Faculty Publications in the Biological Sciences

Papers in the Biological Sciences

1990

The Functional Organization of the Vestigial Locus in *Drosophila melanogaster*

Jim A. Williams
University of Alberta

Audrey L. Atkin
University of Nebraska-Lincoln, aatkin@unl.edu

John B. Bell
University of Alberta

The Functional Organization of the Vestigial Locus in *Drosophila melanogaster*

Jim A. Williams, Audrey L. Atkin, and John B. Bell

Department of Genetics, University of Alberta, Edmonton, Alberta, Canada

Corresponding author – J. B. Bell, Department of Genetics, University of Alberta, Edmonton, Alberta, Canada T6G 2E9

Current address – Jim A. Williams, Department of Zoology, University of British Columbia, Vancouver, BC, Canada G6T 2A9

Abstract

Vestigial mutants are associated with imaginal disc cell death which results in the deletion of adult wing and haltere structures. The vestigial locus has previously been cloned, and mutational lesions associated with a number of *vg* alleles were mapped within a 19 kb DNA region defined as essential for *vg* function. Herein we report the identification and characterization of a developmentally regulated 3.8 kb *vg* transcript which is spliced from exons distributed throughout the essential interval defined above. All the characterized classical alleles have predictable effects on this transcription unit, and the severity of this effect is directly proportional to the severity of the wing phenotype. A repetitive domain within this transcription unit was identified and may serve as a tag to isolate other genes with functions related to *vg*. We also report an exceptional *vg* allele (vg^{83b27}) that produces an extreme wing and haltere phenotype, but which defines a second *vg* complementation unit. This allele is associated with a 4 kb deletion entirely within a 4.5 kb *vg* intron as defined by the 3.8 kb transcription unit. Molecular and genetic evidence indicates that the vg^{83b27} mutation has a functional 3.8 kb transcription unit, thus accounting for its ability to complement classical alleles. The results indicate that sequences within a *vg* intron are essential for normal wing and haltere development.

Keywords: *Drosophila*, vestigial, transcripts, complementation, cDNA

Introduction

A wild-type vestigial (*vg*) gene is required for normal wing imaginal disc development, since the absence of a *vg⁺* gene product results in extensive cell death in this disc (Fristrom 1969). This results in concomitant complete loss of adult wing margin structures in strains containing null alleles, while hypomorphic alleles have less severe wing margin loss. Thus, the phenotypes produced by *vg* alleles range from those which are homozygous wild type through nicked, notched, or strap wing phenotypes to the classical, more extreme alleles (described in Lindsley and Grell 1968; Williams and Bell 1988). Bownes and Roberts (1981b) have proposed that the cell death observed in *vg* mutants may be the consequence of abnormal positional information in *vg* wing imaginal discs. Thus, a molecular analysis of *vg* is important since it may help to elucidate the mechanisms by which positional information in imaginal discs is established.

In addition to wing margin loss, the extreme *vg* alleles also exhibit haltere reduction, erect postscutellar bristles, female sterility, and other less well-defined phenotypes (i.e., extended duration of first and third larval instars, pupal lethality, and leg and abdominal abnormalities; Erk and Podraza 1986; Bownes and Roberts 1981a; Borot and Goux 1981). All but two extant extreme *vg* alleles are completely recessive and define a single complementation group in that they are noncomplementing in *trans* and affect all four phenotypes associated with such *vg* alleles. Of course, hypomorphic weaker *vg* alleles do not show each of these four phenotypes so complementation of these phenotypes in *trans* (i.e., female sterility) is possible (Lindsley and Grell 1968).

Three *vg* alleles exist which differ from those described above. Two of these behave as strong dominants (*vg^U* and *vg^W*). These dominant alleles both show more severe wing and haltere deficiencies when heterozygous with other *vg* alleles, indicating that they likely produce an antimorphic product. The third exceptional *vg* allele (*vg^{83b27}*) is unusual in that it defines a second *vg* complementation unit. The *vg^{83b27}* allele when homozygous produces an extreme wing phenotype, but complete or nearly complete complementation occurs when it is in *trans* with all recessive fertile and viable *vg* alleles (Alexandrov and Alexandrova 1987). The recessive sterile or lethal *vg* alleles are partially complemented or not complemented at all by *vg^{83b27}*. Moreover, homozygous *vg^{83b27}* flies affect only the wing and haltere phenotypes, and have normal female fertility and postscutellar bristles. Thus, *vg^{83b27}* appears to have a complex (or at least different) genetic basis (Alexandrov and Alexandrova 1987; and our unpublished results) than all other previously described *vg* alleles.

In previous studies (Williams and Bell 1988; Williams et al. 1988), we have reported the cloning of the *vg* locus and the physical mapping of the lesions associated with various *vg* alleles, including *vg^{83b27}*. The relevant respective physical lesions affecting *vg* function were localized to a 19 kb stretch of DNA within the cloned region. In the present study, cDNAs corresponding to a 3.8 kb *vg* transcript were isolated and located on the physical map within the 19 kb sequence of DNA previously shown to be involved in *vg* function. An exon map was established for the cDNA and correlated with respect to the lesions associated with the classical *vg* viable and lethal alleles previously placed on the physical map. Our results indicate that the 3.8 kb transcript is the functional *vg* transcript affected

by classical *vg* alleles. A novel repetitive internal region was identified which may serve as a valuable tag to clone other *Drosophila* genes functionally related to *vg*. Analysis of the vg^{83b27} allele indicates that it does not remove this transcription unit and defines further sequences, within a major *vg* intron, which are required for normal wing development.

Material and methods

Stocks and recombination analysis

Drosophila melanogaster stocks were grown at 24°C and maintained on a synthetic medium (Nash and Bell 1968). The phenotypes and origins of the alleles used in this study are described in Williams and Bell (1988) and Williams et al. (1988). Recombination analysis was utilized to demonstrate that both the complementation behavior and the wing phenotype of vg^{83b27} are closely linked to other *vg* alleles. Female vg^{83b27}/*b cn vg^{nw}* heterozygotes were mated to homozygous vg^{83b27} males and wild-type recombinants were selected. Two phenotypically wild-type flies were isolated from 1000 flies scored, but neither was recombinant. Rather, they were due to the weak (or occasional) complementation between vg^{83b27} and vg^{nw}. These results indicate that the mutant phenotype of vg^{83b27} is very closely linked to a classical *vg* allele (vg^{nw}). Heterozygous females of vg^{83b27}/vg^{79d5} and vg^{83b27}/vg^{21} genotypes (which are complementing combinations) were mated to homozygous vg^{79d5} and vg^{21} males respectively, to select recombinants that had lost the ability to complement the above alleles. No genetic recombinants were isolated from among the 5000 F_2 flies scored for each cross. These results demonstrate that the ability of vg^{83b27} to complement other alleles is not due to the presence of a specific second-site *vg* suppressor in the vg^{83b27} genetic background.

Materials

Restriction enzymes and other DNA modifying enzymes were obtained from BRL or Pharmacia and used according to the manufacturer's instructions. All radioisotopically labeled compounds were purchased from either New England Nuclear or ICN. Nick-translated and oligolabeled probes were made with $[\alpha^{32}P]$dCTP (3000 Ci/mmol); RNA probes were labeled with $[\alpha^{32}P]$UTP (3000 Ci/mmol), and DNA sequences utilized $[\alpha^{35}S]$dATP (500 Ci/mmol). The third instar larval disc cDNA library was kindly provided by Dr. G. Rubin. The RP49 clone was a gift from Dr. M. Rosbach.

DNA manipulations

All DNA isolations and manipulations were performed as previously described (Williams and Bell 1988).

Genomic libraries

The vg^{79d5} and vg^{83b27} libraries were constructed in λGT10. Genomic DNA was digested with *Eco*RI, size selected on 1% agarose gels, and electroeluted onto a dialysis membrane. Purification, ligation, and packaging were done as previously described (Williams et al.

1988). These genomic libraries, as well as the cDNA library, were plated on C600 Hfr, transferred to Biodyne membranes (Pall), and prepared for hybridization by standard methodologies (Maniatis et al. 1982).

Southern and Northern filter hybridizations

All gels for Southern or Northern hybridization analyses were blotted onto Genescreen Plus membranes using the capillary blot protocol recommended by the manufacturer (DuPont). All RNA samples were extracted by the guanidinium thiocyanate/CsCl method (Berger and Kimmel 1987) and analyzed on 1% formaldehyde agarose gels as described by Gietz and Hodgetts (1985). Hybridization conditions for all plaque lifts, genomic Southerns, and DNA-probed Northerns were according to Klessig and Berry (1983). Preparation of oligolabeled DNA probes and washing of filters were done as previously described (Williams and Bell 1988). RNA probes for Northern hybridization analyses were prepared from restriction fragments cloned into Bluescribe (Vector Cloning Systems) and using the transcription protocol of Melton et al. (1984). Hybridizations and washing of RNA-probed Northerns were as described by Williams et al. (1988). Southern hybridization blots of recombinant λ phage or plasmid DNAs were hybridized to nick-translated probes as described by Maniatis et al. (1982).

DNA sequencing

All sequencing was performed by either the dideoxy method (Sanger et al. 1977) from inserts in M13mp18 and M13mp19 or by double-stranded DNA sequencing (Chen and Seeburg 1985) of inserts cloned into Bluescribe.

Results

cDNA isolation

A restriction map of the *vg* region with the physical locations and the nature of lesions associated with various *vg* alleles is shown in figure 1. These mutations define a previously characterized region comprising approximately 19 kb of DNA, which is required for *vg* function (Williams and Bell 1988). Several DNA probes from within the 0 to + 16 interval (fig. 1) identified a low abundance 3.8 kb transcript, present in post 8–12 h embryos, but undetected in larval stages. An imaginal disc cDNA library derived from third instar larvae was screened and six *vg* cDNAs, whose restriction patterns indicated independent origins, were isolated. All six cDNAs hybridized to *Eco*RI DNA restriction fragments scattered throughout the 0 to + 16 interval on the physical map of genomic *vg* DNA. The hybridization patterns of cDNAs 1–4 were nearly identical except that cDNA1 hybridized to an additional proximal *Eco*RI fragment (fig. 1B). This is consistent with the observation that cDNA1 is the longest cDNA, encompassing two *Eco*RI fragments of 3.2 kb and 0.5 kb in size (fig. 1B). Subcloning, restriction mapping, and hybridization of cDNA1 restriction fragments to cloned genomic restriction digests of DNA from the *vg* region allowed orientation of the cDNA exons with respect to the genomic physical map and resulted in the definition of five exons. This alignment is presented in figure 1A, and a restriction map of

cDNA1 is shown in figure 1B with the extent of cDNAs 2–6 also indicated. cDNA6 is unusual in that it has homology to the 5′ and 3′ exons but not to one internal exon (see fig. 1B). Whether this represents a functionally significant splicing product or merely an aberrant event is unknown. Hybridization of radiolabeled cDNA subclones to Northern blots of RNA obtained throughout ontogeny identified a 3.8 kb transcript expressed in post 8–12 h embryos (see below). Thus, it appears that the isolated cDNAs correspond to the 3.8 kb transcript identified previously. Since cDNA1 is 3.7 kb long and cDNA3 extends a further 100 bp distally (fig. 1B), we feel that we have cDNAs representing essentially the full length of the *vg* transcription unit. Hybridization of single-stranded RNA probes to Northern blots indicates that transcription is from proximal to distal in relation to the physical map shown in figure 1A.
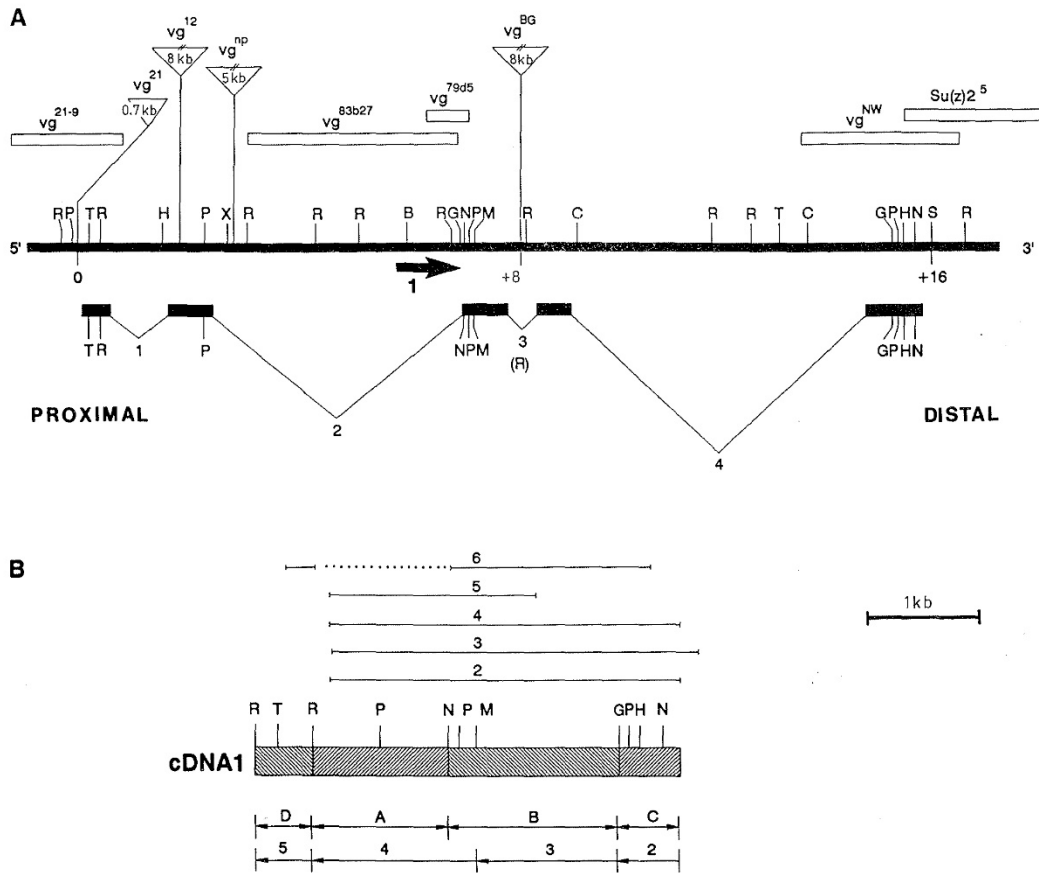

Figure 1A and B. The vestigial locus of *Drosophila melanogaster*. (A) A partial restriction map of the locus. Pertinent restriction sites are indicated: R, *Eco*RI; S, *Sal*I; T, *Sst*I; P, *Pst*I; B, *Bam*HI; G, *Bgl*II; H, *Hin*dIII; X, *Xho*I; C, *Cla*I; M, *Sma*I; N, *Hin*cII. The *open bars* above the restriction map designate deletions associated with specific *vg* alleles, while *triangles* designate insertion alleles. Coordinate 0 is designated as the site of the original P element insert (*vg^{21}*) used to clone the locus, and +16 is 16 kb distal to 0. The data are from Williams and Bell (1988) and Williams et al. (1988). The *bars* below the restriction map denote the

exons of cDNA1, with some pertinent restriction sites indicated. The four mapped intrans are labeled 1 to 4. The location of intron 2 has been conclusively determined by DNA sequencing of genomic and cDNA clones spanning this intron. Similar sequencing distal from the genomic *Eco*RI site (at coordinate +0.4) demonstrated that the cDNA internal *Eco*RI site is genomic and the 5′ splice site of intron 1 is located approximately 150 bp more distal to this site. The localization of intron 3 is inferred by the absence of an *Eco*RI site. This site is not polymorphic in any *Drosophila* stock we have examined. We cannot yet discount the possibility of micro-exons in introns 3 and 4 or of minor introns within the major exons. The splice junctions indicate that transcription is from left to right. The *black arrow* (labeled 1) denotes the orientation and extent of the antisense RNA probe used by Williams et al. (1988). Centromere proximal and distal is also indicated. (B) A restriction map of cDNA1. The terminal *Eco*RI sites are linkers utilized in the cDNA cloning protocol. The extent of cDNAs 2–6 is indicated by *numbered lines* above this map, and these lengths were determined by homology to restriction fragments and partial restriction mapping. Segments A–D (below cDNA1) denote the location and extent of DNA probes, and 2–5 (also below cDNA) the extent and polarity of the RNA probes used in this study. A 1 kb scale bar indicates the relative sizes of the cDNAs.

DNA sequencing of the putative 3′ end of cDNA1 identified two overlapping poly(A) addition sites preceded by a third poly(A) site. DNA sequencing of M13 clones of genomic *vg* DNA from the +16 *Sal*I site indicates that these poly(A) addition sites are located approximately 150 bp proximal to the *Sal*I site, and thus map the 3′ end of the locus to this region (fig. 2A). This is consistent with the distal limits of the locus as described above.

Correlating the genomic locations of exons and the lesions of classical **vg** *alleles*
Essentially, all recessive viable and recessive lethal *vg* alleles are noncomplementing and thus define a single complementation group. A previously characterized classical *vg* allele, *Df(2R)vg^{56}*, was shown to have a deletion endpoint near the centromere distal limit of the sequences required for *vg* function (Williams and Bell 1988). The very weak phenotype of *Df(2R)vg^{56}* (i.e., *Df(2R)vg^{56}/vg^{BG}* shows only a strap-wing phenotype) is likely due to a position effect of the *vg^{56}* breakpoint with respect to the boundaries of the *vg* transcription unit, since the centromere proximal deletion endpoint of *vg^{56}* is now known to be ~3 kb from the 3′ end of the *vg* transcription unit (at +19). The characterized alleles presented in figure 1 were aligned with respect to the exons, to predict how they would affect the 3.8 kb *vg* transcription unit. *Su(z)2^5* exhibits an intermediate *vg* phenotype, stronger than *Df(2R)vg^{56}* but weaker than *vg^{BG}*. Interestingly, the centromere proximal deletion endpoint of *Su(z)2^5* is within the *vg* transcription unit but only ~150 bp from the 3′ end (fig. 1). Thus, the *Su(z)2^5* breakpoint likely affects 3′ end maturation but probably does not affect the product drastically, resulting in only an intermediate *vg* phenotype. The deletion alleles *vg^{nw}* and *vg$^{21-9}$* (at the 3′ and 5′ end of the transcription unit, respectively) are null alleles, consistent with the fact that both deletions remove large exonic regions of the locus. The null allele, *vg^{12}*, has a large insertion into an exon, while the hypomorphic mutants *vg^{BG}* and *vg^{np}* are insertions into intronic regions. Sequence analysis of the insertion sites of these alleles will be required to determine the exact relationship of the respective insertions to the *vg* exons. The *vg^{79d5}* strain is homozygous viable with a strap-wing phenotype, and

the physical lesion associated with it is an ~400 bp deletion (fig. 1). Cloning of the altered region and DNA sequence analysis indicates that the distal endpoint of the vg^{79d5} deletion removes 28 bp of a *vg* exon (see below). The use of more distal alternative splice sites would generate an internally truncated transcript which may retain partial function. A P element insertion allele (vg^{21}) is located near the 5′ end of the *vg* transcription unit (fig. 1). The genomic DNA sequence of the region surrounding the vg^{21} insertion site is presented in figure 2B. These sequence data and those from cDNA1 indicate that vg^{21} is positioned within the promoter region, since the longest cDNA appears to be full length and its 5′ end is close to the 3′ end of the vg^{21} insert (see fig. 2B). However, primer extension mapping of the *vg* promoter will be needed to confirm the above prediction.
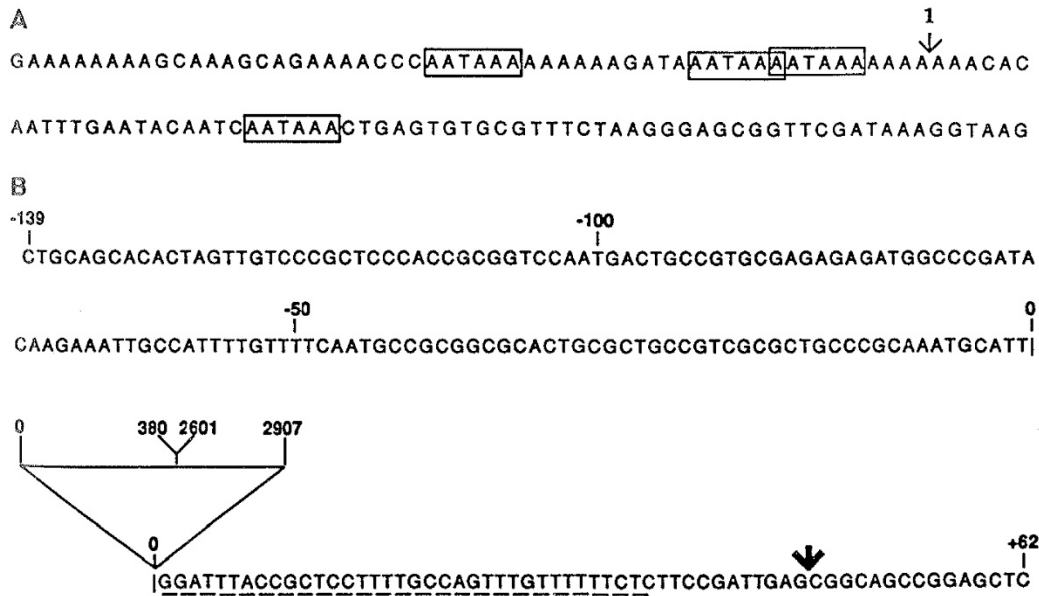

Figure 2A and B. DNA sequence analysis at the 3′ and 5′ ends of the *vg* locus. (A) The genomic DNA sequence of the distal (3′) region of the *vg* locus. This was determined by single-strand sequencing of an OR^R (Oregon-R) derived fragment from the +16 *Sal*I site (fig. 1) located approximately 100 bp 3′ to the sequence shown. *Open boxes* designate polyadenylation signals, and the *arrow* indicates the 3′ end of cDNA1. cDNA3 is approximately 100 bp longer (fig. 1B), but preliminary evidence indicates that this is likely due to a long poly(A) tail in this cDNA. (B) The DNA sequence in the vicinity of the vg^{21} P element insert (near the 5′ end of the gene). The *open triangle* denotes the 687 bp insert of an internally deleted P element which is the physical lesion associated with $vg21$ • Numbers within this insert indicate which P element sequences are present (O'Hare and Rubin 1983). The genomic *Pst*I and *Sst*I sites which flank the vg^{21} insert are at −139 and +62 bp, respectively. The sizing data for vg^{21} are from Williams et al. (1988). The *arrow* indicates where the cDNA and genomic sequences diverge. An additional 180 bp is present at the 5′ end of cDNA1; these sequences are from exon 2 and probably represent a cDNA cloning artefact. DNA sequence analyses of cDNA1 and genomic clones distal to the *Sst*I site indicate that the *Sst*I site in the cDNA corresponds to the genomic site (data not shown). The *dashed line* indicates the extent of the vg^{21-4} deletion (Williams et al. 1988).

Secondary derivatives of *vg²¹* show varying *vg* phenotypes associated with changes in the amount of P element DNA at this site (Williams et al. 1988), consistent with the notion that *vg²¹* is a promoter insertion. The *vg²¹⁻⁴* allele, a homozygous viable phenotypically extreme derivative of *vg²¹*, is associated with a 36 bp deletion of genomic DNA immediately adjacent to the *vg²¹* insert (fig. 2B). This deletion removes part of the putative promoter region consistent with the extreme nature of *vg²¹⁻⁴*. Thus, all of the noncomplementing *vg* alleles that we have examined in detail appear to have explainable *vg* phenotypes based on their predicted effects on the 3.8 kb transcription unit.

Northern analysis

Northern hybridization analysis of RN As collected throughout ontogeny was used to determine the temporal profile of *vg* transcription. Figure 3A shows such a Northern blot probed with cDNA probe A (see fig. 1B). Vestigial is expressed at maximal levels in embryos and pupae, and at a lower level in adults. Thus, *vg* is temporally regulated, with an expression profile characteristic of *Drosophila* developmental genes. Although we have not yet detected the 3.8 kb transcript in third instar larvae, it must be present since the *vg* cDNAs were isolated from a larval imaginal disc cDNA library. We feel that this indicates the transcript is spatially localized in larvae (i.e., only in imaginal tissues); however, tissue in situ analysis will be required to confirm this prediction. The onset of embryonic expression was determined and is shown in figure 3B. The transcript is expressed at very low levels in 0–4 and 4–8 h embryos, but is expressed at much higher levels in 8–12 h embryos. The 0–8 h expression may be due to minor contamination of these samples with later stage embryos. The *vg* transcription unit remains expressed through the remainder of embryogenesis (i.e., 12–16 hand 20–24 h embryos) before decreasing in first instar larvae (data not shown).
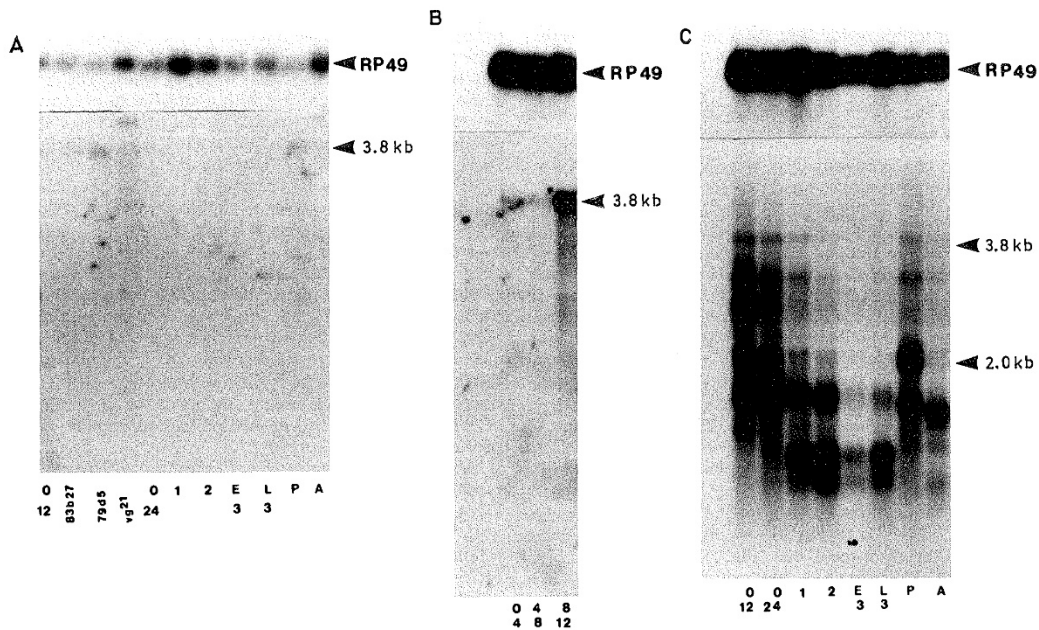
D

```
Phe Asn Gln Asp Leu Tyr Ser Ser Ser Gly10 Val Asn Leu Ala Ala Ser Ser Ser Ala Ser20

Gly Ser Ser His Ser Thr Cys Ser Pro Ile30 Leu Pro Pro Ser Val Ser Ala Asn Ala Ala40

Ala Ala Val Ala Ala Ala Ala His Asn Ser50 Ala Ala Ala Ala Val Ala Val Ala Ala Asn60

Gln Ala Ser Ser Ser Gly Gly Ile Gly Gly70 Gly Gly Leu Gly Gly Leu Gly Gly Leu Gly80

Gly Gly Pro Ala Ser Gly Leu Leu Gly Ser90 Asn Val Val Pro Gly
```

Figure 3A–D. Northern hybridization analysis of *vg* transcription throughout ontogeny. (A) Northern hybridization blot hybridized to DNA probe A (fig. 1B). Size markers are shown in kilobases, with the *arrowhead* indicating the 3.8 kb *vg* transcript. The lane designations are: vg^{21}; 79d5, vg^{79d5}; 83b27, vg^{83b27}, and the RNA was collected from 0–24 h embryos. The remaining lanes contain OR^R RNA from the following stages: 0–12 and 0–24 h embryos; 1, first instar larvae; 2, second instar larvae; E3, early third instar larvae; L3, late third instar larvae; P, brown pupae; A, adult. The size markers were a BRL RNA ladder. (B) Northern hybridization blot of OR^R embryonic RNA samples harvested from samples of the indicated embryonic ages (raised at 23°C). The probe and *arrow* designations are as in A. (C) Northern hybridization blot probed with RNA probe 3 (fig. 1B). The *arrowheads* indicate the 3.8 kb *vg* transcript and the prominent 2 kb transcript which is also detected by cDNA probe B. Ontogenic stage designations are as in A. All three blots were reprobed with a ribosomal protein probe (RP49; O'Connell and Rosbach 1984) to standardize the amount of RNA loaded per lane. The RP49 hybridizations are shown at the top of each panel. All of the above Northern hybridizations were from RNA loaded onto 1% formaldehyde-agarose gels. (D) Predicted amino acid sequence from the open reading frame in the portion of exon 3 for which the DNA sequence is at present known. This comes from a 285 bp portion sequenced proximally from the *Sma*I site at coordinate +7 on the physical map (see fig. 1).

The cDNA probes A, C, and D (see fig. 1B) recognize only the 3.8 kb transcript. However, probe B also recognizes a 2 kb transcript, and several other transcripts very weakly (data not shown). Duplicate filters of the disc cDNA library were screened with probes B and C to isolate B-specific cDNAs. However, the only B-specific cDNA isolated (from > 20 cDNAs identified) was a partial cDNA of the 3.8 kb transcript (cDNA5, fig. 1B). Also, since probes A and D do not recognize the 2 kb transcript, it is unlikely that this transcript represents the cDNA6 splicing product. Antisense RNA probes of A and D also recognize only the 3.8 kb transcript, while an antisense RNA probe of B recognizes the 3.8 kb and 2.0 kb transcripts as well as several other developmentally regulated transcripts (fig. 3C). This is not necessarily unusual, since RNA probes have been demonstrated to detect small regions of cross homology which remain undetected with DNA probes (Cavener et al. 1986). The multiple transcripts are recognized by antisense RNA probes both proximal and distal to the *Sma*I site at +7 (fig. 1). Since these transcripts are not detected with proximal or distal cDNA clones, are not represented in the disc library, and the RNA probes wash off filters at stringencies which do not melt off the probes that hybridize to the 3.8 kb transcript (data not shown), we feel that this exonic region may encode a protein domain which is present in other non-*vg* RNAs (as seen in fig. 3C). We sequenced this exonic region proximally from the *Sma*I site at coordinate +7 (fig. 1). This region comprises 285 bp with only one open reading frame through it. If translated, this region would produce a 95 amino acid protein domain which contains serine, alanine, and glycine rich stretches (fig. 3D). Neither this protein motif nor the DNA sequence is strongly homologous to any known cloned *Drosophila* gene (Bionet). However, weak homologies exist between the *vg* polyserine and polyalanine stretches and the comparable regions seen encoded in the engrailed gene (Poole et al. 1985). However, since these stretches are in the opposite polarity in the respective proteins, the meaning of the homology, if any, is obscure. We are currently extending our DNA sequence analysis of this and other exonic regions.

vg^{83b27} *analysis*

The *vg^{83b27}* allele results in a complex phenotype. Homozygous *vg^{83b27}* flies show extreme *vg* wing and haltere phenotypes but are wild type with respect to the postscutellar bristle and female fertility phenotypes. As well, *vg^{83b27}* completely, or nearly completely, complements *Su(z)2^5*, *Df(2R)vg56*, and all recessive viable *vg* alleles (Alexandrov and Alexandrova 1987; and our unpublished results). Partial complementation is exhibited with *vg^{nw}* and *vg^{12}*. Thus, *vg^{83b27}* identifies a second *vg* complementation unit. Figure 4A shows an example of this complementation ability with *vg^{BG}*. Most of the hybrid heteroallelic flies are wild type, but some heterozygotes fail to exhibit complete complementation. The wing phenotype of these latter flies is unusual, showing bubble wings and unequal wing deficiencies uncharacteristic of classical *vg* wing deficiencies (results not shown). Recombination analysis indicates that both the extreme wing phenotype and the complementation phenotype of *vg^{83b27}* are very closely linked to the *vg* locus (see Materials and methods). This provides evidence that both phenotypes are *vg* specific and not due to a strong suppressor of classical *vg* alleles linked to *vg^{83b27}* which is itself unaffected by this hypothetical suppressor.
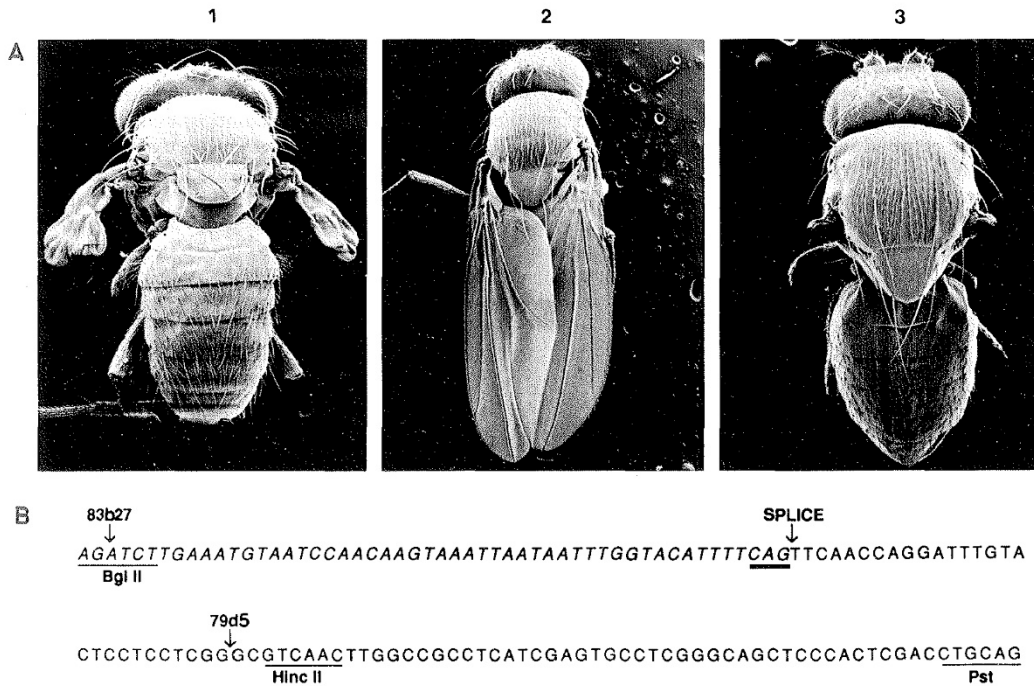
Figure 4A and B. Complementation behavior of vg^{83b27} and DNA sequence analysis of vg^{83b27} and vg^{79d5}. (A) Whole flies of vg^{BG}(1), vg^{BG}/vg^{83b27} (2), and vg^{83b27} (3) genotypes are shown. The vg^{83b27} homozygotes show strong wing and haltere reductions which differ from the corresponding reductions seen in vg^{BG} or other classical vg allele homozygotes. In addition, vg^{83b27} flies have normal postscutellar bristles. (B) The DNA sequence of the 135 bp *Bgl*II-*Pst*I region affected by the vg^{83b27} and vg^{79d5} alleles. The 3' splice acceptor site of cDNA1 is indicated with intronic bases *italicized*. *Labeled arrows* denote the centromere-distal deletion endpoints associated with the vg^{83b27} and vg^{79d5} lesions.

$Df(2R)vg^{56}$ and $Su(z)2^5$ are chromosome deficiency mutations which extend into the vg region (see above), yet are complemented by vg^{83b27}. These deficiency mutants are likely to disrupt local pairing somewhat, which is indicative that the complementation phenotype of vg^{83b27} is unlikely to be due to a transvection-like process (Lewis 1954). However, it is possible that chromosome pairing in the vg region is maintained even in the presence of these deletion alleles. The zeste gene product is required for the manifestation of transvection effects (Kaufman et al. 1973; Zachar et al. 1985) and has been shown to bind cooperatively to sites in the regulatory region of many genes (Benson and Pirrotta 1988). A strong zeste allele (z^{v77h3}; kindly provided by V. Pirrotta) does not alleviate the complementation ability of vg^{83b27} (J. Williams, unpublished data) thus indicating that the behavior of vg^{83b27} is unlikely to be mediated by a transvection-like process.

Previous analyses of vg^{83b27} (Williams and Bell 1988) detected a single 4.0 kb deletion that physically maps within the locus (fig. 1). This deletion removes most of the 4.5 kb intron of the vg transcription unit. This is surprising, since homozygous vg^{83b27} flies have an extreme wing and haltere phenotype more typical of that expected for null alleles that

perturb exons. The 5′ end of this deletion clearly does not overlap a *vg* exon since it maps at least 500 bp from the 3′ end of the upstream exon. However, the 3′ end of the *vg*[83b27] deletion appears to map quite close to the 5′ end of the downstream exon, as does the 3′ endpoint of *vg*[79d5] (fig. 1).

Genomic Southern hybridization analysis of *vg*[83b27] and *vg*[79d5] indicated that both *vg*[83b27] and *vg*[79d5] break within a 70 bp *Bgl*II-*Hinc*II fragment. The *vg*[83b27] deletion appears to break closer to the *Bgl*II site while the *vg*[79d5] deletion endpoint is closer to the *Hinc*II site (data not shown). The cDNA exon/intron junction in question is also within this small DNA restriction fragment. Genomic libraries of *vg*[79d5] and *vg*[83b27] were constructed and clones of the relevant region were isolated (see Materials and methods). The results of DNA sequence analysis of genomic *vg*[83b27], *vg*[79d5], and *vg*+ cloned DNA from this region are shown in figure 4B. The cDNA exon/intron junction is indicated, and the genomic sequence indicates that the junction is a consensus 3′ splice junction. The *vg*[79d5] deletion includes 28 bp of the downstream exon. Northern blot analysis of RNA isolated from *vg*[79d5] embryos indicates that this allele has normal levels of a transcript approximately 100 bp smaller than the wild-type 3.8 kb transcript (fig. 3A). This indicates that *vg*[79d5] uses an alternative splice site ~100 bp within the exon affected by the deletion. A number of putative splice junctions are within this region, in all three frames (data not shown). Since *vg*[79d5] is only an intermediate allele, we feel that the splice is likely to be in the correct reading frame. However, isolation and sequencing of cDNAs from *vg*[79d5] will be required to confirm this. The *vg*[83b27] deletion removes only the proximal two bases of the *Bgl*II site. Thus, the 3′ deletion endpoint of *vg*[83b27] is about 50 bp upstream from the splice site and should not reduce splicing. This is consistent with Northern hybridization analyses which indicate that the 3.8 kb transcript is unaltered in both size and amount in *vg*[83b27] embryos (fig. 3A). Finally, the *vg*[83b27] deletion does not remove a micro exon within the 4.5 kb intron, since sequence analysis of genomic and cDNA clones has conclusively mapped the ends of this intron (fig. 1, legend).

Analysis of the vg[83b27] complementation phenotype

Since the *vg*[83b27] lesion does not remove the 3.8 kb *vg* transcript, then perhaps the complementing behavior of the *vg*[83b27] allele is simply due to the presence of this *vg* transcription unit. This was tested by selecting derivatives of *vg*[83b27] that had lost the ability to complement other *vg* alleles. Males of *vg*[83b27] genotype were treated with 4000 rads γ irradiation (^{60}Co), mated to *vg*[21–7] flies (i.e., a strapwing *vg21* derivative that is normally complemented in combination with *vg*[83b27], Williams et al. 1988) and noncomplementing progeny selected. One such fly was isolated from 1.2×10^4 screened, and it had a phenotype typical of that produced by a *vg* null allele. It differed from *vg*[83b27] in that it had lost the complementation ability and also displayed erect postscutellar bristles and homozygous female sterility. Genomic Southern hybridization analysis (fig. 5) indicates that this allele has the original *vg*[83b27] deletion as well as a second deletion similar in size and location to that of *vg*[nw] (see fig. 1). Since this second deletion removes the 3′ exon of the 3.8 kb transcript, by inference it appears that the complementation ability of *vg*[83b27] is likely due to the ability to make an intact 3.8 kb transcript. Thus, the puzzling feature of this allele is not its complementation ability but its homozygous extreme *vg* wing phenotype. Since the phenotype

of vg^{79d5} homozygotes is much weaker than vg^{83b27} homozygotes and the vg^{79d5} effect can be attributed to alterations of the downstream exon, then the more extreme homozygous phenotype of vg^{83b27} must be due to the deletion of DNA unique to the vg^{83b27} allele and thus proximal to the vg^{79d5} deletion. Since this DNA is entirely intronic with respect to the 3.8 kb transcript, this indicates that there are DNA sequences within the intron which are required for normal wing and haltere development. Thus, the second *vg* functional unit defined by vg^{83b27} resides at least partially within this intron.
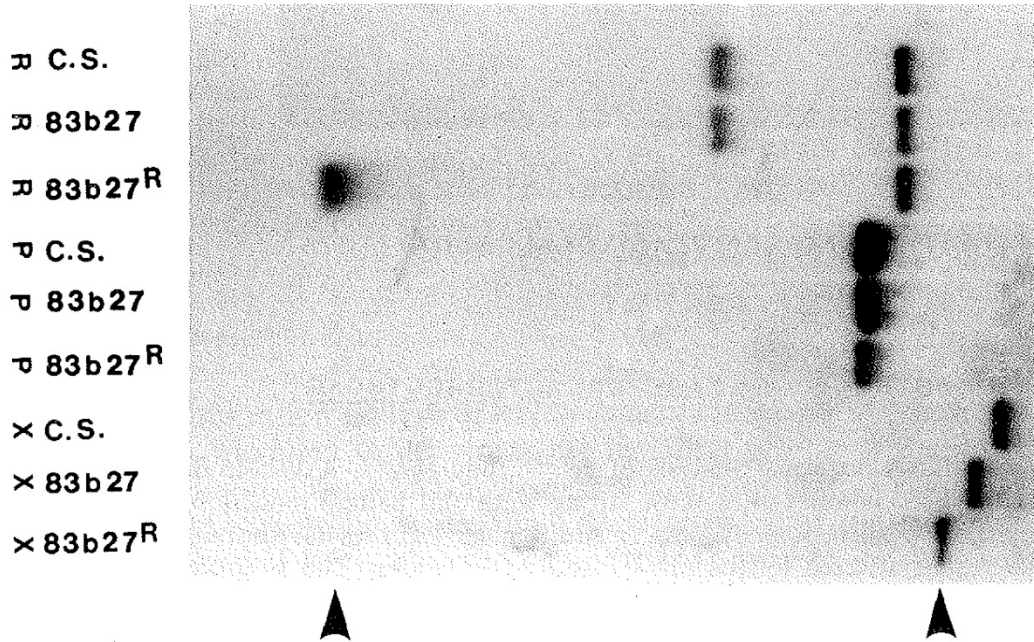

Figure 5. Genomic Southern hybridization analysis of $vg^{83b27-R}$. An autoradiogram of a genomic Southern hybridization blot of *Eco*RI(R), *Pst*I(*P*) and *Xho*I(*X*) digested DNA from Canton S (C.S.), $vg^{83b27-R}$, and vg^{83b27} strains is shown. The probe was a 1.9 kb *Sal*I DNA fragment from coordinates +16 to +18 (fig. 1A). The *arrowheads* indicate the novel deletion fragments in *Eco*RI and *Xho*I digested $vg^{83b27-R}$ DNA. The $vg^{83b27-R}$ lesion removes a *Pst*I site at ~ + 15, creating a fusion fragment the same size as that seen from *Pst*I digested Canton S DNA (6.5 kb). Subsequent stripping and reprobing of the blot with other *vg* region clones indicates the $vg^{83b27-R}$ also contains the expected 4 kb deletion diagnostic of vg^{83b27}, but no other detectable alterations (data not shown).

Discussion

In this study we report the isolation of cDNAs corresponding to a 3.8 kb *vg* transcript. Mutational lesions affecting this transcript appear to be responsible for the phenotype of the classical noncomplementing recessive viable and recessive lethal *vg* alleles for several reasons. The transcript is expressed in third instar larval imaginal discs (i.e., *vg* cDNAs were isolated from an imaginal disc cDNA library), the tissue which undergoes cell death

in *vg* mutants. As well, the cDNA exons are spread throughout the exact region previously defined by deficiency and mutant analyses (Williams and Bell 1988) as essential for *vg* function. The respective mutant phenotypes of all classical *vg* alleles examined are explainable by alterations to this transcription unit. In the case of vg^{79d5}, sequence data indicate that the mutant lesion alters splicing. Indeed, the small exonic deletion predicted with vg^{79d5} sequencing data is observed on Northern hybridization blots. An allele (*vg21*) with a P element insertion into a putative promoter is associated with both a normal and an aberrant larger sized *vg* transcript (fig. 3A). Finally, the severity of the *vg* phenotype of an allele is correlated with how our results predict the respective lesion would affect the 3.8 kb transcript. We feel that our evidence argues strongly that classical *vg* alleles are the result of alterations to, or influences on, the 3.8 kb transcription unit. Null alleles which destroy the integrity and thus the biological activity of this transcript show extreme wing and haltere loss, erect postscutellar bristles, and female sterility as well as other poorly defined phenotypes, including developmental delay, pupal lethality, and leg or abdominal abnormalities.

The temporal and putative spatial localization of the 3.8 kb transcript is consistent with a developmentally important role for the *vg* gene. An internal region of this transcription unit was identified that cross-hybridizes with a number of developmentally regulated non-*vg* transcripts. The region codes for a predicted protein containing polyserine, polyalanine, and polyglycine repeats and shows weak homology to similar repeats in the *Drosophila* engrailed protein. Homoamino acid domains are a common feature of *Drosophila* regulatory genes (Laughon et al. 1985), consistent with the idea that *vg* codes for a regulatory protein required for wing disc pattern formation. This internal domain should serve as a useful "tag" to clone the developmentally regulated genes which cross-hybridize with it. Since no *Drosophila* genes sequenced at present show strong homology to this region (Bionet), it is possible that these related genes represent a new gene family, perhaps functionally related to *vg*. The cloning and analysis of these genes may also serve to increase our understanding of the role of *vg* in wing development.

An exceptional *vg* allele exists (vg^{83b27}) which displays a complex complementation pattern. Homozygous vg^{83b27} flies have severe wing and haltere reduction, indicating loss of *vg* function. However, vg^{83b27} complements all recessive viable *vg* alleles, and weakly complements some recessive null *vg* alleles (i.e., vg^{nw}, vg^{12}). Thus, vg^{83b27} defines a second *vg* complementation group. Interestingly, the molecular lesion associated with vg^{83b27} does not appear to affect the 3.8 kb transcription unit but instead deletes most of a 4.5 kb intron. Indeed, our results indicate that the complementation ability of vg^{83b27} is due to the presence of an intact 3.8 kb *vg* transcription unit. Thus, vg^{83b27} defines additional sequences within a major *vg* intron, which are somehow required for wing and haltere development.

It is of interest to speculate what the mechanism is by which these intronic sequences affect wing and haltere development. Our results do not provide support for models in which the intron is essential in controlling expression of the *vg* 3.8 kb transcription unit. For example, an exon which is required in some discrete developmental stage or tissue but not others, may be present within the relevant intron. We have only matched genomic DNA sequences with cDNA sequences for the cDNAs that we have identified at present. However, since none of the cDNAs isolated from discs contain this putative exon, and no

candidate transcripts have been identified on Northern hybridizations, we feel that this alternative exon explanation is unlikely, although not yet rigorously excluded. A second model implies that the intron deleted by vg^{83b27} is required to modulate expression of the 3.8 kb transcript in a tissue-specific manner. Thus, vg^{83b27} may delete a binding site or regulatory sequence that is essential for normal regulation of vg expression. In this model, the extreme wing and haltere defects seen in vg^{83b27} would be due to a selective absence of the vg product at a developmental stage or tissue where it is required. This would account for the observation that normal 3.8 kb vg transcripts are present in vg^{83b27} embryos, and that vg^{83b27} is mutant for only two of the four classical vg phenes. However, it is difficult to explain the partial complementation that vg^{83b27} exhibits with vg^{nw} and vg^{12} by any model which implies differential splicing, processing, or tissue-specific expression of the same primary transcript since vg^{nw} and vg^{12} grossly disrupt two different exons of the 3.8 kb transcript. Further, our results indicate that this complementation is not likely to be due to transference of *cis*-acting information from vg^{nw} to vg^{83b27} (or vice versa) via a zeste-mediated transvection effect. Thus, if the intron is required to mediate expression of the 3.8 kb transcript, the mechanism remains obscure and is likely to be complex.

We have previously identified a family of transcripts (represented by a predominant 1 kb transcript), which is expressed in adults and quantitatively altered in various vg mutants (Williams et al. 1988). The RNA probe used to identify this transcript is shown in figure 1 and maps entirely within the region deleted in vg^{83b27} (and this in turn is within the second vg intron). The polarity of the probe indicates that the polarity of the intron transcript is opposite that of the 3.8 kb transcript. Thus vg is a nested gene (i.e., as described in Henikoff et al. 1986) containing at least two different transcription units. It is possible that this intron transcription unit is required for wing and haltere development. In this model, the vg^{83b27} deletion removes the intron transcription unit, resulting in extreme wing and haltere loss. Clearly, the exact localization and DNA sequencing of the intronic sequences that apparently are required for wing and haltere development will be required to test this and other model(s). This work is currently in progress.

References

Alexandrov ID, Alexandrova MV (1987) A new *nw* allele and interallelic complementation of the *vg* locus of *Drosophila melanogaster*. Dros Inf Serv 66: 11.

Benson M, Pirrotta V (1988) The *Drosophila* zeste protein binds cooperatively to sites in many gene regulatory regions: implication for transvection and gene regulation. EMBO J 7: 3907–3915.

Berger SL, Kimmel AR (1987) Guide to molecular cloning techniques. Methods Enzymol 152: 219–227.

Borot N, Goux JM (1981) Development larvaire et penetrance du gene vestigial chez *Drosophila melanogster*. 1. Biometrie du development larvaire chez des populations sauvages et mutantes vestigial. Arch Zool Exp Gen 122: 211–218.

Bownes M, Roberts S (1981a) Analysis of vestigial W (*vgw*); a mutation causing homeosis of haltere to wing and posterior wing duplications in *Drosophila melanogaster*. J Embryol Exp Morphol 65: 49–76.

Bownes M, Roberts S (1981b) Regulative properties of wing discs from the vestigial mutant of *Drosophila melanogaster*. Differentiation 18: 89–96.

Cavener D, Corbett G, Cox D, Whetten R (1986) Isolation of the eclosion gene cluster and the developmental expression of the Gld gene in *Drosophila melanogaster*. EMBO J 5: 2939–2948.

Chen EY, Seeburg PH (1985) Laboratory methods—supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA 4: 165–170.

Erk FC, Podraza A (1986) An extreme allele of vestigial in *D. melanogaster*: *vg^x*. Dros Inf Serv 63: 161.

Fristrom D (1969) Cellular degeneration in the production of some mutant phenotypes in *Drosophila melanogaster*. Mol Gen Genet 103: 363–379.

Gietz RD, Hodgetts RB (1985) An analysis of dopa decarboxylase expression during embryogenesis in *Drosophila melanogaster*. Dev Biol 107: 142–155.

Henikoff S, Keene M, Fechtel K, Fristrom J (1986) Gene within a gene: nested *Drosophila* genes encode unrelated proteins on opposite DNA strands. Cell 44: 33–42.

Kaufman TC, Tasaka SE, Suzuki DT (1973) The interaction of two complex loci, *zeste*, and *bithorax* in *Drosophila melanogaster*. Genetics 75: 299–321.

Klessig D, Berry JO (1983) Improved filter hybridization method for detection of single copy sequences in large eukaryotic genomes. Plant Mol Biol Rep I: 551–565.

Laughon A, Carroll SB, Storfer FA, Riley PD, Scott MP (1988) Common properties of proteins encoded by the Antennapedia complex genes of *Drosophila melanogaster*. Cold Spring Harbor Symp Quant Biol 50: 253–262.

Lewis EB (1954) The theory and application of a new method of detecting chromosomal rearrangement in *Drosophila melanogaster*. Am Nat 88: 225–239.

Lindsley DL, Grell EH (1968) Genetic variations of *Drosophila melanogaster*. Carnegie Inst Wash Publ 627.

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Melton DA, Krieg PA, Rebagliati MR, Maniatis T, Zinn K, Green MR (1984) Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res 12: 7035–7055.

Nash D, Bell JB (1968) Larval age and the pattern of DNA synthesis in polytene chromosomes. Can J Genet Cytol 10: 82–92.

O'Connell P, Rosbach M (1984) Sequence, structure, and codon preference of the *Drosophila* ribosomal protein 49 gene. Nucleic Acids Res 12: 5495–5513.

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

Poole SJ, Kauver LM, Drees B, Kornberg T (1985) The engrailed locus of *Drosophila*: structural analysis of an embryonic transcript. Cell 40: 37–43.

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

Williams J, Bell JB (1988) Molecular organization of the vestigial region in *Drosophila melanogaster*. EMBO J 7: 1355–1363.

Williams J, Pappu SS, Bell JB (1988) Molecular analysis of hybrid dysgenesis-induced derivatives of a P-element allele at the *vg* locus. Mol Cell Biol 8: 1489–1497.

Zachar Z, Chapman CH, Bingham PM (1985) On the molecular basis of transvection effects and the regulation of transcription. Cold Spring Harbor Symp Quant Biol 50: 337–346.