April 27, 1984.

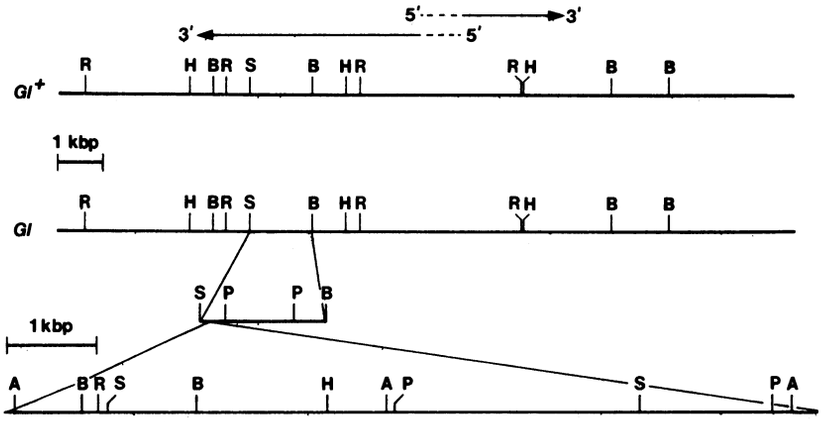

FIG. 1. Organization of genomic DNA from the Glued locus of Oregon R (Gl^+) and the dominant Glued mutant allele (Gl). The analysis of Gl^+ DNA will be described elsewhere. Restriction enzyme sites: R, *Eco*RI; B, *Bam*HI; S, *Sal* I; H, *Hin*dIII; A, *Ava* I; P, *Pst* I. The transcribed regions in the Gl^+ DNA, indicated by arrows above the map, were identified by hybridization of genomic DNA fragments with blots of electrophoretically resolved poly(A)$^+$ RNA, and the polarity of transcription was determined by hybridization tests with cDNA probes and strand-specific RNA probes. Broken lines at 5' ends of arrows indicate that the DNA fragments hybridized with both transcripts.

brary, and these were used to construct a map of restriction sites in *Gl*. A comparison of the restriction maps of *Gl* and Gl^+ shows that the two differ by a 9-kilobase-pair (kbp) insert in *Gl*, located near the 3' end of the region encoding the larger of the two transcripts (Fig. 1).

The effect of the insert on transcription of *Gl* was examined by hybridizing DNA probes for both transcripts with blots of electrophoretically resolved poly(A)$^+$ RNA from homozygous Gl^+ and heterozygous Gl/Gl^+ larvae (Fig. 3). No change was detected in the size of the smaller transcript in the heterozygous poly(A)$^+$ RNA; this result is consistent with the location of the insert outside the region encoding the smaller transcript. The probe for the larger, 6.0-kb transcript also hybridized with a 5.1-kb poly(A)$^+$ RNA, which is present only in the heterozygous poly(A)$^+$ RNA and is therefore encoded by *Gl*. The 5.1-kb poly(A)$^+$ RNA appears to be a truncated form of the larger transcript, resulting from the premature termination of transcription. The termination site in *Gl* was estimated to map near the right end of the insert, based on the lengths of the normal and truncated transcripts and on the map positions of the insertion site and the 3' end of the normal transcript (see Fig. 1). A more precise estimate was obtained by nuclease Sl mapping (8) of a DNA fragment from that region, after hybridization of the DNA with homozygous Gl^+ or heterozygous Gl/Gl^+ poly(A)$^+$ RNA (Fig. 4). The results indicate that transcription of *Gl* terminates about 0.4 kbp within the right end of the insert.

Several *Drosophila* mutations have been associated with the insertion of a transposon within the region of the affected gene (11–16). The insert in *Gl* was analyzed for three proper-

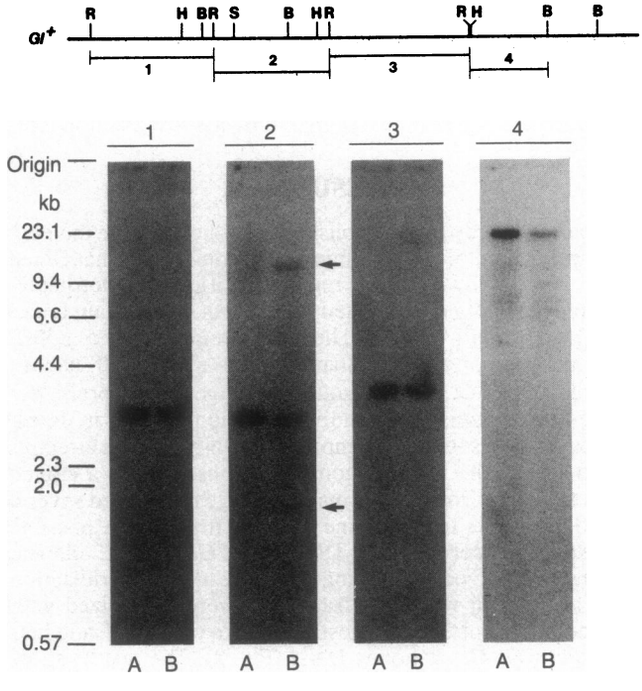

FIG. 2. Hybridization of Gl^+ DNA fragments with Southern blots of genomic DNA from homozygous Gl^+ and heterozygous Gl/Gl^+ flies. DNA fragments used as probes are shown on the map at the top. Homozygous DNA (10 μg, lanes A) and heterozygous DNA (10 μg, lanes B), each digested with *Eco*RI, were fractionated by electrophoresis in 1% agarose gel, transferred by blotting to nitrocellulose, hybridized with the nick-translated ^{32}P-labeled DNA probes, and autoradiographed. The two arrows indicate additional hybridization bands of probe 2 with the heterozygous DNA. The size markers at left correspond to *Hin*dIII fragments of λ phage DNA.

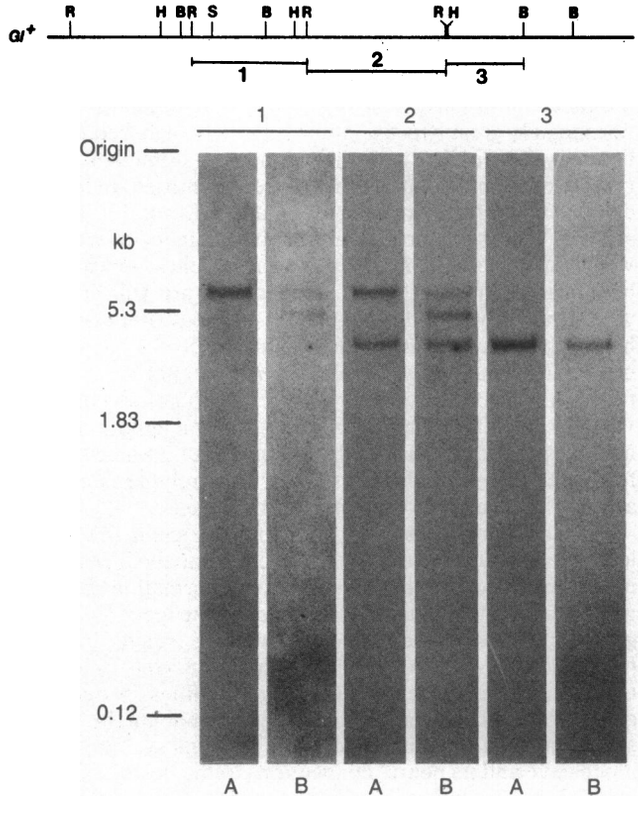

FIG. 3. Blot–hybridization analysis of electrophoretically resolved poly(A)$^+$ RNA from homozygous Gl^+ (lanes A) and heterozygous Gl/Gl^+ larvae (lanes B). The Gl^+ DNA fragments used as probes are shown on the map at the top. Each lane contained about 15 μg of poly(A)$^+$ RNA which was fractionated by electrophoresis in 1.5% agarose gel containing 2.2 M formaldehyde, transferred by blotting to nitrocellulose, hybridized with the ^{32}P-labeled DNA probes, and autoradiographed. Size markers at left correspond to mouse rRNAs; a *Hin*dIII digest of λ DNA was also used for estimating transcript sizes.

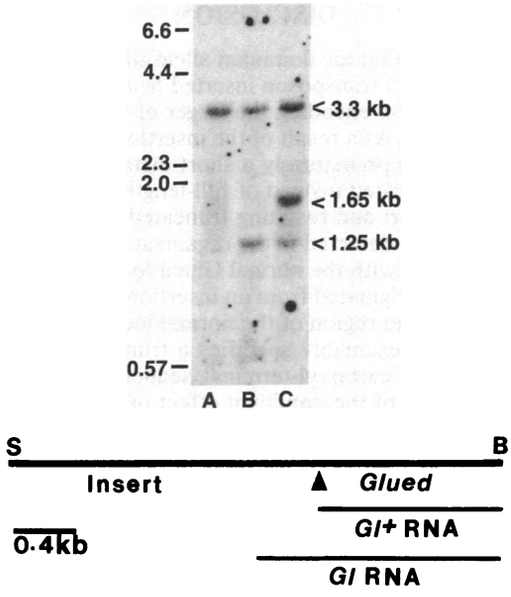

FIG. 4. Nuclease S1 mapping of the transcription termination site in *Gl*. The 3.3-kbp *Sal* I/*Bam*HI fragment, containing 2.05 kbp from the right end of the insert and 1.25 kbp of flanking *Gl* DNA, was hybridized with the following RNA samples: *Escherichia coli* tRNA (lane A), homozygous Gl^+ poly(A)$^+$ RNA (lane B), and heterozygous Gl/Gl^+ poly(A)$^+$ RNA (lane C). Hybridization and subsequent digestion with nuclease S1 were done as described (8). The sizes of the protected DNA fragments were determined by electrophoresis in 1% neutral agarose gel, transfer of the DNA to a nitrocellulose filter, hybridization with a ^{32}P-labeled Gl^+ DNA probe (fragment 2 shown at the top of Fig. 2), and autoradiography. Size markers (in kb) at left correspond to *Hin*dIII fragments of λ DNA. A diagram of our interpretation of the data is shown below the autoradiogram. ▲, Site of insertion in the *Sal* I (S)/*Bam*HI (B) fragment. The protected DNA region that hybridized with each poly(A)$^+$ RNA sample is indicated by a line below the map.

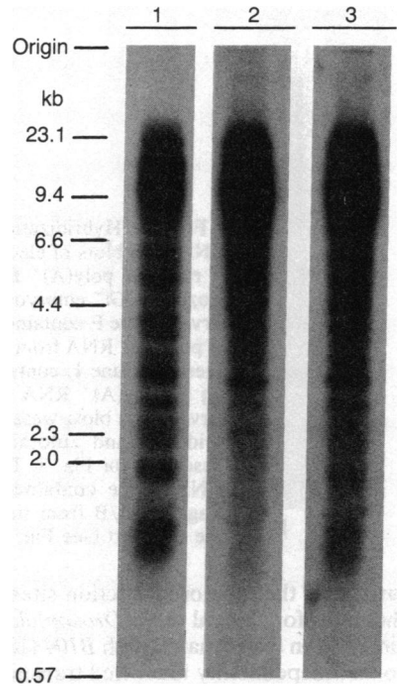

FIG. 5. Hybridization of *Gl* insert DNA with Southern blots of genomic DNA from homozygous Gl^+ flies. The following fragments, reading from left to right on the *Gl* map in Fig. 1, were used as hybridization probes: probe 1, 1.25-kbp fragment S/S at the left end of the insert; probe 2, 5.9-kbp fragment S/S spanning most of the insert; probe 3, 0.67-kbp fragment P/P at the right end of the insert. Each lane contained 10 μg of *Eco*RI digested genomic DNA; the Southern blots and hybridizations were done as described in *Materials and Methods*. Size markers at left correspond to *Hin*dIII-digested λ DNA.

ties of transposons: multiple genomic insertion sites which can vary among different strains, terminal repeat sequences, and transcription into poly(A)$^+$ RNA (16). Multiple genomic sites for the *Gl* insert were identified by hybridization of the insert DNA with Southern blots of *Eco*RI fragments from genomic DNA (Fig. 5) and also by *in situ* hybridization of the insert fragment to salivary gland chromosomes from a heterozygous Gl/Gl^+ strain (Fig. 6). The Southern blots show that the insert DNA hybridizes with numerous fragments of the genomic DNA, which map by *in situ* hybridization to at least 100 sites distributed throughout the *Drosophila* genome. There are differences between the insert hybridization sites in the two strains examined, as evident in regions where the homologs of the heterozygous chromosomes have separated, revealing a hybridization site on only one homolog (Fig. 6).

The tests for terminal repeat sequences in the *Gl* insert involved hybridization of the *Pst* I/*Pst* I fragment from the right end and the *Sal* I/*Bam*HI fragment from the left end (see Fig. 1) with fragments spanning the entire insert. Each end fragment hybridized with its cognate fragment and cross-hybridized with the other end fragment but did not hybridize with any of the intervening fragments (data not shown). These results are consistent with the presence of terminal repeat sequences within the end fragments.

The tests for transcribed sequences in the *Gl* insert involved hybridizing a 3.3-kbp *Sal* I/*Bam*HI fragment, which contains 2.0 kbp from the right end of the insert and 1.3 kbp of flanking *Gl* DNA, with blots of electrophoretically resolved poly(A)$^+$ RNA isolated from embryos and climbing larvae. This fragment produced a complex pattern of hybridization (Fig. 7), in contrast to the genomic DNA flanking the insert, which hybridizes only with the larger Glued transcript (cf. Fig. 3). These results indicate that the insert DNA is transcribed, probably at different genomic sites which might specify transcripts of different sizes.

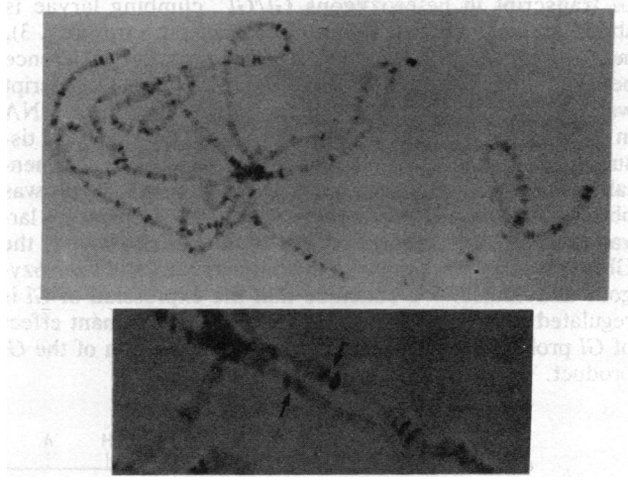

FIG. 6. *In situ* hybridization of *Gl* insert DNA with salivary gland chromosomes from heterozygous Gl/Gl^+ climbing larvae. The *Gl* insert DNA was the 5.9-kbp fragment S/S shown in Fig. 1. The DNA was labeled with biotin and hybridized with the chromosomes (17), after which the hybridization sites were visualized by reacting with avidin/peroxidase and staining with 4,4′-diaminobenzidine (Enzo Biochemicals, New York). (*Upper*) Multiple hybridization sites for an entire nucleus. (*Lower*) Region of homolog separation in another nucleus, with different hybridization sites on the two homologs, as indicated by the arrows.

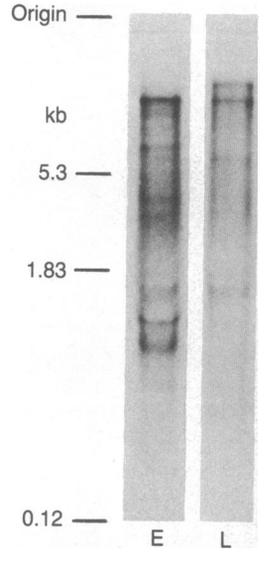

FIG. 7. Hybridization of *Gl* insert DNA with blots of electrophoretically resolved poly(A)$^+$ RNA from homozygous *Gl*$^+$ embryos and climbing larvae. Lane E contained about 15 μg of poly(A)$^+$ RNA from embryos of all ages, and lane L contained about 15 μg of poly(A)$^+$ RNA from climbing larvae. The blots were prepared, hybridized, and autoradiographed as described for Fig. 3. The ^{32}P-labeled DNA probe contained the 3.3-kbp fragment S/B from the right end of the *Gl* insert (see Fig. 1).

A comparison of the map of restriction sites in the *Gl* insert with the maps for several other *Drosophila* transposons revealed virtually an exact match with *B104* (18), which corresponds to an independently identified transposon *roo* (19). According to the polarity of transcription proposed for *B104* (18), the *B104* insert in *Gl* is oriented with the same polarity as the flanking *Gl* DNA (Fig. 8). Transcription of *Gl* should continue along the insert coding strand until a termination site is reached. There is a putative polyadenylylation sequence located 0.370–0.375 kbp from the 5′ end of the *B104* terminal repeat (18), about the position at which transcription of *Gl* terminates (Fig. 4). Since the *B104* terminal repeat can be translated to some extent in all three possible reading frames before a stop codon is reached (18), the polypeptide encoded by the truncated transcript *Gl* should have a carboxyl-terminal amino acid sequence specified by the *B104* terminal repeat.

The dominant effect of *Gl* could be caused by an abnormal regulation of *Gl* expression or by an abnormal function of the *Gl* product (see the Introduction). The amount of truncated *Gl* transcript in heterozygous *Gl*/*Gl*$^+$ climbing larvae is about the same as the amount of normal transcript (Fig. 3), indicating that there is no apparent quantitative difference between the expression of *Gl* and *Gl*$^+$. The normal transcript was shown, by *in situ* hybridization of RNA probes to RNA in tissue sections, to be widely distributed in virtually all tissues of homozygous *Gl*$^+$ individuals at various developmental stages (to be published elsewhere). A similar result was obtained with sections of heterozygous *Gl*/*Gl*$^+$ climbing larvae (Fig. 9), suggesting that the tissue distribution of the Glued transcript is similar in the homozygous and heterozygous individuals. We conclude that the expression of *Gl* is regulated normally and, therefore, that the dominant effect of *Gl* probably results from an abnormal function of the *Gl* product.

DISCUSSION

We have shown that the dominant allele *Gl* in the *Drosophila* locus Glued has a transposon inserted near the 3′ end of the transcribed region encoding the larger of two polyadenylylated transcripts. As a result of the insertion, transcription of *Gl* is terminated prematurely a short distance inside the insert, and a truncated instead of full-length transcript is produced. The insert and resulting truncated transcript are the only differences detected in the organization and function of *Gl* as compared with the normal Glued locus, providing evidence that *Gl* originated from an insertion of the transposon into a transcribed region of the normal locus. The truncated *Gl* transcript presumably specifies a truncated polypeptide with an altered carboxyl-terminal sequence. Although the molecular basis of the dominant effect of *Gl* is not fully understood, two important clues can be deduced from genetic studies. Since Glued is a haplo-sufficient locus, a dominant allele would have to interfere with the Glued function in heterozygotes (see the Introduction). Furthermore, since the *Gl*$^+$ function is essential early in development, as indicated by the early death of Glued deletion homozygotes (5, 6), a dominant allele would have to alter rather than block that function; otherwise viable heterozygous adults could not be produced. A mechanism for such a *trans* interaction between allelic genes might involve the association of normal and mutant monomers to form a heteromeric protein (20). This mechanism could exert a dominant effect for two reasons: (*i*) the heteromeric protein could have an altered function that interferes with the function of the normal protein and (*ii*) the incorporation of normal monomers into heteromeric molecules could reduce the amount of normal protein below the level required for normal development. Accordingly, the dominant effects of *Gl* could result from the formation of a heteromeric protein containing the Gl polypeptide.

The 9-kbp DNA insert in *Gl* has the properties of a *Drosophila* transposon, including numerous insertion sites widely dispersed in the genome, different genomic insertion sites in different strains, homologous terminal sequences, and polyadenylylated transcripts. The restriction map sites of the insert in *Gl* correspond with those of the retrovirus-like *Drosophila* transposon *B104*, which has identical direct terminal repeats, with putative promoter and polyadenylylation signals. Transcription of *B104* DNA probably initiates in the 5′ terminal repeat and terminates in the 3′ terminal repeat (18). However, premature termination of the *Gl* transcript occurs in the 5′ terminal repeat of the *B104* insert, apparently because the polarity of transcription of *B104* and *Gl* is the same, enabling a polyadenylylation signal in the 5′ terminal repeat to function during *Gl* transcription. This mechanism of transposon-induced mutagenesis should be general for all transposons and retroviruses that have a polyadenylylation signal in the 5′ terminal repeat.

Note Added in Proof. An insertion of *B104* in the white recessive allele *w*^{bf} of *Drosophila* was also found to cause premature termination of transcription (21).

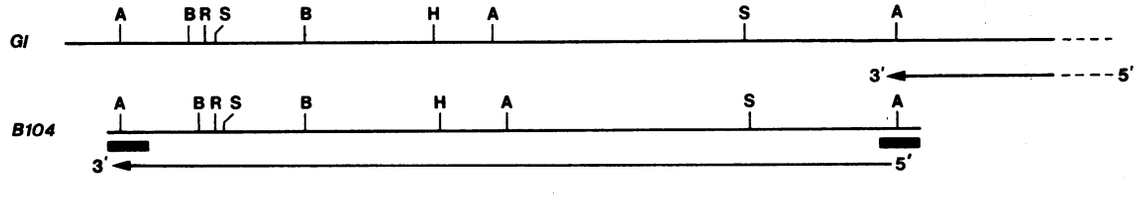

FIG. 8. Maps of restriction sites in *Gl* insert DNA and transposon *B104*. The map for *B104* was drawn as described elsewhere (18). Solid bars at each end of *B104* indicate direct terminal repeats, and the proposed direction of transcription is indicated by the arrow under the *B104* map. The terminal repeats are identified as 5′ or 3′ in the text, according to the direction of transcription. The arrow under the *Gl* insert map indicates the direction of transcription and the premature termination site of the larger *Gl* transcript. The symbols for restriction enzymes are as in Fig. 1.

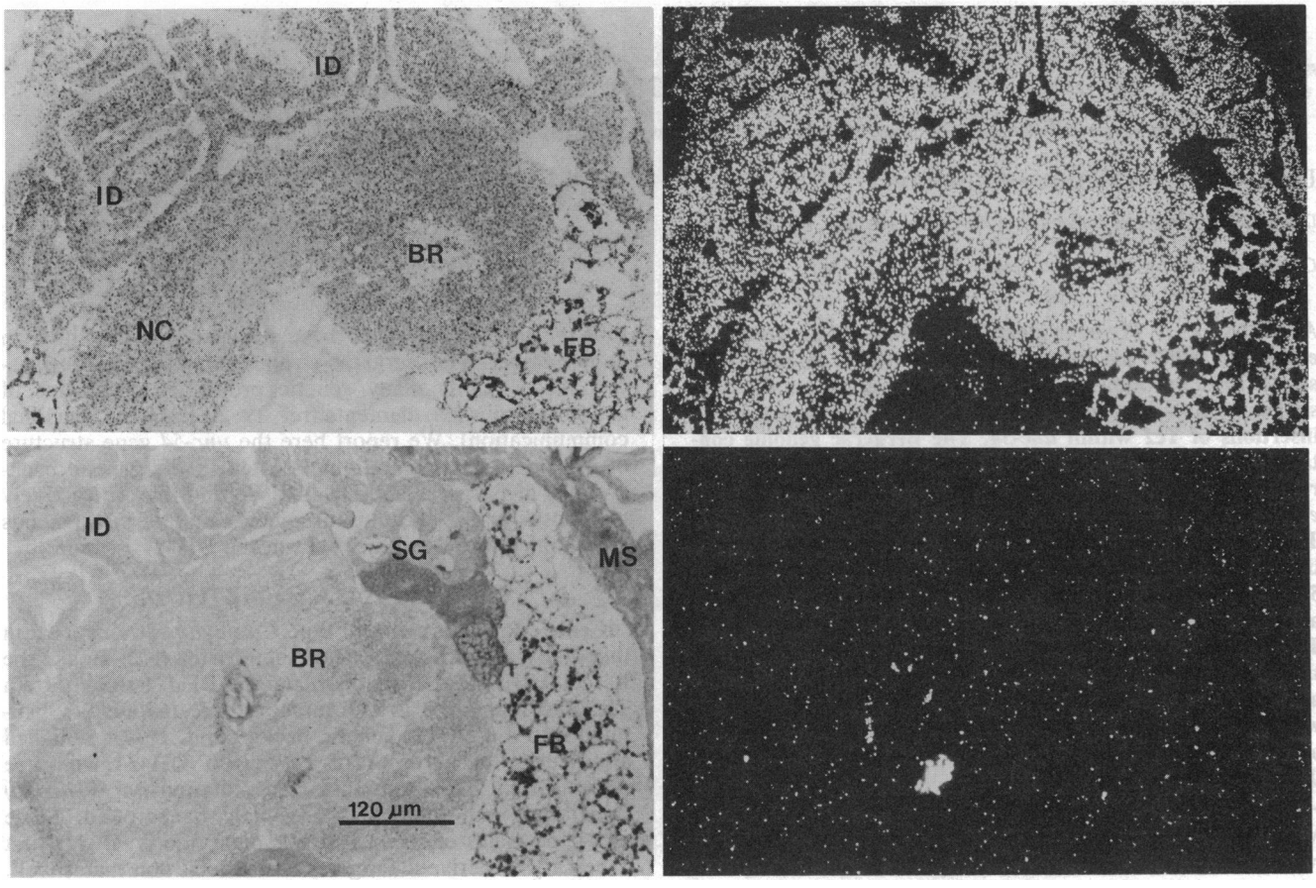

FIG. 9. *In situ* hybridization of Gl^+ RNA probes with RNA in a tissue section from a heterozygous Gl/Gl^+ climbing larva. The probes were prepared with the 2.7-kbp E/H fragment from the transcribed Gl^+ region encoding the 6-kb transcript, using the SP6-derived transcription system (Promega Biotec, Madison, WI), as will be described elsewhere. (*Upper*) Photographs in bright field (*Left*) and dark field (*Right*) of a larval section hybridized with a probe prepared from the Gl^+ DNA inserted in vector pSP65. (*Lower*) Corresponding photographs after hybridization with a probe prepared from the Gl^+ DNA inserted in vector pSP64. The RNA probes are transcribed from opposite strands of the Gl^+ DNA fragment in the two vectors. BR, brain; FB, fat body; NC, notochord; ID, imaginal disc; SG, salivary gland; MS, musculature.

The technical assistance of Mrs. Manju Swaroop and Mrs. Barbara Miller was an important contribution to these experiments. Support was provided by grants from the National Institute of General Medical Sciences, U.S. Public Health Service, and the American Cancer Society.

1. Lindsley, D. L. & Grell, E. H. (1968) *Carnegie Inst. Washington Publ.*, 627.
2. Lindsley, D., Sandler, L., Baker, B. S., Carpenter, A., Denell, R. E., Hall, J. C., Jacobs, P. A., Miklos, G. L. G., Davis, B. K., Gethmann, R. C., Hardy, R. W., Hessler, A., Miller, S. M., Nozawa, H., Perry, D. M. & Gould-Somero, M. (1972) *Genetics* **71**, 157–184.
3. Meyerowitz, E. M. & Kankel, D. R. (1978) *Dev. Biol.* **62**, 112–142.
4. Garen, S. H. & Kankel, D. R. (1983) *Dev. Biol.* **96**, 445–466.
5. Harte, P. J. & Kankel, D. R. (1982) *Genetics* **101**, 477–501.
6. Garen, A., Miller, B. R. & Paco-Larson, M. L. (1984) *Genetics* **107**, 645–655.
7. Levine, M., Garen, A., Lepesant, J. A. & Lepesant-Kejzlarova, J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2417–2421.
8. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
9. Artavanis-Tsakonas, S., Muskavitch, M. A. T. & Yedvobnick, B. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1977–1981.
10. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
11. Spradling, A. & Rubin, G. M. (1981) *Annu. Rev. Genet.* **15**, 219–264.
12. Rubin, G. M., Kidwell, M. G. & Bingham, P. M. (1982) *Cell* **29**, 987–994.
13. Lewis, R., Collins, M. & Rubin, G. M. (1982) *Cell* **30**, 551–565.
14. Bucheton, A., Paro, R., Sang, H. M., Pelisson, A. & Finnegan, D. J. (1984) *Cell* **38**, 153–163.
15. Mattox, W. W. & Davidson, N. (1984) *Mol. Cell. Biol.* **4**, 1343–1353.
16. Shapiro, J. A., ed. (1983) *Mobile Genetic Elements* (Academic, New York).
17. Langer-Safer, P. R., Levine, M. & Ward, D. C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4381–4395.
18. Scherer, G., Tschudi, C., Perera, J., Delius, H. & Pirrotta, V. (1982) *J. Mol. Biol.* **157**, 435–451.
19. Meyerowitz, E. M. & Hogness, D. S. (1982) *Cell* **28**, 165–176.
20. Garen, A. & Garen, S. (1963) *J. Mol. Biol.* **7**, 13–22.
21. Lewis, R., O'Hare, K. & Rubin, G. M. (1984) *Cell* **38**, 471–481.