The Mutant Phenotype Associated With *P*-Element Alleles of the *vestigial* Locus in *Drosophila melanogaster* May Be Caused by a Readthrough Transcript Initiated at the *P*-Element Promoter

Ross B. Hodgetts and Sandra L. O'Keefe

Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Manuscript received September 27, 2000
Accepted for publication January 2, 2001

ABSTRACT

We report here the isolation of a new *P*-element-induced allele of the *vestigial* locus vg^{2a33}, the molecular characterization of which allows us to propose a unifying explanation of the phenotypes of the large number of *vestigial P*-element alleles that now exists. The first *P*-element allele of *vestigial* to be isolated was vg^{21}, which results in a very weak mutant wing phenotype that is suppressed in the P cytotype. By destabilizing vg^{2a33} in a dysgenic cross, we isolated the vg^{2a33} allele, which exhibits a moderate mutant wing phenotype and is not suppressed by the P cytotype. The new allele is characterized by a 46-bp deletion that removes the 3′-proximal copy of the 11-bp internal repeat from the *P* element of vg^{21}. To understand how this subtle difference between the two alleles leads to a rather pronounced difference in their phenotypes, we mapped both the *vg* and *P*-element transcription units present in wild type and mutants. Using both 5′-RACE and S1 protection, we found that *P*-element transcription is initiated 19 bp farther upstream than previously thought. Using primer extension, the start of *vg* transcription was determined to lie 435 bp upstream of the longest cDNA recovered to date and upstream of the *P*-element insertion site. Our discovery that the *P* element is situated within the first *vg* exon has prompted a reassessment of the large body of genetic data on a series of alleles derived from vg^{21}. Our current hypothesis to explain the degree of variation in the mutant phenotypes and their response to the P repressor invokes a critical RNA secondary structure in the *vg* transcript, the formation of which is hindered by a readthrough transcript initiated at the *P*-element promoter.

IN Drosophila, transposable elements have a propensity to insert in the 5′ regions of genes and to cause mutations as a result. Examples of such mutations include a *copia* insertion upstream of the *Adh* adult promoter (Aquadro *et al.* 1986), *gypsy* insertions upstream of the *achaete* promoter (Campuzano *et al.* 1985) and into the 5′-untranslated leader of *yellow* (Parkhurst and Corces 1986), and numerous cases of *P* elements inserted into a number of genes, including *rudimentary* (Tsubota *et al.* 1985), *yellow* (Chia *et al.* 1986; Geyer *et al.* 1988), *singed* (Roiha *et al.* 1988), and *vestigial* (Williams and Bell 1988). The properties of a series of *P*-element alleles of the *yellow* locus have been reported (Geyer *et al.* 1988). These alleles all contain variously deleted *P* elements inserted into the 5′-untranslated leader sequence 76 bp downstream of the transcription initiation. In most of the alleles, the *P* element is inserted in opposite orientation to the target gene. An exception is the allele $y^{1\#7}$, whose transcription parallels that of the host gene. Similarly, *P* transcription parallels that of the host gene in the vg^{21}mutant, bearing the first *P*-element-induced allele of the *vestigial* gene (Williams and Bell 1988). However, unlike the *y* mutation that created a null allele, the vg^{21} mutant exhibits a nearly wild-type phenotype, even though an internally deleted *P* element is inserted within the 5′ region of *vg*. Many derivatives of the original vg^{21} allele in which alterations confined to the *P* element have occurred have now been recovered (Williams *et al.* 1988a; Staveley *et al.* 1994). These alleles result in a wide range of phenotypes. We have shown that derivatives of the weak vg^{21} allele that are missing the 3′-polyadenylation signal/site (P_T) show a moderate to strong mutant phenotype (Staveley *et al.* 1994). In contrast, alleles in which the *P*-element promoter/transcription start site (P_P) has been deleted are wild-type revertants of vg^{21}. The only revertant of $y^{1\#7}$ to be characterized also has P_P deleted. These observations suggest that the level of transcription initiated within the *P* element is in some way responsible for the mutant phenotypes at both the *yellow* and *vestigial* loci.

Staveley *et al.* (1994) explained the origin of the mutant phenotypes of *vestigial* by hypothesizing that *P* transcription not terminated within the *P* element was somehow interfering with *vg* transcription. The transcription of *vg* was assumed to begin downstream of the

Corresponding author: Ross Hodgetts, Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada. E-mail: ross.hodgetts@ualberta.ca

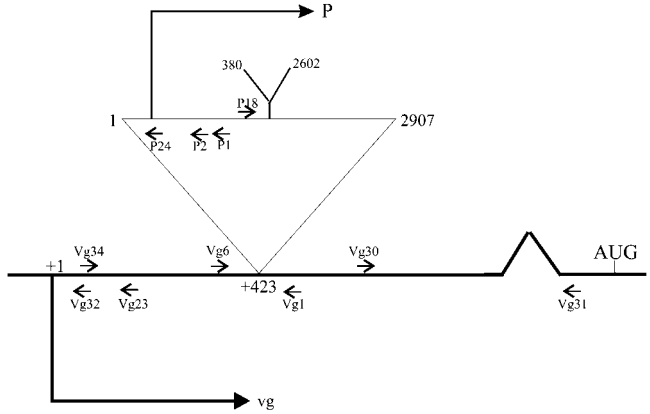
P element, since none of the extant *vg* cDNAs extended beyond the *P*-element insertion site. The discovery, reported here, that *vg* transcription in fact begins upstream of the *P* insertion site has prompted us to reassess this hypothesis.

Another perplexing aspect of some of the *P*-induced alleles of *vg* is that their associated mutant phenotype is dependent on the cytotype. Two types of cytotype can be distinguished: M and P (Engels 1983). In molecular terms, cells exhibiting the M cytotype are devoid of the *P*-element repressor and intact *P* elements introduced into these cells actively transcribe their transposase/repressor gene. The P cytotype defines a cellular environment in which *P* elements are quiescent. A repressive effect of the P cytotype on *P* transcription was demonstrated by showing that reporter genes whose transcription was driven by a *P*-element promoter were completely repressed when introduced into the P cytotype (Lemaitre and Coen 1991; Lemaitre *et al.* 1993; Roche *et al.* 1995).

The dependency of mutant phenotypes on cytotype was first described for several *P*-element-induced alleles of *vestigial* (Williams *et al.* 1988b), the mutant wing phenotype of which was suppressed in the P cytotype. Such dependence on cytotype has also been shown for the mutant bristle phenotype of three *P*-element alleles of *singed* (Robertson and Engels 1989). We are investigating whether the suppression of mutant phenotypes in the P cytotype can be explained by the repression of *P* transcription that occurs in the P cytotype. To establish a molecular basis for the suppression of the *vg* alleles, we are comparing the cytotype-independent allele, vg^{2a33}, with its cytotype-dependent progenitor, vg^{21}. Our observations suggest an explanation for the effects of the P cytotype on the *vestigial* mutant phenotypes.

MATERIALS AND METHODS

Drosophila stocks: The vg^{2a33} allele arose from a dysgenic cross that was carried out to effect a targeted transposition of *P*[inv*Ddc*], an X-linked *P* element bearing the *Ddc* gene, into the *P* element in vg^{21} [see Figure 1 in Heslip and Hodgetts (1994) for a similar crossing scheme]. The new allele was recovered over the very weak vg^{21} allele and identified by the moderate wing nicking seen in the heterozygote. As a homozygote, vg^{2a33} exhibits a mutant phenotype of 2–3 on the scale of 1–6 developed by Williams (1989) from the series of 31 *vestigial* wing phenotypes described by Nakashima-Tanaka (1967). On this scale, 1 = wild type and 6 = extreme mutant. Both wings of vg^{2a33} always exhibit abnormalities that range from a double nick on the distal wing margin to a moderate "strap wing" in which substantial parts of both the anterior and posterior margins are missing. The other *P*-element-induced *vg* alleles, vg^{21} and $vg^{21\text{-}3}$, which are described in Williams and Bell (1988), were obtained from Dr. John Bell. Neither vg^{21} nor vg^{2a33} harbors any *P* elements in addition to the element found at *vg*. A strain containing the *Df*(2R)vg^{B} chromosome, which has the *vestigial* region deleted (Lindsley and Zimm 1992), as well as the π_2 and Harwich strains, which are naturally occurring P strains that contain both complete and internally deleted *P* elements, were obtained from the Bowling Green Stock Center. The wild-type M strain used in this study was Oregon-R. Initially, all stocks were maintained on a synthetic food (Nash and Bell 1968) at ~22°, but later we switched to a standard yeast/molasses/cornmeal food.



Primer designation	Sequence (5' to 3')
Vg1	ATCCCGCGCGGCGGTGAGAG
Vg6	AATCAAGTGGGCGGTGCTTG
Vg23	TGGAGTGCGTCTCAATTGTG
Vg30	CGATTTCAGCGTAAACGAGA
Vg31	TTCCGACGAAGAATTGCCAA
Vg32	GGAATTCGCATTCATCCGGGGCC
Vg34	GGCCCCGGATGAATGCGAATTCC
P1	GAGTCGCCTCCTTTTAAATG
P2	TCTGCTCCCACAATTTTCTC
P18	CCTGCAAAAGGTCAGACATT
P24	CCGCACACAACCTTTCCTCTCAAC-AAGCAAACGTGCACTCAATTTAAG

Figure 1.—Locations of the oligonucleotide primers used in this study. The genomic region near the 5′ end of the *vestigial* gene in the vg^{21} allele is shown. The *P* element is inserted 423 bp downstream of the *vg* transcription start site and contains an internal deletion of nucleotides 381–2601 (Williams *et al.* 1988a).

Cytotype dependence of the *vg* alleles: The three *P*-element alleles of *vestigial* described above were passaged through either the M or P cytotype and the mutant alleles were scored in heterozygous combination with the vg^{B} null allele. The cross to establish the M cytotype was $vg^{B}/Balancer \times vg^{*}$ (where * is one of the three alleles); it was set up reciprocally and the phenotypes of the vg^{B}/vg^{*} progeny were identical. To establish the P cytotype, we crossed $vg^{B}/Balancer1$ (males) × π_2 (females). Progeny males carrying *Balancer1* were again crossed to π_2 females. *Balancer1*/+ female progeny were then crossed to $vg^{B}/Balancer2$ males and P cytotype female progeny of the genotype $vg^{B}/Balancer1$ were crossed to males of the three *vg* mutants. The wing phenotypes of vg^{*}/vg^{B} progeny were scored on the scale of 1–6.

Molecular characterization of the vg^{2a33} allele: The *P* element present in the vg^{2a33} allele was amplified by priming mutant genomic DNA with two primers that flanked the insertion site, Vg6 and Vg1 (see Figure 1 for all the primers used in this study). The PCR product was purified by agarose gel electrophoresis and cloned into pGEM-T (Promega, Madison, WI). The entire insert was then sequenced using an ABI 373 automated sequencing system following dye-terminated cycle sequencing (Amersham/Pharmacia, Piscataway, NJ).

5′ and 3′ RACE: The *P*-element transcription start site in both the vg^{21} and Harwich strains was determined using RNA from adults and 24-hr pupae. Total RNA was extracted from 150- to 200-mg samples of live or frozen organisms using the TRIZOL reagent (GIBCO BRL, Gaithersburg, MD), and then resuspended in 100–200 μl of water and treated for 20 min with DNAase I (10 units, Amersham) in 40 mM Tris-HCl (pH 7.5) containing 6 mM $MgCl_2$ at 37°. The enzyme was inactivated at 65° for 15 min; 5′ rapid amplification of cDNA ends (RACE) was performed on 2-μg samples of the RNA using the GIBCO BRL 5′ RACE System according to the manufacturer's instructions. Reverse transcription of the RNA was primed with P1. The product of first strand synthesis was tailed with dCTP; second strand synthesis was initiated from the 5′ RACE anchor primer and amplified with *Taq*I using the nested primer P2. Cycle 1 consisted of 5 min at 95°, 1 min at 58°, and 2 min at 72°; cycles 2–29 consisted of 2 min at 93°, 1 min at 58°, and 2 min at 72°. The reaction product was fractionated on a 2% agarose gel, after which the appropriate band was isolated and cloned into pGEM-T (Promega). The sequence of at least two different plasmid clones that contained the desired product was obtained as described above.

The 3′ end of the *P*-element transcript was obtained using 3′ RACE, as described above, but with the following modifications. Reverse transcription was initiated on RNA from the Harwich strain using the adapter primer (GIBCO BRL). The second strand was amplified using the universal primer (GIBCO BRL) and primer P18 (Figure 1). The reaction products were cloned into pGEM-T directly and then sequenced.

Nuclease protection: Confirmation of the *P*-element transcription start site was obtained using the oligonucleotide P24 in a nuclease protection assay (GIBCO BRL). Three picomoles of P24 were end labeled with T4 kinase and fractionated on a 12% acrylamide gel containing 8 M urea. The oligonucleotide was recovered in 350 μl of Probe Elution Buffer (Ambion, Austin, TX) and 10^5 cpm were annealed to 50 μg of DNAase-treated total RNA from Harwich, π_2, or vg^{21} flies at 30° for 18 hr in 0.75× Hybridization Buffer (Ambion). In control reactions, 10 μg of yeast RNA was substituted for the Drosophila RNA. Unhybridized probe was removed with an appropriately diluted nuclease mixture. The protected fragment was fractionated on a 12% urea gel and visualized by autoradiography.

Primer extension: The transcription start site for the *vestigial* gene was determined by primer extension, using the oligonucleotide Vg32. Five picomoles of the primer were end labeled and purified as described above. Total RNA (15–50 μg) isolated from wild-type embryos and adult flies was incubated with 10^6 cpm of Vg32 in a 15-μl reaction containing 0.15 M KCl, 10 mM Tris-HCl (pH 8.3), and 1 mM EDTA at 90° for 10 min and then 65° for 30 min and, finally, 37° for 18 hr. The hybridization reaction was added to 30 μl of 30 mM Tris-HCl (pH 8.3), 15 mM $MgCl_2$, 8 mM dithiotreitol, and 0.2 mM dNTPs; the primer was then extended with 200 units of Superscript II reverse transcriptase (GIBCO BRL). Each reaction was incubated at 42° for 50 min and then at 95° for 10 min. The cDNA was precipitated, resuspended in a denaturing loading buffer (Sambrook *et al.* 1989), fractionated on a 12% acrylamide gel containing 8 M urea, and visualized by autoradiography.

mRNA detection using comparative reverse transcriptase-PCR: Total RNA was extracted from adults of the different genotypes and samples of 200 ng were reverse transcribed with 200 units of Superscript II (GIBCO BRL) following the manufacturer's instructions. Synthesis of the *vg*-specific first strand was initiated from primer Vg 23 or primer Vg31 (Figure 1). Each of these was cotranscribed along with a primer (5′-TCTTCTTGAGACGCAGGCGA-3′) for ribosomal protein 49 (RP49) to allow for normalization of sample-to-sample variation. Three microliters of each cDNA sample was amplified with *Taq*I using primer pairs Vg34/23 and Vg30/31 under the following conditions: 1 cycle of 95° for 5 min, 58° for 1 min, and 72° for 2 min, followed by 27 cycles consisting of 93° for 1 min, 58° for 1 min, and 72° for 2 min. At the same time 0.5 μl of each sample was amplified with the RP49 primer pair (5′-AGCATACAGGCCCAAGATCG-3′ and 5′-AGTAAAC GCGGGTTCTGCAT-3′) using the same conditions as above except for only 22 cycles. The products were fractionated on a 2% agarose gel and visualized by staining with ethidium bromide. The amounts of both the *vg* and *rp49* products responded linearly to RNA input.

TABLE 1

The effects of cytotype on the severity of mutant wing phenotypes of various *vestigial* alleles

Allele	Homozygous phenotype	M cytotype (vg^b/vg^B)	P cytotype (vg^b/vg^B)
vg^+	1[a]	—	—
vg^{21}	1	3	1
$vg^{21\text{-}3}$	5	6	1
vg^{2a33}	3	6	6

[a] Wing phenotypes were scored on a scale of 1–6, where 1 = wild type and 6 = extreme mutant.
[b] Signifies the *vg* allele shown in column 1.

RESULTS

Response of the vg^{21}, $vg^{21\text{-}3}$, and vg^{2a33} mutant phenotypes to cytotype: Phenotypes of the three mutant alleles were scored in heterozygous combinations with the null allele vg^B and the results are summarized in Table 1. The phenotypes scored in the M cytotype were uniform in all the progeny of a given cross. The phenotypes in the P cytotype of both vg^{21} and $vg^{21\text{-}3}$ were more variable; ~60% had normal wings (as reflected by the score of 1 in the table) and the remainder had wings that we scored as 2–3. These observations confirm those of Williams *et al.* (1988b), who showed that both vg^{21} and $vg^{21\text{-}3}$ were suppressed in the P cytotype. In contrast, Table 1 shows that vg^{2a33}/vg^B flies exhibited an extreme mutant phenotype (6) in both the P and M cytotypes.

Molecular characterization of vg^{2a33}: The *P* element in this allele was amplified from genomic DNA using the *vestigial* primers Vg1 and Vg6, which flank the P insertion site (Figure 1). The amplified product was cloned into pGEM-T and sequenced. Comparison of vg^{2a33} with its progenitor allele, vg^{21} (Figure 2), shows that the two alleles differ in a 50-bp region near the 3′ end of the *P* element where vg^{2a33} has an abnormal run of four C's between positions 2754 and 2759, followed by a deletion of 46 bp of the normal P sequence. Included within the deletion specific to vg^{2a33} is the 11-bp internal repeat, a binding site for the KP repressor polypeptide (Lee *et al.* 1996) and an enhancer of transposition (Mullins *et al.* 1989).

The transcription start site of *vestigial*: Preliminary

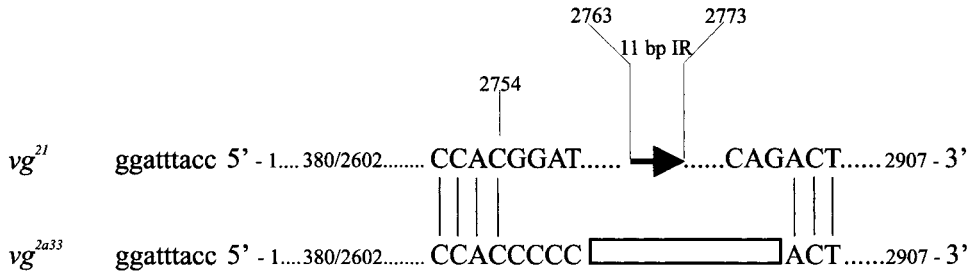


FIGURE 2.—Alignment of the *P*-element sequence in vg^{21} with that in vg^{2a33}. The structure of vg^{21} was derived by WILLIAMS *et al.* (1988a). The vg^{2a33} allele was cloned and sequenced as described in the text. The lowercase letters denote the 9-bp target site duplication. The 11-bp internal repeat (solid arrow) is included within the vg^{2a33} deletion (open rectangle). The polarity of the nontranscribed strand (numbered from 5′ to 3′) of the *P* element is shown.

data from reverse transcriptase (RT)-PCR suggested that the *vg* transcription start site lay upstream of the *P*-element insertion site (data not shown). To confirm these observations, RNA samples from embryos or adults of the wild-type strain were subjected to primer extension; the results are shown in Figure 3. The identical transcription start site was found on two independent samples of adult RNA (lanes 2 and 4) and on the sample of embryo RNA (lane 3). Significantly, this site was located on the genomic DNA 435 nucleotides upstream of the 5′ end of cDNA1, which is the longest cDNA reported to date (WILLIAMS *et al.* 1990). The entire upstream region of *vestigial* was found on the P1 clone D387, the sequence of which has been determined by the Berkeley Drosophila Genome Project. It matches the genomic sequence of *vestigial* that we obtained in the region of the transcription start site using the λ8 clone (WILLIAMS and BELL 1988) as a template (data not shown). The sequence of the transcription start site as determined from the data in Figure 3 is shown in Figure 4. Although it does not fit with the consensus transcription start site in Drosophila $PyPyA_{+1}NT/APyPy$ (ARKHIPOVA 1995; LO and SMALE 1996), a rather good fit (5/7) can be obtained if transcription is initiated at the adjacent A.

We did not determine the *vg* transcription start site in the *vg* mutant alleles. However, we did obtain an RT-PCR product on vg^{21} RNA using primers Vg34 and Vg23 (see Figure 5A, lanes 2 and 3). Both of these primers lie upstream of the *P*-element insertion site (Figure 1), suggesting that transcription is initiated at the normal start site. Further, from a negative RT-PCR result obtained using a primer, the 5′ end of which was 70 nucleotides upstream of the wild-type start site, we know that no transcript exists beyond this point in the mutants. Taken together, these results indicate that it is very likely that the start site of *vg* transcription is the same in both the wild type and the mutants.

The transcription start site of the *P* element: Using 5′ RACE, we determined the transcription start site in embryos and adults from both a naturally occurring *P*-element strain (Harwich) and the vg^{21} mutant. Sequencing the duplicate clones of all four RNA samples revealed that the start of transcription occurred at position 67 on the *P* element. This start site is a very good match to the Drosophila consensus (6/7); however, it lies upstream of the site at position 87, determined by KARESS and RUBIN (1984) who used S1 protection of RNA from π_2 embryos. Within the vicinity of their start site, the best fit (4/7) with the consensus actually occurs at A_{86}. These authors did report a minor start site at position 67 that coincides exactly with the site we report.

Confirmation of the assignment of the P start site was obtained using nuclease protection. The oligonucleotide (P24) was designed so that transcripts initiated at either of the two start sites (67 and 87) could be detected. Figure 6 shows that the labeled P24 (lane 2) is totally unprotected by yeast RNA following nuclease digestion (lanes 3 and 4). In contrast, adding RNA obtained from Harwich adults (lanes 5 and 6) resulted in the protection of a major fragment of 39 bases, which is expected from initiation at position 67. RNA from vg^{21} gave identical results (lanes 7 and 8). No utilization of the P start site reported for π_2 could be detected in any of our RNA samples since this start site would have resulted in a protected fragment of 19 bases in the figure. In fact, nuclease protection analysis carried out

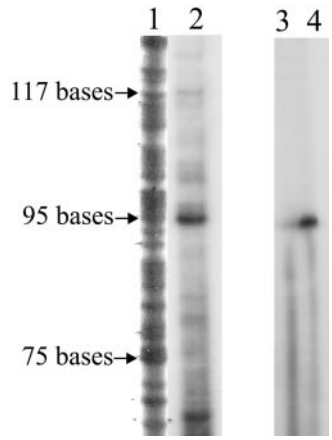


FIGURE 3.—Determination of the *vestigial* transcription start site by primer extension. RNA from Oregon-R embryos and adults was extended and the product detected by autoradiography. Lane 1, size markers; lanes 2 and 4, different preparations of adult RNA; lane 3, 22-hr embryo RNA.

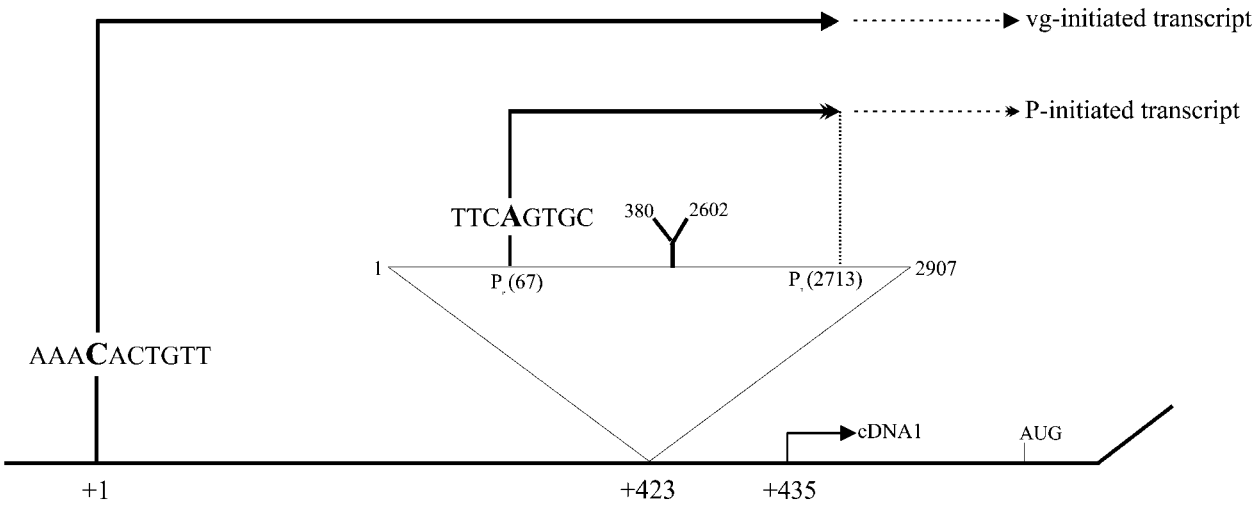


FIGURE 4.—Sequences of the transcription start sites in the vg^{21} allele. The *vg* start site was determined from primer extension of Oregon-R RNA (Figure 3) and is assumed to be the same in vg^{21}. The *P*-element start site (P_P) was determined by 5′ RACE and confirmed by nuclease protection (Figure 6). The polyadenylation site shown was determined by sequencing the cloned product obtained using 3′ RACE on RNA extracted from the Harwich P strain; it is assumed to be the same in vg^{21}.

on π_2 RNA yielded a start site at position 67 (data not shown).

The poly-A^+ addition site within the *P* elements in the Harwich strain was determined using 3′ RACE on RNA from adult flies. The site was located at position 2714 or 2715; the two A residues in the genomic sequence make assignment of the exact position impossible. Our results agree closely with the estimate of 2710 made by LASKI *et al.* (1986).

Figure 4 summarizes what we believe is the set of transcription units in the vg^{21} allele of *vestigial*. P_T, the

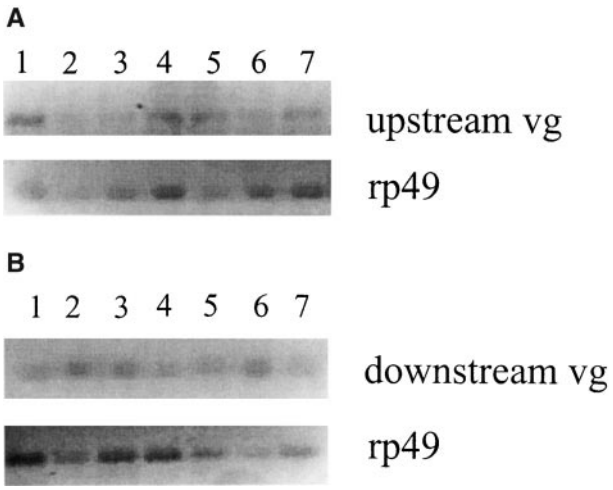


FIGURE 5.—Analysis of the transcription units in wild-type and mutant strains of *vg*. Comparative RT-PCR was carried out on adult RNA samples of 200 ng that were reverse transcribed in reactions containing both a *vg*-specific primer and a primer for *rp49*. Transcription in the region upstream of the *P*-element insertion site was detected by the primer pair Vg34/Vg23 (Figure 1). Transcription of the region downstream of the *P* element was detected with the pair Vg30/Vg31 (Figure 1). Lane 1 in both A and B shows the upstream and downstream products, respectively, that were produced by reverse transcription and amplification of Oregon-R RNA. The upstream products for two independent cDNA samples of mutant RNA are shown in lanes 2 and 3 (vg^{21}), lanes 4 and 5 ($vg^{21\text{-}3}$), and lanes 6 and 7 (vg^{2a33}) of A. The downstream products on the same cDNA samples are shown in B and the lanes are numbered as in A.

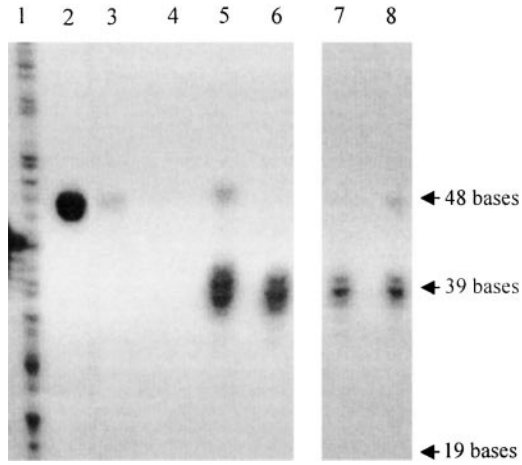


FIGURE 6.—Determination of the *P*-element start site by nuclease protection. The arrows indicate the position of intact P24 (48 bases), the major fragment protected by RNA from the Harwich strain and vg^{21} (39 bases), and the position at which the protected fragment would be expected had the start site found in the π_2 strain (KARESS and RUBIN 1984) been used (19 bases). Lane 1, a sequencing ladder derived from a M13 template; lane 2, labeled P24; lane 3, P24 incubated with yeast RNA and digested with a 1/800 dilution of the nuclease mixture; lane 4, as for lane 3 but digested with a 1/200 dilution of nucleases; lane 5, P24 incubated with 50 μg of Harwich RNA and digested with a 1/800 dilution of nucleases; lane 6, as for lane 5 but digested with a 1/200 dilution of nucleases; lane 7, P24 incubated with 50 μg of vg^{21} RNA and digested with a 1/100 dilution of nucleases; lane 8, as in lane 7 but with a 1/200 dilution of nucleases.

transcription termination signal within the *P* element, is leaky (see below); therefore, both the *vg*-initiated transcript and the transcript initiated at P_P read through into the *vestigial* coding region, as shown by the dashed lines in the figure. The most distal *vg* transcription previously reported (Williams *et al.* 1990) was cDNA1, the 5′ terminus of which is shown.

Transcription units in wild-type and mutant strains: In an attempt to understand the molecular basis of the vg^{21}, $vg^{21\text{-}3}$, and vg^{2a33} mutant phenotypes, we analyzed the transcription units present in wild-type and mutant strains by using comparative RT-PCR. One set of primers (Vg34 and Vg23) was designed to detect transcription upstream of the *P*-element insertion site (see Figure 1); a second set (Vg30 and Vg31) was used to amplify transcripts that read through the *P* element. The products generated by the amplification of cDNA molecules initiated at primers Vg23 and Vg31 on wild-type and mutant RNA are shown in Figure 5. Both primer sets amplify RNA from wild-type adults (lane 1 in Figure 5, A and B). Each of the three mutants shows a product with Vg34/Vg23 (Figure 5A, lanes 3–7), indicating the presence of a transcript initiated at the *vg* start site. In all three mutants, we found clear evidence of readthrough transcription downstream of P_T (Figure 5B, lanes 3–7). In all likelihood, this transcription is the sum of that initiated at the *vg* and *P* promoters.

DISCUSSION

The recovery of vg^{2a33} is an important addition to the derivatives of vg^{21} that have been studied previously (Staveley *et al.* 1994). Derivatives of the weak vg^{21} allele that are missing the 3′-polyadenylation signal /site (P_T), such as $vg^{21\text{-}6.1}$ and $vg^{21\text{-}74.2}$, show a moderate to strong mutant phenotype. In contrast, alleles in which just the *P*-element promoter/transcription start site (P_P) has been deleted (such as $vg^{21\text{-}IR8}$) are wild-type revertants of vg^{21}. This led to the hypothesis that *P* transcription not terminated within the *P* element was somehow interfering with *vg* transcription, which was assumed to be initiated downstream of the *P* element. This assumption was based on the fact that none of the extant cDNAs extended beyond the *P*-element insertion site (Williams *et al.* 1990). The discovery that *vg* transcription in fact begins upstream of the *P*-element insertion site (Figure 3) requires a reassessment of the original explanation of the genetic data.

From what is now known about the *vg* and *P* transcription units, the only explanation for the near absence of mutant effects in vg^{21} is that readthrough of the *P*-element transcription termination site (P_T) is occurring in this allele. This has previously been reported for other *P* elements (Laski *et al.* 1986) and the existence of transcripts downstream of the *P*-element insertion site (Figure 5B) confirms that readthrough does occur in vg^{21}, $vg^{21\text{-}3}$, and vg^{2a33}. Two transcripts therefore reach P_T (see Figure 4); following readthrough, one or both of these must be translated into a normal *vg* polypeptide beginning at the AUG initiation codon that lies in the second *vg* exon (Williams *et al.* 1991). Since the loss of P_P in an allele ($vg^{21\text{-}IR8}$) that retains P_T results in a wild-type phenotype (Staveley *et al.* 1994), we attribute the slight mutant effects in vg^{21} to transcription initiated at P_P. We postulate that the higher the level of this *P*-initiated transcript *relative* to the *vg*-initiated transcript, the more extreme will be the mutant phenotype. With the exception of vg^{2a33}, P_T is missing in all the mutant derivatives of vg^{21} in which a 3′ deletion of the *P* element has occurred. This would lead to more readthrough transcription of both the *P* and *vg*-initiated transcripts, but would not be expected to alter the *ratio* of the two transcripts. How then are we to explain the enhanced mutant phenotype of these alleles? An explanation lies in vg^{2a33}, which differs in a subtle yet important way from all *P* alleles of *vg* that have been described to date. This allele retains P_T but still exhibits a moderate mutant phenotype. The only significant feature that has been identified within the 46-bp 3′ region deleted in vg^{2a33} is the 11-bp internal repeat shown to be an enhancer of transposition (Mullins *et al.* 1989). This region is also missing in all of the 3′-deleted derivatives of vg^{21} that have been recovered (Staveley *et al.* 1994). If the loss of the 11-bp repeat results in a relative increase in the amount of the *P*-specific transcription relative to *vg* transcription, then the mutant effects can be explained if the elevated *P* transcription somehow interferes with the production or function of the readthrough *vg* transcript.

Two important questions are raised by the above explanation of the *P*-element-induced *vg* alleles. First, how does the presence or absence of the 11-bp repeat affect the rate of *P* transcription and, second, how does *P*-initiated transcription interfere with the functioning of the readthrough *vg* transcript? To answer the first question, we hypothesize that *P*-element DNA forms the stem-loop structure that is shown in Figure 7. A structure of this configuration has been proposed by Sakagouchi (1990), who termed it a racket frame. The 5′ and 3′ ends of the molecule are shown tightly paired in the regions where the inverted 31-bp terminal repeats and the 11-bp repeats occur; we assume that nonhistone proteins that bind to the two repeats might stabilize the stem. For example, the 31-bp terminal repeat is known to bind to the host-encoded IRBP protein (Rio and Rubin 1988) and the 11-bp internal repeat is known to bind a truncated form of the P transposase, known as the KP repressor (Lee *et al.* 1996). Since the 11-bp repeat is known to act as an enhancer of transposition (Mullins *et al* 1989), but does not bind the transposase (Kaufman *et al.* 1989), it may also bind an as yet unidentified host protein. The transcription start site, which is shown at nucleotide 67 in accordance with the data in this article, lies between the two halves of the stem of the

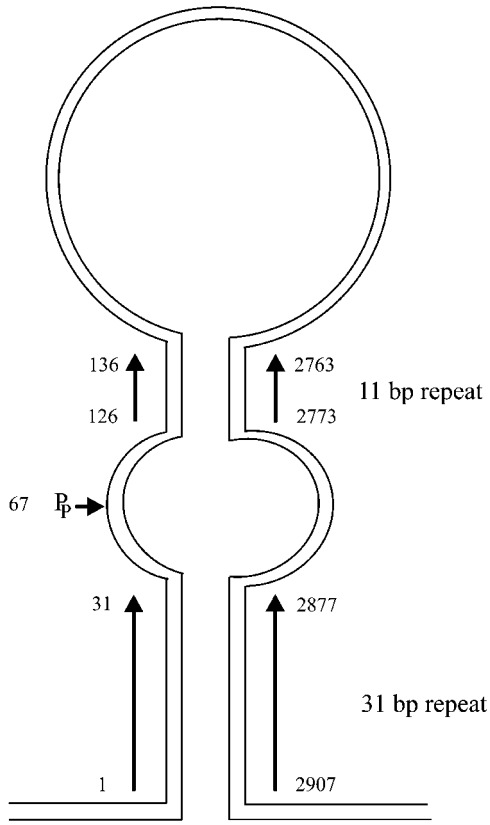


FIGURE 7.—Secondary structure model of *P*-element DNA. The stem of the structure shown may involve a tetraplex structure in which localized denaturation of the two helices, stabilized by nonhistones, allows pairing to occur between the inverted terminal repeats and the inverted 11-bp repeats. The transcription start site P_P is shown at position 67, as determined in this study.

racket frame. We suggest that the stem's configuration interferes with the initiation of transcription at P_P and that a deletion of the 3′ copy of the 11-bp repeat leads to an increase in the amount of *P* transcription. This increased transcription from P_P may occur because the protein(s) that normally bind to the 11-bp repeat can no longer do so, thus removing a steric block to the assembly of the transcription complex. While this explanation predicts that a loss of the 5′ copy of the 11-bp repeat would also enhance transcription from P_P, unfortunately *vg* alleles that have sustained a deletion that removes just this repeat and not also Pp do not exist.

In the P cytotype, it is possible that the region of pairing in the vicinity of the 11-bp repeats could be stabilized by repressor binding. According to our model, this would reduce transcription from P_P and may explain the repression of *P* transcription that has been observed in the P cytotype (LEMAITRE and COEN 1991; LEMAITRE *et al.* 1993; ROCHE *et al.* 1995) as well as the suppression of vg^{21} and $vg^{21\text{-}3}$ by the P cytotype. We assume that the loss of the 3′ copy of the 11-bp repeat in vg^{2a33} destabilizes this region, regardless of whether the repressor is present; thus its mutant phenotype is not influenced by cytotype (Table 1). The stability of the region may also be influenced by the length of *P*-element DNA between the two sets of repeats (see Figure 7). In a mutant such as vg^{21}, which has only 406 bp of DNA between the 11-bp repeats, the stem on occasion may spontaneously form and reduce transcription from the P promoter. As a result, a weak phenotype is observed. With $vg^{21\text{-}3}$, however, the two halves of the stem region are separated by 2300 bp, which could destabilize the pairing and lead to more transcription from P_P. In fact, the wing defect in $vg^{21\text{-}3}$ is significantly more severe than that in vg^{21} (Table 1).

To answer the second question of how the level of *P* transcription might interfere with *vg* expression, we imagine two possibilities. First, the *vg*-initiated transcript may be prematurely terminated in the region near P_P because of the steric effects of the transcriptional apparatus assembling there. Second, the 5′-untranslated leader sequence of the transcript of a wild-type *vg* allele may assume a folded structure that ensures its stability or is required for its translation. The transcript initiated at P_P in a *P*-element mutant allele may interfere with the formation of this folded domain, perhaps by forming an RNA-RNA hybrid molecule with the *vg*-initiated transcript. We tend to favor the second possibility at present because the data in Figure 5A, while not strictly quantitative, do show that the transcript level upstream of the *P* element in the three mutant alleles is comparable with that in the wild type (compare lane 1 with lanes 2–7). If a truncated *vg*-initiated transcript were produced, it is reasonable to assume it would be unstable and present at lower levels in the mutants. Furthermore, using RT-PCR with the primer pair P1 and Vg6 (Figure 1), we have been able to demonstrate that a *vg*-initiated transcript does pass through P_P (our unpublished observations).

All the vg^{21} derivatives with enhanced mutant phenotypes that were isolated previously (STAVELEY *et al.* 1994) are missing the 3′ copy of the 11-bp internal repeat. We now explain their mutant phenotype, not by the loss of P_T, but rather by the absence of the 11-bp internal repeat. One critical prediction of this model is that mutations in the P_P of vg^{2a33} that result in decreased transcription relative to that initiated at the *vg* promoter should exhibit less severe mutant phenotypes. We are now using chemical mutagenesis of vg^{2a33} in an attempt to recover such revertants.

Michael Veeman contributed to this work by cloning the vg^{2a33} allele. Gary Ritzel was consulted on numerous technical details. M. Simmons and one anonymous reviewer suggested a number of improvements to the manuscript; Elaine Street provided detailed editorial comments. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

LITERATURE CITED

AQUADRO, C. F., S. F. DESSE, M. M. BLAND, C. H. LANGLEY and C. C. LAURIE-AHLBERG, 1986 Molecular population genetics of the

alcohol dehydrogenase gene region of *Drosophila melanogaster*. Genetics **114:** 1165–1190.

Arkhipova, I. R, 1995 Promoter elements in *Drosophila melanogaster* revealed by sequence analysis. Genetics **139:** 1359–1369.

Campuzano, S., L. Carramolino, C. V. Cabrera, M. Ruiz-Gomez, R. Villares *et al.*, 1985 Molecular genetics of the "achaete-scute" gene complex in D. melanogaster. Cell **40:** 327–338.

Chia, W., G. Howes, M. Martin, Y. B. Meng, K. Moses *et al.*, 1986 Molecular analysis of the *yellow* locus of *Drosophila*. EMBO J. **5:** 3597–3605.

Engels, W. R., 1983 The P family of transposable elements in Drosophila. Annu. Rev. Genet. **17:** 315–344.

Geyer, P. K., K. L. Richardson, V. G. Corces and M. M. Green, 1988 Genetic instability in *Drosophila melanogaster*: P element mutagenesis by gene conversion. Proc. Natl. Acad. Sci. USA **85:** 6455–6459.

Heslip, T. R., and R. B. Hodgetts, 1994 Targeted transposition at the *vestigial* locus of *Drosophila melanogaster*. Genetics **138:** 1127–1135.

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

Kaufman, P. D., R. F. Doll and D.C. Rio, 1989 Drosophila P element transposase recognizes internal P element DNA sequences. Cell **59:** 359–371.

Laski, F. A., D. C. Rio and G. M. Rubin, 1986 Tissue specificity of Drosophila P element transposition is regulated at the level of mRNA splicing. Cell **44:** 7–19.

Lee, C. C., Y. M. Mul and D. C. Rio, 1996 The *Drosophila* P-element KP repressor protein dimerizes and interacts with multiple sites on P-element DNA. Mol. Cell. Biol. **16:** 5616–5622.

Lemaitre, B., and D. Coen, 1991 P regulatory products repress *in vivo* the P promoter activity in *P-lacZ* fusion genes. Proc. Natl. Acad. Sci. USA **88:** 4419–4423.

Lemaitre, B., S. Ronsseray and D. Coen, 1993 Maternal repression of the P promoter in the germline of *Drosophila melanogaster*: a model for the P cytotype. Genetics **135:** 149–160.

Lindsley, D. L., and G. G. Zimm, 1992 The genome of *Drosophila melanogaster*. Academic Press, New York.

Lo, K., and S. T. Smale, 1996 Generality of a functional initiator consensus sequence. Gene **182:** 13–22.

Mullins, M. C., D. C. Rio and G. M. Rubin, 1989 *Cis*-acting DNA sequence requirements for P element transposition. Genes Dev. **3:** 729–738.

Nakashima-Tanaka, E., 1967 The effect of temperature and genetic background on the phenotypic expression of several vestigial strains of *Drosophila melanogaster*. Genetica **38:** 447–458.

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

Parkhurst, S., and V. Corces, 1986 Interactions among the gypsy transposable element and the yellow and the suppressor of hairy-wing loci in *Drosophila melanogaster*. Mol. Cell. Biol. **6:** 47–53.

Rio, D. C., and G. M. Rubin, 1988 Identification and purification of a Drosophila protein that binds to the terminal 31-base-pair inverted repeats of the P transposable element. Proc. Natl. Acad. Sci. USA **85:** 8929–8933.

Robertson, H. M., and W. R. Engels, 1989 Modified *P* elements that mimic the P cytotype in *Drosophila melanogaster*. Genetics **123:** 815–824.

Roche, S. E., M. Schiff and D. C. Rio, 1995 P-element repressor autoregulation involves germ-line transcriptional repression and reduction of third intron splicing. Genes Dev. **9:** 1278–1288.

Roiha, H., G. M. Rubin and K. O'Hare, 1988 *P* element insertions and rearrangements at the *singed* locus of *Drosophila melanogaster*. Genetics **119:** 75–83.

Sakagouchi, K., 1990 Invertrons, a class of structurally and functionally related genetic elements that includes linear plasmids, transposable elements, and genomes of adeno-type viruses. Microbiol. Rev. **54:** 66–74.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Staveley, B. E., T. R. Heslip, R. B. Hodgetts and J. B. Bell, 1994 Protected *P* element termini suggest a role for inverted-repeat binding protein in transposase-induced gap repair in *Drosophila melanogaster*. Genetics **139:**1321–1329.

Tsubota, S., M. Ashburner and P. Schedl, 1985 P-element-induced control of mutations at the *r* gene of *Drosophila melanogaster*. Mol. Cell. Biol. **5:** 2567–2574.

Williams, J. A., 1989 A molecular analysis of the vestigial locus in *Drosophila melanogaster*. Ph.D. Thesis, University of Alberta, Edmonton, Alberta, Canada.

Williams, J. A., and J. B. Bell, 1988 Molecular organization of the *vestigial* region in *Drosophila melanogaster*. EMBO J. **7:** 1355–1363.

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

Williams, J. A., S. S. Pappu and J. B. Bell, 1988b Suppressible P-element alleles of the vestigial locus in *Drosophila melanogaster*. Mol. Gen. Genet. **212:** 370–374.

Williams, J. A., A. L. Atkin and J. B. Bell, 1990 The functional organization of the vestigial locus in *Drosophila melanogaster*. Mol. Gen. Genet. **221:** 8–16.

Williams, J. A., J. B. Bell and S. B. Carroll, 1991 Control of *Drosophila* wing and haltere development by the nuclear *vestigial* product. Genes Dev. **5:** 2481–2495.

Communicating editor: M. J. Simmons